Exploring Model Defects

Using Linear Analysis

A GMDSI worked example report

by Chris Nicol and John Doherty

DRAFT

Groundwater Modelling Decision Support Initiative

PUBLISHED BY

The National Centre for Groundwater Research and Training
C/O Flinders University
GPO Box 2100
Adelaide SA 5001
+61 8 8201 2193

DISCLAIMER

CITATION

For bibliographic purposes this report may be cited as: Nicol, C. and Doherty J. (2020). Exploring Model Defects Using Linear Analysis: A GMDSI worked example report. National Centre for Groundwater Research and Training, Flinders University, South Australia.

ISBN: 978-1-925562-52-1

DOI: 10.25957/ew1j-9g75. DOI as a weblink: https://doi.org/10.25957/ew1j-9g75

COPYRIGHT

Preface

The Groundwater Modelling Decision Support Initiative (GMDSI) is an industry-funded and industry-aligned project focused on improving the role that groundwater modelling plays in supporting environmental management and decision-making. Over the life of the project, it will document a number of examples of decision-support groundwater modelling. These documented worked examples will attempt to demonstrate that by following the scientific method, and by employing modern, computer-based approaches to data assimilation, the uncertainties associated with groundwater model predictions can be both quantified and reduced. With realistic confidence intervals associated with predictions of management interest, the risks associated with different courses of management action can be properly assessed before critical decisions are made.

GMDSI worked example reports, one of which you are now reading, are deliberately different from other modelling reports. They do not describe all of the nuances of a particular study site. They do not provide every construction and deployment detail of a particular model. In fact, they are not written for modelling specialists at all. Instead, a GMDSI worked example report is written with a broader audience in mind. Its intention is to convey concepts, rather than to record details of model construction. In doing so, it attempts to raise its readers' awareness of modelling and data-assimilation possibilities that may prove useful in their own groundwater management contexts.

The decision-support challenges that are addressed by GMDSI worked examples include the following:

- assessing the reliability of a public water supply;
- protection of a groundwater resource from contamination;
- assessment of mine dewatering requirements;
- assessing the environmental impacts of mining; and
- management of aquifers threatened by salt water intrusion.

In all cases the approach is the same. Management-salient model predictions are identified. Ways in which model-based data assimilation can be employed to quantify and reduce the uncertainties associated with these predictions are reported. Model design choices are explained in a way that modellers and non-modellers can understand.

The authors of GMDSI worked example reports make no claim that the modelling work which they document cannot be improved. As all modellers know, time and resources available for modelling are always limited. The quality of data on which a model relies is always suspect. Modelling choices are always subjective, and are sometimes made differently with the benefit of hindsight.

What we do claim, however, is that the modelling work which we report has attempted to implement the scientific method to address challenges that are typical of those encountered on a day-to-day basis in groundwater management worldwide.

As stated above, a worked example report purposefully omits many implementation details of the modelling and data assimilation processes that it describes. Its purpose is to demonstrate what can be done, rather than to explain how it is done. Those who are interested in technical details are referred to GMDSI modelling tutorials. A suite of these tutorials is being developed specifically to assist modellers in implementing workflows such as those that are described herein.

We thank and acknowledge our collaborators, and GMDSI project funders, for making these reports possible.

Dr John Doherty, GMDSI Project, Flinders University and Watermark Numerical Computing.
Dr Phil Hayes, GMDSI Project, University of Queensland.

Glossary

Anisotropy

A condition whereby the properties of a system (such as hydraulic conductivity) are likely to show greater continuity in one direction than in another. At a smaller scale it describes a medium whose properties depend on direction.

Bayesian analysis

Methods that implement history-matching according to Bayes equation. These methods support calculation of the posterior probability distribution of one or many random variables from their prior probability distributions and a so-called “likelihood function” – a function that increases with goodness of model-to-measurement fit.

Boundary condition

The conditions within, or at the edge of, a model domain that allow water or solutes to enter or leave a simulated system.

Boundary conductance

The constant of proportionality that governs the rate of water movement across a model boundary in response to a head gradient imposed across it.

Time-variant specified head (CHD) package

A Dirichlet (i.e. “fixed head”) boundary condition implemented by MODFLOW in which the head can vary with time on a stress-period-by-stress-period basis.

Covariance matrix

A matrix is a two-dimensional array of numbers. A covariance matrix is a matrix that specifies the statistical properties of a collection of random variables - that is, the statistical properties of a random vector. The diagonal elements of a covariance matrix record the variances (i.e. squares of standard deviations) of individual variables. Off-diagonal matrix elements record covariances between pairs of variables. The term “covariance” refers to the degree of statistical inter-relatedness between a pair of random variables.

Ensemble

A collection of realisations of random parameters.

Drain (DRN) package

A one-way Cauchy boundary condition implemented by MODFLOW. Water can flow out of a model domain, but cannot enter a model domain through a DRN boundary condition.

Evapotranspiration (EVT) package

MODFLOW’s implementation of water withdrawal from a groundwater system whereby the extraction rate can increase, up to a user-supplied maximum, as the head approaches a user-prescribed level from below.

General head boundary (GHB) package

This is MODFLOW parlance for a Cauchy boundary condition. Water flows into or out of a model domain in proportion to the difference between the head ascribed to the boundary and that calculated for neighbouring cells. The rate of water movement through the boundary in response to this head differential is governed by the conductance assigned to the boundary.

Hydraulic conductivity

The greater is the hydraulic conductivity of a porous medium, the greater is the amount of water that can flow through that medium in response to a head gradient.

Jacobian matrix

A matrix of partial derivatives (i.e. sensitivities) of model outputs (generally those that are matched with field measurements) with respect to model parameters.

Matrix

A two-dimensional array of numbers index by row and column.

MODFLOW

A family of public-domain, finite-difference groundwater models developed by the United States Geological Survey (USGS).

MODFLOW package

An item of simulation functionality that describes one aspect of the operation of a groundwater system, for example recharge or a boundary condition. The word “package” describes the computer code that implements this functionality, as well as its input and output file protocols.

Null space

In the parameter estimation context, this refers to combinations of parameters that have no effect on model outputs that are matched to field observations. These combinations of parameters are thus inestimable through the history-matching process.

Objective function

A measure of model-to-measurement misfit whose value is lowered as the fit between model outputs and field measurements improves. In many parameter estimation contexts the objective function is calculated as the sum of squared weighted residuals.

Parameter

In its most general sense, this is any model input that is adjusted in order to promulgate a better fit between model outputs and corresponding field measurements. Often, but not always, these inputs represent physical or chemical properties of the system that a model simulates. However there is no reason why they cannot also represent water or contaminant source strengths and locations.

Phreatic surface

The water table.

Pilot point

A type of spatial parameterisation device. A modeller, or a model-driver package such as PEST or PEST++, assigns values to a set of points which are distributed in two- or three-dimensional space. A model pre-processor then undertakes spatial interpolation from these points to cells comprising the model grid or mesh. This allows parameter estimation software to ascribe hydraulic property values to a model on a pilot-point-by-pilot-point basis, while a model can accept these values on a model-cell-by-model-cell basis. The number of pilot points used to parameterize a model is generally far fewer than the number of model cells.

Prior probability

The pre-history-matching probability distribution of random variables (model parameters in the present context). Prior probability distributions are informed by expert knowledge, as well as by data gathered during site characterisation.

Posterior probability

The post-history-matching probability distribution of random variables (model parameters in the present context). These probability distributions are informed by expert knowledge, site characterisation studies, and measurements of the historical behaviour of a system.

Probability density function

A function that describes how likely it is that a random variable adopts different ranges of values.

Probability distribution

This term is often used interchangeably with “probability density function”.

Quadtree mesh refinement

This term refers to a means of creating fine rectilinear model cells from coarse rectilinear model cells by dividing them into four. Each of the subdivided cells can then be further subdivided into another four cells. However it is a design specification of a quadtree-refined grid that no cell within the domain of a model be connected to more than two neighbouring cells along any one of its edges.

Realisation

A random set of parameters.

Regularisation

The means through which a unique solution is sought to an ill-posed inverse problem. Regularisation methodologies fall into three broad categories, namely manual, Tikhonov and singular value decomposition.

Residual

The difference between a model output and a corresponding field measurement.

Singular value decomposition (SVD)

A matrix operation that creates orthogonal sets of vectors that span the input and output spaces of a matrix. When undertaken on a Jacobian matrix, SVD can subdivide parameter space into complementary, orthogonal subspaces; these are often referred to as the solution and null subspaces. Each of these subspaces is spanned by a set of orthogonal vectors. The null space of a Jacobian matrix is composed of combinations of parameters that have no effect on model outputs that are used in its calibration, and hence are inestimable.

Solution space

The orthogonal complement of the null space. This is defined by undertaking singular value decomposition on a Jacobian matrix.

Specific storage

The amount of water that is stored elastically in a cubic metre a porous medium when the head of water in which that medium is immersed rises by 1 metre.

Specific yield

The amount of accessible water that is stored in the pores of a porous medium per volume of that medium.

Stochastic

A stochastic variable is a random variable.

Stress

This term generally refers to those aspects of a groundwater model that cause water to move. They generally pertain to boundary conditions. User-specified heads along one side of a model domain, extraction from a well, and pervasive groundwater recharge, are all examples of groundwater stresses.

Stress period

The MODFLOW family of models employs this terminology to describe each member of a series of contiguous time intervals that collectively comprise the simulation time of a model.

Tikhonov regularisation

An ill-posed inverse problem achieves uniqueness by finding the set of parameters that departs least from a user-specified parameter condition, often one of parameter equality and hence spatial homogeneity.

Vector

A collection of numbers arranged in a column and indexed by their position in the column.

Executive Summary

Groundwater flows through an unseen, heterogeneous medium. Sources and locations of its recharge and discharge are often only vaguely known. Numerical simulation of groundwater movement and processes is therefore inexact. Nevertheless, if properly implemented, numerical simulation can provide a uniquely powerful mechanism for supporting the management of groundwater systems.

A numerical model cannot predict the future behaviour of groundwater under an existing or altered management regime. However it can bracket a prediction of this behaviour within uncertainty limits that reflect all currently available knowledge of the groundwater system. This knowledge includes that which is available through site characterisation and hydrogeological mapping, as well as that which is available from present and historical measurements of system behaviour. To achieve its decision-support potential, the modelling process must be capable of assimilating these two different kinds of information in order to reduce the uncertainties of management-critical predictions. It must also be capable of informing managers of the uncertainties that remain after assimilation of this information has taken place.

These decision-support imperatives create conflicting demands on groundwater model design. On the one hand, a model must possess sufficient complexity to represent system processes and properties to which predictions of management interest may be sensitive. Furthermore, because these processes and properties are often only vaguely known, they must be represented stochastically. On the other hand, a groundwater model should run quickly and be numerically stable. This is because fulfillment of its data assimilation and uncertainty analysis responsibilities requires that the model be run many times under the control of software that facilitates flow of information from field measurements and expert knowledge to the model, and then from the model to decision-critical predictions. It follows that model design must eschew redundant complexity in order to facilitate passage of information into and out of the model. At the same time it must ensure that complexity retained by the model is sufficient to provide receptacles for outside information, and to quantify uncertainties that arise from a deficit of such information.

Design of a decision-support groundwater model is therefore a compromise. While the model itself may be “physically based”, some boundary conditions may embody abstract representations of complex processes and properties whose details are purposefully omitted from the model in order to reduce its run time and enhance its numerical health. However a modeller, as well as those who are affected by decisions that are supported by a model, must be sure that strategic model simplification undertaken in this way does not impair a model’s decision-support integrity.

Inappropriate simplicity can impair a model’s decision-support potential in three ways.

1. If a model is too simple, it cannot replicate the historical behaviour of the system that it attempts to simulate. It is therefore incapable of assimilating information that is resident in this behaviour.
2. Removal from a model of properties and processes to which a prediction is sensitive may reduce its capacity to associate a valid uncertainty interval with that prediction. This, in turn, impairs the ability of the modelling process to assess the risks associated with contemplated courses of management action.
3. Model simplifications may bias certain decision-critical model predictions. Paradoxically, potential errors in these predictions that are incurred by model

simplicity may be exacerbated by attempts to lower predictive uncertainty through history-matching.

The last two of these unwanted outcomes of model simplification may not be easy to detect, both during calibration of the model, and during subsequent deployment of the model to make predictions. However if the action of the model on its parameters is represented, through model linearisation, as the action of a matrix on a vector, all three of these potential impacts can be investigated.

Linear analysis requires firstly that model constructs that constitute potential sources of predictive error (for example questionable boundary conditions) be endowed with adjustable parameters that characterize possible errors in their representation of unsimulated processes. Next, the sensitivities of model outcomes to these, and other, model parameters must be computed. This normally requires that the model be run many times; during each of these runs, one of its parameters is varied incrementally. Finally, calculations embodied in matrix equations are made to assess all three of the above potential impacts of inappropriate model simplification.

This process is illustrated using a model that was built to evaluate ongoing extraction from two wellfields that are situated near the south western margin of the Great Artesian Basin. This report discusses methodologies and public domain software that were used in this investigation, and provides a few details of how this study was done.

The study focusses on a number of boundary conditions that were assigned to this model. Linear analysis establishes that the potential for bias instilled in management-salient predictions by simplifications embodied in these boundary conditions is small in comparison with the uncertainties of these predictions. It is further established that the model's ability to quantify these uncertainties is virtually unimpaired by present boundary condition design.

The study that is documented herein also demonstrates the importance of endowing model features such as these boundaries with flexible, adjustable parameters that allow potential errors in their design to contribute to the assessed uncertainties of decision-critical predictions, while minimizing their potential to bias them. This principle is embodied in the design of a model which replaces that which is the subject of the present report.

CONTENTS

1. INTRODUCTION

1.1 General

It was George Box, the famous British statistician, who first coined the often-repeated phrase that “all models are wrong but some are useful”. While undoubtedly true in the groundwater modelling context, this statement is not in itself very useful, as it does not define “useful”. Nor does it provide a basis for evaluating the utility (or otherwise) of a particular model – a matter which is often the subject of heated debate by competing experts.

Another common phrase – a phrase that adorns the executive summaries of many, if not most, modelling reports - is that a model should be viewed as “fit for purpose”. However the words “fit” and “purpose” are seldom, if ever, defined. Presumably this phrase implies that use of the model in support of groundwater management will not precipitate the making of a wrong decision. However rarely is model fitness expressed in such direct terms. Even more rarely is justification for its fitness provided.

In this GMDSI worked example report, we demonstrate how linear analysis can be used to explore whether a groundwater model can indeed be useful while being wrong, and under what circumstances it can actually be described as being “fit for purpose”. However, before doing this, we explore the metrics on which these descriptions must rest.

1.2 Conceptual and Numerical Models

A numerical model comprises an abstract, simplified representation of a complex, natural reality. In doing so, it gives numerical voice to an underlying conceptual model.

A conceptual model is a cognitive device for identifying what is important (and implicitly what is unimportant) to the operation of a natural system. It supports dissection of a complex system into a finite number of parts which can be individually understood. Once these parts are identified, calculations can focus on how they interact.

In the groundwater management context, the basis of a conceptual model often includes at least some of the following premises.

- The importance of system boundary specifications diminishes with distance from the focus of management interest.
- The direction and amount of groundwater flow is dominated by certain recharge processes and certain discharge processes.
- The interaction of groundwater with processes that prevail at the land surface can be characterized by relatively simple equations. Though approximate, errors arising from their use are small.
- The medium through which groundwater flows can be subdivided into a number of discrete geological units. These form a basis for characterizing the hydraulic properties of the system.
- An adequate description of vertical groundwater movement rests on subdivision of the groundwater domain into a discrete number of layers. These layers are often based on stratigraphy.

The details and nuances of natural processes are, of course, far more complex than can be expressed by a conceptual model, and are far too detailed for replication by a numerical model. However, this does not erode the benefits of numerical simulation, especially if a model is employed to explore the large scale repercussions of system management rather than its details. Just as importantly, numerical simulation can provide a unique basis for assimilation of information that is encapsulated in historical measurements of groundwater status and behaviour. This information often pertains to the scale at which management-salient predictions must be made – a scale that eclipses nuances of system behaviour that are unimportant for system management. The uncertainties of decision-critical model predictions can often be reduced through assimilation of this information. For a model to assimilate the information that is contained in historical measurements of system states and fluxes, it must be capable of reproducing them, at least approximately. This, as much as anything else, defines the importance of numerical simulation for environmental decision-support.

So how much do the simplifications embodied in construction of a conceptual model matter? And how much do the approximations encapsulated in numerical simulation matter? They matter if they erode a model's ability to support the making of decisions that management of a system requires.

1.3 System Management

Ideally, management of a groundwater system should promulgate beneficial use of that system while forestalling the occurrence of unwanted events, particularly those that are suffered by environmental assets. Conceptually, by simulating a system numerically, the efficacy of a particular management strategy in achieving these outcomes can be tested before it is actually implemented.

Unfortunately, groundwater models cannot simulate groundwater behaviour very well. Any prediction that is made by a groundwater model is accompanied by uncertainty. Depending on the prediction, this uncertainty may be considerable. Much of this uncertainty arises from the fact that the hydraulic properties of media through which groundwater flows are heterogeneous and variable (sometimes over orders of magnitude). The magnitudes of system drivers such as recharge are also often only poorly known. So too are facets of the conceptual model on which a numerical model is based.

Predictive uncertainty does not erode a model's ability to support environmental management unless the model is used in a way that ignores this uncertainty. Nor should predictive uncertainty be construed as a model's fault; it is an outcome of information inadequacy. It is not a model's task to overcome this inadequacy. In contrast, it is a model's task to quantify the repercussions of this inadequacy as it pertains to predictions on which system management rests. Management of the system can then account for the risks that attend information inadequacy, and incorporate precautions that accommodate these risks.

With this in mind, the goals of decision-support modelling can now be defined. These are to:

- make predictions of management interest;
- quantify the uncertainties associated with these predictions;
- reduce these uncertainties through assimilation of pertinent data.

These, then, are the "purposes" for which a groundwater model must be "fit". These must be reflected in the specifications of a model whose predictions are necessarily "wrong" while the model itself remains "useful".

1.4 Seeking Purpose

A model is able to quantify the uncertainty of a decision-critical prediction only if it represents characteristics and properties of a natural system in ways that embody their partially known status. When it is used to make a prediction, it may need to make that prediction many times, each with a different realisation of these characteristics and properties. To make matters more complicated, these characteristics and properties must be constrained such that they allow the model to replicate historical system behaviour while still reflecting their partially-known status. The repetitive making of management-salient predictions, and the imposition of constraints on partially-known model characteristics and properties, are model-run-intensive procedures. They require that a model runs quicky; they also requires that it be numerically stable. Both of these qualities are abetted by simplicity of model design.

At the same time, a modeller must include in his/her model, in either an explicit or abstract way, all aspects of a system which he/she considers may contribute to the uncertainty of a decision-critical prediction. Integrity of uncertainty assessment is therefore abetted by complexity of model design. This is especially the case if a model is required to make many different predictions.

Because of these two conflicting requirements, the design of a decision-support numerical model must be a compromise. A model cannot serve the imperatives of decision-support if it is too simple to quantify the uncertainties of decision-critical predictions. Nor can it serve these imperatives if it is too complex to assimilate data which can reduce these uncertainties.

A numerical model can possess many moving parts. A modeller is called upon to make many design decisions that pertain to these parts. He/she must decide which parts to include, which parts to exclude, and which parts to represent in a simplistic or abstract manner. These decisions are subjective. Some are debatable. All of them must be defensible.

Defence of any aspect of model design must ultimately address its ability to support the goals of decision-support modelling that are set out above. These pertain to quantification and reduction of the uncertainties of decision-critical predictions. Model design justification therefore requires methodologies for rapid, approximate quantification of predictive uncertainty, for assessment of contributions made to this uncertainty by different parts of a model, and for identifying, and then avoiding, design simplifications that may introduce unquantifiable bias to decision-critical model predictions.

1.5 The Present Worked Example

The GMDSI worked example which is the subject of the present report shows how linear analysis can be used to explore the integrity of model design.

The model on which this report is based was commissioned by BHP. It was built in 2015 to evaluate changes to hydrogeological conditions along the south western margin of the Great Artesian Basin (GAB) caused by groundwater extraction near this margin. Water is extracted from GAB aquifers by BHP to satisfy the requirements of its Olympic Dam mine. The model was updated in 2019 in order to provide a superior numerical basis for re-assessing water supply security and potential risks to springs which are sustained by artesian water. The improved model retains much of the conceptual basis of the existing model. However, its parameterisation was revised in order to enhance its ability to assimilate newly acquired information. At the same time, its ability to quantify post-data-assimilation uncertainties of management-critical predictions, particularly those pertaining to pumping-induced depressurisation beneath springs, was enhanced.

Prior to model enhancement, some effort was devoted to examining whether reliance should continue to be placed on certain aspects of the conceptual model. These aspects simplified model construction. As such, they increased the amenability of the existing model to parameter estimation and uncertainty quantification. However, BHP wished to ensure that these simplifications did not compromise the integrity of important model predictions. Of particular interest were specifications of model boundaries.

This short report is organized as follows. Chapter 2 is devoted to outlining the concepts on which linear analysis rests. Chapter 3 describes the model which forms the subject matter of this report – that is, the model on which linear analysis was undertaken. Chapter 4 documents some of the outcomes of that analysis. Chapter 5 finishes the report with some concluding remarks.

2. LINEAR ANALYSIS

2.1 Parameters and Stochasticity

The task of decision-support modelling is to quantify and reduce the uncertainties of predictions of management interest.

Quantification of the uncertainty of a management-salient prediction requires that aspects of a system that contribute to this uncertainty be represented in a model in ways that reflect their unknown status. That is, they should be represented as stochastic quantities that can assume any value within a range that is decreed to be reasonable based on so-called “prior knowledge”. According to this knowledge, some of these values may be more likely than others. This aspect of prior knowledge is respected by awarding prior probability distributions to parameters. The purpose of these distributions is to assign greater likelihood to some values over others.

We denote a model feature that can assume a range of possible values as a model “parameter”. Parameters include properties of the subsurface such as hydraulic conductivity and specific storage. Notionally, these properties can actually be measured at discrete points within a model domain. If enough measurements are made, an empirical prior probability distribution can be assembled. As a reflection of prior knowledge, this probability distribution should include the notion that hydraulic properties at nearby points are more likely to be similar to each other than properties at points that are far apart. This tendency for closeness-based hydraulic property similarity is often referred to as “spatial correlation”.

A model’s parameters may also include system drivers such as recharge. They may even include historical pumping rates if these are only approximately known. More abstract model specifications, such as the values ascribed to boundary conductances, may also be denoted as parameters. The conductance ascribed to all or part of a model boundary determines how easily water can move in or out of the modelled system through that boundary to a broader groundwater system, or to the land surface. Conductances can rarely be directly measured in the field. However their values may be inferable from system behaviour. Their representation as model parameters recognizes that these inferences can only be approximate.

Generally a groundwater model is endowed with many parameters – hundreds, thousands or even tens of thousands of parameters – all describing different aspects of a system whose behaviour we wish to manage. Their representation as stochastic (i.e. probabilistic) quantities embodies the fundamental truth that our knowledge of the system that we wish to manage is far from complete. Acceptance of this truth is a prerequisite for decision-support groundwater modelling. The modelling process thereby strives to represent that which is known, while quantifying the repercussions for management risk of that which cannot be known.

The inclusion of thousands, or even tens of thousands, of parameters in a model may appear to be unnecessarily cumbersome. Indeed, there are numerical and cognitive costs associated with the use of so many parameters. Furthermore this number of parameters is far greater than that which can be estimated uniquely through history-matching (see below). Nevertheless, the following should be born in mind.

- As Freeze et al (1990), Doherty and Simmons (2013) and other authors make clear, a model best serves the decision-making process by associating uncertainties with decision-critical model predictions.
- For this to occur, it is just as important for a model to include parameters whose values cannot be uniquely estimated as it is to include those whose values can be

uniquely estimated, for it is the former that may contribute most to the uncertainties of decision-critical predictions.
- Modern methods of history-matching and uncertainty analysis, such as those provided by the PEST and PEST++ suites, can readily accommodate large numbers of parameters.

2.2 History-Matching

2.2.1 Data Assimilation

By definition, stochastic variables such as model parameters cannot be assigned a single value. Instead, they can only be populated by using "realisations" of their values. These are samples drawn from their joint probability distribution. The word "joint" when applied to parameters implies that not all parameters are statistically independent of each other. Some parameter types show spatial correlation. Hence, while the value of a parameter type at one location is not completely determined by its value at another, the former is nevertheless influenced or constrained by the latter. Therefore patterns of heterogeneity that emerge in any realisation of parameter values drawn from their joint probability distribution will reflect their spatial correlation structure.

Suppose that a model is populated by a single realisation of parameter values drawn from the prior parameter probability distribution; this is the probability distribution that emerges from expert knowledge and site characterisation. Suppose further that the model is then run over an historical period of time over which measurements of system state (such as heads or spring flows) were made. Pertinent model outputs can be compared with these measurements. This process can be repeated for other prior parameter realisations. Because model parameters are stochastic quantities, so too are model outputs. It will generally be found that model-calculated system states for any realisation differ widely from those that were actually measured, and that collectively, over all realisations, they span a far greater range than measured values.

Suppose that we now insist that the values assigned to model parameters must be such that its outputs are reasonably close in value to corresponding field measurements. The metric for "reasonably close" is governed by:

- the credibility of the measurements themselves (i.e. the extent to which these measurements are contaminated by possible errors); and
- the extent to which a numerical model can be expected to reproduce measured nuances of system behaviour.

Normally, for the sake of simplicity, modellers group these two sources of potential model-to-measurement misfit into a single stochastic term which they denote as "measurement noise", even though the latter source of error is more correctly described as "structural noise". The act of tuning model parameters so that a model can reproduce past system behaviour is known as "history-matching".

If the conceptual basis for a numerical model is sound, then it should be possible to find a set of parameters which allows the model to replicate (within limits set by measurement noise) the historical behaviour of the simulated system at the discrete number of points at which this behaviour was observed. In fact, it is generally possible to find an uncountable number of realistic parameter sets which achieve this match. A "realistic" parameter set is one for which parameter values are such that it could have been sampled from the prior parameter probability distribution. However, our insistence that parameters with which we endow a model for the purpose of predicting the future must also allow it to respect what we know about the past may significantly reduce the range of parameter options that are available to us through their prior probability distribution. This reduction of parameter uncertainty that is accrued through history-matching is often referred to as "data assimilation".

History-matching-induced alterations to the prior parameter probability distribution affect the uncertainties ascribed to model predictions. Depending on the contents of the history-matching dataset, the uncertainties of some predictions may be reduced through history-matching more than those of others. To the extent that predictions of management interest are sensitive to parameters whose uncertainties are significantly reduced through data assimilation, their uncertainties are correspondingly reduced. This is schematised in Figure 2.1.

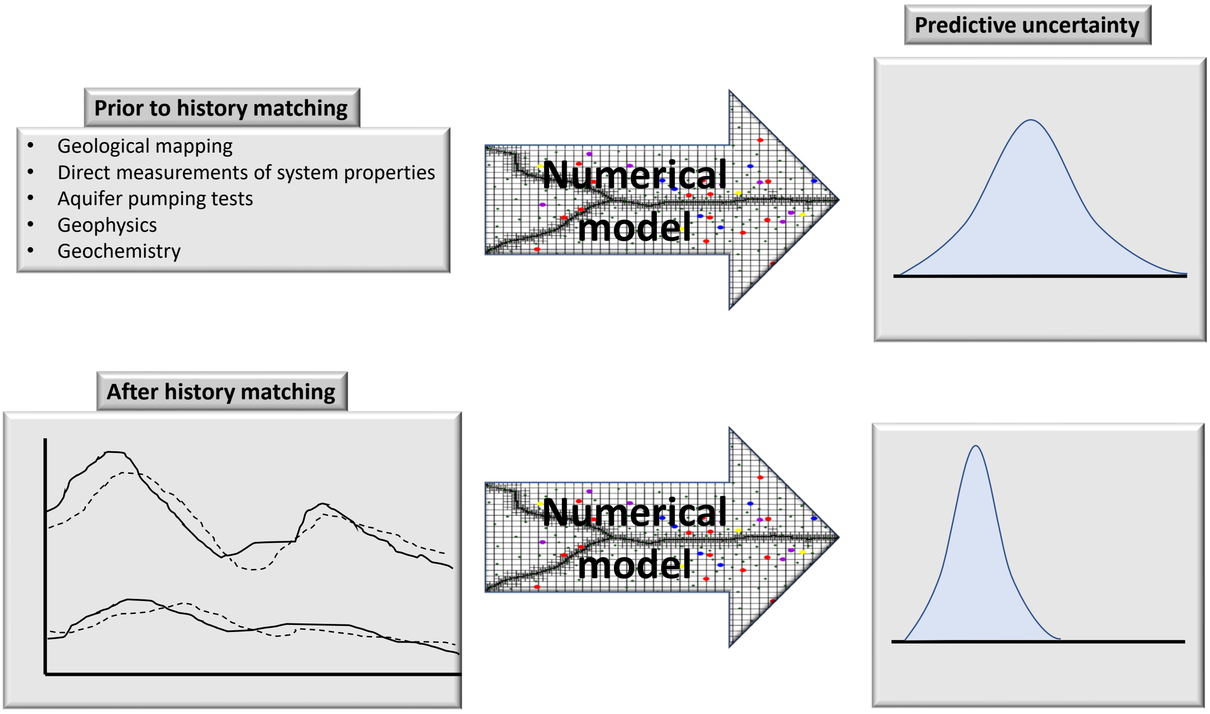


Figure 2.1 An ability to replicate the past improves a model's ability to forecast the future.

2.2.2 Bayes Equation

The imposition of constraints on parameter values by history-matching is described by Bayes equation. This equation has inspired a field of mathematical and numerical endeavours whose purpose is directed toward one end - characterizing the so-called "posterior probability distribution" of parameters. This is the name given to the stochastic description of joint parameter uncertainty that respects the constraints imposed on parameters born of our insistence that model outputs replicate measurements of system behaviour to within limits set by measurement noise.

This terminology can be extended to predictions of management interest made by a model. The prior probability distribution of a model prediction describes the range of predictive possibilities that emerge from the prior probability distribution of model parameters. The posterior probability distribution of a prediction is often (but not always) narrower than its prior probability distribution. This reflects the fact that any set of parameters that is used to make that prediction must earn its place in the posterior parameter probability distribution by allowing the model to replicate field measurements of system behaviour.

Unfortunately, while the above concepts are easy to describe (and are somewhat obvious), they are difficult to implement. Implementation problems include the following.

- It is difficult to characterize the prior probability distribution of parameters that pertain to geological media whose origins and dispositions are the outcomes of somewhat chaotic natural processes.
- It is difficult to characterize the stochasticity of “measurement noise” that includes model-to-measurement misfit arising from the many simplifications and abstractions that beset any numerical model.
- It is difficult to calculate the stochasticity of model predictions from that of model parameters when the relationship between the two is so complex that it must be represented by a numerical model.

It follows that mathematical expressions for posterior parameter and predictive probability distributions cannot be derived. Hence these distributions must be defined by sampling them. Software packages such as the PESTPP-IES ensemble smoother (White; 2018) are designed to achieve this outcome. First PESTPP-IES generates realisations of parameter values by sampling an approximation of the prior parameter probability distribution. It then attempts to adjust each of these realisations by the minimum amount required for model outputs to match field observations. Once adjusted, each such realisation constitutes a sample from the posterior parameter probability distribution. If enough parameter sets are subjected to this process, and if a model prediction of management interest is then made using all of these parameter sets, model-calculated values of the prediction collectively define its posterior probability distribution.

Sampling of the posterior parameter and predictive probability distributions in this way can be a numerically demanding process. However if certain approximations are made, the prior and posterior probability distributions of parameters, and of predictions which depend on them, can be characterized using mathematical expressions that can be evaluated rather easily. This is useful in itself. However, even more enticing are the possibilities that this offers for undertaking other informative, uncertainty-related calculations. These include evaluation of data worth, discovery of the principle sources of prior and posterior predictive uncertainty, and assessment of model predictive integrity.

2.3 Linear Analysis

The theory and benefits of linear analysis, when undertaken in concert with groundwater modelling, have been discussed by authors such as Moore and Doherty (2006), James et al (2009), Dausman et al (2010), and White et al (2016). It can be implemented using utility software supplied through the PEST, PEST++ and PyEMU suites.

2.3.1 Assumptions

Linear uncertainty analysis is also known as “first order second moment” (or “FOSM”) analysis. It provides approximate mathematical characterisation of prior predictive probability distributions, and of posterior parameter and predictive probability distributions. It can be used to demonstrate how the history-matching process bestows worth on data. It can also be deployed to track the flow of information from field measurements of system state to parameters, and ultimately from parameters to model predictions. It does this by implementing Bayes equation under the following assumptions.

- The prior probability distribution of parameters is multiGaussian.
- “Measurement noise” (including structural noise) is also characterized by a Gaussian distribution.
- The relationships between model outputs that correspond to measurements of system state and parameters employed by a model can be approximated by the action of a matrix on a vector.

- Model outputs that correspond to predictions of management interest can be calculated using another matrix that acts on model parameters.

Let us briefly examine each of these assumptions in turn.

2.3.2 MultiGaussian Distribution

A Gaussian distribution is a normal distribution. This can be characterized with mathematical ease. It is the simple "bell shaped" probability distribution with which we are all familiar.

If a single random variable has a normal probability distribution, then its stochastic properties can be completely described by its mean and standard deviation. The former denotes its central value, while the latter denotes its propensity to vary about this central value; for a Gaussian distribution, this propensity is symmetrical.

If *m* random variables (for example a parameter set with *m* members) have a multiGaussian distribution, then each of them can be awarded a mean and a standard deviation. However this does not complete the characterisation of their joint probability distribution; something must be said about whether the value of one variable is affected by that of another, that is if one variable tends to be above its mean or below its mean if the other is above its mean or below its mean. If two variables are not statistically independent, then they are said to possess a covariance (or to be correlated). To complete the stochastic characterisation of a set of random variables, we must depict the amount of covariance that exists between each pair of random variables that collectively comprise the set.

Covariances between pairs of random variables can be collected into a matrix; unsurprisingly, this matrix is referred to as a "covariance matrix". If there are *m* random variables in a set, then the covariance matrix that describes this set has *m* rows and *m* columns. Because a covariance matrix is symmetric, half of its elements are duplicated. Within this matrix, off-diagonal elements describe covariances between pairs of individual variables. Meanwhile, the diagonal elements of this matrix (of which there are *m*) denote the variances of individual variables; variance is the square of standard deviation.

Generally, the covariance matrix of measurement noise is assumed to be diagonal. That is, the values of all of its off-diagonal elements are assumed to be zero. This means that the "noise" ascribed to any one measurement (which, as stated above, quantifies a model's license to eschew a perfect fit with that measurement) is independent of its licence to eschew a perfect fit with any other measurement. In reality, this is not actually an accurate description of the proclivity of models to misfit real world measurements. However in most groundwater modelling circumstances, errors in the uncertainties attributed to decision-critical predictions incurred by this assumption are relatively small.

In contrast, a covariance matrix that describes prior parameter uncertainties is rarely diagonal. Because the hydraulic properties of the heterogeneous material through which groundwater flows are likely to show spatial correlation, the prior parameter covariance matrix must be endowed with off-diagonal elements which reflect the distance over which parameter value similarity is likely to prevail. Representation of earth property heterogeneity using concepts that are based on spatial correlation is actually an extreme simplification of the complex patterns of heterogeneity that characterize real geological media. Nevertheless, it is adequate for linear analysis, the purpose of which is to obtain approximate estimates of parameter and predictive uncertainty, and to calculate value-added quantities such as data worth and parameter contributions to predictive uncertainty.

2.3.3 Model as a Matrix

The matrix that replaces a model when conducting linear analysis is a matrix of sensitivities. When the model is deployed to reproduce field measurements, the elements of this matrix are the sensitivities of corresponding model outputs to all of the model's parameters. This matrix must therefore possess *n* rows and *m* columns, where *n* is the number of field measurements and *m* is the number of model parameters. The sensitivity matrix which describes the action of the model under historical conditions is sometimes referred to as the "Jacobian matrix".

Where a model is deployed to make predictions of management interest, the matrix which replaces the model contains sensitivities of these predictions to all model parameters. Where only a single prediction is of interest, this matrix becomes a vector, for it possesses only one row; however it retains *m* columns.

Filling of historical and predictive sensitivity matrices is the most expensive part of linear analysis. Generally it requires that the model be run *m* times under each of these conditions, with a single parameter varied incrementally on each occasion. Differences in model outputs are divided by the difference in the varied parameter in order to calculate their sensitivities with respect to this parameter. Obviously, filling of historical and predictive sensitivity matrices can be very time-consuming where parameter numbers are large and where model run times are long.

2.4 Some Outcomes of Linear Analysis

2.4.1 General

As has been described, linear analysis relies on matrices – covariance matrices and sensitivity matrices. Using equations that are built from these matrices, quantities that support and illuminate the role of data assimilation in decision-support modelling can be readily calculated. We now briefly discuss some of these.

2.4.2 Parameter Estimation

Under the assumptions that underpin linear analysis, Bayes equation is separated into two independent matrix equations. One of these equations evaluates the posterior mean of all parameters. These are linear approximations to values that emerge from the process of "model calibration". As is discussed in texts such as Doherty (2015), calibration is a form of history-matching that seeks a unique set of parameter values. However "uniqueness" does not mean "correctness". Pseudo-uniqueness of a calibrated parameter field is achieved through pursuit of parameter values which lie at the centres of their respective posterior probability distributions.

Equations that are very similar to the linearized form of Bayes equation form the numerical engines of programs from the PEST and PEST++ suites. Of course, these equations are only approximate because groundwater processes are nonlinear with respect to parameters. Hence estimates of posterior parameter means obtained using these equations must be iteratively refined by these software packages through re-calculation of Jacobian matrices (or approximations thereto) as these estimates are improved.

2.4.3 Parameter Uncertainty

The second matrix relationship that emerges from linearisation of Bayes equation calculates the posterior parameter covariance matrix. Posterior parameter correlations are readily obtained from the off-diagonal elements of this matrix. Because history-matching rarely, if ever, promulgates parameter uniqueness (except when it is purposefully, and artificially, sought through model calibration), many subsets of parameters exhibit high amounts of posterior correlation.

Diagonal elements of the posterior parameter covariance matrix denote post-history-matching parameter variances. (Recall that variance is the square of standard deviation.) Similarly, prior parameter variances form the diagonal elements of the user-supplied prior covariance matrix. Armed with both of these matrices, a modeller can readily compare prior and posterior parameter uncertainties. If desired, this ratio can be mapped. A modeller, or modelling stakeholder, can thus see at a glance where, in a model domain, parameter uncertainties have been reduced through history-matching and where they have not been reduced.

2.4.4 Predictive Uncertainty

Simple matrix equations allow evaluation of the prior and posterior uncertainties of any prediction from prior and posterior parameter covariance matrices. Use of these equations requires that the sensitivity of a prediction to all model parameters be known. As stated above, these can be obtained through repeated predictive model runs based on finite parameter differences. Once these sensitivities are available, the capacity of the history-matching process to reduce the uncertainty of a prediction can be readily assessed.

2.4.5 Data Worth

The worth of field measurements rises in proportion to their ability to reduce the uncertainties of decision-critical predictions. As is discussed above, pre- and post-history-matching predictive uncertainties are readily evaluated using matrix equations that linearize Bayes equation. Once these equations have been formulated, it is a simple matter to include or exclude subsets of a measurement dataset in order to test the effects of their inclusion or exclusion on the uncertainty of a prediction of interest. Their ability to reduce (or not) the uncertainty of that prediction can thereby be assessed. So too can the uniqueness of the information that is resident in different subsets of an entire measurement dataset. If the omission of a data subset from the complete history-matching dataset precipitates a significant rise in the uncertainty of a prediction of interest, this indicates that the information that it contains is unique to that subset.

A remarkable characteristic of the matrix equations that are used to calculate both prior and posterior predictive uncertainties is that these equations do not feature the values of parameters, the values of field measurements, nor the values of the predictions themselves. They feature only sensitivities of pertinent model outputs to parameters employed by the model. This has an extremely useful consequence. It means that the worth of data can be assessed before these data are actually gathered. Optimized strategies for future data acquisition can therefore be developed using linear analysis.

2.4.6 Contributions to Uncertainty

A further use of linear analysis (that to which the present report is partly devoted) is that of assessing the importance of different model specifications to the task of making particular predictions of management interest. These assessments can guide a modeller toward implementation of model design and simplification strategies that facilitate use of a model in nonlinear data assimilation and predictive uncertainty quantification. These assessments can also resolve arguments about whether a certain aspect of model design requires correction or improvement before a model is deployed in a particular decision support context.

In analyses of this type, parameters can be used as surrogates for different facets of a model's design. For example, the parameters associated with a certain boundary condition can act as surrogates for the boundary condition itself, and/or for real-world hydraulic processes that the boundary condition replaces or simplifies. In order to assess the integrity of simplifications that are encapsulated in the boundary, a modeller can first assume that its parameters are free of error. The posterior uncertainties of predictions of management interest can then be calculated using matrix equations discussed herein. The modeller can then repeat this calculation after assigning to these boundary parameters prior uncertainties that reflect the potential for error that is incurred by use of these boundaries in place of a more complex reality. If the posterior uncertainties of decision-critical model predictions increase by amounts that are small compared to their overall uncertainties, it is thereby established that the simplification strategy embodied in the boundary condition has integrity for the decision-support role that the model is intended to play. If the model's speed of execution and numerical stability are well served by this simplification, this may enhance use of the model in nonlinear parameter estimation and Bayesian analysis. Meanwhile linear analysis has demonstrated that the benefits of simplification outweigh its costs.

2.4.7 Singular Value Decomposition

If the Jacobian matrix (i.e. the matrix of sensitivities of model outputs used in the calibration process to parameters employed by a model) is subjected to singular value decomposition, parameter space can be subdivided into two orthogonal subspaces. One of these is referred to as the "calibration null space", while its orthogonal complement is referred to as the "calibration solution space". The null space is comprised of combinations of parameters that are completely uninformed by the history-matching dataset. The individual components of each of these parameter combinations are thus completely correlated with each other. Therefore, if they are varied in ratios that are defined by these combinations, they have no effect on model outputs that correspond to field measurements.

In contrast, each of the linear combinations of parameters that collectively comprise the calibration solution space is uniquely estimable on the basis of the calibration dataset. (PEST jargon sometimes refers to each such parameter combination as a "super parameter".) This does not mean that this combination of parameters can be estimated without uncertainty. It means that its uncertainty is an outcome of random errors associated with measurements that comprise the calibration dataset, and not of an information deficit in the calibration dataset.

In most groundwater history-matching contexts, the dimensionality of the null space is far greater than that of the solution space. The dimensionality of the latter space can be viewed as the number of individual pieces of information that are accessible to model parameters through the history-matching process. Each such piece of information supports unique estimation of one super-parameter.

2.5 Traditional Sensitivity Analysis

We conclude this section with a few words on sensitivity analysis in general.

The requirement that sensitivity analysis accompany model construction and deployment is deeply engrained in groundwater modelling culture. However the purpose, and desired outcomes, of sensitivity analysis are rarely stated in modelling proposal requests and responses. Similarly, repercussions of the outcomes of sensitivity analysis for model usage in decision support are rarely addressed in reports that accompany delivery of a model.

The term "sensitivity analysis" means different things to different people. So-called "global sensitivity analysis" is a field of mathematical endeavour on which books and papers have been written; see, for example, Pianosi et al (2020) and Saltelli et al (2004; 2007). Meanwhile, in the groundwater modelling context, "local sensitivity analysis" has been recommended as a means of separating parameters whose values can be estimated through history-matching from those whose values cannot. Texts such as Hill and Tiedemann (2005) advise a modeller to exclude the latter parameters from the history-matching process, as attempts to estimate them could promulgate over-fitting of model outputs to field measurements. Sensitivity analysis is also recommended as a mechanism for gaining qualitative insights into parameter and predictive uncertainties.

Modern-day parameter estimation and uncertainty analysis is numerically untroubled by large numbers of parameters, nor by attendant parameter nonuniqueness. In fact, as is discussed above, the integrity of uncertainty analysis requires the inclusion of all parameters to which a management-salient prediction may be sensitive, regardless of their estimability. Old-fashioned local sensitivity analysis, conducted as a precursor to excluding parameters from the history-matching process, or to gain qualitative insights into model parameter and predictive uncertainty, is therefore unnecessary.

The analyses described herein fall under the ambit of local sensitivity analysis because their implementation requires the filling of matrices which characterize sensitivities of model outputs to model parameters. These analyses yield estimates of parameter and predictive uncertainty in contexts where parameter numbers are as high as they need to be for these analyses to have integrity. The matrix equations that implement them are slightly more complicated than those that are employed in traditional local sensitivity analysis. Furthermore, they require that a modeller provide a prior parameter covariance matrix that specifies prior parameter uncertainties and correlations. Nevertheless, in spite of their slightly greater complexity, all of the analyses described herein are easily implemented using public domain software which was written for this purpose.

3. THE MODEL

3.1 General

3.1.1 Purpose of the Present Study

The model which is the subject of the present GMDSI report was developed in 2015 for BHP by a consulting company. It is referred to herein as the “ODGAB model”. It replaces a previous model which was developed to explore similar issues. In the recent past, it has itself been superseded by an improved model.

The short study that is described herein employs the ODGAB model. It was undertaken in order to explore issues which are salient to the design of the improved model. Of particular interest are the specifications of some of the new model’s boundary conditions. Boundary condition specification is a problem that decision-support modelling must face on an everyday basis. It is incumbent on those who design and build such a model to ensure that predictions of stakeholder interest made by that model either suffer no bias, or that any (hopefully small) potential bias that emerges from abstractions in conceptual or numerical model design that are necessary for an analysis of predictive uncertainty, are included in the uncertainty margins themselves.

A plan of the ODGAB model domain is provided in Figure 3.1, together with place names that feature in the following discussion. Our description of the model is brief. It is restricted to those aspects of its design that are necessary for an understanding of linear analysis that was accomplished using this model. The latter is the focus of the study reported herein.

3.1.2 The Problem

BHP’s Olympic Dam mine, and the township of Roxby Downs which services it, are located approximately 520 km north north west of Adelaide, South Australia. The mine operates two wellfields for water supply; these are named “Wellfield A” and “Wellfield B”. Five abstraction wells and 42 monitoring wells comprise Wellfield A, while three abstraction wells and 47 monitoring wells comprise Wellfield B. Government monitoring bores have also been drilled in the general area. Wellfield A is located approximately 100 km north of the Olympic Dam mine near the south western margin of the Great Artesian Basin (GAB). Wellfield B is located about 80 km north east of Wellfield A, also within the GAB but further from its margin. At the time of writing, about 27 Ml/d of water is extracted from these wellfields.

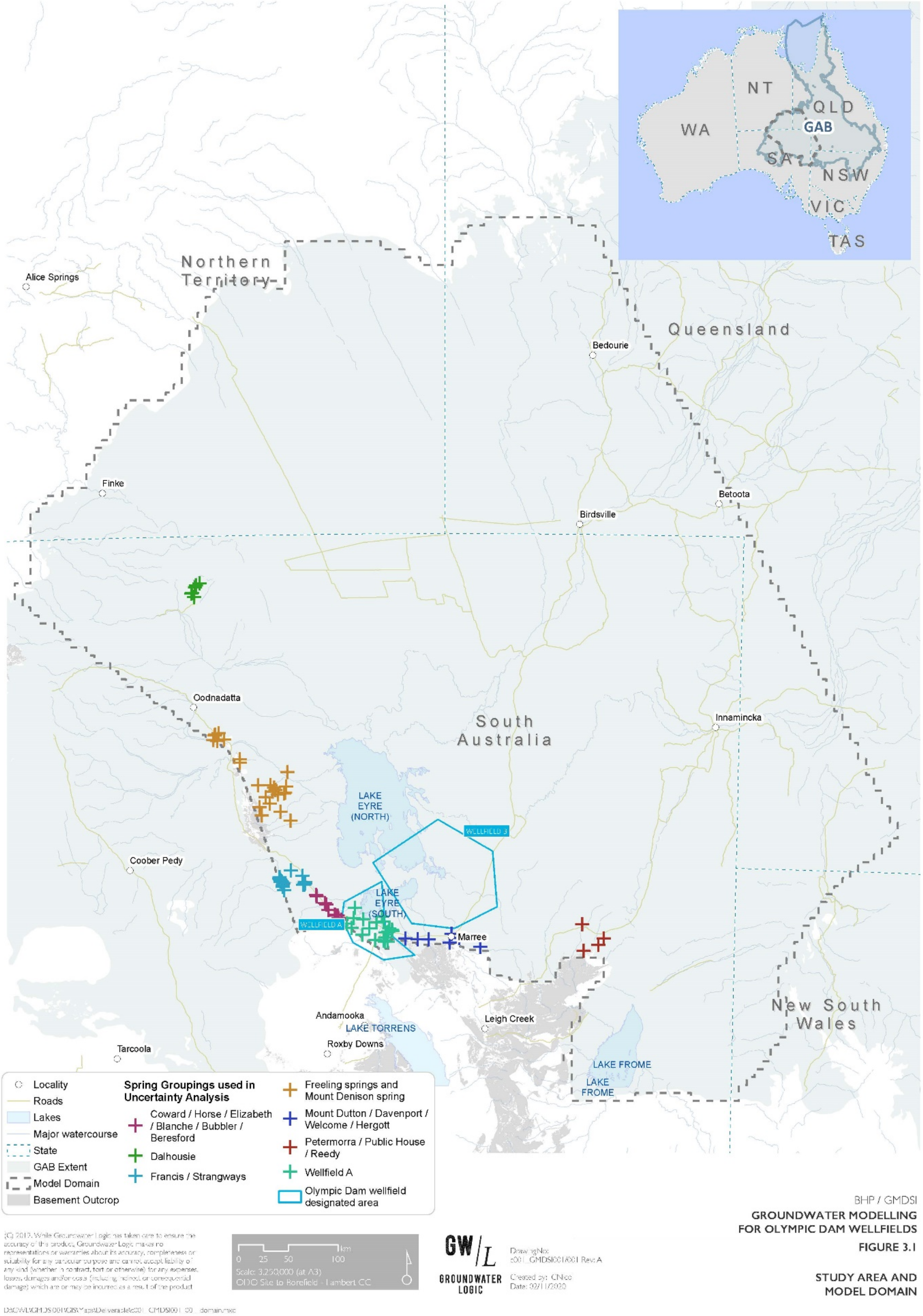
BHP wish to ensure that numerical modelling supports management of water extraction from these wellfields to the best extent possible. It is hoped that this will be achieved if modelling is open for stakeholder discussion, and recognizes the need to quantify the uncertainties of predictions of management and stakeholder interest, while reducing these uncertainties as much as possible through state-of-the-art data assimilation.

3.1.3 Geology and Hydrogeology

With an area of about 1.7 million square kilometres, the GAB is one of the largest underground freshwater resources in the world. See Figure 3.1. It includes the Eromanga, Surat and Carpentaria basins. Parts of the GAB lie in the Northern Territory, Queensland, South Australia and New South Wales.

Groundwater flows through sandstones of Triassic, Jurassic and early Cretaceous age from high ground at the eastern edge of the basin in Queensland and New South Wales. A much smaller amount of water also recharges the basin along its western margin in arid central Australia within the domain of the ODGAB model. Much of this latter recharge is diffuse.

However that which results from ephemeral flows in the Finke and Plenty Rivers is concentrated in time and space.



BHP / GMDSI
GROUNDWATER MODELLING FOR OLYMPIC DAM WELLFIELDS
FIGURE 3.1
STUDY AREA AND MODEL DOMAIN

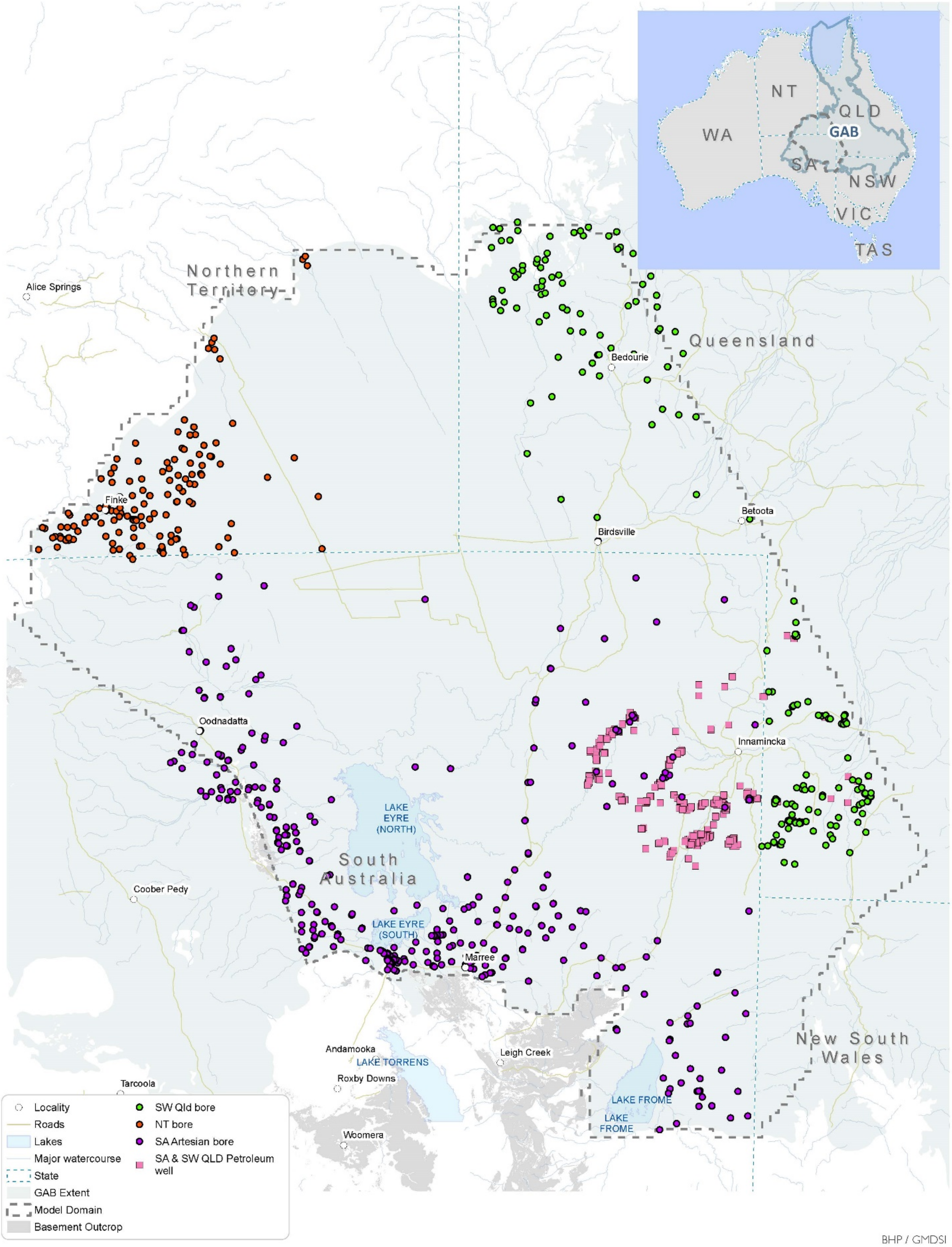
3.1.4 Elements of the Regional Water Balance

Water leaves the south western portion of the GAB through a number of springs and seeps. Details of water movement from GAB aquifers to springs are not well understood. Spring locations are thought to be controlled by faulting, and/or by abutment of sandstone aquifers with impermeable rocks from neighbouring basins. Springs sustain their own ecosystems. They were an important source of water for Aboriginal communities; as such, they have a high cultural value. Springs that are pertinent to the present study are depicted in Figure 3.1. Up to 75 Ml/d of water (possibly more) is thought to flow from these springs. However this estimate is uncertain because of difficulties in measuring spring flows and the paucity of flow records. Some of this water is nearly 2 million years old (Mudd, 2000).

Within the domain of the ODGAB model, somewhere between 20 and 300 Ml/d of water is thought to be lost from GAB aquifers as diffuse, evaporative discharge. Towards the south western margins of the GAB, where overlying aquitards are relatively thin, water migrates upwards through these aquitards in response to the prevailing vertical head gradient.

Since the late 19th century, artesian GAB waters have sustained the pastoral industry in large parts of central and southern Australia which are otherwise bereft of water. Water from free-flowing bores has been wasted, and pressures have fallen, as have flows from springs. A government funded well capping program is rectifying this situation. Pastoral bores located within the domain of the ODGAB model are shown in Figure 3.2. Rates of water extraction from these bores have fallen from about 200 Ml/d in 1980 to about 100 Ml/d at the present time. Pastoral water use is expected to continue at the latter rate into the indefinite future.

Large amounts of water are also extracted from GAB aquifers by the fossil fuel industry. In eastern parts of the GAB, water is extracted in order to depressurise coal measures, from which methane gas is desorbed and extracted. Within the domain of the ODGAB model, about 30 Ml/d of water was co-produced with conventional gas in the Moomba wellfield during 2013 and 2014. Extraction has since risen to 60 Ml/d in the Western Flank; see Figure 3.2.



BHP / GMDSI

GROUNDWATER MODELLING FOR OLYMPIC DAM WELLFIELDS

FIGURE 3.2

GAB GROUNDWATER USERS PASTORAL / TOWN / OIL & GAS

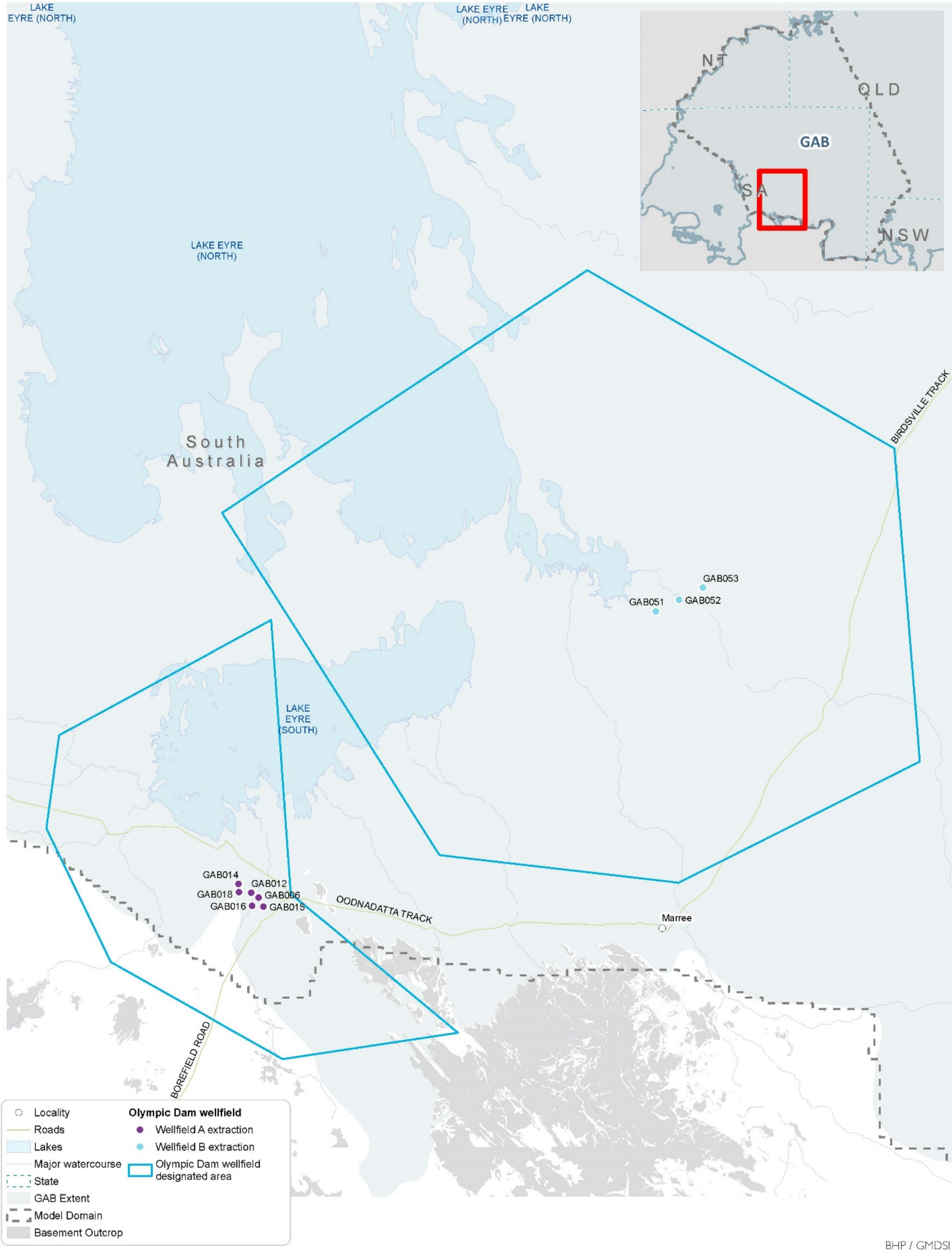
GW/L GROUNDWATER LOGIC

DrawingNo: c001_GMDSI001_002 Rev: A
Created by: CNicol
Date: 02/11/2020

D:\GWL\GMDSI001\GIS\Maps\Deliverable\c001_GMDS001_002_water_users.mxd

Groundwater extraction from the GAB in the southern part of the model domain to meet the requirements of the Olympic Dam mine and Roxby Downs township began in the early 1980s. The rate of extraction varies; it has averaged about 32 ML/d over the period 2000-2020. As stated above, this water is taken from Wellfields A and B which currently employ eight production bores. See Figure 3.3.

GAB groundwater flows into the ODGAB model area, primarily from Queensland, at an estimated rate of between 200 Ml/d and 600 Ml/d. The lower estimate implies a net loss of groundwater from the GAB aquifer within the ODGAB model area (because outflows exceed inflows); evidence for this includes broadly declining pressures. However this trend is being reversed because of pastoral bore capping programs.



BHP / GMDSI

GROUNDWATER MODELLING FOR OLYMPIC DAM WELLFIELDS

FIGURE 3.3

GAB GROUNDWATER USERS
OLYMPIC DAM / ROXBY DOWNS WATER SUPPLY

0 5 10 20 km
Scale: 600,000 (at A3)
ODO Site to Borefield - Lambert CC

GROUNDWATER LOGIC

DrawingNo: c001_GMDSI00 /003 Rev: A
Created by: CNicol
Date: 07/11/2020

D:\GWL\GMDSI001\GIS\Maps\Deliverable\c001_GMDSI001_003_water_users_OD.mxd

3.2 MODEL DESIGN

3.2.1 The Model Grid

The ODGAB model simulates groundwater flow using MODFLOW-USG (Panday et al, 2013). Despite the fact that high temperatures affect the density of groundwater, simulation assumes a uniform water density. To compensate for this, measured heads that are employed for model calibration are density-corrected.

The model grid is rectilinear; it is refined in areas of interest such as Wellfields A and B, and around springs that may be impacted by extraction from these wellfields. The smallest model cells have dimensions of 1 km × 1 km, while the largest model cells are 8 km × 8 km in size. Cell dimensions change by a factor of 2 at cell refinement boundaries. See Figure 3.4.

3.2.2 Layering

The ODGAB model employs three layers. The deepest layer (layer 3) represents the Jurassic Algebuckina Sandstone and Lower Cretaceous Cadna-owie Formations. Collectively, these comprise the principle GAB aquifer in the region; it is referred to as the "J-K Aquifer" herein. The Algebuckina Sandstone is comprised of terrestrial and fluvial sediments while the Cadna-owie Formation is comprised of terrestrial to marginal marine sandstone with finer grained interbeds. Layer 3 outcrops over part of the north west boundary of the model domain, where it is able to directly receive recharge waters.

Layers 1 and 2 of the ODGAB model are aquitards; they are the confining agents of layer 3. Nevertheless, they allow vertical transmission of water from the J-K Aquifer to the surface over parts of the model domain, where it evaporates. Layer 1 represents Quaternary to Cretaceous surficial sediments, as well as the Winton and Mackunda Formations. Layer 2 represents the Cretaceous Oodnadatta Formation, Coorikiana Sandstone and Bulldog Shale.

All layers of the ODGAB model are designated as confined at all places within its domain, even where they outcrop. Failure to designate layers as unconfined in outcrop areas has very little effect on model outcomes as groundwater head variations are small in these places.

3.2.3 Simulation Time

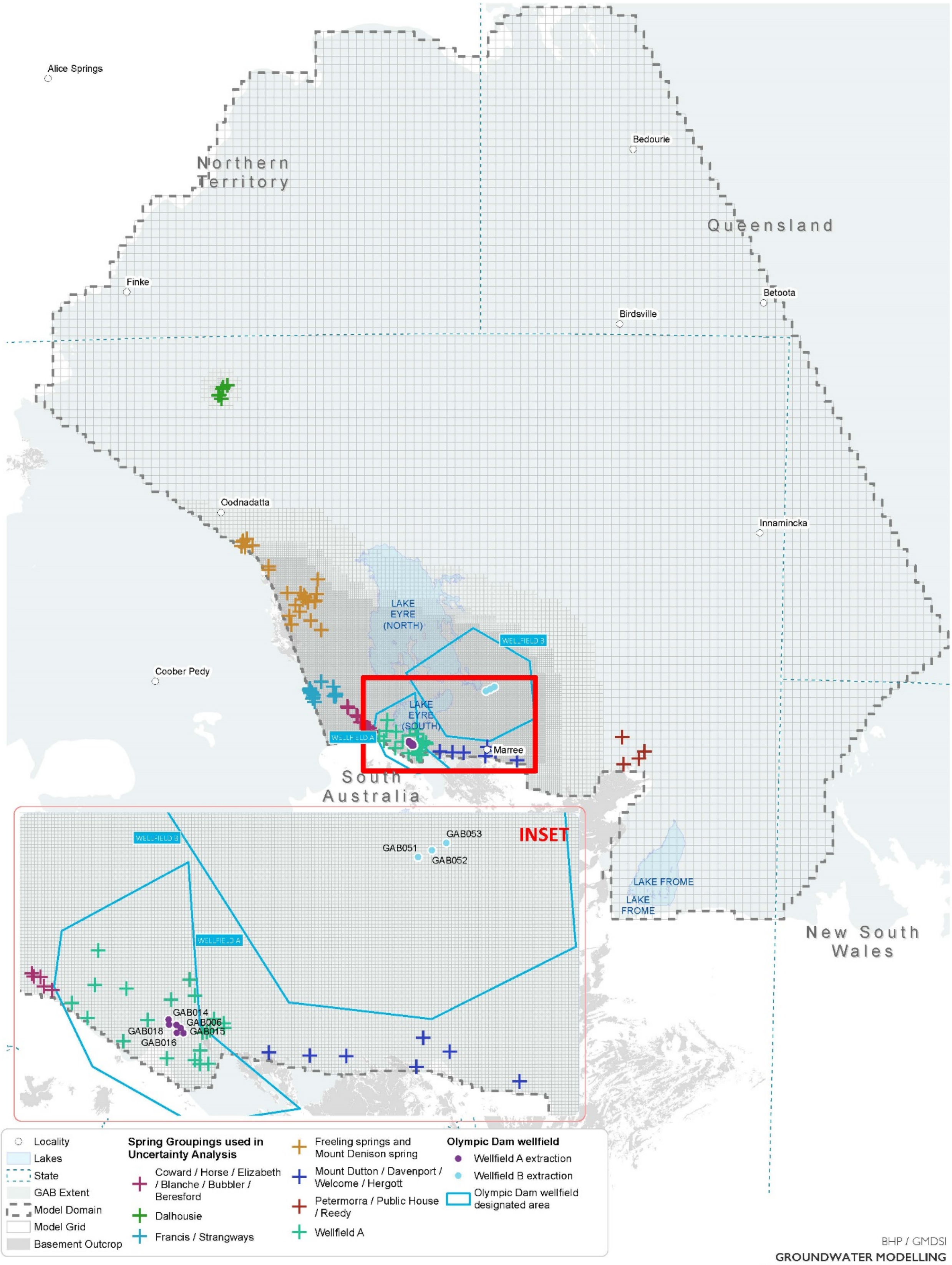
The ODGAB model runs for 196.5 years, of which 96.5 years span the period leading up to 2014, and 100 years comprise the period over which the model is required to make predictions. The simulation time is subdivided into 170 stress periods. A "stress period" is MODFLOW jargon for a period of time over which boundary conditions (including recharge and pumping rates) do not vary. The simulation begins with a steady state stress period; this represents conditions before any extraction took place from the J-K aquifer. Within the confines of the ODGAB model domain, extraction of water for pastoral purposes began in July, 1918.

3.2.4 Boundary Conditions

In this report, we employ the term "boundary condition" to describe any mechanism through which water can enter of leave a simulated groundwater system.

3.2.4.1 Recharge

Recharge of between 2.5 mm/yr and 5.4 mm/yr is applied to layer 3 of the ODGAB model where it outcrops; see Figure 3.5.



0 25 50 100 km
Scale: 1,000,000 (at A3)
ODO Site to Borefield - Lambert CC

GW/L GROUNDWATER LOGIC

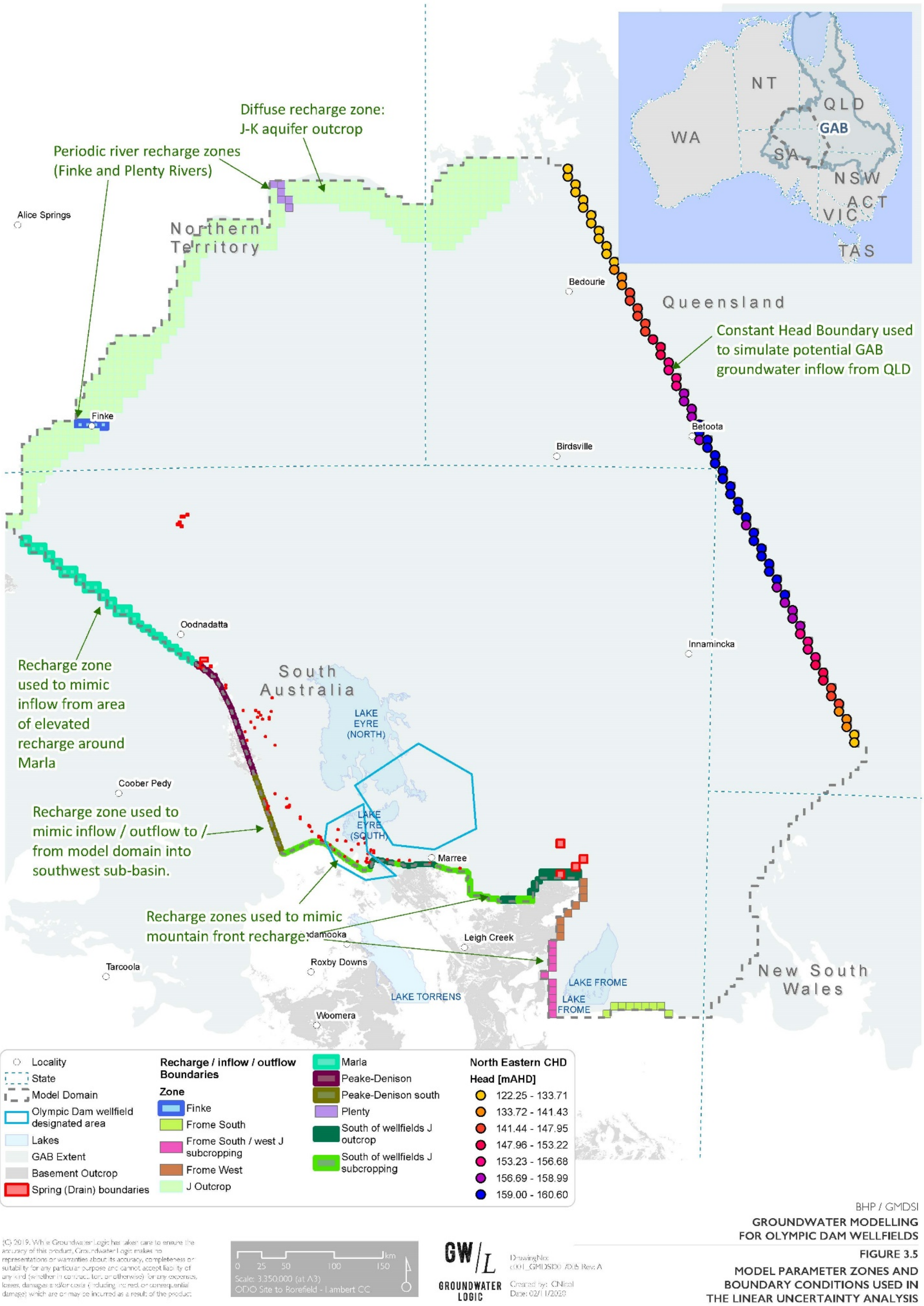
DrawingNo: c001_GMDSI00_004 Rev A
Created by: CNicol
Date: 07/11/2020

D:\GWL\GMDSI001\GIS\Maps\Deliverable\c001_GMDSI001_004_grid.mxd

BHP / GMDSI
GROUNDWATER MODELLING
FOR OLYMPIC DAM WELLFIELDS

FIGURE 3.4

MODEL GRID



D:\GWL\GMDSI001\GIS\Maps\Deliverable\c001_GMDS 001_005_BCs.mxd

BHP / GMDSI
GROUNDWATER MODELLING
FOR OLYMPIC DAM WELLFIELDS

FIGURE 3.5

MODEL PARAMETER ZONES AND
BOUNDARY CONDITIONS USED IN
THE LINEAR UNCERTAINTY ANALYSIS

3.2.4.2 Flow from the Wider GAB

It has been estimated that 200 Ml/d to 600 Ml/d of water flows into the domain of the ODGAB model from the eastern GAB. The ODGAB model represents this connection as a fixed head boundary. Heads assigned to cells along this boundary were inferred from head measurements in a small number of nearby wells. Measured heads were density-corrected in order to accommodate the high temperatures that prevail deep within GAB aquifers, as well as the elevated salinities of GAB waters.

Figure 3.5 depicts fixed head cells along the eastern boundary of the model domain. These are all emplaced in layer 3 of the ODGAB model. A no-flow condition prevails for all other layer 3 boundary cells, and for all boundary cells in all other model layers.

3.2.4.3 Springs

The ODGAB model represents springs in an approximate manner, as its focus is on calculating artesian pressures beneath springs, rather than flows from the springs themselves. The latter are dependent on local conditions that are beyond the ability of a regional model to represent. Nevertheless, the ODGAB model allows water to escape from the system in areas where springs occur, for this affects local groundwater heads. This is simulated using the MODFLOW DRAIN boundary condition. In construction of the ODGAB model, the conductances associated with these drains were adjusted so that losses of water through different groups of nearby springs is respected.

Model cells to which DRAIN boundary conditions are assigned are depicted in Figure 3.5.

3.2.4.4. Diffuse Discharge

In those parts of the model domain where model layers 1 and 2 are relatively thin, water migrates upward through these layers from the J-K aquifer to the surface, where it evaporates. The EVT package of MODFLOW-USG simulates this mechanism of water loss. Like the DRAIN package, the EVT package is a one-way boundary condition whereby water is removed at a rate that is governed by proximity of its head to the surface. However a user-defined maximum potential rate of water loss is specified for the EVT package, as is the depth range over which evapotranspired water can be taken from the model.

3.2.4.5 Well Extraction

A total of 68 individual wells are represented in the ODGAB model. These include pastoral wells, BHP extraction wells, and Moomba gas field extraction wells. In all cases, water is extracted from layer 3 of the model at user-specified rates using the MODFLOW-USG WELL package.

3.3 System Properties

3.3.1 Model Parameters

In the present discussion, we use the word “parameters” to denote values ascribed to model representations of hydraulic and boundary properties that govern flow of water within the model domain, and between the model domain and the outside world. These can be adjusted through the history-matching process in order to allow model-generated heads and fluxes to replicate field measurements of these quantities.

3.3.2 Hydraulic Conductivity

Rocks are heterogeneous, and must be represented as such in a groundwater model. Local measurements of aquifer and aquitard properties are generally sparse; furthermore, such local measurements are generally of limited use in characterizing regional groundwater movement.

While assignment of properties to different parts of a model domain can be illuminated by knowledge of prevailing rock types, generally hydraulic properties must be inferred (albeit with considerable uncertainty) from historical observations of system heads and fluxes.

The horizontal hydraulic conductivity of an aquifer is of greater interest than its vertical hydraulic conductivity, as groundwater flows parallel to its upper and lower boundaries. In contrast, the vertical hydraulic conductivity of an aquitard is of more interest than its horizontal hydraulic conductivity. In recognition of this, the vertical anisotropy of layer 3 of the ODGAB model was set to a uniform, and rather arbitrary, value of one tenth of its spatially variable horizontal hydraulic conductivity. The horizontal hydraulic conductivities of layers 1 and 2 were set to uniformly low values.

Because layers 1 and 2 act in concert to confine layer 3, while providing a pathway for slow migration of water to the surface where it can be evaporated, both of these layers were treated as a single layer from a parameterisation point of view. Thus at any location within the model domain, a single value of vertical hydraulic conductivity is assigned to both of these layers. Spatial variability of vertical hydraulic conductivity is characterised using pilot points; their disposition is shown in Figure 3.6.

Spatial variability of horizontal hydraulic conductivity in layer 3 is also characterized using pilot points; see Figure 3.7. The inset of this figure depicts subparallel lines of pilot points oriented in a north north westerly to south south easterly direction in the vicinity of Wellfield A. These were deployed to accommodate the presence of barriers to groundwater flow that are known to prevail in these areas.

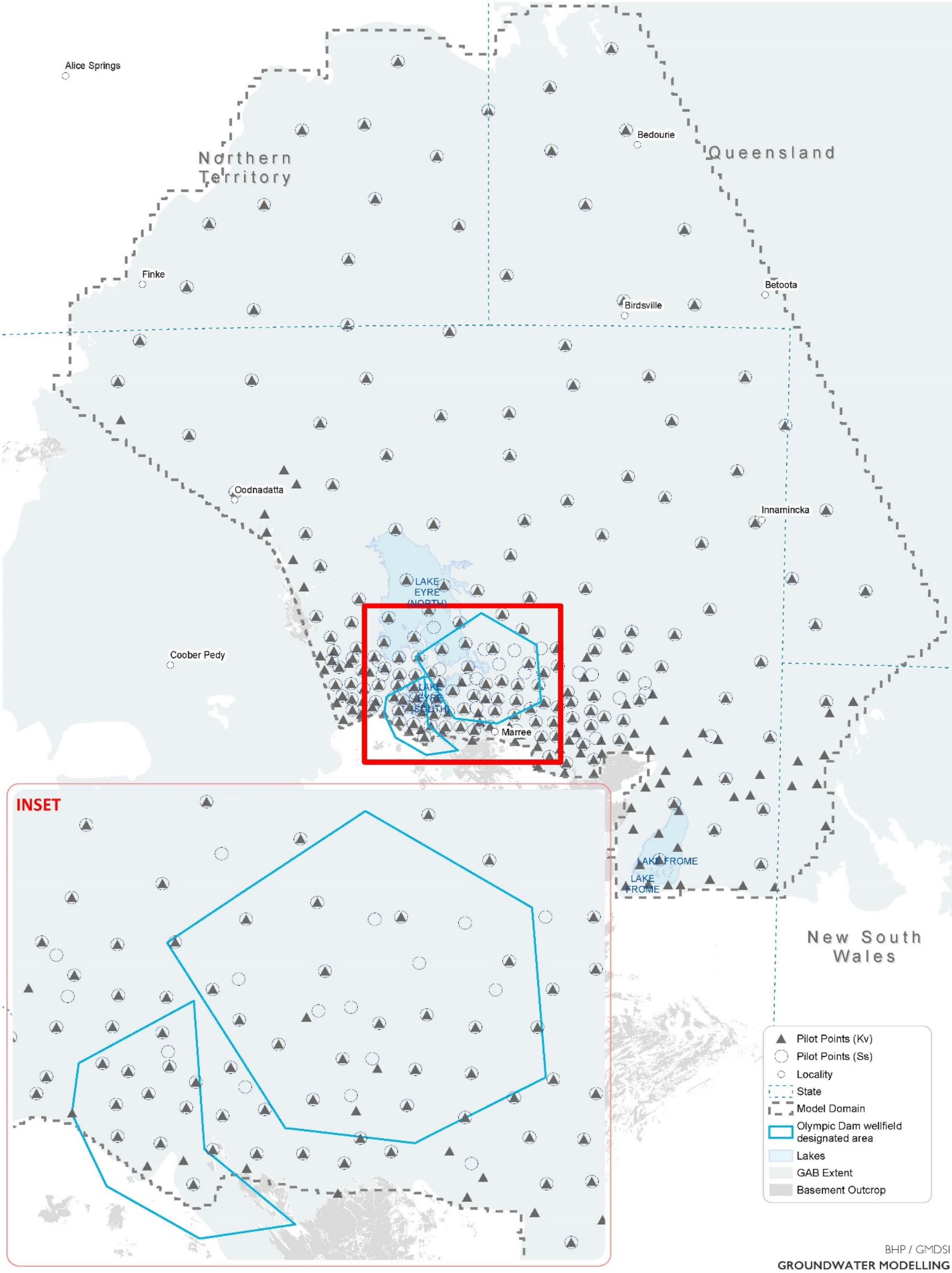
3.3.3 Specific Storage

Pilot points are also used to describe spatial variability of specific storage in layer 3 of the ODGAB model. Figure 3.6 plots the disposition of pilot points that are used for this purpose. In contrast, the specific storage of layers 1 and 2 was ascribed a calibration-adjustable, uniform value. Spatial variation of this property has little consequence for either replication of field measurements, or for predicting future system behaviour.

3.3.4 Drain Conductance

The rate of water emergence at springs is governed by conductances ascribed to MODFLOW DRAINs that connect springs to the GAB aquifer. These conductances can be adjusted in order for the model to respect observed spring flow and groundwater heads.

In the ODGAB model, springs are collected into groups. Groups are defined by proximity, and by the fact that measured flow rates comprising the model calibration dataset pertain to groups of springs rather than to individual springs. All DRAINs within a spring group are assigned the same conductance value. Collectively, 91 individual springs are amalgamated into 32 groups for which 32 values of DRAIN conductance require assignment.



BHP / GMDSI

GROUNDWATER MODELLING FOR OLYMPIC DAM WELLFIELDS

FIGURE 3.6

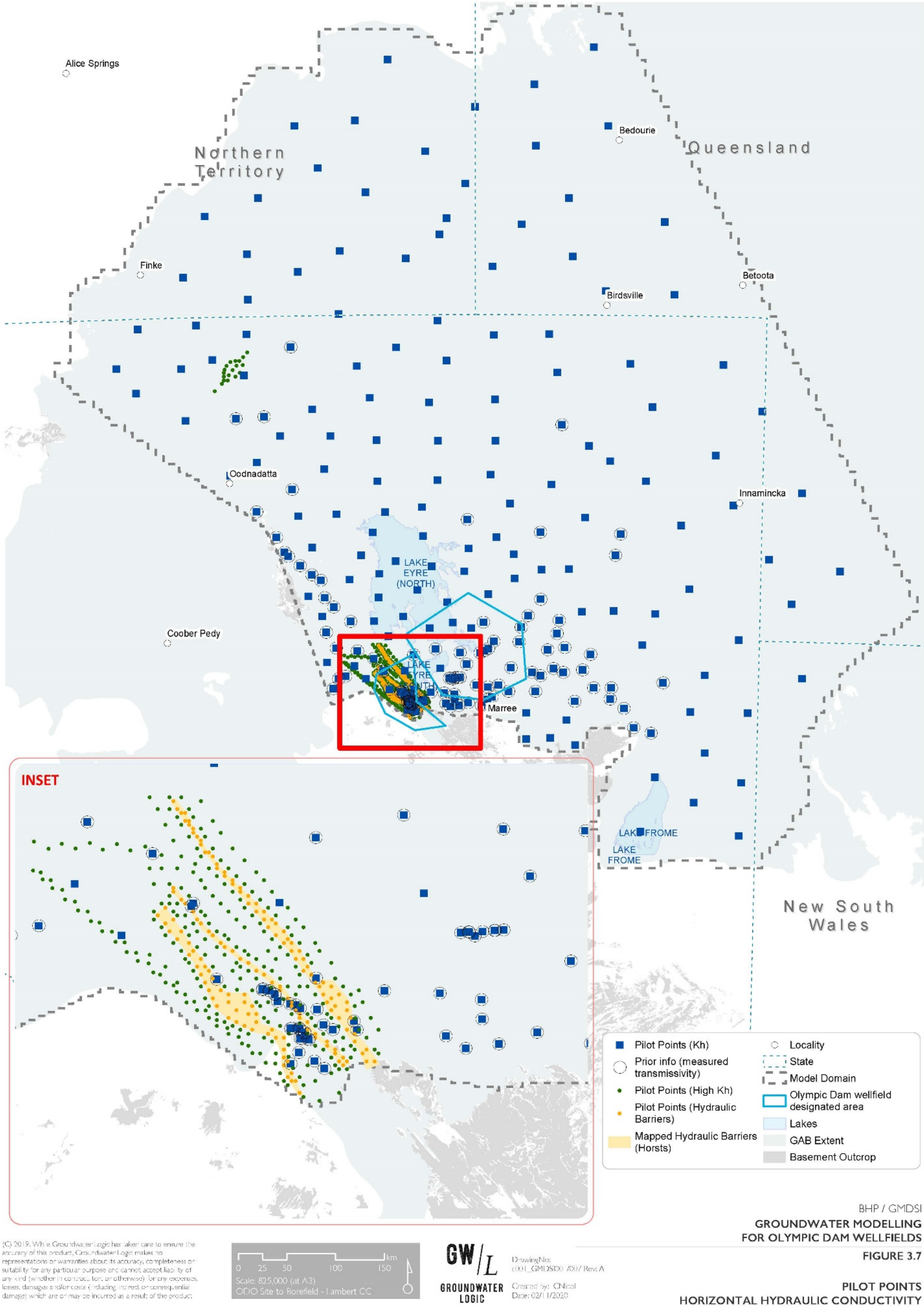
PILOT POINTS - SPECIFIC STORAGE AND VERTICAL HYDRAULIC CONDUCTIVITY

0 25 50 100 150 km
Scale: 1:1,110,000 (at A3)
ODO Site to Borefield - Lambert CC

GROUNDWATER LOGIC

Drawing No: c001_GMDSI001_006 Rev: A
Created by: CNicol
Date: 02/11/2020

D:\GWL\GMDSI001\GIS\Maps\Deliverable\c001_GMDS001_006_PilotPoints_Kv.mxd



GROUNDWATER LOGIC

Drawing No: c001_GMDSI001_007 Rev: A
Created by: CNicol
Date: 02/11/2020

D:\GWL\GMDSI001\GIS\Maps\Deliverable\c001_GMDSI001_007_PilotPoints_Kh.mxd

BHP / GMDSI
GROUNDWATER MODELLING FOR OLYMPIC DAM WELLFIELDS

FIGURE 3.7

PILOT POINTS
HORIZONTAL HYDRAULIC CONDUCTIVITY

3.3.5 Pilot Points as a Parameterisation Device

Where a hydraulic property varies spatially over a model layer, a value for this property must be assigned to every model cell within that layer. A model layer may contain thousands, or even tens of thousands, of cells. However, the use of pilot points to characterize a particular hydraulic property allows representation of its spatial variability over that layer using only a few hundred parameters; these are ascribed to pilot points rather than to cells. The assignment of hydraulic properties to model cells then becomes a two-step process. Firstly hydraulic properties are assigned to pilot points. These properties then undergo spatial interpolation to cells which comprise a model layer.

As has already been discussed, history-matching rarely, if ever, promulgates unique estimation of groundwater model parameters if these parameters are represented at the spatial scale at which prediction-salient heterogeneity of hydraulic properties is likely to exist. In the past, it was not uncommon for groundwater model parameters to be defined in a parsimonious manner specifically to accomplish uniqueness of their estimation. This was often achieved by subdividing each model layer into a small number of zones in each of which hydraulic properties are assumed to be uniform. Unfortunately, this strategy is beset by a number of disadvantages. These include the following.

- Geological media are not piecewise homogeneous with sharp property discontinuities at polygonal boundaries. The artificial nature of zone-based parameterisation can induce bias in some model predictions. This bias is inherent in the zonation scheme itself; however its magnitude can be amplified by a history-matching process that employs this scheme.
- It is often a difficult (and subjective) task for a modeller to design a zonation scheme that encompasses enough zones to support good model-to-measurement fit, but embodies few enough parameters to be uniquely estimable. Different manifestations of this subjectivity will result in different zonation patterns. These will introduce different levels of bias to model predictions.
- As has been discussed, part of the role of decision-support modelling is to quantify the uncertainties of decision-critical model predictions. These uncertainties arise from a modeller's inability to assign unique values to parameters, either through expert knowledge or through history-matching. Hence while adoption of a parsimonious parameterisation scheme may promulgate parameter uniqueness, it eliminates a modeller's ability to quantify predictive uncertainty. This removes from the modelling process one of its fundamental decision-support roles.

Where pilot points are employed as a parameterisation device, there is no need to deploy them in a parsimonious manner. Modern regularisation methods promulgate calibration uniqueness in ways that express the inexact nature of expert knowledge as it applies to the properties of a particular natural system. Numerical regularisation is better implemented in contexts of parameter superfluity than in contexts of parameter parsimony.

In summary, use of pilot points as a spatial parameterisation device reduces the potential for predictive bias at the same time as it supports parameter and predictive uncertainty analysis.

3.4. Model Calibration

3.4.1 General

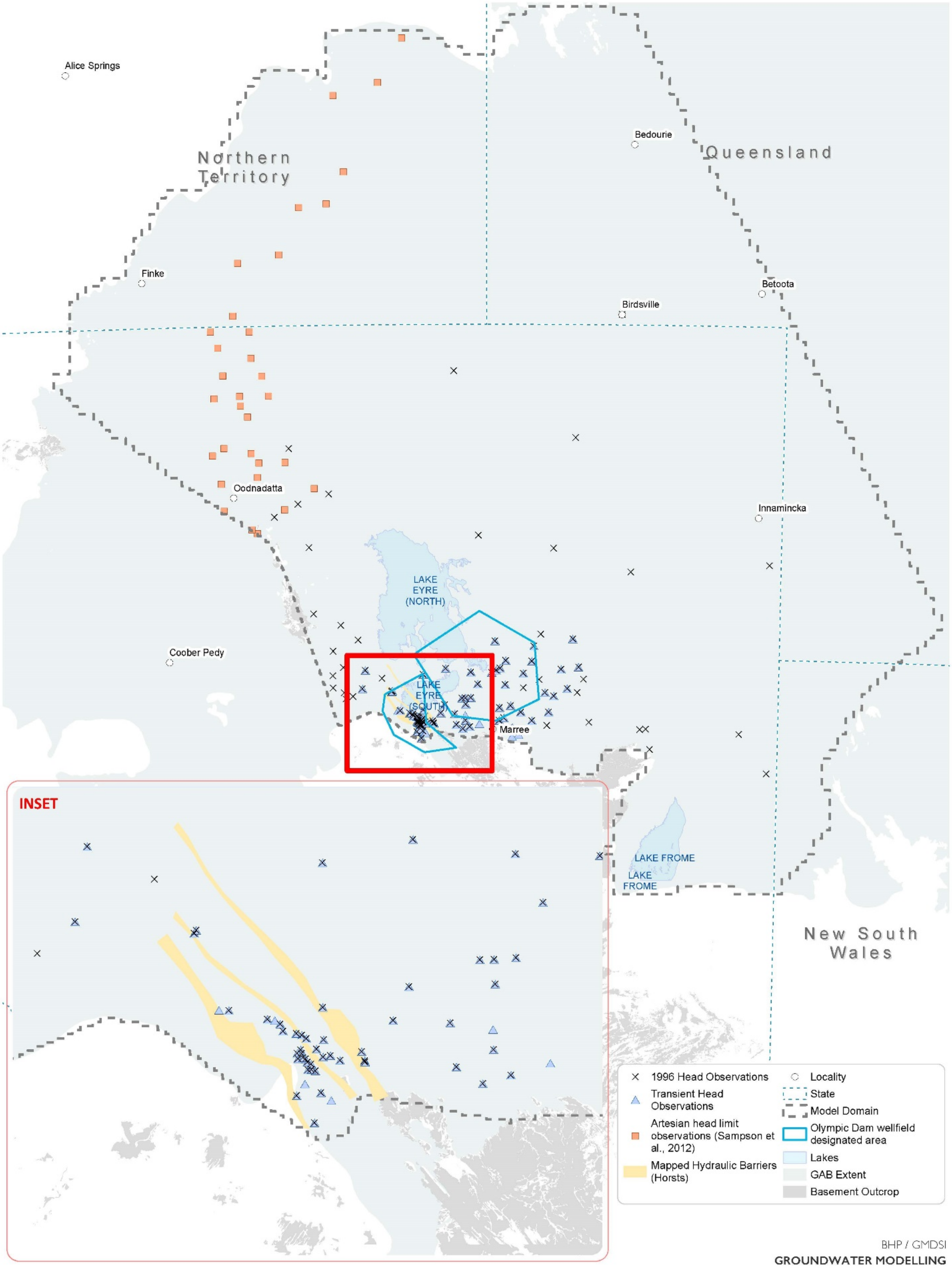
Prior to assessing the impacts of BHP pumping on J-K aquifer drawdown beneath GAB springs, the ODGAB model was calibrated. As was discussed in Chapter 2 of this document, the term "calibration" refers to back-calculation of hydraulic property values from historical measurements of system states and fluxes. Where uniqueness is sought (as is implied by the term “calibration”), estimates obtained in this way pertain to averaged, or upscaled, system properties. The nature of the averaging process is a function of the composition of the calibration dataset, and of the regularisation process that is employed to attain parameter uniqueness. Because the same model that is used for estimation of these properties is then used to make predictions of management interest, it is hoped that the upscaling process that is implemented through model calibration is also suitable for the making of predictions. At the same time, it is hoped that predictions made by the calibrated model will lie somewhere near the centres of their posterior probability distributions. This is the basis on which parameter uniqueness is sought.

3.4.2 Components of the Calibration Dataset

3.4.2.1 Heads in 1996

Head measurements that are recorded for 102 bores reflect J-K Aquifer conditions over the model domain before the commencement of BHP pumping. A single measurement from each of these bores was attributed to 1996. Collectively, they reflect pressure depletion from pastoral water extraction. The bores in which these measurements were made are depicted in Figure 3.8.

In order to increase spatial coverage, these borehole-measured heads were supplemented by 33 heads sampled from a line drawn by Sampson et al (2012) which marks the limit of artesian pressures. This limit line is also featured in Figure 3.8.



BHP / GMDSI

GROUNDWATER MODELLING FOR OLYMPIC DAM WELLFIELDS

FIGURE 3.8

GROUNDWATER OBSERVATIONS USED IN MODEL CALIBRATION

GROUNDWATER LOGIC

DrawingNo: c001_GMDSI001/008 Rev: A
Created by: CNicol
Date: 02/11/2020

D:\GWL\GMDSI001\GIS\Maps\Deliverable\c001_GMDSI001_008_Head_Obs.mxd

3.4.2.2 Transient Heads

Figure 3.8 shows the locations of wells from which transient head information has been collected. The number of available measurements varies widely from bore to bore. Collectively, they span the period from August 1965 to June 2014. A total of 9829 measurements from 82 bores were included in the ODGAB model calibration dataset.

Transient head data were used twice in formulation of the calibration dataset. A fit was sought between model outputs and the heads themselves. Fits were also sought between the drawdown in each bore relative to the first measurement in that bore, and model-calculated counterparts to these drawdown measurements. The fitting of specific aspects of a calibrated dataset (in this case drawdown relative to the first measurement) in addition to the data themselves, can facilitate passage of information from that data to model parameters. If studiously done, it can also reduce the propensity for parameters to adopt roles which compensate for model defects during the calibration process; these roles may bias some decision-critical model predictions. See Doherty and Welter (2010), White et al (2014) and Doherty (2015) for more details.

3.4.2.3 Spring Flows

The model was required to match long term flow estimates from the 32 groups of springs discussed above.

3.4.2.4 Evapotranspiration

Estimates of long-term evaporative losses from the J-K Aquifer have been provided by a number of authors, these including Rousseau-Gueutin et al (2012). Parameter estimation required that these estimates be respected.

3.4.2.5 Independent Estimates of Transmissivity

Nearly 120 estimates of J-K Aquifer transmissivity are available from hydraulic testing of production wells that are spread throughout the domain of the ODGAB model; see Figure 3.9. Model-employed transmissivities, spatially interpolated to measurement sites, can be compared to these estimates; differences between them were minimized through model parameter adjustment.

3.4.3 Strategy and Outcomes

3.4.3.1 Calibration Specifications

Values were estimated for a total of 1409 parameters by history-matching against a calibration dataset comprised of 21382 observations. Observations were divided into groups based on their different types. They were ascribed weights such that each observation group is visible in an overall objective function that increases with misfit between (processed) observations and their model-calculated counterparts. Model calibration reduces this objective function.

While observations considerably outnumber parameters, their information content was insufficient to promulgate parameter uniqueness without the assistance of mathematical/numerical regularisation. As has already been discussed, cogently-implemented regularisation pursues parameter uniqueness by estimating values for parameters that lie roughly at the centres of their posterior probability distributions.

3.4.3.2 Regularisation

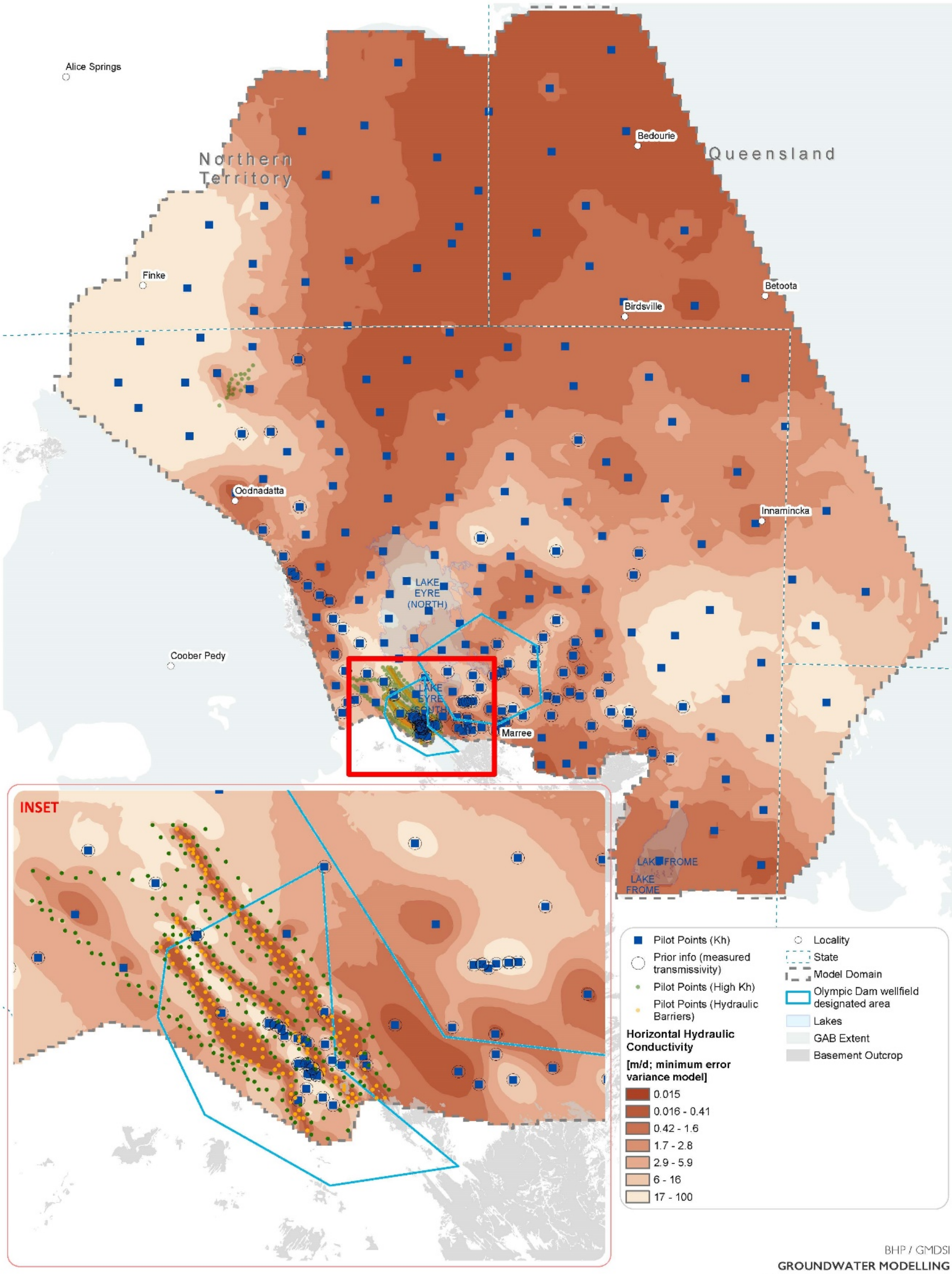
Model calibration was implemented using PEST (Watermark Numerical Computing, 2015). A Tikhonov regularisation strategy sought parameter estimates which allow model outputs to replicate field measurements while differing to the minimum extent possible from a modeller-assigned parameter field. Theoretically, posing the history-matching problem as a minimisation problem in this way achieves parameter uniqueness at the same time as it achieves centrality of estimated parameter values with respect to their joint posterior probability distribution.

3.4.4 Calibration Results

Presentation in this report of all of the outcomes of the calibration process would serve little purpose. However, by way of example, we present some of them.

Figure 3.9 depicts calibrated hydraulic conductivity for layer 3, while Figure 3.10 depicts calibrated specific storage for the same layer.

We also provide two plots that compare model outputs to field observations. Both of these plots pertain to observation bores in Wellfield B. Figure 3.11(a) depicts a bore for which model calibration achieved a relatively good fit between model outputs and observed heads. In contrast, Figure 3.11(b) shows a case where pumping-induced drawdowns are well fit, but where heads are not. Unfortunately, the second of these figures is more representative of ODGAB model calibration outcomes than the first.



BHP / GMDSI

GROUNDWATER MODELLING FOR OLYMPIC DAM WELLFIELDS

FIGURE 3.9

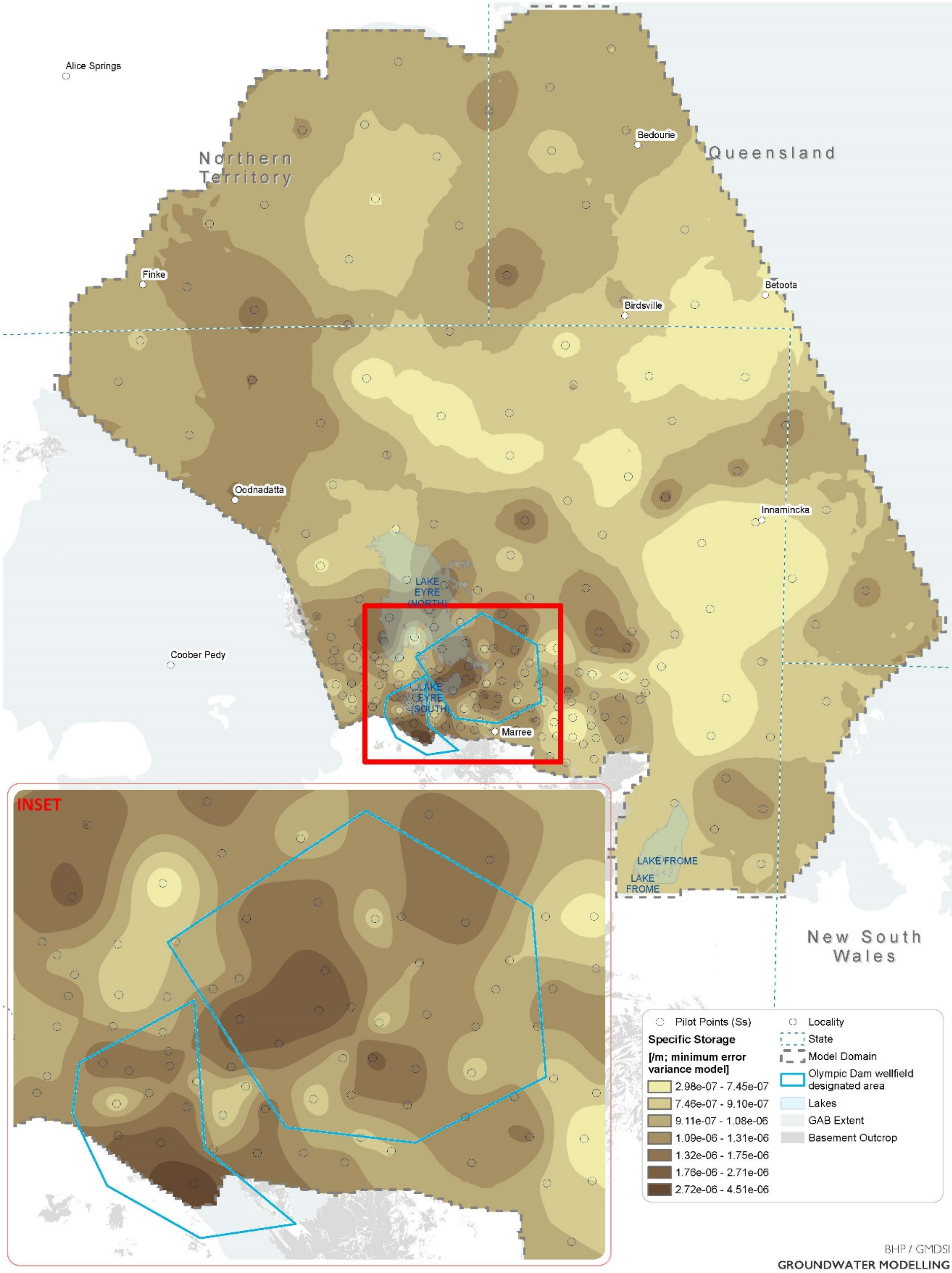
MINIMUM ERROR VARIANCE MODEL HORIZONTAL HYDRAULIC CONDUCTIVITY

0 25 50 100 150 km
Scale: 825,000 (at A3)
ODO Site to Borefield - Lambert CC

GROUNDWATER LOGIC

Drawing No: c001_GMDSI001/009 Rev: A
Created by: CNicol
Date: 02/11/2020

D:\GW_\GMDS.001\GIS\Maps\Deliverables\c001_GMDSI001_009_Calibrated_Kh.mxd



GROUNDWATER LOGIC

Drawing No: c00 GMDSI001/010 Rev: A

Created by: CNicol
Date: 02/11/2020

BHP / GMDSI

GROUNDWATER MODELLING FOR OLYMPIC DAM WELLFIELDS

FIGURE 3.10

MINIMUM ERROR VARIANCE MODEL SPECIFIC STORAGE

D:\GW_\GMDS001\GIS\Maps\Deliverables\001_GMDSI001_010_Calibrated_Ss.mxd

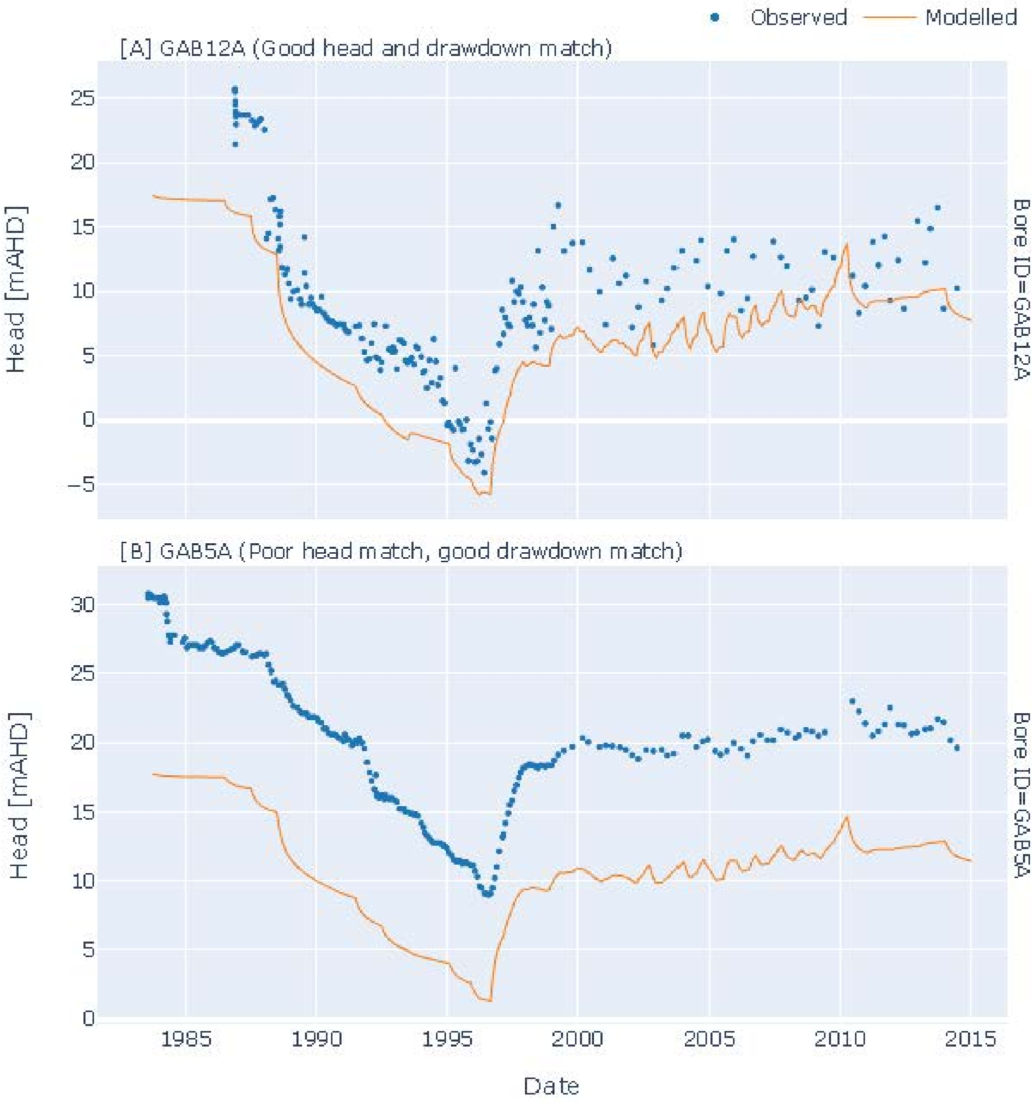


Figure 3.11 Selected model calibration hydrographs

3.5 Improving the ODGAB Model

3.5.1 General

The ability of the ODGAB model to replicate historical pumping-induced drawdowns reasonably well suggests that it may also possess the ability to predict future drawdowns resulting from an altered pumping regime reasonably well. In theory, its ability (or inability) to predict future drawdowns can, and should, be quantified through predictive uncertainty analysis. In the following chapter of this report we document how linear analysis was employed for this purpose. A more sophisticated, fully nonlinear, uncertainty analysis based on parameter ensembles has been undertaken using the ODGAB replacement model. This model is named "ODGABv2".

Prior to construction of ODGABv2, BHP personnel and modelling reviewers assessed concepts on which the ODGAB model is based, and specifications on which its construction rests. It was suggested that some aspects of its construction could be improved in order to preclude the possibility of unquantifiable predictive bias. Three issues whose exploration was assisted using linear analysis are now briefly described.

3.5.2 Recharge Proximal to the North Western Model Boundary

As shown in Figure 3.5, the J-K Aquifer is recharged in areas that are proximal to the north western boundary of the ODGAB model domain. The rate of recharge varies spatially between 2.5 mm/yr and 5.4 mm/yr within the area that is depicted in this figure. However local recharge is considerably higher than this for short periods of time when water flows in the Finke and Plenty Rivers, the locations of which are also shown in Figure 3.5. See Love et al (2013) for more details.

3.5.3 Inflow from the Greater GAB

Groundwater enters layer 3 of the ODGAB model from the wider GAB along its north eastern boundary. This connection is represented by a line of model cells in which heads are fixed; see Figure 3.5.

Potential inadequacies with this approach are as follows.

- By design, a fixed head boundary cannot be impacted by drawdown. It represents an unlimited supply of water from outside the model domain. As such, it may artificially mitigate drawdowns induced by pumping which takes place inside the model domain.
- Heads ascribed to this ODGAB model boundary may not be correct. Density corrections applied to borehole-measured heads, from which boundary heads were inferred, are uncertain.

3.5.4 No-Flow Status of Southern and Western Model Boundaries

The no-flow status of the southern and western boundaries of the ODGAB model have been questioned by reviewers. The hydrogeology of this area suggests that there is a potential for inflow and outflow to neighbouring basins along different segments of these boundaries, and that there is potential for small amounts of mountain front recharge in places. See Figure 3.5.

3.5.5 Consequences of Existing Boundary Conceptualizations

Most of the areas where the design of ODGAB model boundaries are examined using the linear methods documented herein are far from places where model predictions are of most interest, the latter being the locations of springs. Furthermore, it can be argued that the no-flow assumption as it pertains to the southern and western boundaries of the ODGAB model is conservative, as it prevents inflowing water from mitigating pumping-induced drawdown; it can be argued that this may lead to over-prediction of pumping-induced depressurisation of GAB aquifers in the vicinity of springs rather than to under-prediction of depressurisation. It may therefore be concluded that the decision-support role of the ODGAB model is unimpaired by conceptual errors that may be encapsulated in these boundaries.

Unfortunately, these arguments have both psychological and theoretical shortcomings. Their psychological shortcomings arise from their subjective nature. Hence they are unlikely to assuage the concerns of those who hold another point of view. Their theoretical shortcomings are a consequence of the fact that distance is not the only factor that determines whether a particular model simplification has the potential to inculcate bias in a particular model prediction, for the post-calibration relationship between a boundary condition and a prediction can be quite complex. Nor can the argument that certain model boundary condition specifications induce a conservative predictive bias be sustained unless it can be established that calibration of the model does not result in parameter estimates that correct this bias, or even induce an opposing bias in some model predictions. White et al (2014) and Doherty (2015) show that this is far from impossible.

Similar problems beset many, if not all, decision-support groundwater models. As was discussed in Chapter 2 of this report, model design is always a compromise. Quantification and reduction of predictive uncertainty requires numerical stability and speed of execution. However if model simplifications and abstractions that enable these qualities induce predictive biases that are commensurate with the uncertainties that the modelling process must attempt to quantify, then the cost of simplification exceeds its benefits.

4. EXPLORING MODEL ERRORS

4.1 Concepts

4.1.1 Simplification-Induced Model Errors

Modelling supports decision-making by quantifying and reducing the uncertainties of management-salient predictions. These tasks require a minimum level of model complexity. A model, together with its parameterisation, must be of sufficient complexity to represent aspects of a system to which decision-critical model predictions are sensitive, regardless of their estimability. As is discussed in Chapter 2 of this document, acknowledgement of their lack of estimability is essential for predictive uncertainty quantification. A decision-support model must also be complex enough to assimilate decision-pertinent information through history-matching. Complexity beyond that which serves both of these imperatives serves little purpose. In fact, it may degrade a model's decision-support utility by incurring long run times and by providing fertile numerical ground for simulator solution convergence difficulties.

There is no sharp cut-off between "useful complexity" and "useless complexity". Conversely, the boundary between "enabling simplicity" and "damaging simplicity" is not easy to recognize. Hence a modeller is faced with many subjective choices. In making these choices, he/she must bear in mind that the omission of certain parameters and processes from a model may have the following undesirable consequences.

1. It may impair a model's ability to replicate past system behaviour, thereby impairing its ability to assimilate information that is encapsulated in that behaviour.
2. Calculated predictive uncertainty limits may be narrower than actual uncertainty limits, this compromising a model's ability to quantify the risks associated with contemplated courses of management action.
3. It may induce bias in certain model predictions as parameters adopt roles that compensate for model defects during the history-matching process.

In the following discussion, we refer to the above model shortcomings as "Type 1", "Type 2" and "Type 3" shortcomings, to reflect the order in which they are listed above.

Type 1 shortcomings are relatively easy to detect during the history-matching process as they impair a model's ability to replicate the historical behaviour of a system. At the same time, minimized model-to-measurement misfit exhibits a high degree of temporal and/or spatial correlation, this being indicative of so-called "structural noise".

Unfortunately Type 2 and Type 3 shortcomings are harder to identify.

Type 3 problems are a particularly worrisome outcome of inappropriate model simplification for the following reasons.

- They do not necessarily incur model-to-measurement misfit. Moreover, the better is the fit between model outcomes and historical system behaviour, the greater are the surrogate roles that some parameters may need to adopt to achieve this fit.
- The repercussions of Type 3 shortcomings are prediction-specific. For some predictions, model defects can be "calibrated out". For other predictions made by the same model, calibration-induced predictive bias can be considerable.

4.1.2 Exploration of Simplification-Induced Model Shortcomings

Linear analysis can be employed to explore the existence and ramifications of the above model shortcomings. The first two are easily explored using linearized Bayes equation. The third requires a slightly more complex form of linear analysis based on singular value decomposition of the Jacobian matrix computed during the model calibration process. All of these analyses can be performed using software available through the PEST and PyEMU suites.

In many modelling contexts, including that which is the subject of the present report, boundary conditions are employed to simplify representations of processes which a modeller judges to be of secondary importance to a model's primary decision-support role. Parameters associated with these boundary conditions are often calibration-adjustable. The relationships between some types of boundary parameters and real-world hydraulic properties are often somewhat abstract. The conductance of a MODFLOW-USG GHB (general head) or DRAIN boundary is one such parameter type. In contrast, other boundary specifications (for example the head ascribed to a fixed head boundary, or recharge rates associated with different land uses and soil types) may not be easily adjusted through model calibration. Their omission from the calibration process reduces the complexity of that process, the number of model runs that are required to implement it, and the threats to model numerical stability that adjustment of some parameters may pose.

Fortunately, it is generally an easy matter to include boundary specifications in linear analysis, even if they are not estimated during model calibration. This is because they are not actually adjusted when undertaking linear analysis. Instead, each of them is varied incrementally during a sequence of model runs that is dedicated to calculating sensitivities of model outputs to these boundary specifications. If model run times are not too long, calculation of a suite of parameter and boundary condition sensitivities in this way is not a numerically demanding task. Nor is incremental variation of their values likely to precipitate model solution convergence failure.

As is discussed in Chapter 2 of this document, once sensitivities have been calculated in this manner, the action of a model on its parameters (including boundary specifications that would not otherwise be considered as parameters) can be replaced by the action of a matrix on a vector. If sensitivities to these parameters have been acquired for model outputs that correspond to field measurements, and for model outputs that correspond to predictions, the linearized form of Bayes equation can be used to quantify the uncertainties of these predictions, with uncertainties in boundary specifications taken into account. Predictive uncertainties that prevail prior to history-matching (prior uncertainties), and those that prevail after history-matching (posterior uncertainties) are readily evaluated. Meanwhile, a related type of linear analysis can be used to explore predictive bias incurred through fixing boundary parameters at possibly erroneous values instead of estimating them. In undertaking all of these analyses, it is incumbent on a modeller to assign prior uncertainties to boundary condition parameters that reflect their roles in the simulation process, including their possibly abstract representation of more complex hydraulic processes.

As has already been discussed, neither parameter values, observation values, nor the values of model outputs that correspond to observations or predictions, appear in the matrix equations that implement linear analysis. These equations include only model-output-to-parameter sensitivities. Hence the repercussions of fixing a parameter can be examined without specifying the value at which it is fixed. The repercussions of model-to-measurement misfit can be examined without actually specifying the measured values that the model is unable to fit. The uncertainty of a prediction can be quantified without actually making the prediction.

We now briefly describe how linear analysis can be turned to exploration of the above types of simplification-induced model shortcomings.

4.1.2.1 Type 1 Shortcomings

Type 1 shortcomings express the inability of a simplified model to fit a calibration dataset as well as a more complex model.

Once a model has been calibrated, a modeller is aware of the fit that can be attained between the outputs of his/her model and the components of a calibration dataset. Normally model-to-measurement misfit far exceeds that which would be expected to arise from measurement errors alone; misfit is more reflective of model inadequacies than it is of measurement noise. After calibrating a model, a modeller can back-calculate the amount of collective measurement/structural noise that engenders this same level of model-to-measurement misfit. This can then be included in linearized Bayes equation. Predictive uncertainties that arise from suboptimal model-to-measurement fit can then be calculated using this equation. So too can predictive uncertainties that accompany optimal model-to-measurement fits. By comparing the two, an estimate of the impact of Type 1 model shortcomings on the quantified uncertainties of decision-critical model predictions can be made.

This strategy is approximate, as it does not account for temporal and spatial correlations exhibited by structural noise. Nevertheless it is adequate in many groundwater modelling circumstances. This is because the uncertainties of many model predictions are a product of information inadequacy rather than contamination of information by measurement or model errors.

4.1.2.2 Type 2 Shortcomings

A Type 2 shortcoming denotes a simple model's inability to evaluate the full uncertainty of a prediction because of its failure to represent all facets of the system to which the prediction is sensitive.

The impact of this shortcoming is easily assessed through inclusion in linear uncertainty analysis of parameters that are not included in more computationally demanding nonlinear uncertainty analysis. Using linearized Bayes equation, the posterior uncertainty of a decision-critical prediction can first be calculated using parameters that are employed by the actual model. It can then be evaluated using the expanded set of parameters that is only available to linear analysis. The difference between these two uncertainties is a measure of the impact of Type 2 model shortcomings on the model's ability to quantify the uncertainty of that particular prediction.

4.1.2.3 Type 3 Shortcomings

Type 3 shortcomings reflect the compensatory roles that some model parameters must play in order to accommodate the fact that other model parameters cannot be adjusted in order to achieve a good fit with a calibration dataset. In a decision-support model these fixed parameters may be real or notional. In the latter case they can be viewed as errors that are hardwired into the construction of a simplified model that would otherwise be represented as processes and associated parameters in a more complex model. As has been discussed above, the effect of these missing processes and parameters can often be imitated by appropriate parameterisation of simple model boundary conditions for the purpose of linear analysis.

Singular value decomposition (SVD) is a process that, when undertaken on a Jacobian matrix, subdivides parameter space into two orthogonal subspaces. These subspaces are occupied by linear combinations of parameters rather than by individual parameters. Linear parameter combinations that belong to the "solution subspace" are adjusted during the history-matching process. Those that belong to the orthogonal-complementary "null subspace" do not need to be adjusted, as the calibration dataset does not inform them. It can be shown that calibration based on singular value decomposition yields a parameter field of minimized error variance, and that SVD comprises an optimal decomposition of parameter space.

Doherty (2015) and White et al (2014) show that model simplification can also be viewed as a form of parameter space decomposition. However, in contrast to SVD, this decomposition is unlikely to be optimal, for it requires that some combinations of parameters be adjusted that should not be adjusted according to the precepts of SVD, while other combinations of parameters cannot be adjusted which actually should be adjusted according to the precepts of SVD. Adjustment of the former set of parameters leads to “null space parameter entrainment”; this induces bias in some model predictions.

Linear analysis based on SVD allows the extent of calibration-induced predictive bias to be assessed. As for linear analysis based on Bayes equation, only sensitivities and parameter covariance matrices are required for this analysis. The values of model parameters, model outputs and field measurements do not figure in the equations on which this form of linear analysis is based. All operations involve matrices and vectors; they are easily programmed and readily evaluated.

4.1.2.4 Utility Software

The analyses discussed above can be implemented using programs available through the PEST and PyEMU suites. Utility programs from the PEST suite were used in the analyses described herein. These programs are listed in Table 4.1.

Program	**Role**
PPCOV_SVA	Computes a covariance matrix based on a spatially variable variogram.
PWTADJ2	Adjusts weights in a PEST control file to be commensurate with model-to-measurement misfit achieved through a preceding calibration process.
PREDUNC1	Employs the linearized form of Bayes equation to calculate the uncertainty of a model prediction.
PREDUNC4	Computes contributions to the uncertainty of a prediction made by different parameter groups.
PREDUNC7	Computes the posterior parameter covariance matrix. The diagonal elements of this matrix are the squares of posterior parameter standard deviations.
PREDVAR1B	Assesses calibration-induced predictive bias.
SUPCALC	Computes the dimensionality of the calibration solution space.

Table 4.1 PEST-Suite utility programs used to perform linear analyses described herein.

4.2 Application to the ODGAB Model

4.2.1 General

Features of the ODGAB model for which improvements were being considered in an upgraded model are outlined in Section 3.5. As has already been discussed, strategic simplification is an important constituent of decision-support model design, for dispensing with non-essential complexity can facilitate critical decision-support tasks such as data assimilation and uncertainty quantification. However the benefits of any compromise are associated with costs. The extent to which strategic simplifications are also shortcomings that may degrade the decision-support potential of a model can be analysed using the methodologies described above. This analysis can then provide a basis for improved model design, should this prove necessary. (It is worthy of note that the costs and benefits associated with a particular model design feature do not remain static as technology and computing power increase, for data assimilation and uncertainty quantification can now be undertaken with more complex models than were possible in the past.)

Implementation of linear methods that can analyse the potential limitations of the abovementioned ODGAB model features required the following steps.

1. Parameters which represent potential model defects were introduced to the ODGAB model specifically for the purpose of linear analysis. These are referred to as “defect parameters” in the discussion that follows.
2. These new parameters were assigned prior uncertainties that account for possible errors incurred by their use in place of more complex processes and accompanying parameters.
3. The Jacobian matrix attained during the preceding model calibration process was expanded to include sensitivities of calibration-pertinent model outputs to these defect parameters.
4. Sensitivities of decision-salient predictions with respect to existing and defect parameters were calculated using finite differences.
5. Bayes equation was used to calculate the uncertainties of management-salient predictions with and without inclusion of defect parameters. The stochasticity attributed to measurement noise in this equation was evaluated to reflect model structural noise exposed as model-to-measurement misfit during the preceding calibration process.
6. Linear analysis based on singular value decomposition (SVD) was used to evaluate the consequences for predictive bias of omission of defect parameters from the calibration process. (This is equivalent to fixing them at values that are possibly erroneous, as was effectively done when the ODGAB model was calibrated.)

Some implementation details as they pertain to features of the ODGAB model that were discussed above are now briefly presented.

4.2.2 Recharge Proximal to the North Western Model Boundary

Diffuse recharge near the north western boundary of the ODGAB model where the J-K aquifer outcrops was declared as adjustable. Three parameters were introduced to the ODGAB model to represent this recharge. A single zonal parameter encapsulated most of this outcrop area; for the purpose of linear analysis it was assigned a prior uncertainty that supports a maximum recharge rate of about 5 mm/year.

Two smaller model recharge zones were also defined within the J-K Aquifer outcrop area - one for the Finke River (5 model cells) and the other for the Plenty River (6 model cells); see Figure 3.5. Uncertainties assigned to these zones allow a maximum recharge rate of 85 mm/yr in the former case and 9 mm/yr in the latter case. These recharge rates are smaller than are possible under the wet-season beds of the rivers themselves; they are scaled to account for the size of model cells which contain these beds.

4.2.3 Inflow from the Greater GAB

A new parameter was introduced to the ODGAB model whose function is to uniformly raise or lower heads along the fixed head boundary that defines the north eastern edge of the model domain. This boundary simulates inflow of upgradient GAB waters. This new parameter was assigned a prior standard deviation of 5 m. (Recall that in linear analysis this boundary is not actually raised or lowered by this amount. It is raised incrementally for the purpose of calculating sensitivities of model outputs to its elevation. The prior standard deviation of 5 m is used in matrix equations which implement linear analysis.)

4.2.4 No-Flow Status of Southern and Western Model Boundaries

Recharge zones were introduced along what were previously no-flow boundaries at the south and south western edges of the ODGAB model. Parameters associated with these zones simulate the possibility of water movement across these boundary segments to neighbouring basins. These zones are depicted in Figure 3.5. The prior uncertainty ascribed to each zonal recharge parameter was informed by expert-knowledge of maximum possible inflows or outflows along respective boundary segments.

It is reiterated that linear analysis does not require that these inflows/outflows be actually provided to the model, and that the model be re-calibrated to accommodate them. It requires only that the sensitivities of model outputs with respect to these possible inflows/outflows be evaluated, and that the range of possible water exchange rates be reflected in the prior uncertainties that are assigned to pertinent parameters in the matrix equations which implement linear analysis. Accommodation by other parameters of these inflows in notional model re-calibration is implicit in these linear equations.

4.2.5 Other Parameters

All of the calibration-adjustable parameters that are described in the previous chapter of this report were retained in the ODGAB model for the purpose of linear analysis. They were all assigned prior uncertainties that are in accordance with hydrogeological expectations. Covariance matrices were ascribed to pilot point parameters in order to account for spatial correlation between them. For the sake of brevity, details of these prior parameter uncertainties are omitted from the present discussion.

4.3 Some Outcomes of Linear Analysis

4.3.1 Dimensionality of The Solution Space

The PEST SUPCALC utility employs singular value decomposition to compute the dimensionality of the calibration solution space. This is equivalent to the number of items of useable information that are resident in a calibration dataset. It is also equivalent to the number of combinations of parameters that are uniquely estimable on the basis of that dataset.

Despite the fact that the calibration dataset is comprised of 21382 observations, SUPCALC evaluated a solution space dimensionality of about 121.

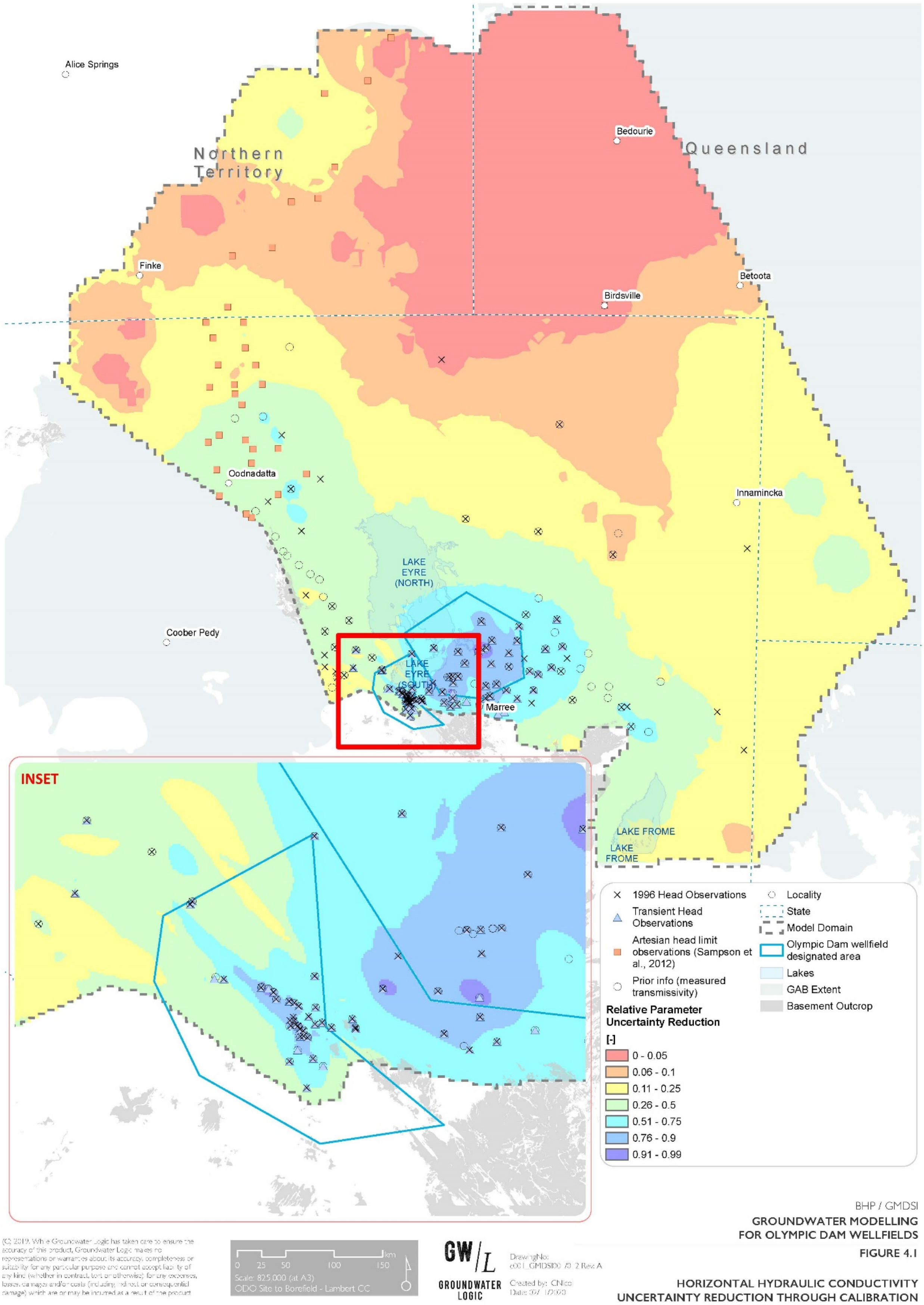
The size of the solution space that can be informed by a measurement dataset is influenced by the amount of noise that accompanies this dataset. In calculations comprising the present analysis, this was inflated by structural noise whose magnitude was gleaned from the level of model-to-measurement misfit that remained after calibration of the ODGAB model. SUPCALC was also asked to estimate the dimensionality of the calibration solution space using estimates of measurement noise that omit contributions from structural noise. Its revised estimate of solution space dimensionality is 160.

A higher dimensional solution space promulgates lower posterior uncertainties of some model parameters and some model predictions. The above result shows that Type 1 model shortcomings can reduce the dimensionality of the solution space. This can deliver increased uncertainties for some model parameters and some model predictions. However it is important to bear in mind that Type 1 model shortcomings affect some predictions more than others. They have little effect on predictions whose uncertainties reflect lack of information in a calibration dataset, rather than contamination of that information by structural noise.

4.3.2 Parameter Uncertainty

Estimates of posterior parameter uncertainty can be readily made using linearized Bayes equation. Doherty and Hunt (2009) describe a mappable statistic which they refer to as "relative parameter uncertainty variance reduction" (RPUVR). (Recall that variance is the square of standard deviation, and that standard deviation is a measure of uncertainty.) Every model parameter can be assigned its own RPUVR value; it varies between 0.0 and 1.0. A value of 0.0 indicates that the corresponding parameter enjoys no reduction of uncertainty through history-matching; its posterior uncertainty is thus the same as its prior uncertainty. In contrast, an RPUVR value that approaches 1.0 indicates a small posterior parameter uncertainty relative to its prior uncertainty. The calibration dataset is therefore rich in information pertaining to this parameter.

Figures 4.1 and 4.2 map RPUVR (with structural noise taken into account) of horizontal hydraulic conductivity and specific storage of layer 3 of the ODGAB model. The locations of points which contributed head, drawdown and transmissivity observations to the calibration dataset are superimposed on these figures. A strong relationship between data density and parameter uncertainty reduction is obvious from these figures.



BHP / GMDSI

GROUNDWATER MODELLING
FOR OLYMPIC DAM WELLFIELDS

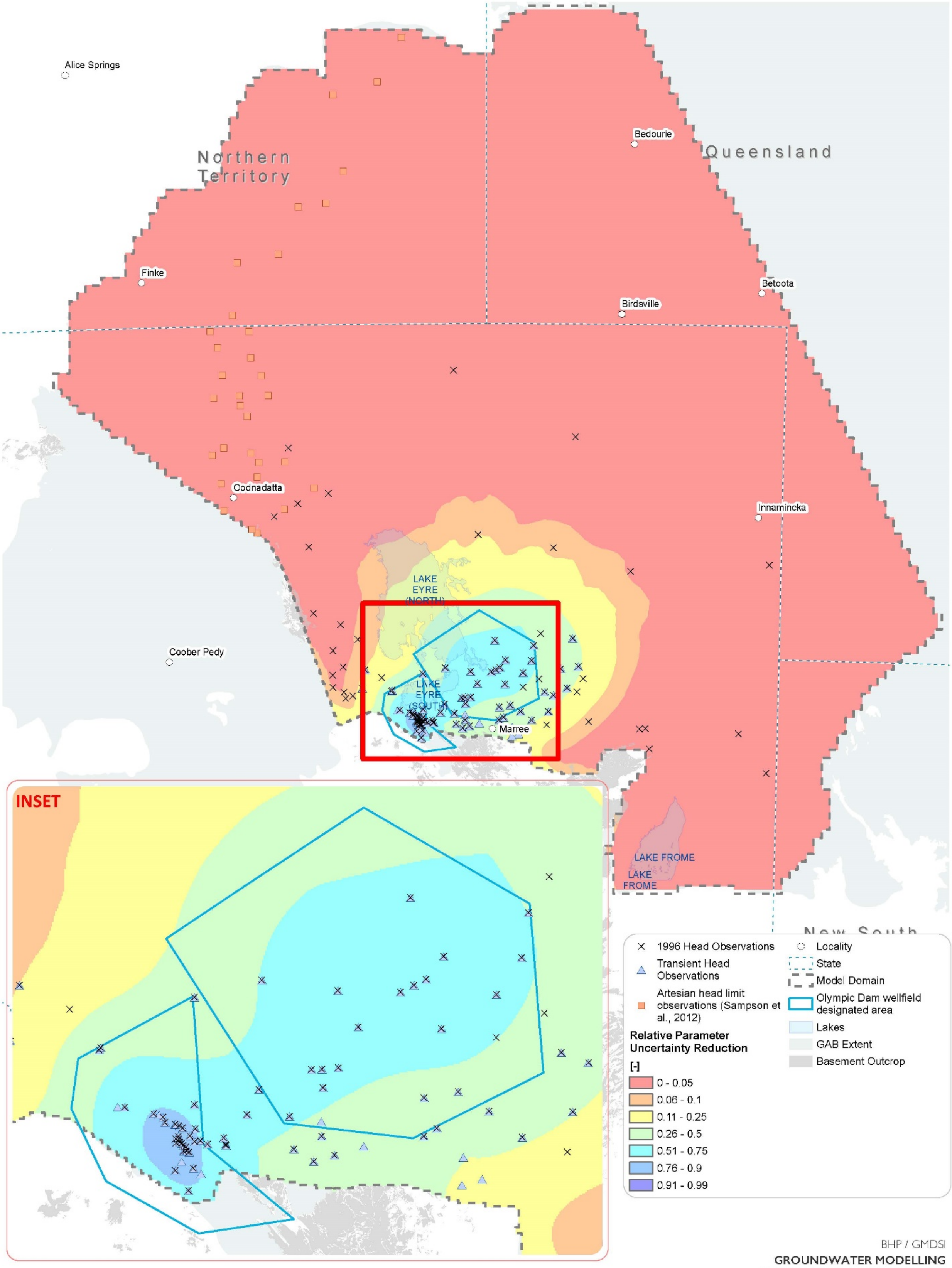
FIGURE 4.1

HORIZONTAL HYDRAULIC CONDUCTIVITY
UNCERTAINTY REDUCTION THROUGH CALIBRATION

GROUNDWATER LOGIC

DrawingNo: c001_GMDSI001_012 Rev: A
Created by: CNicol
Date: 07/11/2020

D:\GWL\GMDSI001\GIS\Maps\Deliverable\c001_GMDSI001_012_Kh_uncert_reduction.mxd



0 25 50 100 150 km
Scale: 1,110,000 (at A3)
ODO Site to Borefield - Lambert CC

GROUNDWATER LOGIC

DrawingNo: c001_GMDSI00_013 Rev: A
Created by: CNicol
Date: 07/11/2020

BHP / GMDSI
GROUNDWATER MODELLING FOR OLYMPIC DAM WELLFIELDS

FIGURE 4.2

SPECIFIC STORAGE
UNCERTAINTY REDUCTION THROUGH CALIBRATION

D:\GWL\GMDSI001\GIS\Maps\Deliverable\c001_GMDSI001_013_Ss_uncert_reduction.mxd

4.3.3 Predictive Uncertainty

In this and subsequent subsections of this report we examine a single prediction. This prediction is the maximum drawdown (relative to pre-development conditions) in the J-K Aquifer under any spring in 2084. This is the time at which modelled extraction from Olympic Dam wellfields ceases. The value of this prediction made by the calibrated model is 3.27 m.

Linearized Bayesian analysis (with structural noise taken into account) yields the outcomes presented in Table 4.2.

Quantity	**Value**
Value of prediction	3.27 m
Prior standard deviation	1.55 m
Prior variance	2.39 m^2
Posterior standard deviation	0.63 m
Posterior variance	0.39 m^2

Table 4.2 Outcomes of linear uncertainty analysis as they pertain to maximum layer 3 drawdown under any spring.

We take this opportunity to remind the reader of the assumptions on which linearized Bayesian analysis is based. These are:

- Model behaviour with respect to its parameters is linear.
- All probability distributions (those pertaining to parameters, predictions and measurement noise) are multiGaussian.

All of these assumptions are violated by the behaviour and properties of natural systems. Furthermore, because drawdown has a lower bound of 0.0, it is unlikely to possess a probability distribution that is symmetrical with respect to its mean. The uncertainties tabulated in Table 4.2 must therefore be interpreted as approximate.

Using linearized Bayes equation, it is a simple matter to repeat the above calculations under the assumption that measurement/structural noise is zero. This establishes the extent to which the uncertainty of the above prediction is an outcome of lack of information within the calibration dataset, rather than contamination of that information by measurement noise and/or model imperfections. It is found that the no-noise posterior standard deviation of maximum subspring drawdown is 0.36 m. Hence more than half of the posterior uncertainty of this prediction arises from information insufficiency.

A repetition of the above calculation with measurement noise awarded stochastic properties that are in accordance with those expected from field measurements yields a standard deviation of predictive uncertainty of about 0.5 m.

4.3.3 Parameter Contributions to Predictive Uncertainty

Figure 4.3 shows contributions made by different parameter types to the uncertainty variance of maximum predicted subspring drawdown in 2084. Recall from Table 4.2 that the prior variance of this prediction is 2.39 m^2. Contributions to variance, rather than contributions to standard deviation, are graphed in Figure 4.2 because contributions to prior predictive variance are additive. The same does not apply to prior standard deviations. Nor does it apply to either posterior variances or posterior standard deviations.

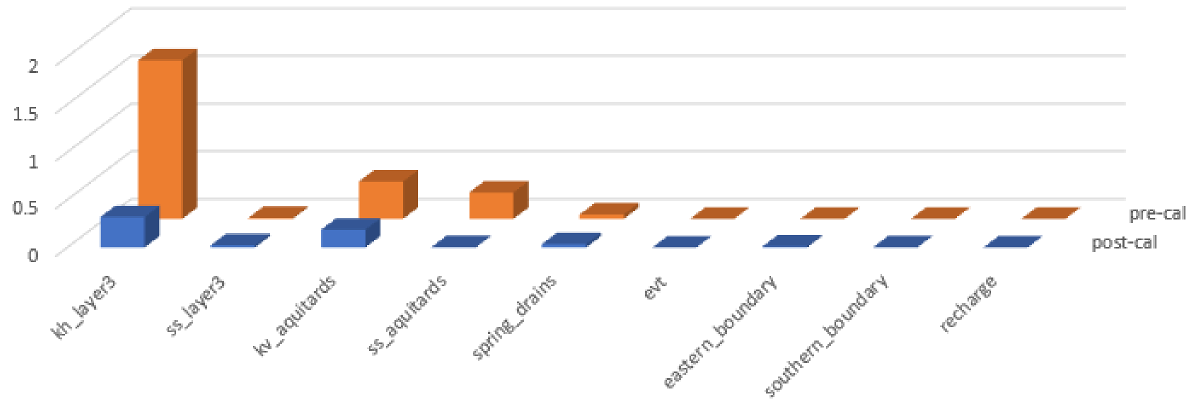


Figure 4. Contributions made by different parameter types to the uncertainty variance of maximum predicted subspring drawdown in the J-K Aquifer in 2084. The back row pertains to the prior uncertainty of the prediction while the front row pertains to its posterior uncertainty.

As is explained earlier in this report, we define the “contribution to predictive uncertainty” of a parameter group as the uncertainty of a prediction of interest as calculated using all adjustable parameters minus the uncertainty of that same prediction when calculated under the assumption that the values of all members of that parameter group are perfectly known; that is, all members of the parameter group are assumed to possess prior uncertainties of zero.

The parameter groups that are featured in Figure 4.3 are, from left to right:

- horizontal hydraulic conductivity of the J-K Aquifer (that is, model layer 3);
- specific storage of the J-K Aquifer;
- vertical hydraulic conductivity of aquitards overlying the J-K Aquifer (that is, model layers 1 and 2);
- specific storage of aquitards overlying the J-K Aquifer;
- conductances ascribed to MODFLOW-USG DRAIN cells which govern outflow to springs from the J-K Aquifer, and affect artesian pressures within the J-K Aquifer around discharge areas;
- maximum rate of evapotranspiration of water which migrates vertically from the J-K aquifer to the surface in the north western part of the ODGAB model domain;
- heads ascribed to the eastern fixed head boundary that connects the J-K Aquifer to the wider GAB;
- potential recharge/discharge along no-flow southern and western margins of the ODGAB model domain;
- diffuse and river recharge to the J-K aquifer where it outcrops in proximity to the north western boundary of the ODGAB model domain.

The last three of the above parameter types are so-called “defect parameters” for the purpose of the current analysis.

It is obvious from Figure 4.3 that by far the largest contributor to the uncertainty of predicted subspring drawdown is the hydraulic conductivity of the J-K Aquifer. This arises from the influence of hydraulic conductivity on drawdown propagation, its propensity for spatial variability, and the range of values that it may take. Specific storage of the J-K aquifer makes a much smaller contribution to the uncertainty of this prediction, possibly because heads near springs have begun to stabilize by 2084. Hydraulic properties of the aquitard which overlies the J-K aquifer govern movement of water to the surface where it is lost to evapotranspiration.

Water that is lost from the J-K Aquifer through this mechanism is not available for spring flow. This explains the contribution made by aquitard parameters to the uncertainties of subspring drawdown predictions.

The following is also apparent from Figure 4.3.

- Contributions to subspring drawdown uncertainty made by J-K Aquifer properties, and by the hydraulic properties of the overlying aquitard, are considerably reduced through history-matching.
- Model defect parameters contribute little to either the prior or posterior uncertainties of subspring drawdown predictions.

Table 4.2 lists contributions made by model defect parameters to the uncertainty of the 2084 subspring drawdown predictions in terms of standard deviation rather than variance. (Note that these are not additive; hence the value assigned to "all of the above" is not the sum of its parts.) The small values of these contributions are, once again, abundantly clear.

Parameter type	**Contribution to posterior uncertainty standard deviation of predicted drawdown (m)**
Eastern fixed heads	0.01
Southern and western inflows/outflows	0.0023
North western recharge	0.00083
All of the above	0.013

Table 4.2 Contributions made by model defect parameters to maximum subspring drawdown uncertainty in 2084. The total standard deviation of this prediction is 0.63 m.

4.3.4 Calibration-Induced Predictive Bias

The PEST PREDVAR1B utility evaluates the potential for error in user-specified predictions induced by inappropriate model simplification. These are the outcomes of Type 3 model shortcomings. Predictive bias can occur when some parameters adopt compensatory roles during model calibration in order for model outputs to attain good fits with a calibration dataset when model defect parameters are fixed at possibly erroneous values.

Table 4.3 tabulates the standard deviation of potential error incurred through this mechanism in the prediction of maximum subspring drawdown in 2084. (Note that these contributions to predictive error standard deviation are not additive; hence the value assigned to "all of the above" is not the sum of the parts.)

Parameter type	**Contribution to posterior error standard deviation of predicted drawdown (m)**
Eastern fixed heads	0.001
Southern and western inflows/outflows	0.0008
North western recharges	0.017
All of the above	0.018

Table 4.3 Standard deviations of possible predictive error incurred by fixing model defect parameters at possibly erroneous values.

Like the numbers that are recorded in the second column of Table 4.2, those that are recorded in the second column of Table 4.3 are small. They are, in fact, calculated using the same quantities. These are sensitivities of model outputs to parameters, prior parameter standard deviations, and a statistical characterisation of measurement noise that also reflects structural noise exposed through the model calibration process. However these quantities are used in different ways in the calculations on which these two tables are based. The contents of Table 4.2 are calculated using a linearized form of Bayes equation. In contrast, those of Table 4.3 emerge from singular value decomposition of the Jacobian matrix that is a linearized representation of the action of the model on its parameters as the model is calibrated.

4.3.5 Computational Burden

We conclude this chapter with a few words on the computational cost of computing the quantities that are tabulated and graphed herein.

As has been described, matrix equations which implement linear analysis are based on sensitivities of model outputs to model parameters. Model outputs of interest are those which correspond to measurements which comprise a calibration dataset, as well as predictions of future system behaviour. Sensitivities are calculated by varying each parameter incrementally from its calibrated or (for model defect parameters) fixed value, and then running the model. For the purpose of linear analysis, the ODGAB model was endowed with 1423 parameters. Calculation of sensitivities therefore required that 1424 model runs be undertaken. (The extra model run featured calibrated and fixed parameter values.) All of these model runs included a steady state stress period, followed by a transient simulation spanning the period 1918 to 2084.

Once sensitivities are available, the computational cost of evaluating linearized Bayes equation, and of undertaking singular value decomposition on a sensitivity matrix, is small. Calculations that are pictured and tabulated in this chapter are the outcome of a few minutes numerical work.

Because the complex environmental processes that are simulated by a numerical model are not actually linear, analyses that are based on matrix operations are approximate. This is the price to be paid for numerical efficiency. However, it is salient to ponder how much numerical work would have been required to undertake the analyses that are documented herein using nonlinear methods.

Evaluation of the prior uncertainty of a prediction would have required that the model be run many times. The parameter set employed on each model run would have comprised a sample of the prior parameter probability distribution. Evaluation of parameter contributions to uncertainty would have required repetition of this process with different groups of parameters held fixed. Differences in predictive uncertainties, with and without parameters held fixed, would then have been used to calculate parameter contributions to these uncertainties. It would have been necessary for the number of model runs employed for predictive uncertainty evaluation to be sufficient for the evaluated uncertainty, and for uncertainty differences, to stabilize.

For evaluation of parameter contributions to the post-calibration uncertainty of a prediction, it would have been necessary to repeat the above process using samples from the posterior parameter probability distribution rather than from the prior parameter probability distribution. Sampling of a posterior parameter probability distribution is a numerically difficult undertaking. It requires that random parameter fields (samples of the prior parameter probability distribution, or samples of a linear approximation to the posterior parameter probability distribution) be adjusted until model outputs match field measurements. While methodologies such as ensemble smoothers can undertake this task, the model run burden can be high. Repetition of the parameter adjustment process with certain parameters held fixed would have increased this burden dramatically. Subsequent evaluation of predictive uncertainty differences would have been approximate.

5. CONCLUSIONS

5.1 Repercussions for the ODGAB Model

Linear analyses that are documented in this report demonstrate that while simplifications that were embodied in the design of the ODGAB model advance its utility support potential in some ways, they may degrade that potential in other ways. Fortunately, linear analysis also demonstrates that degredation is only slight, and that it can be relatively easily addressed in next generation modelling.

In particular, linear analysis shows that the design of some ODGAB model boundary conditions increases the uncertainties of some decision-critical model predictions. At the same time, they impair the capacity of the modelling process to quantify these uncertainties. This impairment arises from two sources, namely:

- failure to endow these boundaries with a flexible parameterization scheme that characterizes uncertainties in their specifications; and
- a propensity for calibration-adjustable non-boundary parameters to adopt compensatory roles during history-matching in order to accommodate the fixing of boundary specifications at possibly erroneous values.

Because the consequence of these ODGAB model design features for management-salient predictions are only slight, they do not invalidate use of the model. This in itself is a useful outcome of the analyses discussed herein. Nevertheless, design of these boundary conditions can, and will, be improved in next generation modelling. Meanwhile, analyses documented herein show that the decision-support utility of an improved model will not suffer if these boundary conditions retain their somewhat abstract roles. The need to ensure that this is done in a way that allows quantification of abstraction-incurred predictive uncertainties is a challenge that faces the design of the improved model. This challenge is likely to be forgiving, as their contributions to the overall uncertainties of management-salient predictions do not appear to be large.

5.2 General Conclusions

The primary purpose of the present GMDSI report is to demonstrate tools and methodologies that are available to all modellers. These tools enable a modeller to inquire whether various aspects of a model's design that may have been introduced to expedite its capacity to serve the imperatives of decision-support modelling, may actually impair it.

For the ODGAB model, application of these methodologies suggests improvements in representation of some boundary conditions in subsequent versions of the model. At the same time, it shows that shortcomings in those aspects of its design which are the focus of the present report do little to impair its decision-support utility.

There will be other modelling contexts where this is not the case. A modeller must then decide on an appropriate course of action. Where model simplification increases the uncertainties that a model ascribes to decision-critical predictions in ways that are visible and quantifiable through linear or nonlinear analysis, the modeller can decide whether he/she wishes to trade the cost and effort of increased model complexity against the reduced uncertainties that these will yield.

Model simplifications which impair a model's ability to quantify predictive uncertainty, while possibly increasing this uncertainty through introduction of unquantifiable predictive bias, are of greater concern. The propensity for model simplifications to introduce these kinds of problems can be examined through analyses described herein.

We began this report with the often-repeated assertion that “all models are wrong, but some are useful”. We pointed out that a designation of “fit for purpose” often replaces “useful” in reports which describe specific models.

Modern methods of analysis allow us to go beyond this somewhat sententious statement of the obvious – that a numerical model constitutes a gross simplification of a real world system, and that this may erode the credibility of at least some of its predictions.

Predictions made by groundwater models are indeed wrong. This is because these predictions are uncertain. However their uncertainties can, and should, be quantified. Indeed this is one of the goals of decision-support modelling. Attainment of this goal requires that a model be complex enough to represent facets of system behaviour that contribute to the uncertainties of decision-critical predictions. At the same time, the model must be simple enough, and numerically stable enough, to undertake the many model runs that are required for assimilation of information that has the capacity to reduce the uncertainties of these predictions.

Model simplification may induce bias in some of its predictions. If possible, the potential for simplification-induced bias should be included in the uncertainties that are ascribed to these predictions. These, together with uncertainties that arise from simplicity-induced model-to-measurement misfit, then becomes a cost of simplification – a cost that may erode its decision-support utility. However if simplicity-induced uncertainty is small compared with quantifiable uncertainties that arise from gaps in expert knowledge, and if a simplified model can more readily reduce these latter uncertainties through data assimilation than a complex model, the benefits of model simplification outweigh its costs.

Decision support modelling should attempt to quantify uncertainties arising from all sources. These include information inadequacy, contamination of information by measurement error, and contamination of the process of model-based data assimilation by model inadequacies. This requires that a model include parameters that represent not only the properties of a simulated system, but also aspects of its design that may be subject to error, particularly those that replace processes that are not explicitly simulated by the model. Modern methods of data assimilation and uncertainty quantification can readily accommodate the large number of adjustable parameters that this may require. The numerical cost of quantifying real and model-induced uncertainties may therefore be relatively small.

There will be occasions, however, where calibration-adjustable parameterisation of model simplifications may not be easy to achieve. On these occasions, linear analysis methodologies that are exemplified in this report offer a modeller an alternative means of achieving similar outcomes. Analysis of the costs incurred by various model simplification strategies then requires that pertinent model design features, such as boundary conditions that replace more complex environmental processes, be endowed with parameters that are varied incrementally during model runs that are dedicated to calculating sensitivities of model outputs to these features. The costs of these simplifications can then be evaluated by replacing the action of the model on its parameters by that of a matrix on a vector. At the same time, these costs can be put into a context – the context that matters most from a decision-support modelling point of view. This context is defined by the uncertainties of decision-critical predictions as they depend on availability and quality of data.

6. REFERENCES

Dausman, A.M., Doherty, J., Langevin, C.D. and Sukop, M.C., 2010. Quantifying data worth toward reducing predictive uncertainty. *Ground Water*, 48(5): 729-740.

Doherty, J. 2015. Calibration and Uncertainty Analysis for Complex Environmental Models, 227. Brisbane: Watermark Numerical Computing. ISBN: 978-0-9943786-0-6. Available from: http://www.pesthomepage.org.

Doherty, J. and Hunt, R.J., 2009. Two statistics for evaluating parameter identifiability and error reduction. *Journal of Hydrology*. 366, 119-127.

Doherty, J. and Simmons, C.T., 2013. Groundwater modelling in decision support: reflections on a unified conceptual framework. *Hydrogeology Journal*. 21: 1531-1537.

Doherty, J. and Welter, D., 2010, A short exploration of structural noise, *Water Resour. Res*., 46, W05525, doi:10.1029/2009WR008377.

James, S.C., Doherty, J.E., and Eddebbarh, A., 2009. Practical postcalibration uncertainty analysis: Yucca Mountain, Nevada. *Ground Water*, 47 (6): 851-869.

Love, A.J., Wohling, D., Fulton, S., Rousseau-Gueutin, P. and De Ritter, S (eds), 2013. Allocating Water and Maintaining Springs in the Great Artesian Basin, Volume II: Groundwater Recharge, Hydrodynamics and Hydrochemistry of the Western Great Artesian Basin. *National Water Commission*, Canberra

Moore, C. and Doherty, J., 2006. The cost of uniqueness in groundwater model calibration. *Advances in Water Resources*, 29: 605-623.

Mudd, G.M., 2000. Mound springs of the Great Artesian Basin in South Australia: a case study from Olympic Dam". *Environmental Geology*. 39 (5). doi:10.1007/s002540050452.

Freeze, R.A., Massmann, J., Smith, L., Sperling, T., and James, B., 1990. Hydrogeological decision analysis: 1, a framework. *Ground Water* 28(5):738–766

Hill, M.C. and Tiedeman, C.R., 2005. *Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty.* John Wiley & Sons, Ltd.

Panday, S., Langevin, C.D., Niswonger, R.G., Ibaraki, M. and Hughes, J.D., 2013. MODFLOW-USG version 1: An unstructured grid version of MODFLOW for simulating groundwater flow and tightly coupled processes using a control volume finite-difference formulation. *U.S. Geological Survey Techniques and Methods. Book 6.*

Pianosi, F., Sarrazin, F., and Wagener, T., 2020. How successfully is open-source research software adopted? Results and implications of surveying the users of a sensitivity analysis toolbox. *Environmental Modelling and Software*., 124, https://doi.org/10.1016/j.envsoft.2019.104579

Rousseau-Guetlin, P, Love, A.J., and Smerdon, B.D., 2012. Chapter 7. Regional Water Budget. In: Smerdon, B.D., Welsh, W.D. and Ransley, T.R. (eds). Water resource assessment for the Western Eromanga region. A report to the Australian Government from the CSIRO Great Artesian Water Resource Assessment. *CSIRO Water for a Healthy Country Flagship*, Australia.

Saltelli, A., Tarantola, S., Campolongo, F. and Ratto, M., 2004. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models.* John Wiley & Sons, Ltd.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S., 2007. *Global Sensitivity Analysis. The Primer.* John Wiley & Sons, Ltd.

Sampson, L., Wohling, D., Jensen-Schmidt, B. and Fulton, S., 2012. South Australia and Northern Territory Great Artesian Basin (Eromanga Basin) Hydrogeological Map – part 2. *Department of Environment, Water and Natural Resources, South Australian Government*.

Watermark Numerical Computing, 2015. PEST: Model-Independent Parameter Estimation. Downloadable from http://www.pesthomepage.org.

White, J. T., 2018. A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions. *Environmental Modelling and Software.* 109:191-201.

White, J.T., Doherty, J.E. and Hughes, J.D., 2014. Quantifying the predictive consequences of model error with linear subspace analysis. *Water Resour. Res, 50 (2): 1152-1173.* DOI: 10.1002/2013WR014767

White, J.T., Fienen, M.N. and Doherty, J., 2016. A python framework for environmental model uncertainty analysis. *Environmental Modelling & Software*. 85, 217-228.

gmdsi

Groundwater Modelling
Decision Support Initiative

gmdsi.com

CRICOS NO 00114A

BHP

Flinders University

NATIONAL CENTRE FOR GROUNDWATER RESEARCH AND TRAINING

RioTinto