Rapid Data Assimilation

using an appropriately complex model

A GMDSI worked example report

DRAFT

by Neil Manewell and John Doherty

gmdsi
Groundwater Modelling
Decision Support Initiative

BHP
Flinders UNIVERSITY
NATIONAL CENTRE FOR GROUNDWATER RESEARCH AND TRAINING
RioTinto

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

The National Centre for Groundwater Research and Training, Flinders University advises that the information in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice.

CITATION

For bibliographic purposes this report may be cited as: Manewell, N. and Doherty J. (2021). Rapid Data Assimilation Using an Appropriately Complex Model. A GMDSI worked example report. National Centre for Groundwater Research and Training, Flinders University, South Australia.

ISBN: 978-1-925562-53-8

DOI: 10.25957/0vp9-a807

DOI as a weblink: https://doi.org/10.25957/0vp9-a807

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 various GMDSI worked examples include the following:

- assessing the reliability of a public water supply;
- protection of a groundwater resource from contamination;
- estimation 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 often 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 parameterise 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

This GMDSI report addresses a number of related issues. They include:

- appropriate model complexity;
- appropriate parameterisation complexity;
- efficient model-based assimilation of information-rich data; and
- linear analysis.

The focus of modelling work that is reported herein is BHP's Orebody 31 (OB31) situated in the Pilbara region of Western Australia. The environs of this mine have been the focus of a number of generations of modelling, some of which is described in the present report. Mining of OB31 commenced in 2016; however data collection and modelling took place for a number of years prior to that.

The area is geologically complex. Iron ore is found in steeply dipping beds which are transected and offset by local faulting.

One of the tasks with which pre-mine modelling was charged was that of predicting how much water must be pumped from the ground in order to keep the OB31 pit dry. This prediction has ramifications for the design of mine-support infrastructure. This infrastructure includes a pipeline to convey pumped water to a nearby dam; the diameter of the pipe must be adequate for the water that it must convey.

Because the local geology is so complex, dewatering rate predictions are likely to be accompanied by considerable uncertainties. However, data were acquired that have the capacity to reduce these uncertainties to at least some extent. In late 2014, a series of 6 constant-rate pumping tests of between 5 and 11 days duration were undertaken in order to gain some insights into the disposition and variability of subsurface hydraulic properties. Drawdowns and recoveries were measured in 21 observation wells. Collectively this dataset is referred to as "the CRT" (for "constant rate test").

Unfortunately, modelling which preceded that which is described herein had difficulty in assimilating these data. Part of the reason for this lay in their complexity of model design. Among their design criteria was the need for "accurate" three-dimensional portrayal of complex geology, as far as it was understood at the time of modelling. Two previous models that are discussed in this report each employed 7 layers for this reason; this resulted in lengthy run times. At the same time, parameterisation of these numerically complex models was based on a limited number of zones of piecewise hydraulic property constancy whose geometry reflected the presumed dispositions of known geological units. However, in order to accommodate evidence embedded in the CRT and other pumping-based datasets that hydraulic connections may cross the prevailing geological strike, a small number of elongate zones that transect local geology were introduced to the domains of these models at locations where it was thought that such structures may exist.

Attempts at data assimilation based on this model design philosophy were somewhat disappointing. Model-to-measurement fits attained with the CRT dataset were not very good; hence information which was resident in that dataset was denied the opportunity to inform model parameters. The latter were too few in number, and were defined too inflexibly, to receive this information. Model predictions of pit dewatering requirements were too low. While attempts were made to quantify the post-calibration uncertainties associated with these predictions, failure to attain good fits with the calibration dataset, and parameter insufficiency, reduced the credibility of calculated uncertainty limits.

In late 2020, GMDSI personnel “went back in time” in order to establish whether an alternative approach to decision-support modelling based on data assimilation may have allowed better use of CRT data, and may have promulgated better predictions of pit dewatering requirements.

There are two major differences between the modelling approach that is described in the present report and that which preceded it. The first is the philosophy of model design. The model described herein was designed with history-matching in mind. Though simple in its construction, the model employs many parameters. Embodiment of prevailing hydrogeological concepts was not its primary parameterisation design criterion. Rather, its parameterisation scheme was designed to respond to the information content of a calibration dataset in flexible ways that support assimilation of these data by providing them with a blank canvas on which to draw.

The second departure from previous modelling approaches is use of an efficient inversion methodology that enables data assimilation to be carried out at a relatively small cost in terms of model runs. This supports the use of many parameters, adding to the ability of the modelling workflow to assimilate information.

Calibration of the new OB31 model introduced patterns of hydraulic property heterogeneity to the model domain that enable model outputs to match nuances of CRT drawdowns and recoveries very well indeed. Some aspects of these patterns can be readily associated with aspects of the prevailing geological model. Other aspects of these patterns have shapes and dispositions that suggest a structural origin. Other emergent parameter value patterns are more difficult to interpret. Nevertheless the inversion process from which they arose, and post-calibration linear analysis, suggest that these patterns are worthy of being taken seriously. Although the locations and dispositions of three-dimensional heterogeneity projected onto a single model layer do not reproduce geological reality, the effects of this projected heterogeneity on pumping-induced drawdowns and recoveries are nevertheless real.

The calibrated model was used to predict inflow to a hypothetical OB31 pit. Because the pit is dewatered by pumping from nearby bores, this prediction cannot be directly compared with real-world measurements. Nevertheless, it is somewhat gratifying to note that model-predicted inflows are similar to current pit dewatering rates. This may indeed be a coincidence. What is of greater importance, however, are suggestions arising from modelling described herein that large volumes of stored water have the ability to flow large distances through zones of connected permeability towards the pit dewatering system.

Linear analysis was employed to associate uncertainties with predictions of pit inflow. These are no-doubt understated because of the relative simplicity of the model, and because of the proximity of model boundaries to the pit. Nevertheless, linear analysis effectively demonstrates the high information content of the CRT dataset with respect to predictions of pit dewatering requirements. Specifically, it suggests that this dataset carries about 200 separate items of information. (Each such item of information has the capacity to assign a value to a unique linear combination of model parameters.) Furthermore, this information has the potential to impose significant reductions on uncertainties associated with dewatering rate predictions from those which arise from expert knowledge alone.

CONTENTS

1. INTRODUCTION

1.1 Background

This GMDSI report begins by describing modelling work that was undertaken during 2014 and 2015 to predict dewatering requirements for a planned open cut mine. (At the time of writing - early 2021 - the mine has been operating for over 4 years.) Despite attempts to faithfully represent prevailing geology, and to undertake rudimentary history-matching against pumping-induced drawdowns, dewatering rates were under-predicted by the model.

The report then goes on to document modelling that was dedicated to extracting as much information as possible from time-varying drawdowns induced by pumping from 6 production bores. This modelling was able to expose permeable pathways that are likely to have a significant impact on mine dewatering rates. This revised modelling strategy employs inversion methodologies that were not available in 2015, but are available now.

BHP extracts iron ore from the "Orebody 31" (i.e. OB31) deposit in the Pilbara region of Western Australia. Much of this ore lies below ambient groundwater levels in rocks of high permeability. The pit must be kept dry by extracting water from a series of nearby bores. Prior to the commencement of mining operations, predictions of dewatering requirements were required by BHP in order to obtain a pumping license, and in order to design a pipeline to convey extracted water to a dam which is situated 21 km away.

Over a period of about six months, starting in December 2014, 6 pumping tests were undertaken in the vicinity of the proposed pit. The durations of these tests varied between 5 days and 11 days. Drawdowns and recoveries were monitored in up to 21 observation wells. Collectively, these tests are referred to as "the CRT" (constant rate test), in technical reporting and in the present report.

In 2014 and 2015, drawdown and recovery measurements comprising the CRT dataset were used to refine the parameterisation of a large, multi-purpose groundwater model. This model was built in order to assess pit dewatering requirements, and in order to assess the impacts of this dewatering on receptors of interest, some of which are up to 15 km from the mine. In order to reflect what was known of the prevailing geology, the model was complex, with its grid comprising 1,366,827 active cells disposed across seven layers. The geology is marked by steeply dipping beds, tightly folded into neighbouring anticlines and synclines. The lithologies which comprise these beds have markedly different hydraulic properties which vary along strike. The region is pervaded by a series of faults, the dispositions and hydraulic significance of which are only partially known.

This complex local geology was represented rather awkwardly in the model because of the discretised nature of its finite-difference grid. Nevertheless, representation of geology, however cumbersome, was considered a fundamental precursor to good model performance. Like the model grid, parameterisation of the model was also designed to respect current understandings of local geology. Parameterisation was based on zones ascribed to discrete geological units and to a handful of discrete structural features that were represented in the model. To ease the burden of calibration, parameter numbers were kept low. During the calibration process, new parameters were introduced to the model by subdividing existing zones where it was thought this was needed in order to improve model-to-measurement fit.

Model calibration was pursued through manual adjustment of parameter values. Unfortunately, it was not possible to obtain a particularly good fit between model outputs and CRT drawdowns and recoveries. The fit with this important component of the calibration dataset was deemed to be good enough when its broad features were replicated. While it was recognised that some details of the drawdown and recovery time series may contain important information on connected near-pit permeability, these details could not be fit by the model. Part of the reason for this was deployment of a coarse parameterisation scheme based on an approximate representation of known and supposed geology. Though not openly stated, acceptance of failure to fit fine, but important, details of CRT drawdown and recovery was justified by the premise that a numerical groundwater model cannot be expected to fit every nuance of system behaviour, and that attempts to do so would constitute “overfitting”. If a choice had to be made between approximate replication of known and implied complex geology on the one hand, and fitting of a complex calibration dataset on the other hand, the former was implicitly decreed to be of greater importance.

Further refinement of model parameterisation took place after conduction of a “hydrodynamic trial” (HDT) in the vicinity of the proposed pit. Three wells were pumped continuously over a period of about 3 months while drawdown was monitored in 26 observation wells; recovery was monitored for a further month. Observed drawdowns were much larger than for the CRT. Attempts to fit HDT drawdowns, and particularly recoveries, suggested the presence of connected transmissivities that cross prevailing geological strike. In doing so, they link the OB31 orebody to a regional aquifer. Their presence has the potential to markedly increase estimates of pit dewatering requirements.

By introducing cross-strike linear zones of enhanced transmissivity, and by employing a combination of manual and machine-based history-matching, reasonable fits with the HDT dataset were attained using a model of reduced areal footprint but which retained the same 7 layers as the original model. However model-to-measurement fits were far from perfect. Nevertheless, they allowed the model to make estimates of pit dewatering requirements that are better aligned with those that are now being experienced.

1.2 The Present Study

At the suggestion of BHP personnel, GMDSI personnel reinterpreted 2014 CRT data using a different modelling approach from that which had previously been implemented at the OB31 site. Specifically, the aims of the study were as follows:

1. To investigate whether information that is resident in the CRT dataset can shed light on aspects of local hydrogeology that impact mine dewatering requirements if these data are interpreted using modern inversion technology;
2. To inquire whether decision-support modelling workflows at similar sites may be in need of revision. Of particular relevance is the relationship between a conceptual model and a numerical model.

Conventional wisdom dictates that development of a detailed conceptual model precedes that of a numerical model, and that the role of the latter is to give numerical voice to the former. The present study suggests that conceptual and numerical model development should be seen as parallel rather than serial activities. Attempts to interpret CRT data using a numerical model that embodied current hydrogeological concepts proved fruitless because the history-matching process did not allow the latter to be questioned. This, and an unstated premise that underfitting of data constitutes a valid means of avoiding predictive bias, precluded the possibility that the history-matching process could result in anything other than vindication of an inadequate conceptual model.

The present study suggests that the concept of a “conceptual model” should be broadened. As is discussed herein, a primary specification of decision-support modelling is that it must provide the means to assimilate data that has the potential to reduce the uncertainties of management-critical predictions. The design of a modelling workflow, and of the numerical model through which that workflow is implemented, must include the means by which such data can indeed be assimilated. The numerical model must be imbued with a parameterisation scheme that can provide receptacles for information that is resident in these data. At the same time, the history-matching process through which parameters are estimated must be capable of accommodating "surprises". In doing so, it must invite, rather than suppress, the introduction of unusual parameter shapes and values. Where these lack geological credibility, they may indicate that parameters are adopting roles that compensate for numerical (and hence conceptual) model defects; the latter can then be investigated. Alternatively, they may indicate the need to revise concepts which define "geological credibility".

Another important aspect of decision-support model design that is exposed by the current study is the need for a decision-support model to possess many parameters – generally many more parameters than can be estimated uniquely. Inflexible parameterisation schemes based on immovable zones of piecewise parameter constancy that respect questionable geological boundaries that are poorly replicated by a coarse numerical grid should be eschewed in favour of a multitude of pilot points, or even cell-by-cell parameterisation. When using the latter parameterisation schemes, spatial patterns that represent current geological concepts can be suggested (rather than decreed) through design of an appropriate regularisation methodology. Prevailing concepts can then be respected at the same time as they can be challenged.

This study also demonstrates that while the geology of a particular site may be complex, it does not follow that a decision-support numerical model focussed on that site should be complex. Model complexity induces long run times, and provides fertile ground for numerical instability. These erode a model's ability to be used in partnership with programs from the PEST and PEST++ suites that can assimilate information that is resident in historical measurements of system state and fluxes.

However, problems associated with excessive model complexity go deeper than this. Complexity expresses detail – detail that may not be relevant to a particular management issue. Detail is uncertain; hence it must either be represented stochastically or suffer misrepresentation. Stochastic representation of "picture perfect" geology in ways that allow a model to replicate past system behaviour constitutes an extremely demanding numerical task.

An alternative approach to decision-support modelling at a particular site is to forego the expression of "geologically realistic" detail in favour of a more abstract representation of only those aspects of geology that are salient to a particular management problem. If such a model can be populated by a flexible parameter field that enables it to reproduce the past and constrain the future within quantifiable uncertainty limits, then this approach serves the decision support process far better than construction of a "realistic" model which can do neither.

Finally, this study demonstrates that the decision-support process is best served by a modelling workflow that is designed to address a specific management issue. A decision pertaining to the inclusion or omission of a specific subsurface feature or process in a model's construction and parameterisation must follow an assessment of the costs and benefits of its inclusion or omission. Without a defined modelling purpose, costs and benefits have no reference point. In contrast, if a model is burdened with the onerous responsibility of "simulating a system", then it can dispense with little complexity. This will compromise its ability to make predictions of system behaviour whose potential for error has been minimized through assimilation of pertinent data. Even more importantly, it will compromise its ability to quantify the possible magnitude of that error.

The conceptual model on which a decision-support numerical modelling workflow rests cannot ignore these concepts. It must include the notion that it is open to challenge, and subject to revision. It must form the basis for design of a numerical modelling process that enables flexible history-matching that blends information born of expert knowledge with that which is resident in measurements of system behaviour. To render this process tractable, a conceptual model should provide as much guidance on what should be omitted from a numerical model as it does on what should be included.

1.3 This Report

This report is organised as follows.

Chapter 2 expands the discussion that was started above on the role of the conceptual model in decision-support modelling practice. It then discusses history-matching in general, and the role of model calibration in particular. It also provides brief coverage of a number of topics that are related to history-matching and which are pertinent to the work which is reported herein.

Chapter 3 describes the study area, and presents a short history of decision-support modelling that has been undertaken to address its issues.

Chapter 4 describes the model that is the focus of the current study. It also describes its parameterisation, its calibration, and use of the calibrated model to predictions pit dewatering requirements. Using linear analysis, the uncertainty associated with this prediction is examined, as are uncertainties associated with estimated parameters.

Chapter 5 concludes this report with a brief discussion of decision-support modelling principles that are illustrated by OB31 modelling.

2. DECISION-SUPPORT MODELLING: SOME PRINCIPLES

2.1 The Decision Support Imperative

The role of groundwater modelling in decision support is discussed in other GMDSI reports, as well as in papers such as Freeze et al (1990), Doherty and Simmons (2013) and Doherty and Moore (2019). Its task is to:

- Make management-salient predictions of groundwater behaviour;
- Quantify the uncertainties of these predictions so that decision-makers are fully acquainted with the risks that accompany a contemplated course of management action;
- Where necessary, reduce the uncertainties of management-salient predictions by assimilation of pertinent data.

The above imperatives set the context for the many design choices that a modeller must make as he/she builds a decision-support model. It is important to note that these imperatives do not include the necessity for a model to provide faithful replication of underground processes. The decision-support utility of the modelling process rises in proportion to its ability to provide receptacles for decision-pertinent information. The existence of these receptacles is an outcome of a model's ability to simulate, albeit in an approximate and incomplete manner, processes that affect management. The number of these receptacles, and their integrity as repositories of decision-pertinent information, may be related only loosely to a model's fidelity as a simulator.

Simulation requires that values be assigned to parameters. Some of these values determine simulation outcomes. In many cases they reflect hydraulic properties of the subsurface. As such, they may be partially informed by site characterisation studies. They may also be partially informed by history-matching. Knowledge and data which informs management-salient parameters informs management-salient predictions. Parameters of a model can therefore be viewed as receptacles for management-salient information. It is the task of the decision-support modelling process to capture this information and use it. By definition, information is "used" when it reduces uncertainty.

It follows that a model's parameters are key to its decision-support role. A model's awkward attempts at simulation bring parameters into existence. Integrity of simulation need only be such as to support the integrity of parameters as receptacles for decision-relevant information. The task of decision-support modelling is to deliver information to these receptacles, and then from these receptacles to management-salient predictions. This does not happen as a convenient by-product of a model's ability to simulate subsurface processes, for no model can simulate subsurface processes very well. Instead this task should be woven into the fabric of the decision-support modelling workflow as it is applied at any particular site.

2.2 The Conceptual Model

Conventional modelling wisdom dictates that a groundwater model give numerical voice to a conceptual model. Ideally, a conceptual model is the product of hydrogeological investigations. The more extensive are these investigations, the more detailed is likely to be the conceptual model.

At the heart of the conceptual model is a geological model. A geological model specifies (among other things):

- rock types that prevail within a study area;
- juxtapositional relationships between different rock types;
- the nature and depth of weathering;
- the presence or secondary features such as faults that can alter juxtapositional relationships, and that can introduce connected alterations to hydraulic properties;
- hydraulic properties associated with different geological media;
- the locations and mechanisms of recharge and discharge;
- the nature of interactions between surface and subsurface waters.

Where groundwater modelling supports management of mining-related activities, the geological model that is associated with a conceptual model may be detailed. Where geology is complex, so too may be the conceptual model.

It is tempting to assume that detailed characterisation of geology provides the groundwater modelling process with a “head start”, and that its predictions therefore have the potential to be reasonably accurate. It may therefore be surmised that it is incumbent on a modeller to represent what is known of local geology as well as possible in his/her model. Failure to do so may prompt severe criticism of his/her model on the basis that if it does not replicate geometrical relationships between rocks, then it cannot replicate flow of water through them.

There may be modelling contexts where this argument holds true. There are also those where it does not hold true.

Groundwater flow is predominantly horizontal. This can be easy to forget, as most vertical sections through geological and groundwater models undergo significant vertical exaggeration before display. Where horizontal aquitards create vertical head differences, and where vertical head differences are salient to a prediction, then obviously the vertical direction matters. However in other cases, decision-salient predictions, and the uncertainties associated therewith, may be more affected by lateral hydraulic property variations than by vertical hydraulic property variations. If this is the case, the need for a model to provide faithful representation of vertical geological connections fades. This is replaced by the necessity to represent lateral variations in hydraulic properties that can help or hinder the flow of groundwater in the horizontal direction.

It must also be recalled that uniform colours on a geological map or section do not imply uniform hydraulic properties. While shale is less permeable than dolomite, the hydraulic properties of both of these rocks can vary considerably with the extent to which they have been subjected to diagenetic and/or weathering-based alteration. Furthermore, where an area has been subjected to intense deformation, continuity of aquifers or aquitards may be interrupted in unknown ways in unknown places, this impairing their capacity to assist or hinder the flow of water. The presence of faults, dykes and other linear features adds even greater complexity to an already complex situation. Faults can enable flow parallel to their planes while interrupting flow in perpendicular directions.

Because they are incompletely known, hydraulic properties should be represented probabilistically in a groundwater model. Prior to history-matching, the probability distribution that characterises them is referred to as their “prior probability distribution”. Ideally, this distribution is a distillation of expert hydrogeological knowledge. As such, it constitutes an important component of any conceptual model. The prior probability distribution of hydraulic properties reflects their range of possible values, as well as the lengths and directions over which anomalous properties are likely to exhibit spatial correlation. Conceptually, it is not a difficult matter to populate a model with hydraulic property realisations that constitute samples of this prior probability distribution. Using modern history-matching techniques, these realisations can then be modified so that pertinent model outputs conform with measurements of system state. Samples of the prior parameter probability distribution are thus moulded into samples of the posterior parameter probability distribution. If the model is then run using each of these samples, the posterior probability distribution of a decision-salient prediction can be evaluated. The imperatives of decision-support modelling are thereby served.

A conceptual model must be flexible enough to identify contexts where the above approach to decision-support modelling is impractical. This can occur where:

- the geology is extremely complex;
- the continuity of geological units is unknown;
- intra-formational heterogeneity is high;
- there is a high likelihood that structural features may offset geological units, and/or interrupt groundwater flow;
- the disposition of model grid cells is too coarse to properly represent complex geological juxtapositional relationships.

In cases such as these, a more useful conceptual model may comprise a relatively simple stochastic representation of management-pertinent hydraulic property variability based on parameters whose values are readily adjusted through history-matching while retaining the ability to exhibit post-history-matching stochastic variability. This is the approach that is adopted in the present study.

2.3 History Matching

2.3.1 Options

History-matching provides the means through which information that is resident in system behaviour can inform model parameters. However certain tenets must be followed if history-matching is to be maximally effective in achieving this outcome, particularly in complex geological environments where hydraulic property uncertainties are high.

The first tenet is that a model should be endowed with many parameters. This grants the history-matching process flexibility when responding to information that is resident in field measurements of system behaviour. (It was not so long ago that modellers were advised against “over-parameterisation” on the basis that it would lead to over-fitting of field measurements, and that over-fitting would incur predictive bias. Modern methods of regularised inversion make this advice irrelevant at best, and harmful at worst.)

As already mentioned, two different approaches to history-matching are available to groundwater modellers. These are not to be seen as competing, but as complementary. Each relies on the premise that subsurface hydraulic property variability is more complex than can be represented uniquely in a model. History-matching therefore endeavours to capture as much of this variability as the calibration dataset allows. That which cannot be captured must be represented probabilistically. This serves the dual decision-support imperatives of reducing and quantifying predictive uncertainty.

Bayesian methods of history-matching eschew uniqueness. As described above, they require that a modeller generate random samples of the prior parameter probability distribution. These parameter fields are then modified until pertinent model outputs fit the calibration dataset. They thereby comprise samples of the posterior parameter probability distribution. The PESTPP-IES ensemble smoother implements history-matching according to this philosophy. See White (2018) and references cited therein for further details.

In contrast, inversion methods seek a single parameter field which fits the calibration dataset well. Their quest for uniqueness is not, however, a quest for correctness. Instead, uniqueness is attained by purposefully minimizing their potential for incorrectness. This potential can still be high where hydraulic property variability is complex, and where the information content of a calibration dataset is low. Nevertheless, a spatially-simplified parameter field that lies somewhere near the centre of the posterior parameter probability distribution can achieve this status of minimised error variance. Predictions that are based on this parameter field inherit this status. Parameter and predictive error variance can then be quantified through post-calibration uncertainty analysis.

To put it another way, Bayesian methods explore the "heterogeneity which CAN exist", that is compatible with field measurements of system behaviour. In contrast, inversion methods seek to define the "heterogeneity which MUST exist" in order to explain observations of system behaviour. In the geological context which is the subject of the present report, wherein prior parameter probability distributions are only vaguely known, inversion offers a suitable approach to history matching because:

- it often allows better fits to be attained with informative nuances of field datasets than do Bayesian methods;
- it serves the investigative role that history-matching must play in complex geological contexts.

2.3.2 Regularisation

"Regularisation" can be loosely defined as "whatever it takes to attain uniqueness".

In the following pages of this report we discuss calibration of a model that is furnished with 3,526 parameters – many more than can be endowed with unique values. This is an essential decision-support modelling strategy. For the study reported herein, the use of many parameters allows the inversion process to introduce heterogeneity to any location within the model domain that it deems necessary as it fits model-calculated drawdowns and recoveries to those measured during 6 pumping tests. Field data can thus speak freely; freedom of data expression can be achieved only with a superfluity of parameters.

However, in order to avoid parametric chaos, parameter uniqueness must be sought according to a strict set of conditions that are embodied in so-called "Tikhonov constraints". Application of these constraints requires that parameter heterogeneity be introduced to the model domain only if its existence is essential to the attainment of a required level of model-to-measurement fit, and only in ways that are orderly. This allows a modeller to learn from the parameter patterns that emerge from the inversion process, as it challenges him/her to give these patterns meaning. At the same time however, a modeller must be cautious about taking these patterns too literally. They may accurately reflect an underlying reality where data are plentiful, but are only indicative of reality where data are scarce.

2.3.3 Implementation

In the study documented herein, history-matching was undertaken using PEST_HP (Doherty, 2020). PEST_HP improves parameters by adjusting them according to the following equation.

$$\delta\mathbf{k} = (\mathbf{J}^t\mathbf{QJ} + \mu^2\mathbf{k}^t\mathrm{C}^{-1}(\mathbf{k})\mathbf{k} + \lambda\mathbf{I})^{-1}\mathbf{J}^t\mathbf{Qr} \tag{2.1}$$

In equation 2.1, all bold variables that employ lower case letters are vectors; a vector is a column of numbers. Capitalised bold variables are matrices; matrices are arrays of numbers. The "t" superscript denotes the transpose of a matrix or vector; this flips it on its side. The "-1" superscript denotes the inverse of a matrix. Evaluation of the inverse of a large matrix can be a numerically intensive procedure.

In equation 2.1:

δk characterises adjustments that are made to parameters;

r (for "residuals") denotes current model-to-measurement misfit;

Q is a diagonal matrix of observation weights;

C(**k**) is a user-supplied prior parameter covariance matrix whose purpose is to guarantee spatial orderliness of estimated parameters;

J is the Jacobian matrix;

μ^2 and λ are scalar coefficients that are determined by PEST_HP; λ is the so-called "Marquardt lambda".

A modeller provides PEST_HP with an initial set of parameter values. PEST_HP then improves them. It then improves them again, and again, until it can improve them no more. Each iteration of this process gives numerical voice to equation 2.1, drawing on assistance provided by powerful numerical tools such as singular value decomposition to prevent numerical instability – another cause of potential parameter chaos. During each iteration of the inversion process, many different parameter upgrades are calculated and tested; these are calculated using different values of the Marquardt lambda. The parameter set which attains the best fit with the calibration dataset is adopted as the starting point for the next iteration.

The most time-consuming part of the inversion process is the filling of the **J** matrix. A new **J** matrix is required for every iteration. Each column of **J** contains the partial derivatives (i.e. sensitivities) of all model outputs used in the inversion process to a single parameter. Sensitivities comprising the elements of **J** can be calculated by varying each parameter incrementally, running the model, and then dividing alterations to model outputs by parameter increments. Where a model employs 3,526 parameters (as does the model described herein), this requires 3,526 model runs (or twice this number for a more accurate finite difference approximation to true derivatives).

In the worked example described herein, two strategies were implemented to reduce the numerical burden of repeated model runs. The first was to simplify the model until it was no more complex than it needed to be. Model design is discussed in the next chapter. The second strategy was implementation of "randomised Jacobian" functionality accessible through PEST_HP. (Similar functionality is available through PESTPP-IES.)

Instead of running the model 3,526 times during each iteration of the inversion process in order to fill the Jacobian matrix, PEST_HP ran the model 600 times during early iterations and 1,000 times during later iterations. Instead of incrementing only a single parameter during each model run, PEST_HP incremented all parameters in a random fashion. Using model outputs computed from random parameter increments, it then computed an approximation to the Jacobian matrix. This approximation was good enough for use in equation 2.1; furthermore, PEST_HP's implementation of this process ensures that the integrity of the randomised Jacobian matrix as an approximation to the true Jacobian matrix improves as the inversion process progresses.

For those interested, some aspects of the randomised Jacobian process implemented by PEST_HP are listed in Table 2.1. Refer to the PEST_HP manual for more details.

Feature	Explanation
Jacobian retainment	The Jacobian matrix used in each iteration of the inversion process retains part of that calculated during the previous iteration. Rank-deficiency of the Jacobian matrix is therefore reduced as the inversion process progresses.
Autolocalisation	Model-output-to-parameter sensitivities are reduced or zeroed if they are below a certain threshold; this protects them from being assigned spurious non-zero values.
Broyden update	During each iteration of the inversion process, a suite of Marquardt lambdas is used to calculate potential parameter updates. Model outputs calculated using these potential updates are used to improve the Jacobian matrix. The lambda-based parameter update calculation process is then repeated.
Mode of SVD	The PEST SVDMODE variable is set to 2. Singular value decomposition is therefore conducted on $\mathbf{Q}^{½}\mathbf{J}$ instead of $\mathbf{J}^{t}\mathbf{QJ}$. Under these circumstances, the Marquardt lambda flattens the singular value spectrum. This results in improved parameter updates where the filling of **J** is based on random parameter increments.
Approximate regularisation	The PEST REG2MEASRAT variable is set to 0.1. This speeds up calculation of μ^2 in equation 2.1. It also makes the inversion process more immune to imperfections in **J** arising from the use of random parameter increments.

Table 2.1 Some details of the PEST_HP randomised Jacobian process.

2.4 Appropriate Model Complexity

The geological complexity of the OB31 site has already been mentioned. More details are provided in the following chapter of this report.

The model that was originally used to interpret CRT data was designed to serve two primary purposes. These were:

1. assessment of pit dewatering requirements; and
2. assessment of the impact of dewatering-induced drawdown on all possible receptors.

The need to serve these two purposes added to the size and complexity of the model. Its complexity was compounded by the perceived need to provide a “faithful representation” of near-pit hydrogeology.

The original model’s inability to match drawdowns and recoveries induced by 6 pumping tests demonstrated that its representation of near-pit hydrogeology was anything but faithful. The complexities of local geology, the complex relationships that link geology to hydraulic properties, and the discretised nature of a numerical model grid suggest that detailed replication of local hydrogeology may not possible. Nor, indeed, may it be necessary. If this is the case, then the conceptual model on which a groundwater model rests should embody this concept.

As is described below, in order to extract as much information as possible from the CRT dataset, the OB31 model was rebuilt. In particular:

1. A new model grid was constructed using the ALGOMESH grid generation package. This allowed economy of model cells while enabling grid refinement near pumping wells.
2. The domain of the model was reduced to an area that includes all pumping and observation wells as well as geological units to which drawdowns incurred by up to 11 days of pumping may extend. This area included the OB31 pit.
3. The number of model layers was reduced from 7 to 1.

The modelling process was thus focussed unequivocally on data interpretation. Meanwhile, the new model's run time was roughly 1 second. Model stability was irreproachable. The inversion process faced no numerical challenges.

It may be argued that use of a single layer invalidates the simulation process, for it fails to acknowledge that geological layers dip steeply, and that groundwater flow may have a vertical component. White et al (2014) and Doherty (2015) point out that inappropriate model simplification may have the following undesirable consequences.

1. It may prevent a good fit from being attained between model outputs and members of a calibration dataset.
2. Model parameters may adopt roles that compensate for model imperfections as they are adjusted to fit a calibration dataset; this may bias some model predictions.
3. The uncertainties of some decision-pertinent model predictions may be under-stated.

As will be seen, the first of the above consequences is of only minor concern in the present study. The second of the above consequences may be of some concern. However, its effect on predictions of mine dewatering requirements is at least partially mitigated by the similarity of these predictions to observations comprising the calibration dataset. Under these circumstances, the above authors show that predictive bias is actually "calibrated out".

The third of the above consequences may indeed be relevant to the present study. However it should be considered in context. The addition of even a single extra layer to the revised OB31 model that is described in the following chapter of this report would more than double its parameter requirements. While this would not increase the burden of randomised Jacobian matrix construction, it would add greatly to parameter nonuniqueness. This may (or may not) increase the uncertainties ascribed to predictions of pit inflow. Linear analysis described in Chapter 4 reveals that the CRT calibration dataset is rich in information; at the same time, the simplified OB31 model leaves parameters with plenty of "room to move" to express deficits in this information. This suggests that uncertainties ascribed to predictions of pit dewatering requirements would not increase greatly if the number of model layers were doubled, and the number of model parameters were thereby more-than-doubled. The same applies if the number of model layers is trebled or quadrupled.

We close this subsection by noting that simplicity of model design has benefits that extend beyond those of numerical utility.

In the inversion process that is described in Chapter 4 of this report, the simplest parametric explanation for drawdowns and recoveries induced by pumping in 6 wells is sought. As was discussed above, some aspects of the emergent patterns of hydraulic property heterogeneity should not be taken too literally, while others should. However, all of them are important. All of them are informative. All of them are indicative of some aspect of the groundwater system that influences flow of water in and around the planned OB31 pit as it responds to pumping. Their representation as two-dimensional patterns that can be readily superimposed on mapped geology maximises the extent to which a modeller and his/her associates can understand what the CRT dataset has revealed. The didactic outcomes of the data assimilation process may have suffered if these patterns were spread over many model layers in uncertain and arbitrary ways.

3. OREBODY 31

Some text in this chapter is paraphrased from internal BHP reports. Some figures are also extracted from BHP reports. This source of information is gratefully acknowledged.

3.1 Background

3.1.1 General

The Orebody 31 (OB31) iron ore deposit is located approximately 40 km east of Newman Township in the Pilbara region of Western Australia. The deposit is situated to the east of BHP's existing OB17/18 mine, and approximately 7 km north of the Wheelarra Hill/South Jimblebar deposits. See Figure 3.1.

Mining of OB31 commenced in 2016. As approximately 70% of its iron ore reserve lies below the water table, significant dewatering is required to enable pit development.

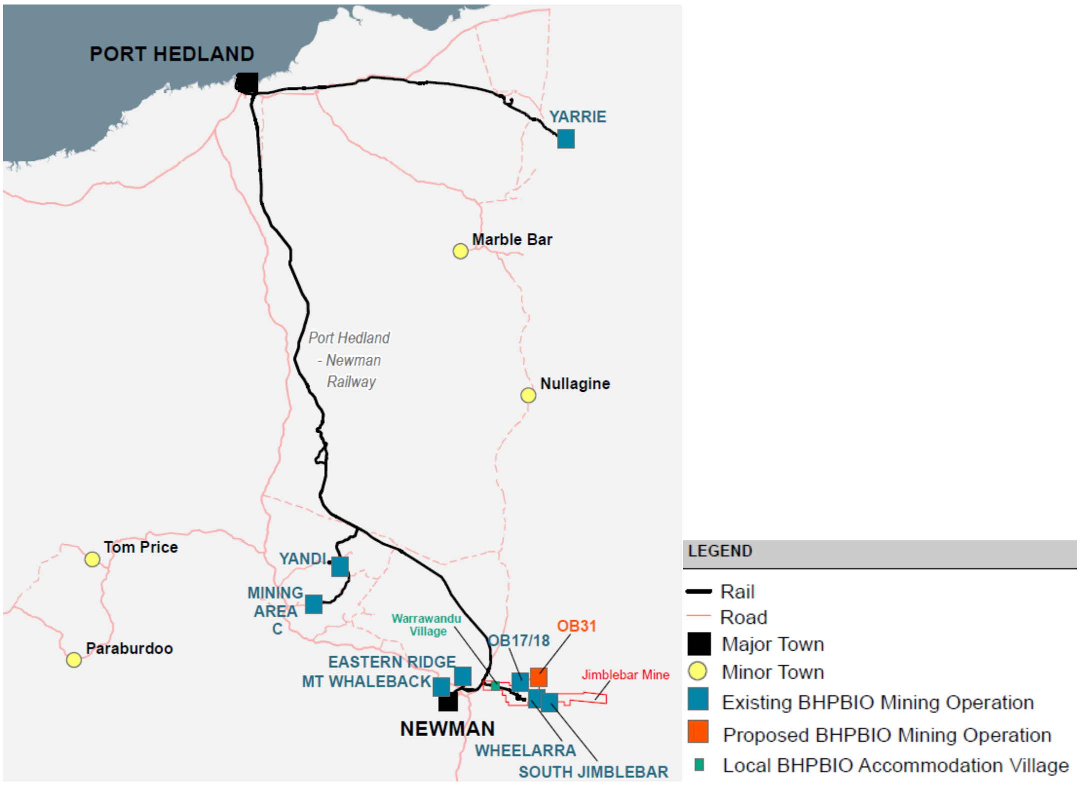


Figure 3.1 Location map (extracted from a BHP internal report).

3.1.2 Climate

The mean annual rainfall in the area which is the focus of the present study is about 300 mm. Most of this rain falls between December and March. The Pilbara is characterised by high evaporation rates (about 3100 mm/yr) and a generally low soil infiltration capacity. Groundwater recharge occurs during major rainfall events. Within the study area, this is thought to occur primarily through leakage from streambeds (McFarlane et al, 2015).

3.1.3 Topography

The OB31 mine is located adjacent to Jimblebar Creek in the upper portion of the Fortescue River catchment which drains into the Fortescue Marsh around 80 km north of OB31. Figure 3.2 summarises the local surface water drainage and topography.

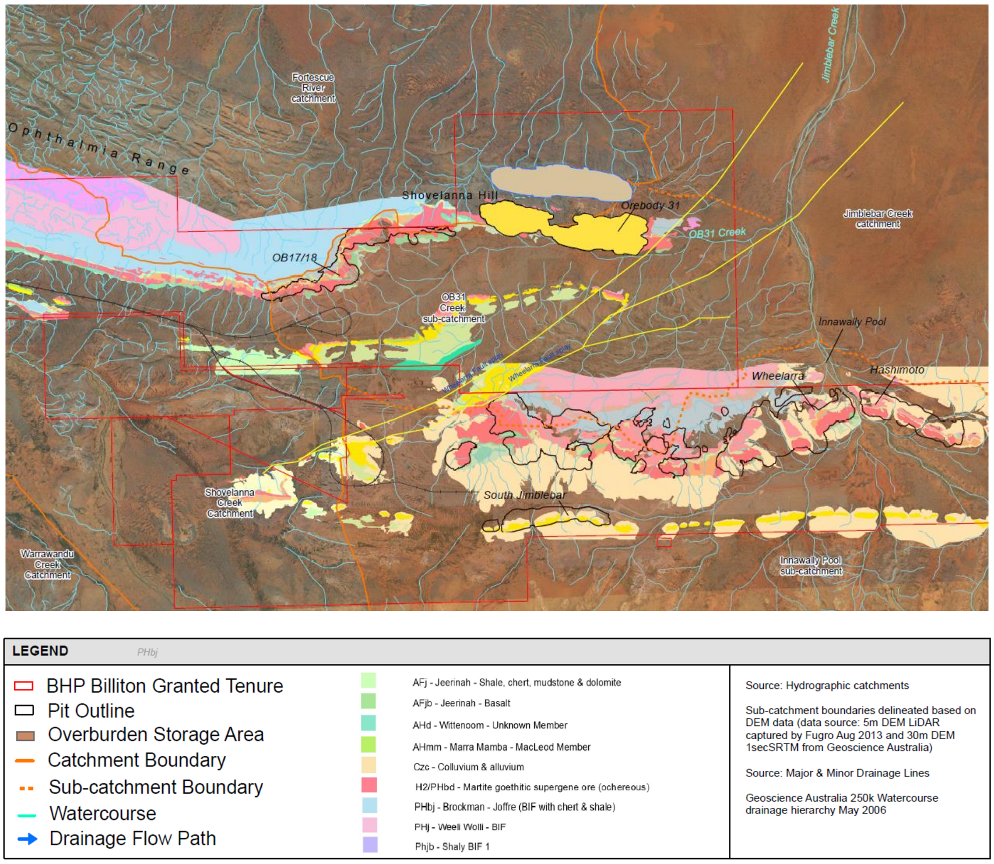


Figure 3.2. Topography and surface water drainage. (Extracted from a BHP internal report.)

Both OB17/18 and OB31 form part of the east-west trending Shovelanna Hill ridgeline, which forms the eastern extremity of the Ophthalmia Range. To the immediate south of OB31 runs an east-west trending valley drained by “OB31 Creek”. This creek has generally developed along the strike of the Wittenoom Formation (see below), which has been eroded and subsequently infilled with Tertiary-aged sediments and more recent valley alluvium. A small sub-catchment area of roughly 85 km² supports OB31 Creek which merges with the much larger Jimblebar Creek at a confluence approximately 4 km to the east of OB31. OB31 and Jimblebar Creeks are ephemeral; their channels are mostly dry.

The land surface elevation in the vicinity of the OB31 mine is about 540 m.

3.2 Geology and Hydrogeology

3.2.1 Geology

Outcrop geology in the vicinity of OB31 is dominated by rocks of the Hamersley Group and to a lesser extent the upper Fortescue Group. These unconformably overlie Archean granite and greenstone of the Pilbara Craton. The Hamersley Basin comprises a sequence of weakly metamorphosed sedimentary and volcanic rocks that have been subjected to a complex tectonic history, this resulting in predominantly ESE-WNW striking regional synclines and anticlines. The large-scale structure at OB31 comprises an open, east-west striking anticline-syncline pair with southerly dipping axial-planes which host the iron ore deposit. The anticline is situated south of the syncline, with the common limb dipping about 35% to the north. Orebody 31 is an elongate east-west deposit that extends about 4.8 km along strike and is about 1 km wide.

The OB31 pit intersects the following stratigraphic units: Brockman Iron Formation (the orebody), Mt Sylvia and Mt McRae Formations (footwall along the southern pit margin), and the Yandicoogina Member and Weeli Wolli Formation (hanging wall along the northern pit margin). A conceptual geological cross-section is provided in Figure 3.3.

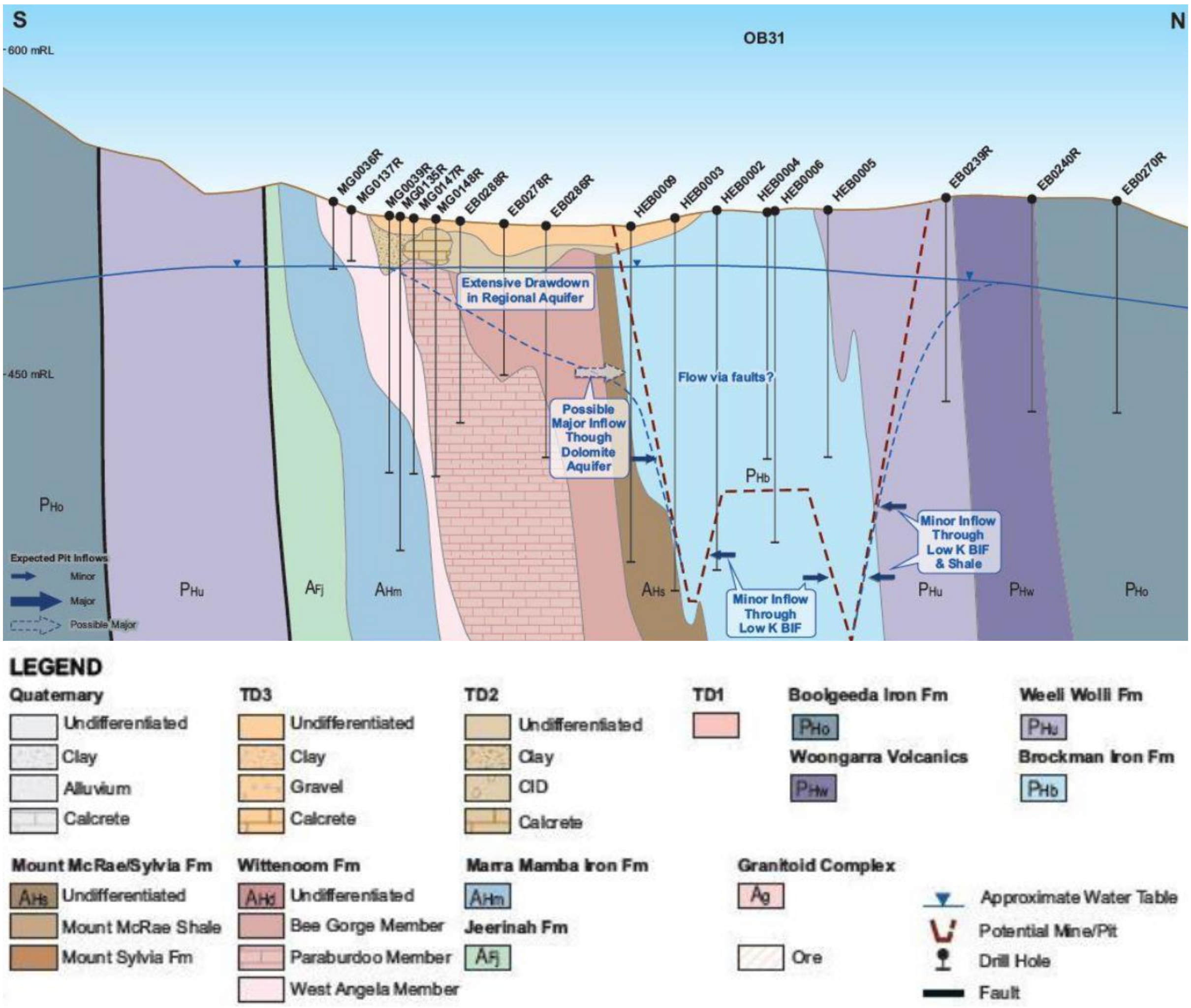


Figure 3.3 Conceptual geological cross section through the OB31 pit. (Extracted from a BHP internal report.)

The region is cross-cut by a series of NE-SW trending fault systems, while a series of much smaller NW-SE trending local fault structures occurs closer to the pit.

3.2.2 Hydrogeology

Within the OB31 mine footprint, the Brockman Iron Formation is subdivided into the Dales George, Whaleback Shale and Joffre Members. The Dales Gorge and Joffre Members are highly permeable, while the Whaleback Shale which separates them is not. In contrast, banded iron and shale formations to the north and south of the Brockman Iron Formation are of low permeability. However high airlift yields have been recorded in bores targeting the Mt McRae formation; these are thought to be related to zones of faulting.

To the south of the OB31 orebody lies the steeply dipping Wittenoon formation. This formation includes weathered and karstic dolomite comprising the Paraburdoo Member. The east-west trending valley which overlies this dolomite is filled with Tertiary sediments that are up to 150 m thick. These sediments, and the dolomite which underlies them, are thought to possess high permeability. The Paraburdoo member is, in fact, a significant regional aquifer.

The hydraulic gradient is eastwards across the OB31 deposit. Groundwater elevations range from 498 m at its west to 496 m at its east. Regional groundwater measurements indicate a 50 m hydraulic “step” between OB31 and the Weeli Wolli Formation/Woongarra Volcanics to the north.

Rainfall recharge to outcropping/subcropping orebody aquifers is relatively rapid, and anticipated to vary between 1% and 2% of mean annual rainfall. Minor groundwater throughflow from the OB31 Creek valley may also contribute to local recharge.

3.3 Investigations

We describe only two investigations in this report. The first of these is the focus of data assimilation that is described in the next chapter.

3.3.1 Constant Rate Test (CRT)

A series of constant rate pumping tests was carried out in late 2014. Collectively, these are referred to as “the CRT” in this and other reports. Water was extracted at rates of between 50 L/s and 100 L/s from 6 production bores over periods of between 5 and 11 days (with only one production bore operating at any one time). Drawdown and recovery were monitored in up to 21 wells (including some of the production wells). Data were acquired over a period of 183 days. Note that drawdowns were not measured in all observation wells during all pumping periods.

The locations of pumping and observation wells are shown in Figure 3.4a and b. Pumping wells are labelled in Figure 3.4a while observation wells are labelled in Figure 3.4b.

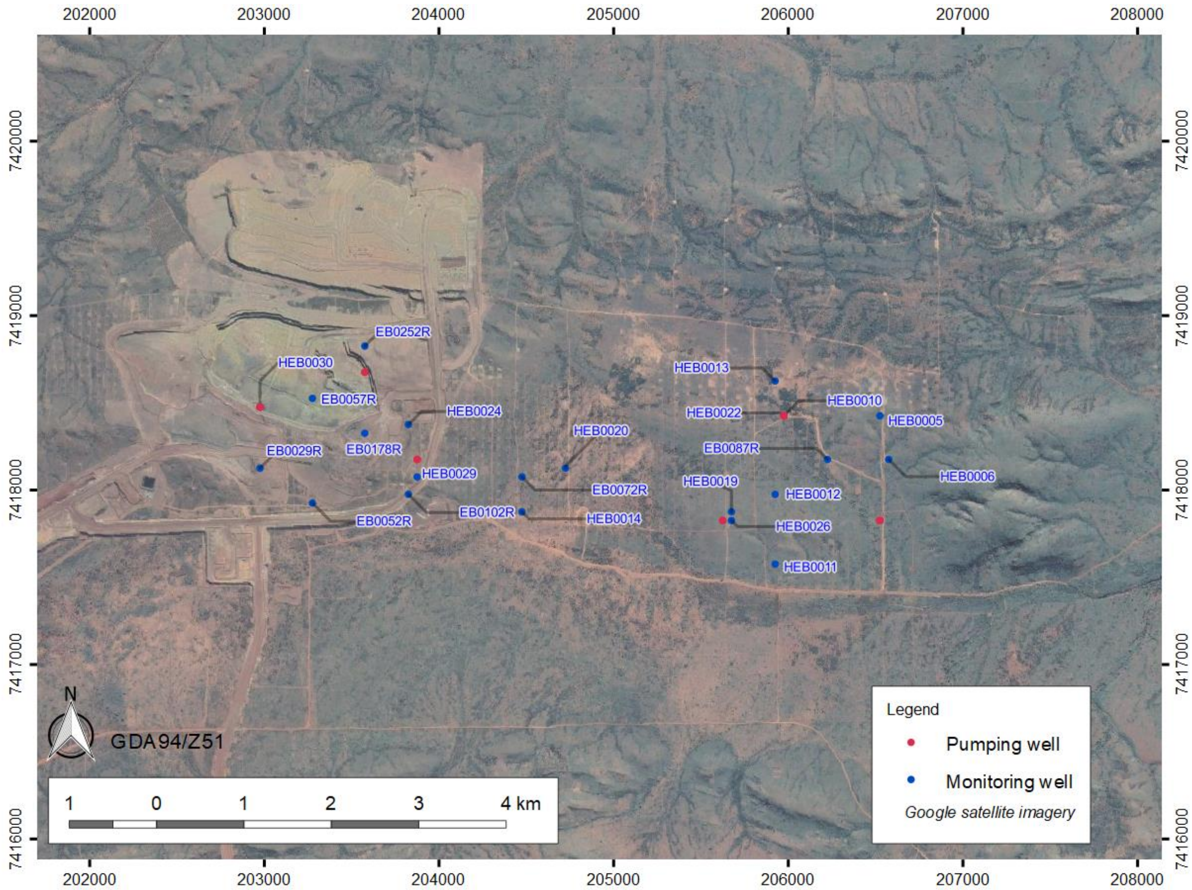


Figure 3.4a. Pumping wells (red) and observation wells (blue) comprising the CRT. Observation wells are labelled.

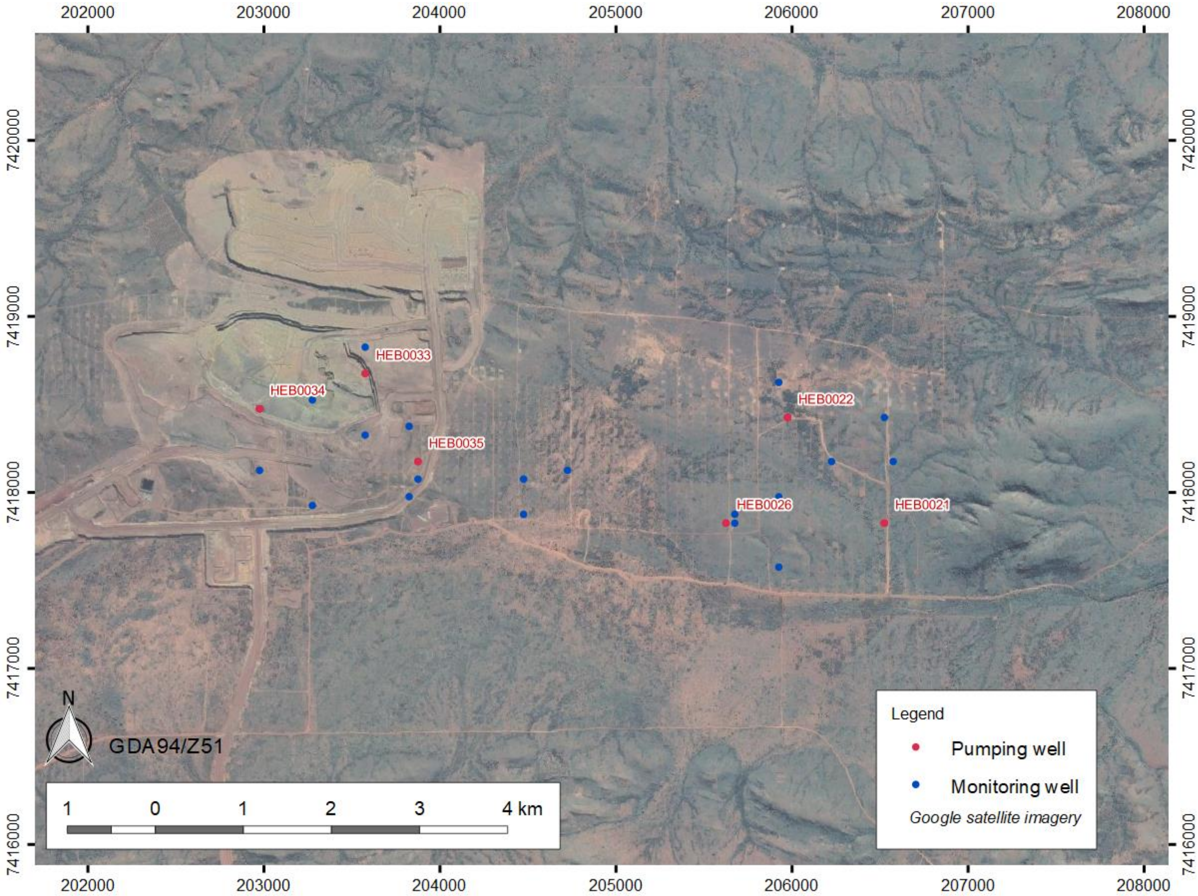


Figure 3.4b. Pumping wells (red) and observation wells (blue) comprising the CRT. Pumping wells are labelled.

Shortly after completion of the CRT, it was reported that a qualitative examination of drawdowns and responses indicates that pumping-induced groundwater flow is anything but radial, and that transmissivity in the area is anything but homogeneous. It was stated that some drawdown/recovery responses to pumping suggest the presence of nearby impervious boundaries, while others illustrate rapid propagation of drawdown over comparatively large distances followed by rapid recovery. The possibility of a hydraulic connection between the OB31 orebody and the regional Paraburdoo Member dolomite aquifer to its south was suggested.

3.3.2 Hydrodynamic Trial (HDT)

During the latter part of 2015, three production bores situated within the Dales Gorge and Joffre Units were simultaneously pumped at rates of up to 60 L/s for a period of about 3 months. Drawdown, followed by one month of recovery, was monitored in 26 observation wells. Interpretation of these data (particularly recoveries) strongly suggested a hydraulic connection between Brockman orebodies and a significant source of water, probably the Paraburdoo Member dolomite. The existence of such a connection increases estimates of pumping rates required for dewatering of the OB31 pit.

3.4 Previous Modelling

3.4.1 The Regional Model

In 2014 (before undertaking the CRT), BHP commissioned the development of a groundwater model to support development and approval of OB31 mining. The model was tasked with predicting OB31 pit dewatering requirements, and predicting the impact of this dewatering on local and regional receptors. The domain of this model (henceforth referred to as "the regional model") is shown in Figure 3.5.

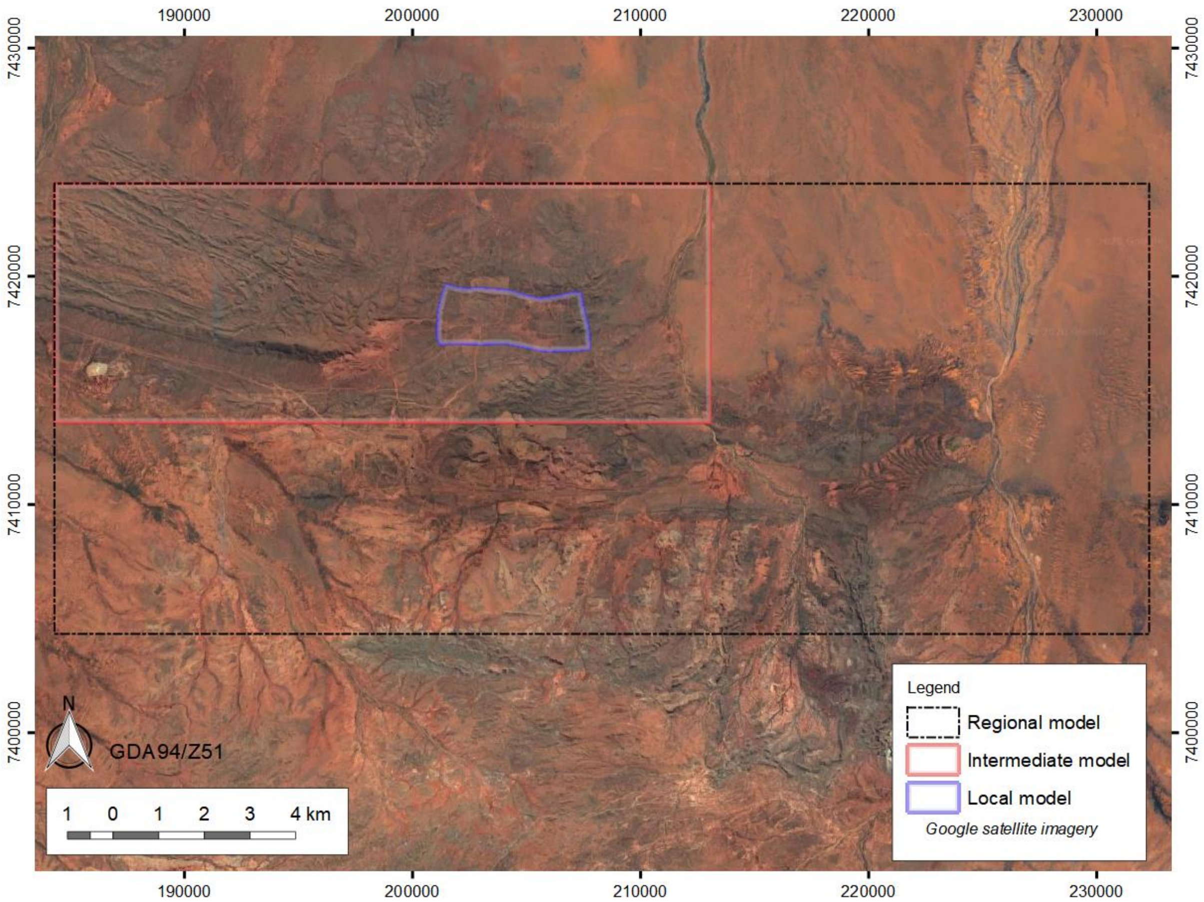


Figure 3.5. The domains of the three models that are discussed in this report, namely the regional model, the intermediate model and the local model.

The regional model employed MODFLOW SURFACT 3.0 (Panday et al, 2007) as its simulator. Its domain covered an area of 19.8 km × 48 km using 291 rows × 671 columns. Its 7 layers extended to a depth of 330 m. It employed 1,366,827 active model cells; these cells had dimensions of 50 m × 50 m in the vicinity of the OB31 mine. Note that MODFLOW SURFACT employs a structured grid.

The regional model superseded a previous model. Considerable care was taken in construction of the regional model to faithfully represent the prevailing geological model.

Parameterisation of the regional model was based on zones of piecewise constancy. Zonation reflected the disposition of geological units as far as these were known. Extra zones were added within orebody formations to reflect different intensities of mineralisation; there is evidence that higher grades of mineralisation sustain higher permeabilities. To assist the model calibration process, a small number of structural features that transect bedding planes were added to the model. However:

- The locations and dispositions of these structural features were not exactly known.
- Model grid discretisation allows only approximate representation of these and other geological entities.

The regional model was calibrated against pre-development water levels, as well as transient responses to pumping from boreholes in the vicinity of the OB18 mine. Upon completion of the CRT, parameters of this model underwent further adjustment in order for the model to fit drawdowns and recoveries emerging from these tests. Although the fit with CRT data was deemed to be "good" at the time, many nuances of CRT drawdown and recovery time series were not replicated by the re-calibrated model. Those charged with its re-calibration reported that a combination of the model's coarse grid and short CRT pumping durations made the attainment of a better fit impossible.

The calibrated regional model was used to predict dewatering requirements for the OB31 mine. Maximum extraction rates of about 25 Ml/day were predicted.

Once HDT data became available, the regional model underwent further re-calibration; at the same time, the disposition of geological units (and hence model zonation) was also revised to reflect revisions of the geological model. Following parameter adjustment, model fits with HDT drawdowns and recoveries were decreed to be reasonable. Care was taken to respect those aspects of drawdown and recovery time series that suggested a connection between the orebody and the regional dolomite aquifer. This required addition to the model of zones representing structural features that may connect the Brockman Iron Formation to the Paraburdoo Member. Predictions of OB31 dewatering rates made with the re-calibrated regional model rose accordingly. Rudimentary post-calibration uncertainty analysis suggested that dewatering requirements may be as high as 50 Ml/day, though would probably be less than this.

3.4.2 The Intermediate Model

Following considerable investments in aquifer testing, and in modelling to interpret testing-acquired datasets, by the end of 2015 a sense was emerging that groundwater modelling was not serving the BHP decision-making process as well as it should be. Concerns included the following.

- It was accepted that hydrogeological complexity imbues mine dewatering rate predictions with considerable uncertainty. Nevertheless, while early predictions of dewatering requirements were low, uncertainty intervals did not include what were later accepted as more likely estimates of these requirements.
- While modelling that had been undertaken up until that time was able to replicate some aspects of CRT and HDT drawdowns and recoveries, fits with informative nuances of these time series were not particularly good. As long as fits between model outcomes and CRT/HDT drawdowns/recoveries remained indifferent, the capacity of model-based history-matching to reduce the uncertainties of decision-critical predictions remained untapped.

To investigate these issues, BHP commissioned further modelling work in late 2016. Modellers were tasked with improving fits between model outputs and CRT/HDT drawdowns and recoveries, and with quantifying the range of uncertainties associated with key model predictions following attainment of these fits.

As these tasks required the use of model-partner software, it was necessary that model run times be reduced from that of the regional model. Hence the size of the model domain was reduced to 28.8 km × 10.5 km, and the number of active model cells was reduced to 288,834. The minimum cell size was still 50 m × 50 m. MODFLOW-SURFACT was replaced by MODFLOW-USG (Panday et al, 2013). The resulting model is referred to herein as "the intermediate model".

The intermediate model retained the same 7 layers as the regional model. Initially, parameterisation continued to be zone-based, with zonal dispositions aligned with known or inferred geological boundaries and/or with known or inferred structural features. The domain of this model is depicted in Figure 3.5, along with those of the other models that are discussed herein. Figure 3.6 shows zonation in layer two of the intermediate model domain. This zonation echoes geological units that are featured in the geological model. Additionally, a number of steeply dipping, narrow zones transect the model domain in NE-SW and NW-SE directions; these represent possible structural features.

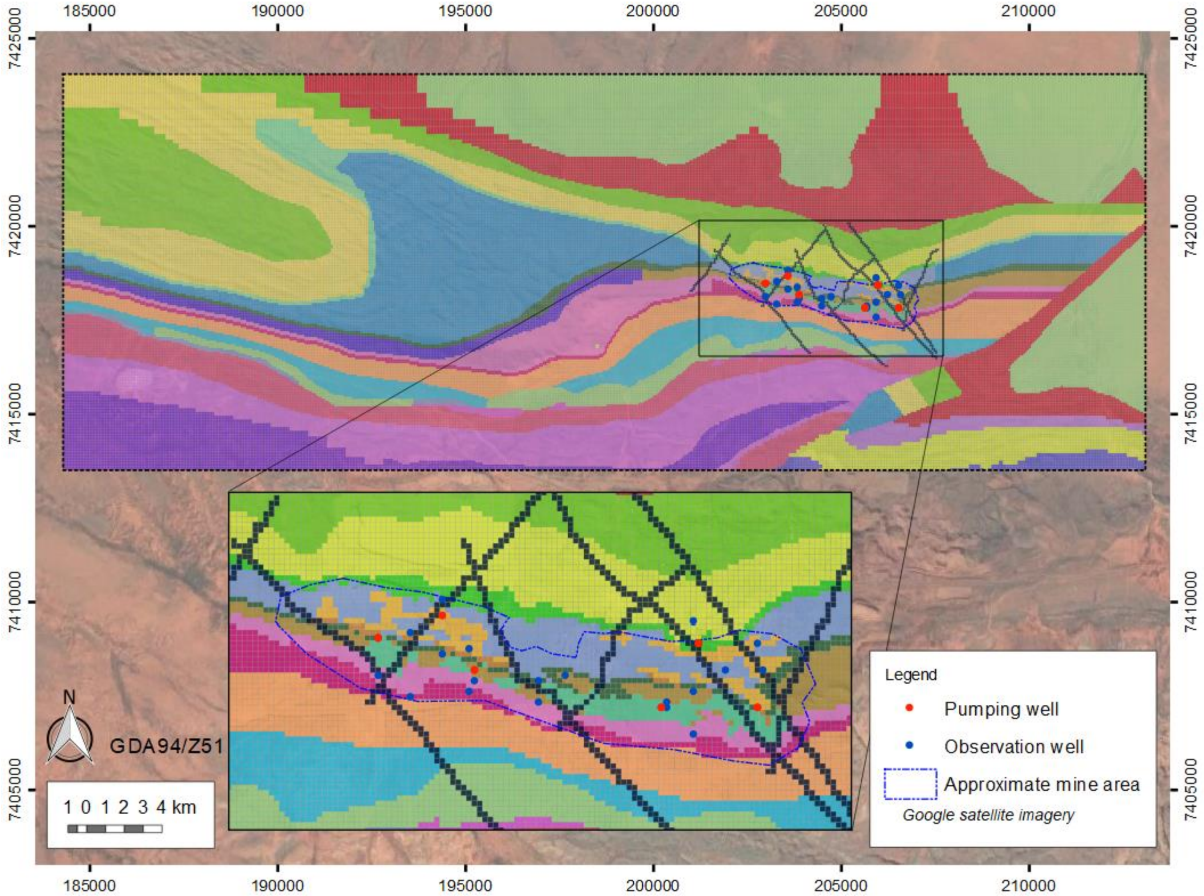


Figure 3.6. Zonation in layer two of the intermediate model. The locations of CRT pumping and observation wells are also shown.

Despite the use of a variety of software-based history-matching methodologies, fits between model outputs and CRT data were only slightly better than those that had been previously attained. Fits with HDT data were somewhat improved, but were still not convincing. It was established however, that the removal of structural zones prevented the attainment of even a mediocre fit with HDT data, especially HDT recessions. It was therefore concluded that hydraulic linkage of the OB31 orebody with the regional dolomite aquifer was highly likely; pit dewatering rates were therefore likely to be high.

Calibration-constrained uncertainty analysis was also attempted by adjusted zone-based parameters. Recognising that the model was under-parameterised, 300 pilot points were added to the intermediate model in the vicinity of HDT pumping and observation wells. The parameters associated with these points acted as multipliers on existing zone-based parameterisation. Unfortunately, model-to-measurement fit underwent little improvement. Pumping rates required for pit-dewatering were estimated to lie between 25 Ml/yr and 50 Ml/yr (a prediction which turned out to be quite reasonable).

4. THE LOCAL MODEL

4.1 Modelling Philosophy

A defining feature of modelling that was undertaken prior to the present study was its reliance on real or imagined geological entities as a basis for its parameterisation. This has a number of undesirable consequences. They include the following.

- The three-dimensional disposition of formation boundaries can be accurately characterised at some locations within the model domain, but not at others.
- The three-dimensional dispositions of structural entities that may enhance or inhibit cross-formational flow of water are only vaguely known.
- The mapping of complex geological units onto a relatively coarse, three-dimensional model grid can only be approximate.
- The hydraulic properties of geological entities are unlikely to be uniform; to the extent that they exhibit spatial variation, boundaries are unlikely to be sharp.
- Relationships between hydraulic properties and mappable properties such as mineralisation intensity are poorly known.
- The desire to represent geological "reality" in three dimensions promulgates long model run times; this makes history-matching and calibration-constrained uncertainty analysis difficult.

Because of all of these factors, attempts at history-matching that are described in the previous chapters of this report led to indifferent fits with datasets that may have been informative of important sources of stored water, and of important connections between these sources and the Brockman orebody.

In late 2020 a more data-focussed approach to OB31 modelling was attempted. The philosophy underlying this approach was "fit first and explain later". Its purpose was to remove all encumbrances to flow of information from pumping-derived data to model parameters.

Only the CRT dataset was fit. This was done in order to ascertain whether this dataset, if subjected to state-of-the-art, model-based interpretation, provides evidence of an orebody-to-regional-aquifer hydraulic connection, or of other hydraulic connections which may influence pit dewatering requirements

4.2 Model Re-Design

4.2.1 Model Grid

The model described in the present chapter is referred to as "the local model". It employs the MODFLOW-USG simulator. Its domain is shown in Figure 3.5. A closer view of its grid is provided by Figure 4.1. This figure also shows the locations of CRT pumping and observation wells.

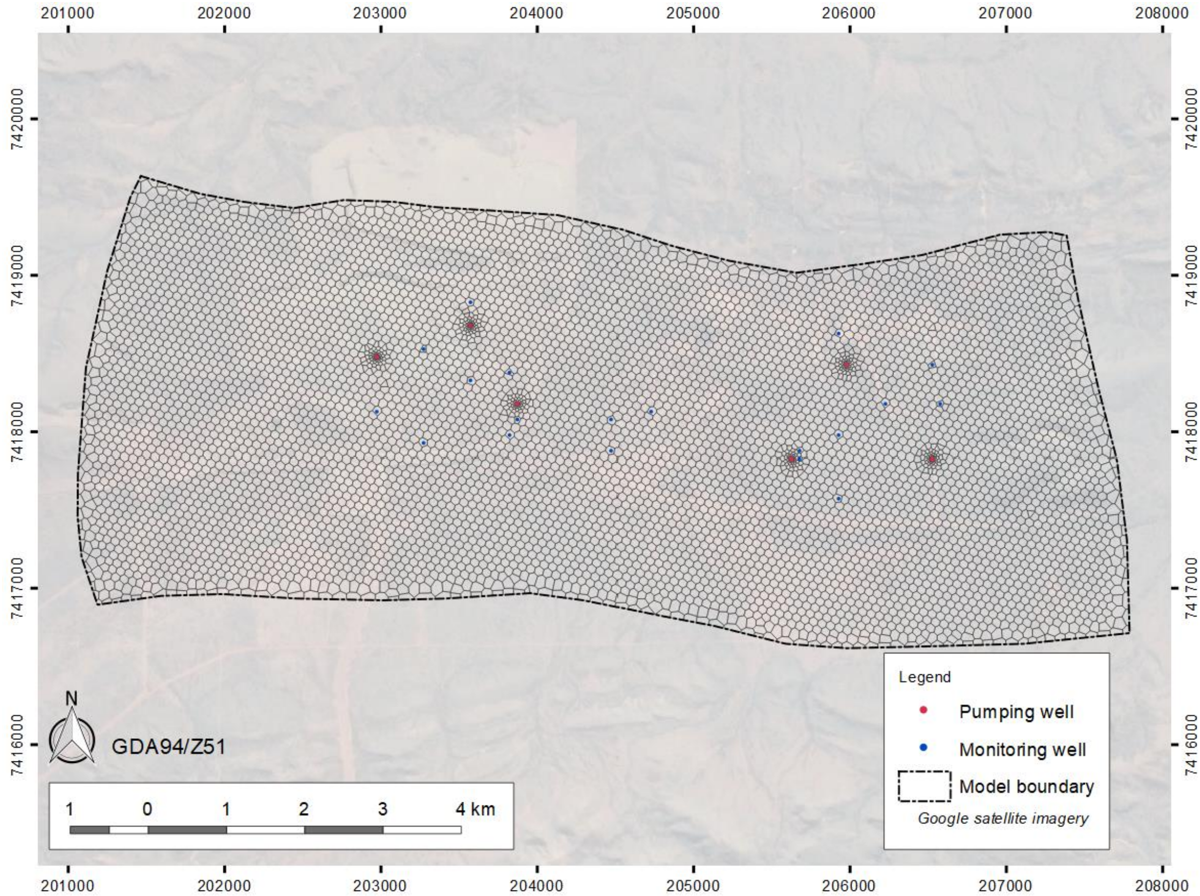


Figure 4.1. Domain and grid of the local model. CRT pumping wells (red) and observation wells (blue) are also shown.

The northern and southern boundaries of the local model domain follow geological strike. The northern boundary coincides with the outcrop of the low permeability Woongarra Volcanics. The southern part of the model domain includes the Paraburdoo Member of the Wittenoon Formation – the regional dolomite aquifer.

The model contains only a single layer. This hastens its execution and renders the model immune to numerical instability. It also respects the notion that features which impede or facilitate flow of water to the OB31 pit do so because they affect horizontal, rather than vertical, movement of water. If information on these features is, in fact, resident in the CRT dataset, opportunities for its emergence are therefore heightened.

The single model layer is designated as "unconfined"; hence transmissivity can vary with water level. However transmissivity variations are minor during the CRT as the elevation of the base of the single layer is set to 300 m, this being about 180 m below the water table. This depth is somewhat arbitrary. Hence, when inspecting calibration outcomes that are presented below, the reader should bear in mind that even though parameters bear the name "hydraulic conductivity" (for this is what the model requires), it is effectively transmissivities that are estimated.

The local model grid was generated using the ALGOMESH package developed by HydroAlgorithmics. Cells which contain pumping wells have an area of 4 m^2. Cell areas increase in size with distance from pumping wells up to about 2500 m^2.

The northern and southern boundaries of the local model are of the "no-flow" type. Water can enter and leave the model domain only along the direction of geological strike through its eastern and western boundaries. These boundaries are of the "general head" (i.e. GHB) type. Heads along these boundaries are set at 486 m. (The small west-to-east pre-development head gradient is ignored, as interpretation of the CRT dataset is based only on pumping-induced drawdowns and recoveries, and not on heads.) The conductance assigned to each GHB instance is calculated automatically using the hydraulic conductivity of the cell to which it is assigned; it represents material lying between the cell centre and the lateral cell boundary.

The model receives no recharge. As is stated above, modelling is dedicated to inference of hydraulic properties from CRT-induced drawdowns and recoveries. Recharge-induced water level variations during the 6 month period over which the CRT was conducted are minor.

4.2.2 Timing

When configured for interpretation of CRT data, the model runs over 15 stress periods. The first of these stress periods is steady state. Six of the other stress periods represent times over which a single extraction well is operating. The other stress periods simulate recovery. The maximum stress period length is 30 days. The minimum stress period length is 5 days.

Under these conditions, the local model takes approximately 1 second to run on a computer that uses an I9-9900KF CPU running at 3.6 Ghz.

4.3 Parameterisation

Pilot points are employed for parameterisation of hydraulic conductivity and specific yield. Parameterisation of each of these properties relies on the same set of pilot points – 1,763 in each case. Their locations are shown in Figure 4.2.

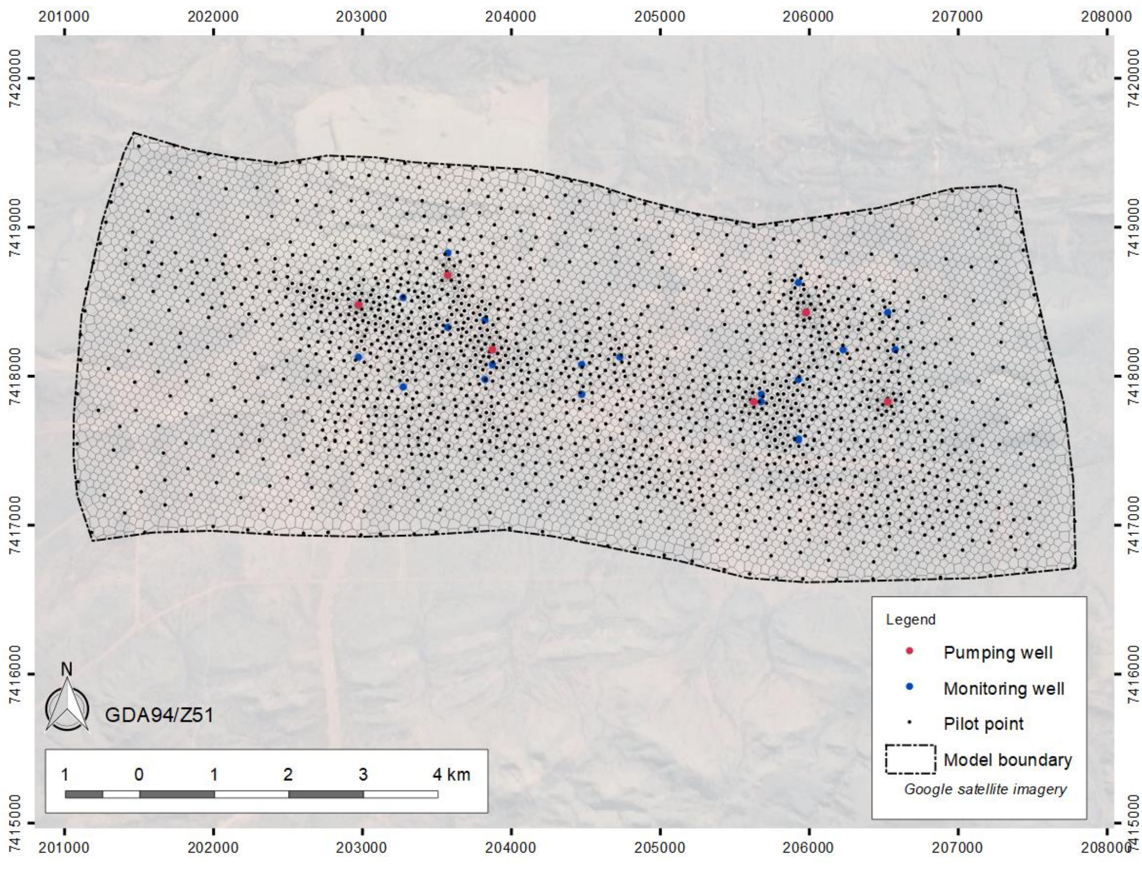


Figure 4.2. Pilot points used for parameterisation of hydraulic conductivity and specific yield.

The operation of pilot points is described in other GMDSI reports and in documentation of PEST utility support software, particularly the PLPROC model preprocessor; see Doherty (2020a). When using pilot points as a parameterisation device for a particular hydraulic property, assignment of hydraulic property values to cells of a model domain becomes a two-step process. First values are assigned to pilot points. Then these values are spatially interpolated to the centres of model grid cells. For the local OB31 model described herein, spatial interpolation is undertaken using kriging. The variogram on which kriging is based is spatially variable; its range is calculated automatically by PLPROC as about 3 times the local inter-pilot-point spacing.

Emplacement of the pilot points depicted in Figure 4.2 was manual. As is apparent from this figure, pilot point spatial density is high in the vicinity of pumping and observation wells, and is lower at greater distances from these wells. This strategy of pilot point emplacement allows the calibration process to introduce appropriate patterns of local hydraulic property heterogeneity to the model domain where observations of drawdown and recovery require this. Meanwhile, the use of lower pilot point densities in observation-poor parts of the model domain maintains parameters at a reasonable number. This facilitates post-calibration linear analysis.

4.4 History-Matching

4.4.1 Software and Methodology

History-matching was undertaken using PEST_HP (Doherty, 2020b). The numerical burden of Jacobian matrix calculation was significantly reduced by basing its construction on random parameter increments; see Section 2.3 of this document for further details. Despite the fact that the inversion process features 3,526 parameters, the number of model runs required per iteration of that process was only 600 for initial iterations, and 1,000 for later iterations. A further 100 model runs per iteration were devoted to parameter upgrade testing and to implementing Broyden improvement of the Jacobian matrix.

The inversion process was allowed to proceed for 27 iterations. However a more-than-satisfactory fit with the calibration dataset was achieved after only 10 iterations at a numerical cost of about 10,000 model runs.

A number of strategies were employed to optimise extraction of information from the CRT dataset through the process of model calibration. Two of these strategies are now briefly described.

4.4.2 Objective Function Definition

Model-to-measurement misfit is measured using an objective function. This is defined as the sum of weighted squared differences between model outcomes and corresponding field measurements. The lower is the objective function, the better is the fit.

Strategic design of an objective function can ensure that appropriate nuances of system behaviour are replicated by the calibrated model, while aspects of system behaviour which are of secondary importance are ignored if they cannot be easily replicated. This can improve the performance of inversion software in minimizing the objective function, at the same time as it can reduce parameter and predictive bias. See Doherty and Welter (2010) and White et al (2014) for details.

Two major components comprise the objective function that was minimised by PEST_HP. The first of these components represents pumping-induced drawdown (including post pumping recovery from drawdown). For any observation well, the drawdown at any time is calculated by subtracting the head measured at that time from the first head that was measured in the well. The time of this measurement varies from well to well. Model-calculated counterparts of measured heads are subjected to the same process before being compared with respective field observations.

During the 183 day period over which CRT data were gathered, heads in some observation wells underwent a small drift. In order to grant the inversion process some immunity from observation drift, temporal differences between subsequent drawdown measurements and their model-calculated counterparts are also matched. These temporal drawdown differences comprise the second major component of the objective function.

Special consideration is given to drawdown measurements that were made in pumped wells. A model-to-measurement penalty is incurred (i.e. the objective function is increased) only if model-calculated drawdowns (and temporal drawdown differences) exceed those that were actually measured. This strategy accommodates the fact that model cells which contain pumping wells have an area of 4 m^2; this is far greater than the cross-sectional area of a pumping well.

4.4.3 Initial Parameter Values

As was explained in Section 2.3 of this document, history-matching is an iterative process. A modeller specifies an initial set of parameter values. PEST_HP improves these values. In doing so, it alters them by the minimum amount that is required for model outputs to match their field-measured counterparts. Minimisation of parameter adjustment is implemented through Tikhonov regularisation. This procedure is intended to nurture reasonableness of estimated parameter values. Ultimately, however, it is the modeller who must decide whether estimated parameter values are, in fact, reasonable.

When assigning initial values to parameters in different parts of a model domain, it is common practice for a modeller to follow hydrogeological advice, as this can encourage the estimation of final parameter values whose minimum error variance status is underscored by hydrogeological respectability. This strategy was not adopted in calibration of the local OB31 model. Instead hydraulic conductivity and specific yield parameters were assigned spatially uniform initial values of 1 m/day and 10^{-3} respectively. This strategy was adopted for the following reasons.

- The somewhat arbitrary depth assigned to the base of the single model layer renders initial and estimated values of hydraulic conductivity somewhat arbitrary.
- Previous attempts to fit the CRT calibration dataset provided no cause for optimism that estimates of hydraulic properties based on mapped and inferred geology are good indicators of actual hydraulic properties.
- Unmapped structural features are likely to exert a significant influence on groundwater flow in the vicinity of the OB31 pit. Spatial uniformity of parameter values imposes no preconditions on estimated parameter values. The inversion process is therefore granted freedom to introduce heterogeneity to the model domain where it has most effect. This is in accordance with the "fit first and explain later" philosophy that guided calibration of the OB31 model.
- The CRT calibration dataset includes incidences of observation wells that incur significant drawdowns from pumping in distant wells, and of observations wells that are minimally impacted by nearby pumping. The need for the inversion process to introduce parameter patterns that include significant, and possibly narrow, connected hydraulic conductivity is clear. However it will be unable to achieve this unless initial hydraulic property values ensure sensitivities of drawdowns in all observation wells to extraction from all pumping wells. This requires that the initial hydraulic conductivity be high and that the initial specific yield be low.

4.5 Outcomes of History Matching

4.5.1 Drawdowns

Measured and model-calculated drawdowns for all observation wells are graphed in the Appendix. As stated above, the drawdown assigned to each well is obtained by subtracting its measured head from the first measured head in the well-specific time series. Hence the first drawdown plotted for each well is zero. For reference, water extraction rates from production bores are also plotted.

In these figures, measured drawdowns are represented as dots joined by straight lines. It is not implied that these straight lines provide a mechanism for temporal interpolation between measurements. However, in some of the figures, they make drawdown trends a little more apparent.

As is described in the previous subsection, the objective function that characterises model-to-measurement misfit is formulated in a way that encourages the inversion process to fit short-term drawdown variations and amplitudes that embody the response of an individual observation well to an individual pumping well, even if it cannot fit drawdowns in a specific well exactly over the entire 183 day duration of the CRT. This should be born in mind when inspecting the figures that are provided in the Appendix. A fit is "good" if the pattern and amplitude of a modelled drawdown response to pumping from a certain production well matches those of the observed drawdown response to pumping from the same production well, even if the two drawdown curves are not superimposed on each other.

A number of measurements were made in pumping wells. In general, recovery measurements were made in these wells, while drawdown measurements were not. Recall that the inversion process does not require that modelled drawdowns and recoveries for these wells replicate measured drawdowns and recoveries, only that the former are smaller in amplitude than the latter.

In general, fits between modelled and measured drawdowns and recoveries are extremely good – better by far than fits attained through any previous attempt at fitting CRT data. The only disappointing aspect of graphs presented in the Appendix is that model-calculated drawdowns in wells HEB0014 and HEB0029 incurred by pumping from well HEB0034 are smaller than those which were actually observed. Both of these observation wells are relatively distant from HEB0034; they are both disposed in a roughly ESE direction from it. The amplitudes of their drawdown responses, taken in combination with more limited drawdown responses in other wells to this same and other pumping, suggests the existence of a narrow band of high conductivity material whose geometry is difficult to represent using pilot points. Alternatively, these drawdown anomalies may be outcomes of three-dimensional flow that cannot be represented using the current model. The former explanation is thought to be more plausible as the inversion process did, indeed, attempt to introduce a narrow band of high permeability material that connects the above pumping and observation wells.

4.5.2 Parameter Fields

The inferred distribution of the log of hydraulic conductivity is shown in Figure 4.3. The inferred distribution of the log of specific yield is depicted in Figure 4.4. In part b of both of these figures, this distribution is overlain on satellite imagery of the study area. In part c of each of these figures, the inferred hydraulic property distribution is overlain on a map in which layer 2 cells of the regional model are coloured according to rock type. Part b of both of these figures feature polylinear structural features whose existence and locations have been inferred with varying degrees of certainty from geological data. Little about these features is known, including any effects that they may have on movement of groundwater. It is not impossible that other unmapped, but more hydraulically relevant, features exist in this area.

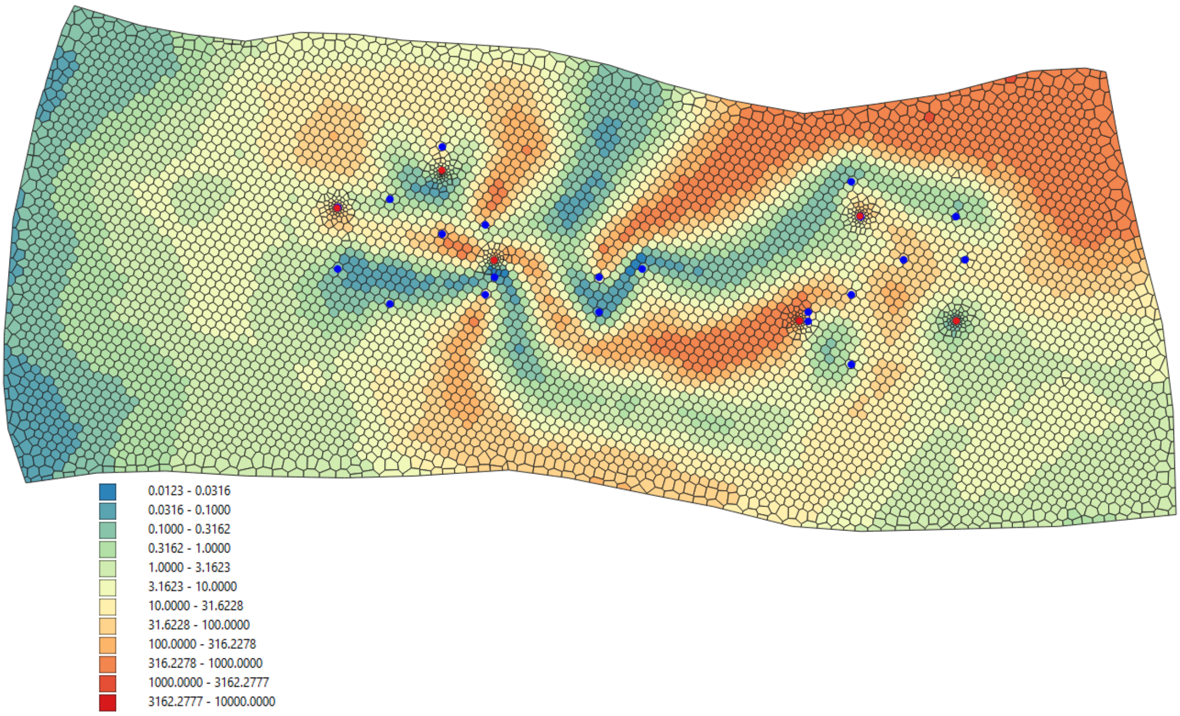


Figure 4.3a. Inversion-inferred hydraulic conductivity (m/day).

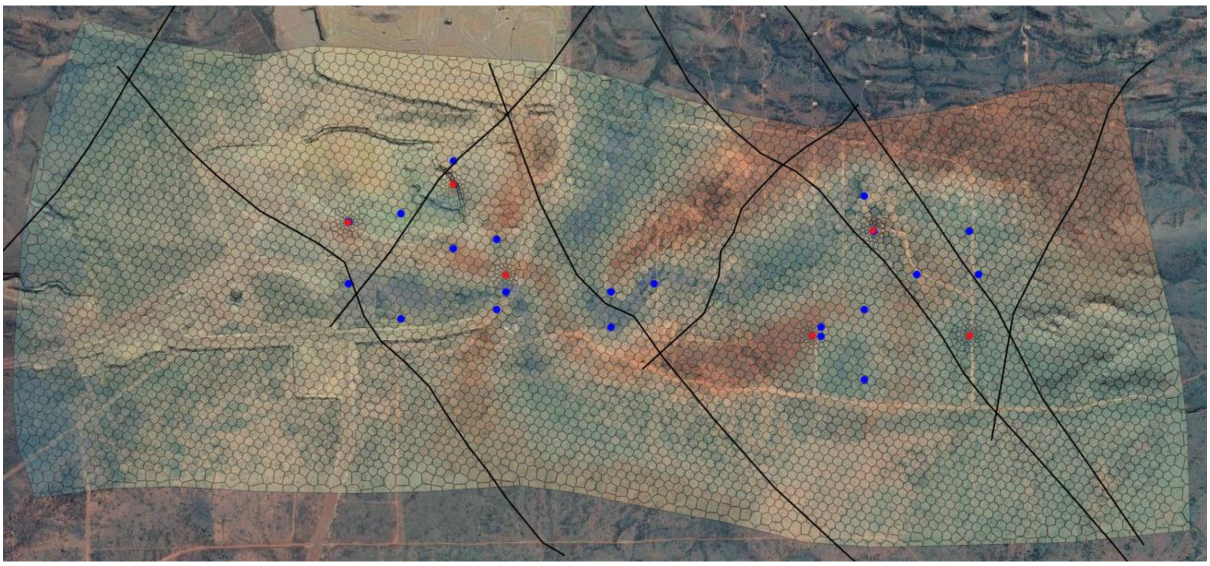


Figure 4.3b. Inversion-inferred hydraulic conductivity overlain on a satellite image.

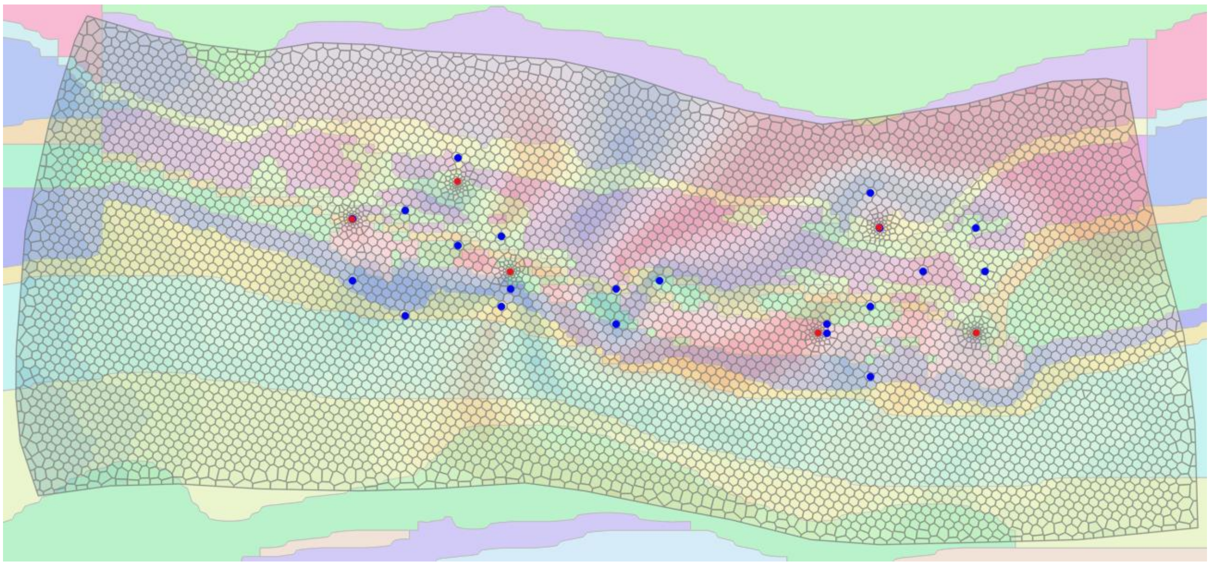


Figure 4.3c. Inversion-inferred hydraulic conductivity overlain on layer 2 zonation of the regional model. Zone colours are arbitrary; zones reflect the geological model of the time.

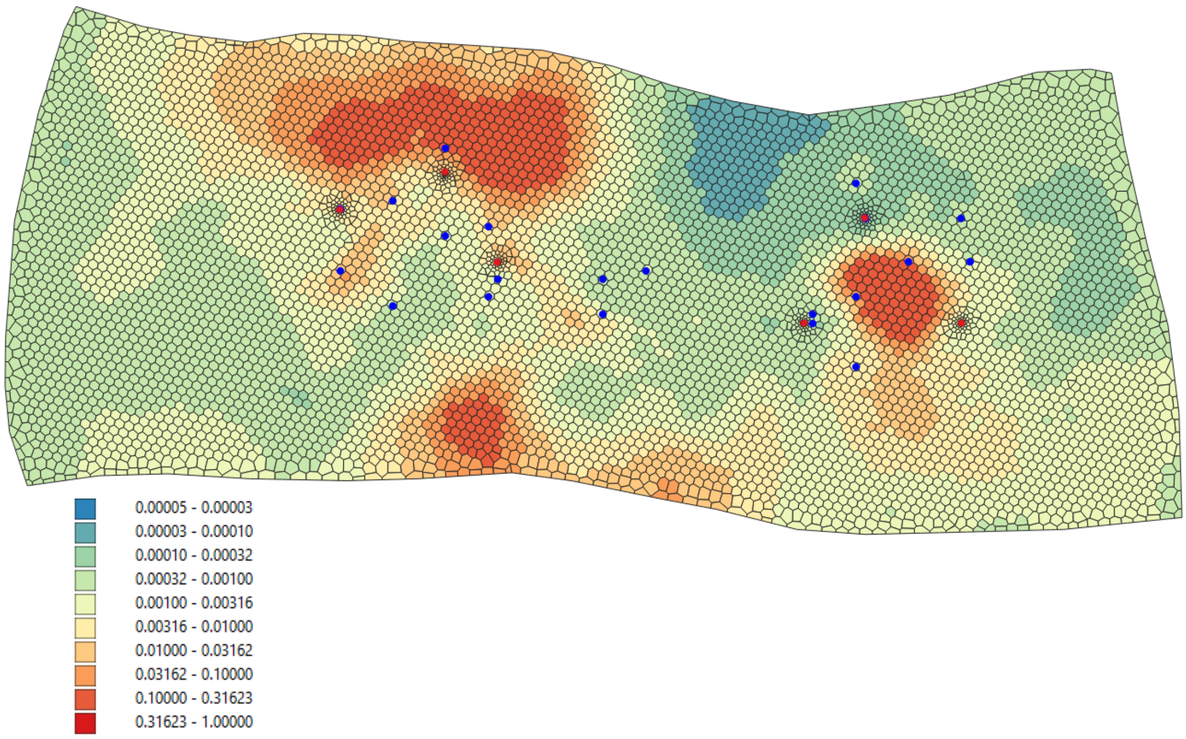


Figure 4.4a. Inversion-inferred specific yield.

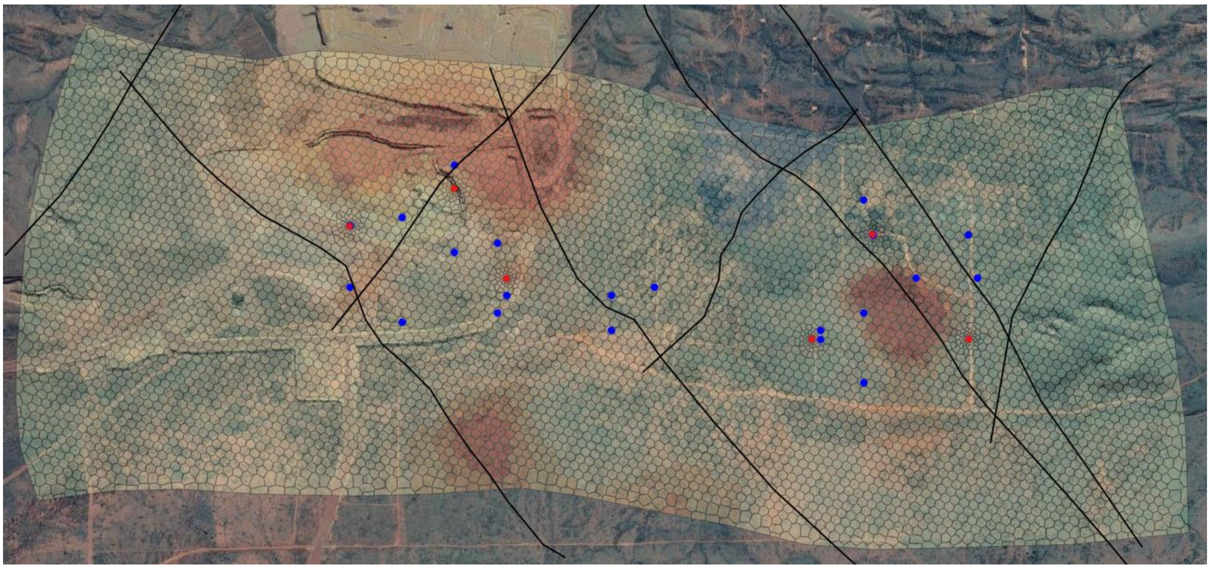


Figure 4.4a. Inversion-inferred specific yield overlain on a satellite image.

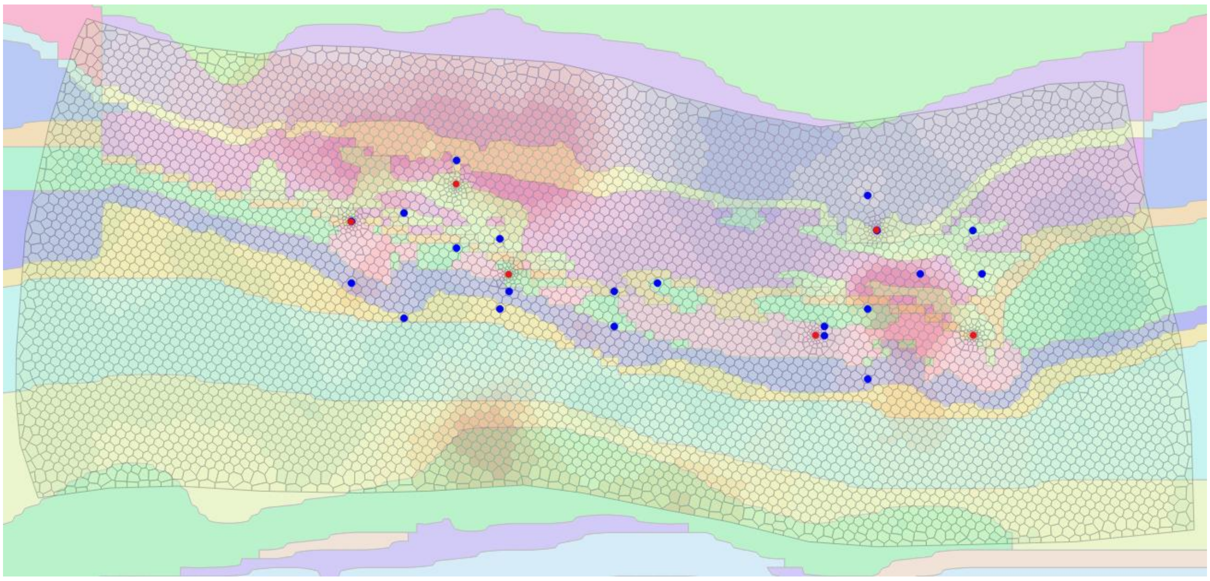


Figure 4.4c. Inversion-inferred specific yield overlain on layer 2 zonation of the regional model. Zone colours are arbitrary; zones reflect the geological model of the time.

Care must be taken in interpreting the parameter fields that are depicted in Figures 4.3 and 4.4. High uncertainties are often associated with inferences of subsurface hydraulic properties from inversion, even where a calibration dataset is relatively large. Therefore, the patterns that are displayed in Figures 4.3 and 4.4 cannot be construed as defining the exact dispositions of material of high and low hydraulic conductivity and specific yield beneath the surface. This is particularly the case for specific yield, for which low values may reflect confined flow in parts of the model domain.

However the patterns of Figure 4.3 and 4.4 indicate that a high degree of hydraulic property heterogeneity is required to explain the drawdowns and recoveries that comprise the CRT dataset. Nevertheless, while the three-dimensional patterns that prevail in the real world may differ from the two dimensional patterns that are displayed in Figures 4.3 and 4.4, their hydraulic consequences are the same. In the vicinity of pumping and observation wells, this probably requires that real-world areas of anomalously low and high hydraulic conductivity exist in roughly the same places as depicted in Figure 4.3. In contrast, in those parts of the model domain in which no pumping or observation wells exist, it is an easy matter for the inversion process to introduce anomalous hydraulic property values in order that drawdowns and responses to CRT pumping are reproduced in other parts of the model domain. Considerably uncertainty therefore surrounds the true dispositions of anomalous property values in these areas.

Taking all of this into account, Figures 4.3 and 4.4 suggest the following.

- There is a tendency for areas of high and low hydraulic conductivity to be aligned in an east-west direction in accordance with the strike of the prevailing geology. High conductivities within the Brockman Iron Formation, and low conductivities within shales of the Mount McRae and Sylvia Formations, are particularly apparent.
- However, strike-aligned bands of high and low hydraulic conductivity are intersected by narrow zones of anomalously low and high hydraulic conductivity respectively. In some cases the strike of these intersecting features is aligned with one of the two directions along which structural features are thought to strike, and are situated close to the suggested locations of these features.
- The inversion process makes it clear that reproduction of CRT drawdowns and responses requires that water be transmitted to areas where pumping wells exist from locations that lie beyond those where observation wells exist. The inversion process also forges a hydraulic connection between some of the former areas and the regional dolomite aquifer that occupies the southern part of the model domain. Moderate values of hydraulic conductivity are awarded to this aquifer. The inversion process also introduces high hydraulic conductivities to the north-eastern part of the model domain which is occupied by the Brockman Iron Formation. This gives the model access to water through its eastern boundary.
- High specific yields are introduced to parts of the Brockman Iron Formation and to parts of the regional dolomite aquifer. This is in accordance with the known hydraulic properties of these two formations.
- Other areas to which the inversion process has introduced anomalously high specific yields appear to be aligned with, and/or coincide with, structural features. Perhaps these structural features provide access to water stored within overlying Tertiary sediments.

In summary, attempts to understand the patterns that characterise Figures 4.3 and 4.4 should not necessarily invoke the existence of bodies of anomalously high or low hydraulic properties at exactly the places indicated by the patterns. In some cases these patterns do indeed indicate anomalous hydraulic properties somewhere within the general area of inversion-emergent anomalies; however the complex, three-dimensional disposition of prevailing geological units and structure must be taken into account when trying to link details of these patterns to known or inferred geology.

The band of high hydraulic conductivity that the inversion process has introduced in the north eastern part of the model domain is perhaps the most curious in terms of location. PEST_HP has conveniently assigned high hydraulic conductivities to an area where observations are lacking in what seems like an obvious attempt to convey water to the central part of the model domain. Recall that conductances of general head boundaries assigned to the OB31 model increase in proportion to cell hydraulic conductivities; this enables the inversion process to access water from the boundary at the same time as it introduces the means to move it.

There is little doubt that locally high hydraulic conductivities, and access to stored or boundary water, are integral to fitting CRT drawdowns and recoveries. However it is possible that the inversion process cannot distinguish between different locations at which these sources of water may exist. This could be readily tested by repeating the inversion process with low, non-adjustable conductances assigned to northern parts of the eastern model boundary. PEST_HP may then introduce an alternative source of stored or boundary water to the model domain, and then convey it to pumping areas through a conductive pathway. Alternatively, model-to-measurement fits may deteriorate; this would constitute evidence that north-eastern elevated hydraulic conductivities are real.

It must be born in mind however, that the existence of sources of water, and pathways to them, may be more significant than their exact locations as far as predictions of pit dewatering requirements are concerned. Their exact locations may be of secondary importance.

4.6 Linear Analysis

4.6.1 General

Following calibration of the OB31 model, linear analysis was undertaken in order to gain some idea of post-calibration parameter and predictive uncertainties. To maintain reader interest, theory is not provided herein. Those who are interested can refer to Doherty (2015) and other GMDSI reports.

Linear analysis is based on the concept that the action of a model on its parameters, under both calibration and predictive conditions, can be represented by the action of a matrix on a vector. The matrix contains sensitivities of pertinent model outputs to all model parameters. Under calibration conditions, this matrix is the Jacobian matrix.

Unfortunately, the approximate, randomised, Jacobian matrix which was used to great effect by PEST_HP to estimate parameter values, does not provide a sound basis for linear analysis. Hence, following model calibration, a high-fidelity Jacobian matrix was computed through individual incremental variation of each of the 3,526 parameters of the OB31 model. Sensitivities of predictions to parameters were computed in the same way.

Post-calibration parameter and prediction uncertainties were then computed using matrix equations which are derived from Bayes equation. Software which carries out these tasks is available through the PEST and PyEMU (White et al, 2016) suites.

4.6.2 Parameters

Figure 4.5 maps relative parameter uncertainty reduction throughout the model domain - for hydraulic conductivity in Figure 4.5a, and for specific yield in Figure 4.5b. The relative uncertainty reduction of parameter *i* (i.e. r_i) is defined using the equation:

$$r_i = 1 - \frac{\sigma_i^{post}}{\sigma_i^{prior}} \tag{4.1}$$

where σ_i^{prior} is the prior uncertainty of parameter *i* (an outcome of expert knowledge and site characterisation) and σ_i^{post} is the posterior uncertainty of the same parameter. r_i varies between 0.0 and 1.0. The closer is r_i to 1.0, the greater is the information content of the calibration dataset with respect to parameter *i*.

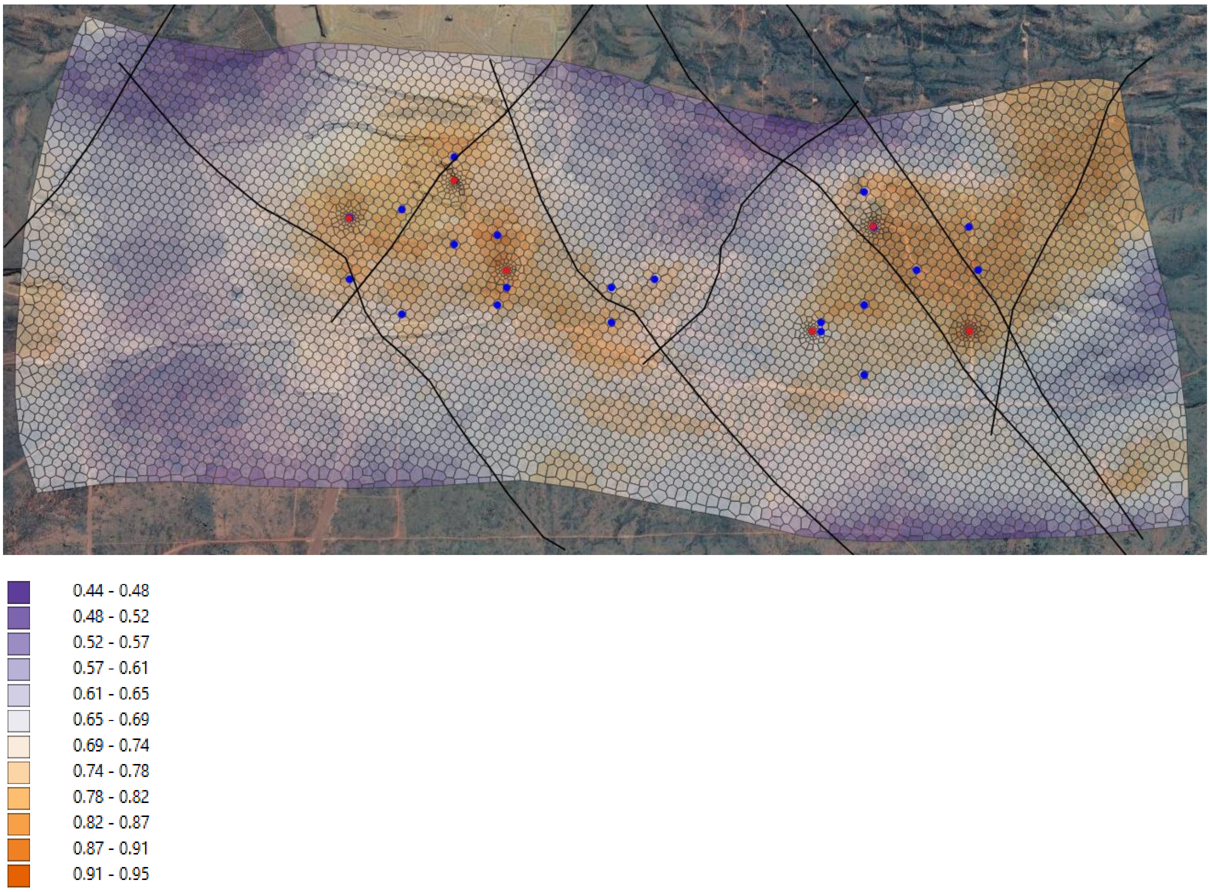


Figure 4.5a. Relative uncertainty reduction of hydraulic conductivity gained through calibration of the OB31 model against the CRT dataset.

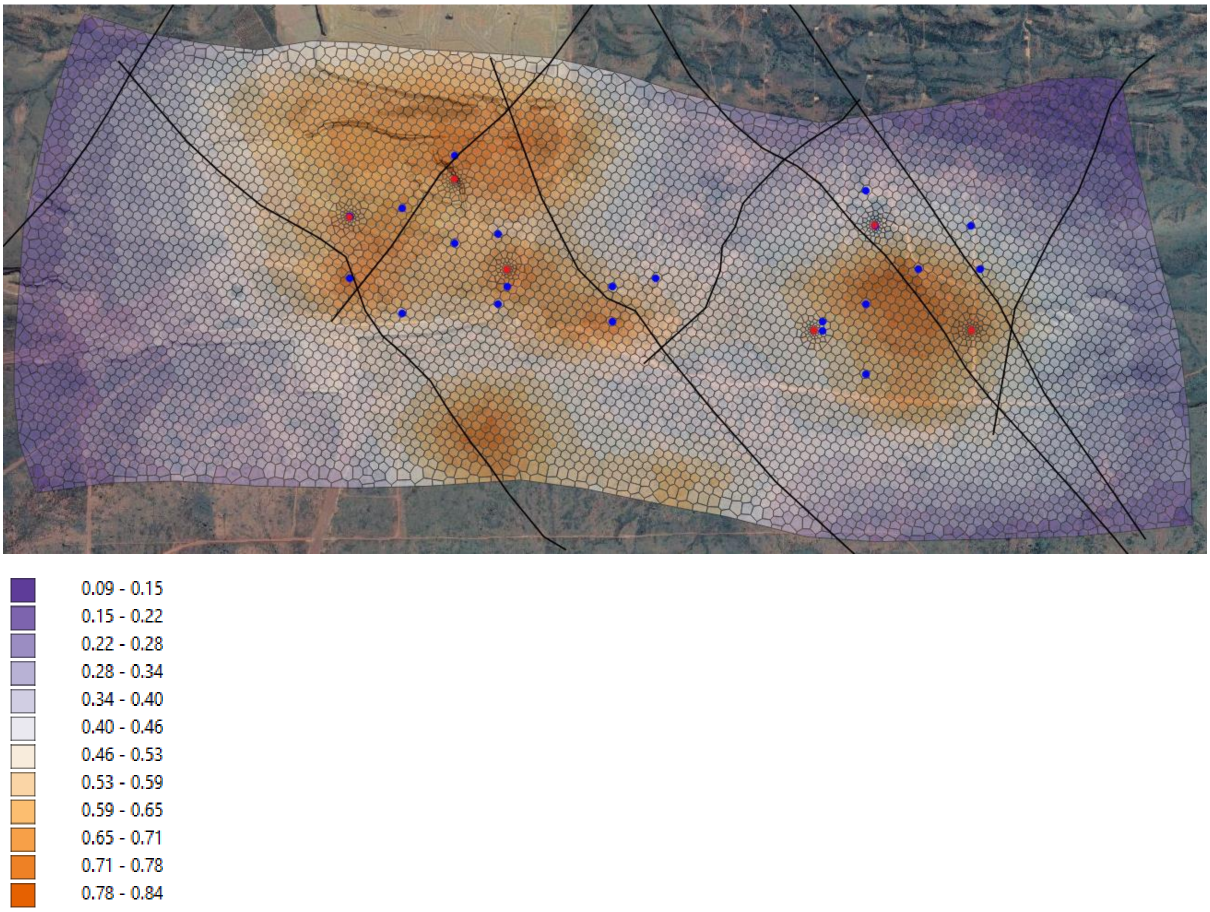


Figure 4.5b. Relative uncertainty reduction of specific yield gained through calibration of the OB31 model against the CRT dataset.

Figure 4.5 shows that parameter uncertainty reduction generally increases with proximity of parameters to pumping and observation wells. However, of particular interest are the high levels of certainty associated with high hydraulic conductivity values ascribed to the north-eastern corner of the model domain. The alignment of this area of high parameter certainty with a possible structural feature is of interest.

Linear analysis suggests that hydraulic conductivities ascribed to the regional dolomite aquifer in the southern part of the model domain are also reasonably certain.

It also appears that the inversion process seems to be relatively certain of the need to assign high values of specific yield to three areas within the model domain. Two of these areas lie in the central north and south of the model domain; a third lies in the eastern part of the model domain in the vicinity of a group of pumping and extraction wells. The central northern area coincides with the Brockman Iron Formation while the central southern area coincides with the regional dolomite aquifer. The eastern area may represent water stored in Tertiary sediments to which deeper geological layers have access through conductive structural features.

4.6.3 A Prediction

Though not built to simulate pit inflows or dewatering requirements, the local OB31 model was used to predict inflow to an artificial pit whose footprint is shown in Figure 4.6 (the footprint of the current pit) as it deepens at a uniform rate of 12 m per year. This provides a crude estimate of pit dewatering requirements. Linear analysis was then used to estimate the pre- and post-calibration uncertainties of this prediction. The ratio of these uncertainties provides an estimate of the information content of the CRT dataset with respect to predictions of pit dewatering requirements.

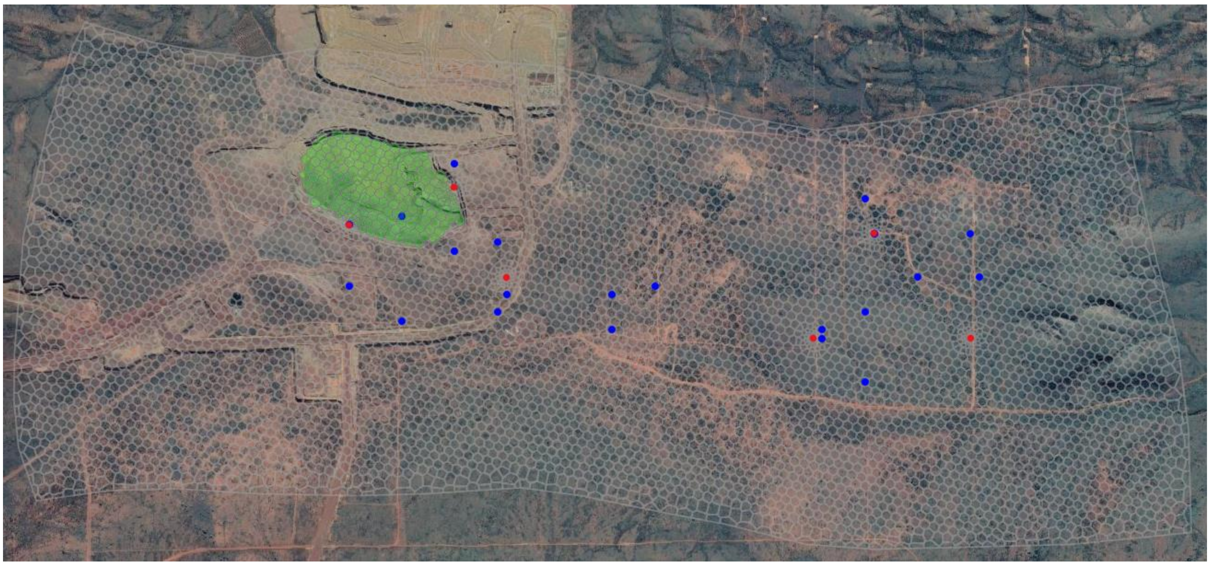


Figure 4.6. MODFLOW-USG DRAIN boundary conditions were emplaced in green-coloured cells to simulate inflow into a deepening pit.

Deepening of the OB31 pit was simulated using the MODFLOW-USG DRAIN package. High conductance DRAIN boundary conditions were introduced to the green cells of Figure 4.6. Model-predicted inflows for 9 years of pit deepening are shown in Figure 4.7. At the time of writing, the pit has been operating for 4 years. According to Figure 4.7, predicted inflows into the artificial pit at this time are 37.8 Ml/day.

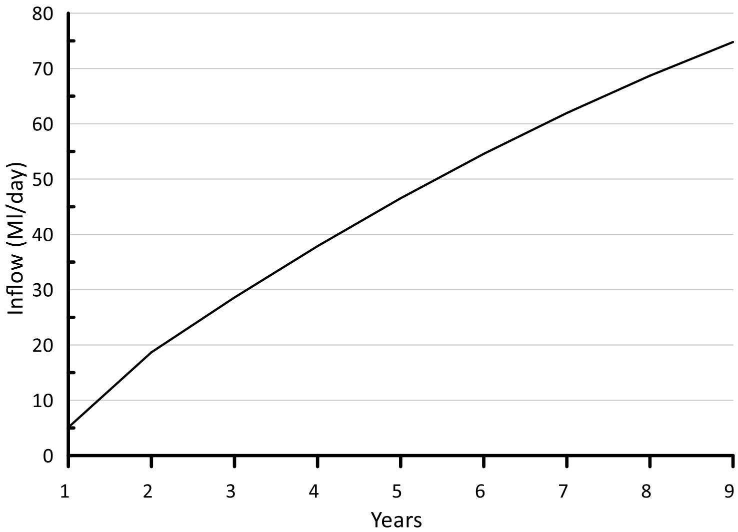


Figure 4.7. Predicted inflow to the artificial pit of Figure 4.6.

Dewatering of the real-world pit is effected through extraction of water from a number of near-pit production wells. Collective extraction rates from these wells have been in the vicinity of 40 Ml/day since dewatering began.

Prior to undertaking linear predictive uncertainty analysis, four new parameters were added to the model described herein. These parameters are factors applied to GHB boundary conductances. The inclusion of these parameters in the analysis, and the assignment of high prior uncertainties to them (4 orders of magnitude), partially accommodates model errors in inflow predictions incurred by the proximity of boundaries to the pit. Hence while predictions of pit inflow made by the current model are probably over-estimated because of the proximity of these boundaries, linear analysis is able to take this into account when assessing the information content of the CRT dataset with respect to predictions of pit dewatering requirements.

Using linear analysis, a post-calibration uncertainty standard deviation of 3 Ml/d can be associated with the 4 year pit inflow prediction described above. Meanwhile, linear analysis associates a prior standard deviation of about 20 Ml/d with this same prediction. It can therefore be concluded that the information content of the CRT dataset with respect to this prediction is high.

As well as enabling approximate quantification of parameter and predictive uncertainty, linear analysis enables a modeller to track the flow of information from observations to parameters. This can be achieved by undertaking singular value decomposition of the weighted Jacobian matrix. In the present instance, analysis of this matrix establishes that the inversion process is able to extract about 200 separate "pieces of information" from the CRT calibration dataset. That is, it is able to estimate about 200 "super parameters". Super parameters are actually combinations of user-defined parameters that comprise the receptacles into which information contained in the calibration dataset is placed. The inversion process defines these combinations of parameters, and then estimates them.

The number of super parameters exposed through calibration of the model described herein is significantly less than the number of parameters (3,526) with which the model is endowed. This is a good thing, for the greater is the size of a model's parameter set, the greater is the freedom granted to the inversion process in defining combinations of parameters that are optimal for estimation. These combinations are such as to minimise the potential for error (i.e. error variance) associated with any prediction that the model must make. Ideally, a prediction made by the calibrated model is therefore roughly at the centre of its posterior uncertainty interval. Meanwhile, model parameters in excess of super parameters are used by linear analysis to associate posterior uncertainties with its predictions.

Before closing this subsection, we repeat the point made above that the model that is described in this report was not built to predict pit dewatering rates. Instead, it was built to analyse data that may be informative of hydrogeological conditions which impact pit dewatering rates. The above predictions of dewatering requirements made by the present model are invalidated by (among other things):

1. the use of DRAIN cells to simulate an artificial pit that deepens at a uniform rate through a uniform footprint;
2. the proximity of general head boundary conditions to the pit.

The second of the above factors exerts a considerable influence on predictions of pit inflow made by the model described herein, for it constitutes an infinite supply of water within hydraulic reach of the pit. As time progresses and the pit deepens, inflow to the pit can draw on this source of water through connections of relatively high conductance.

Nevertheless, notwithstanding its limitations, the making of the above prediction, and the application of linear uncertainty analysis to it, contributes to the present study for reasons which we now explain.

Model-based interpretation of CRT data has exposed sources of stored water, and pathways to this water, in the vicinity of the OB31 pit. These will affect pit dewatering rates. While long term dewatering requirements for a real-world pit will probably be considerably less than the 9-year inflows plotted in Figure 4.7 because general head boundary conditions which constitute the eastern and western extremes of the model do not exist in the real world, it is nevertheless of interest to note that model predictions of inflow through the eastern boundary are much greater than those through the western boundary; the latter boundary is much closer to the pit than the former boundary. This supports the notion that near-pit extraction has the capacity to draw water from considerable distances. The existence, and in some cases the locations, of permeability pathways through which water may flow to the pit dewatering system have been exposed by assimilation of CRT data.

As explained above, uncertainties in pit dewatering requirements that were evaluated using linear analysis do not neglect errors that are encapsulated in model boundary conditions. At the same time, they indicate that information resident in the CRT dataset is highly pertinent to predictions of pit dewatering rates. This is exposed by the high prior-to-posterior uncertainty ratio of inflows into the artificial pit that was the subject of the above analysis. It follows that any model that is focussed on making predictions of OB31 pit dewatering requirements will benefit from assimilation of these same data. In the case of OB31, the more extensive HDT dataset (see above) encompasses much of this same information. These HDT data were, indeed, processed in models that were used to make predictions of OB31 pit dewatering requirements.

5. DISCUSSION AND CONCLUSIONS

The GMDSI worked example that is documented herein demonstrates a number of important features of decision-support modelling.

First of all, it demonstrates that construction of a decision-support numerical model should not be viewed as construction of a numerical device that can accurately replicate the behaviour of a natural system, either in the past or in the future. Emulation of system behaviour is not its purpose. The purpose of decision-support modelling is to assimilate information that can reduce the uncertainties of decision-critical predictions. This information flows through the parameters with which a model is endowed to the predictions that the model must make. Access to much of this information is gained through history-matching. Therefore, where measurements of system behaviour are available, a decision-support modelling workflow must facilitate the history-matching process. This does not mean that it should eschew information born of site characterisation and expert knowledge. Instead, a decision-support modelling workflow should blend information from these two sources in a way that is appropriate for the geological and management context in which decisions are required.

The constant rate test (CRT) conducted at Orebody 31 (OB31) comprises a rich source of decision-pertinent information. Linear analysis establishes an ability to estimate about 200 super-parameters (combinations of model parameters) on the basis of this information. A challenge faced by the OB31 modelling process was design of a modelling strategy that could capture this information.

Assimilation of information requires a fast-running, numerically-stable model. It also requires design of a flexible parameterisation scheme which is “up to the task” of information assimilation. Expert knowledge and site characterisation may suggest values for some of these parameters, and for relationships between them, that should be respected as long as they are not demonstrably incompatible with information which the inversion process reveals. Meanwhile, a model’s parameters must be numerous enough, and flexible enough, to adopt values and spatial patterns that a modeller may not have foreseen as they are adjusted in order for a model to replicate the measured behaviour of a system. This is especially the case in areas of complex geology where the factors that influence subsurface hydraulic behaviour are only vaguely known. In summary, a decision-support model must be designed “from the ground up” to assimilate all available data.

Previous decision-support numerical modelling of the OB31 site attempted to portray the hydrogeological complexity of the site. However, experience gained in the construction, history-matching and deployment of these models suggests that the optimal response to hydrogeological complexity may not necessarily be numerical model complexity. Construction of a complex numerical model runs a high risk of cementing inadequate hydrogeological concepts into the modelling process. Instead, an appropriate response to hydrogeological complexity may be the construction of a model that can support sufficient parameterisation complexity to assimilate decision-pertinent information, while being simple enough to grant that information free expression through patterns of parameter heterogeneity that emerge from the history-matching process. The modeller is thereby informed of things that matter at a particular site. These will not be expressed as picture-perfect, three-dimensional replications of subsurface heterogeneity. Instead they may be expressed as minimalist patterns of heterogeneity that, if properly interpreted, may either confirm or challenge currently-held understandings of hydrogeological processes that are operative at that site.

The concept of “conceptual model” may therefore require some revision if it is to provide a more useful basis for decision-support modelling than it does now. Concepts which guide decision-support modelling as it is implemented at a particular study site should include more than an imperative to build a numerical model that echoes current hydrogeological understanding of that site. They should recognise that currently available data, if properly processed, may alter current hydrogeological understandings. In doing so, they should recognise that the numerical modelling process should be designed to question current hydrogeological concepts as much as it reflects them. This allows modellers and modelling stakeholders to learn as much as possible about the site through the modelling process.

The modelling workflow that is documented in this GMDSI report demonstrates that where a site is hydrogeologically complex, attempts to express hydrogeological detail in a model that must also assimilate information-rich site data may actually diminish the decision-support potential of the modelling process rather than magnify it. Where a site is complex, but the details of that complexity are only vaguely known, the conceptual model on which a numerical model rests should impel the numerical modelling process to provide a canvas on which data can paint. The picture that emerges from assimilation of these data may be somewhat abstract. However this does not devalue the worth of these data. Modellers do not have the luxury of enjoying representational art; any “picture perfect” representation of the subsurface is almost certainly incorrect.

The parameter field that emerged from calibration of the OB31 model requires hydrogeological interpretation – more so in some parts of the model domain than in others. It is a projection onto a two-dimensional modelling landscape of complex, three-dimensional, hydrogeological features. Linear analysis demonstrates that some aspects of these parameter patterns should be taken very seriously. This does not mean that they should be interpreted literally. Rather it means that they are indicative of sources of water, and pathways along which water can flow, that are real, and that may have a serious impact on OB31 pit dewatering requirements. The exact nature and disposition of these water sources and pathways is a matter for local geological expertise to address. However, for some decisions, further details may not be necessary; their existence alone may provide the information on which a decision can be based. For example, hypothetically (as the time has already passed) it may be possible to formulate licencing and pipeline requirements for OB31 pit dewatering based only on assimilation of CRT data. In contrast, the number and placement of pit dewatering wells may require the acquisition of further data, followed by model-based assimilation of these data.

The history of OB31 modelling exemplifies approaches to decision-support modelling that, in the authors’ opinion, leave some room for improvement. The premise that sustained much of that modelling, and indeed sustains much of today’s decision-support groundwater modelling in general, is that movement of groundwater can be simulated accurately by a numerical model. In areas of complex geology, accurate simulation of groundwater movement requires representation of hydrogeological detail. However this detail cannot be known. It must therefore be represented stochastically. Predictive certainty is therefore not an option. The best that can be expected is reduction in the uncertainties of decision-critical predictions through data assimilation effected by history-matching.

It is an inconvenient truth that current technology does not yet allow the information that history matching conveys to a model to be expressed as an extensive suite of picture-perfect, calibration-constrained stochastic realisations of hydrogeological detail. Furthermore, even if this were technically possible, a hydrogeologist may not be aware of some of the complex three-dimensional details that the stochastic history-matching process may need to express. Failure to give stochastic voice to all geological possibilities may therefore bias this process. This is especially the case in an area that has undergone as much tectonic deformation as that which is the focus of the current study.

An alternative decision-support strategy is to design a modelling workflow that grants the history-matching process the freedom that it needs to express bottom-line heterogeneity that is required to explain measurements of system state. If the hydraulic stresses that prevailed when these measurements were made resemble those to which the system will be subjected in the future, then history-matching undertaken in this way may be capable of expressing all of the hydrogeological details that matter to decision-salient model predictions. Uncertainties can still be quantified, and decisions can still be made. Meanwhile the modelling process has served the decision-support imperatives of uncertainty quantification and reduction in a way that is tailored to the decisions that it must support.

6. REFERENCES

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., 2015. Calibration and uncertainty analysis for complex environmental models. Published by Watermark Numerical Computing, Brisbane, Australia. 227pp. ISBN: 978-0-9943786-0-6. Downloadable from www.pesthomepage.org.

Doherty, J., 2020a. PLPROC: A Parameter List Processor. *Watermark Numerical Computing*, Brisbane, Australia.

Doherty, J., 2020b. PEST_HP: PEST for Highly-Parallelized Computing Environments. *Watermark Numerical Computing*, Brisbane Australia.

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

Doherty, J. and Moore, C., 2019. Decision support modeling: data assimilation, uncertainty quantification, and strategic abstraction. *Groundwater* 58(3): 327-337. doi:10.1111/gwat.12969

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

Panday, S. and Huyakorn, P.S., 2007. MODFLOW SURFACT: A state-of-the-art use of vadose zone flow and transport equations and numerical techniques for environmental evaluations. *Vadose Zone Journal.* 7(2):610. DOI:10.2136/vzj2007.0052.

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, chap. A45, 66 p.

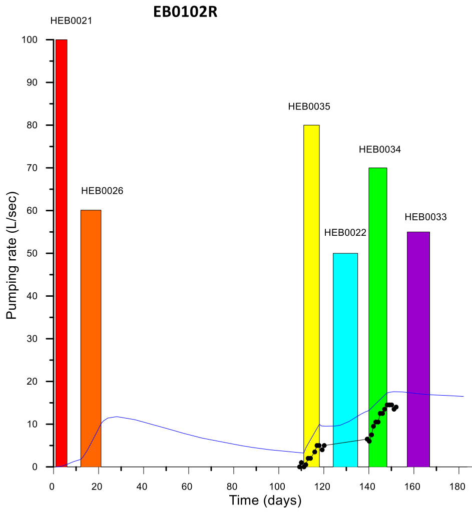
McFarlane, DJ (ed) (2015). Pilbara Water Resource Assessment: Upper Fortescue region. A report to the Government of Western Australia and Industry Partners from the CSIRO Pilbara Water Resource Assessment. CSIRO Land and Water, Australia.

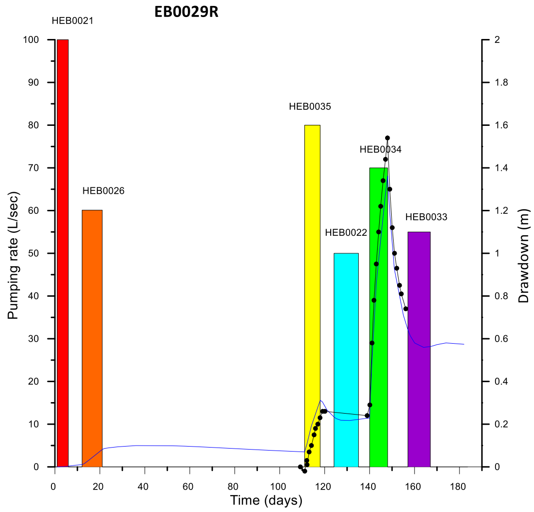
White, J.T., Doherty, J.E. and Hughes, J.D., 2014. Quantifying the predictive consequences of model error with linear subspace analysis. *Water Resour. Res, 50 (2): 1152-1173.* DOI: 10.1002/2013WR014767

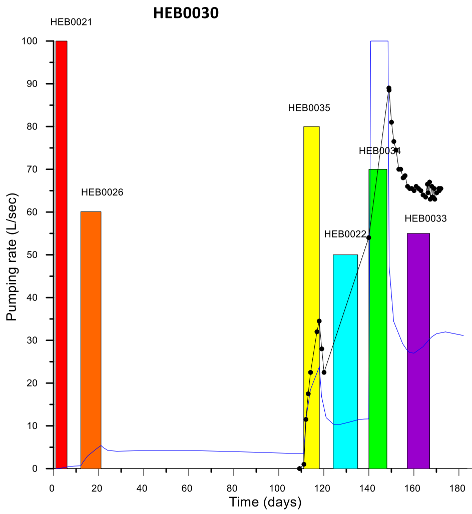
White, J.T., Fienen, M.N. and Doherty, J., 2016. A python framework for environmental model uncertainty analysis. *Environmental Modelling & Software*. 85, 217-228.

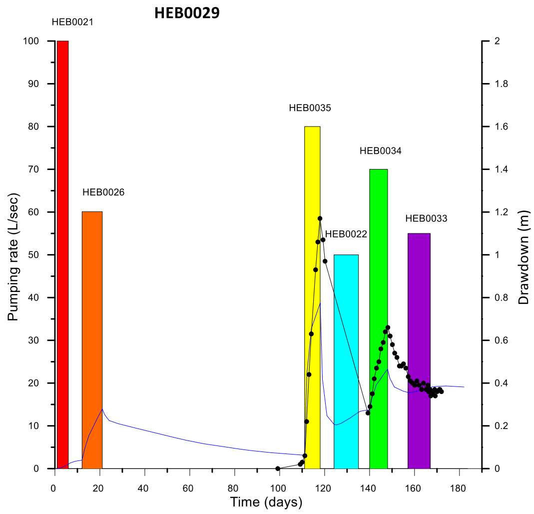
White, J.T., 2018. A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions. *Environmental Modelling & Software*. 109. 10.1016/j.envsoft.2018.06.009. http://dx.doi.org/10.1016/j.envsoft.2018.06.009.

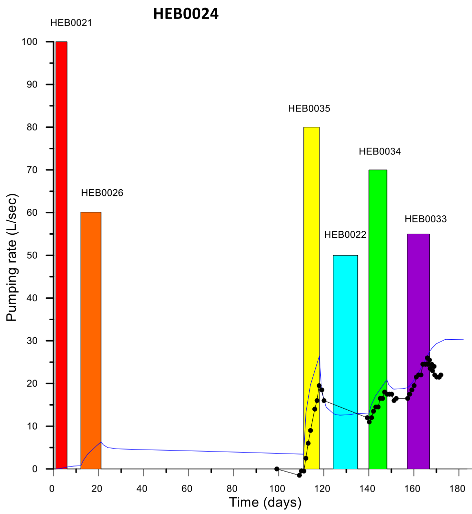
APPENDIX. MODELLED AND OBSERVED DRAWDOWNS

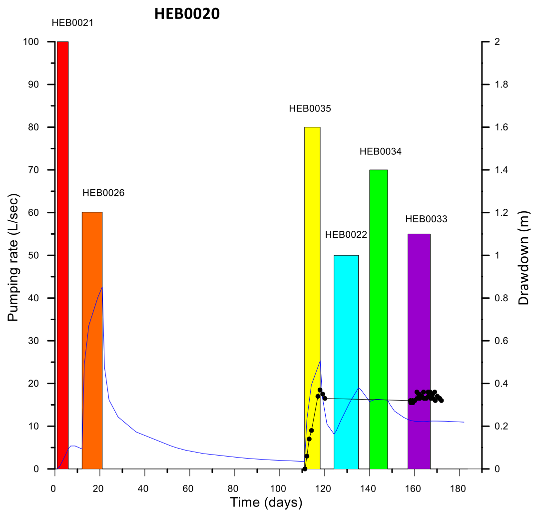


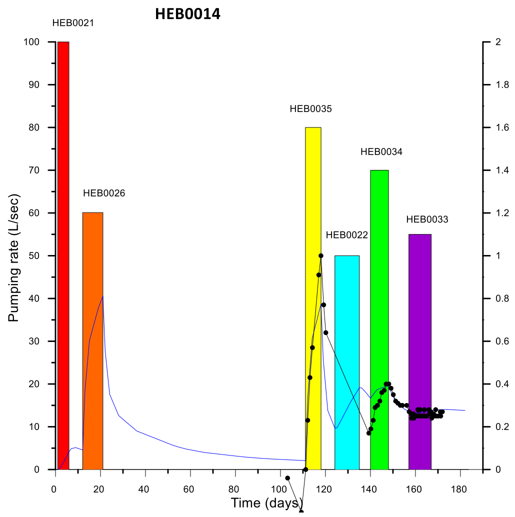


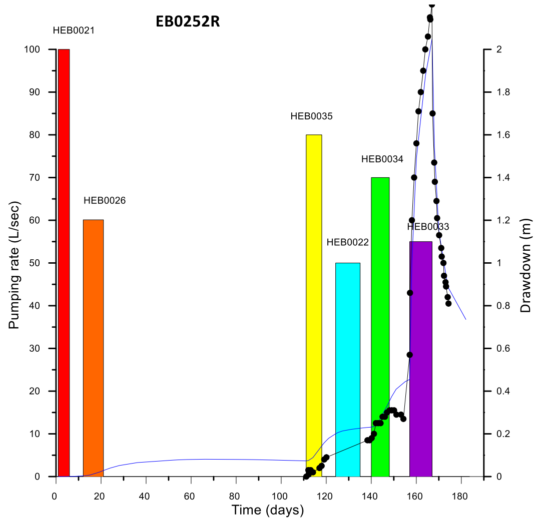


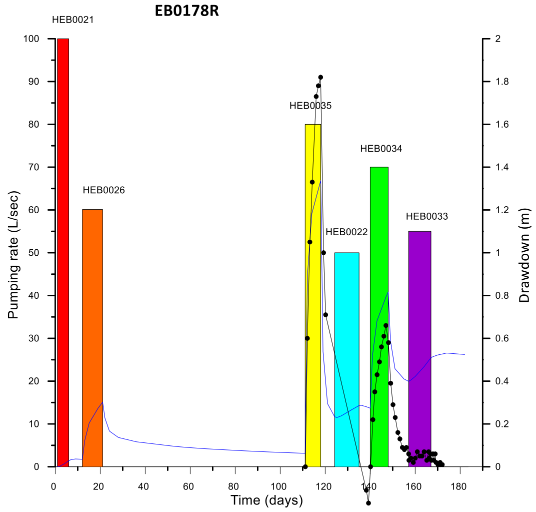


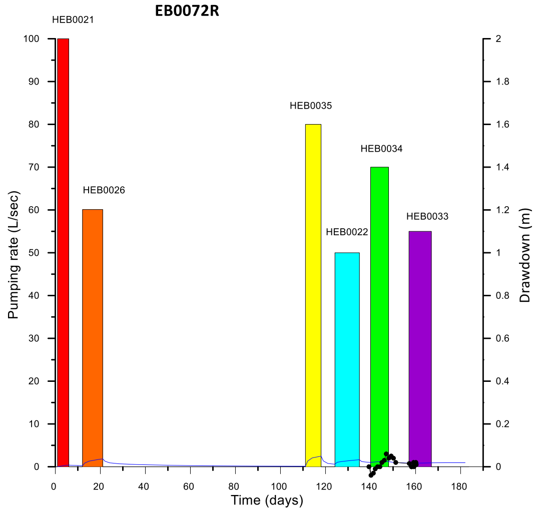


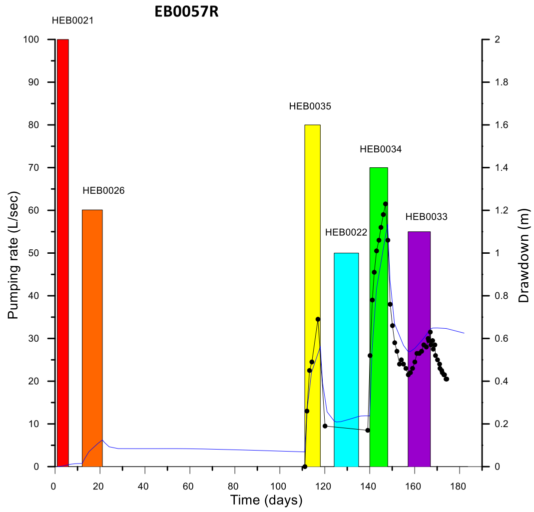


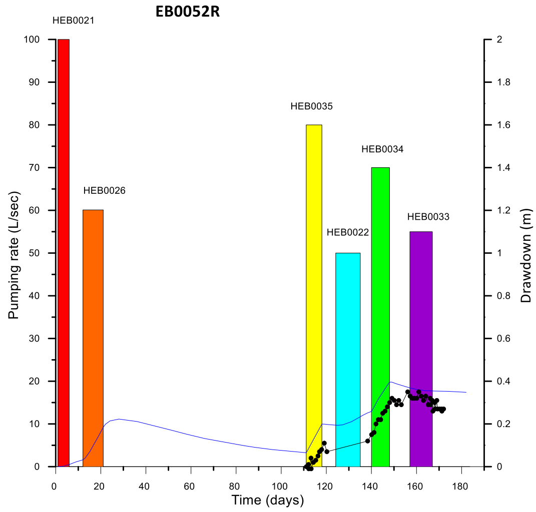

HEB0014
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)

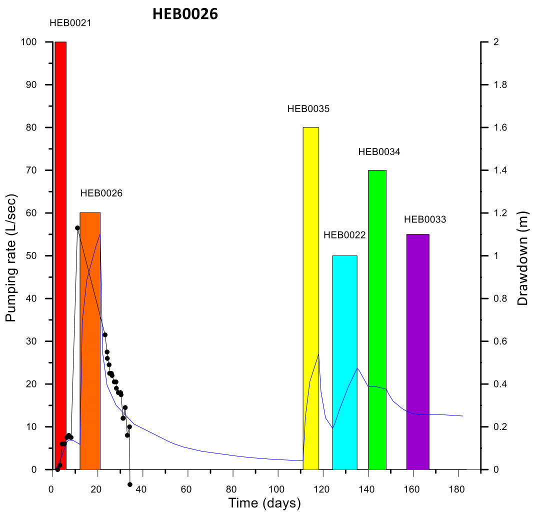

EB0252R
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)

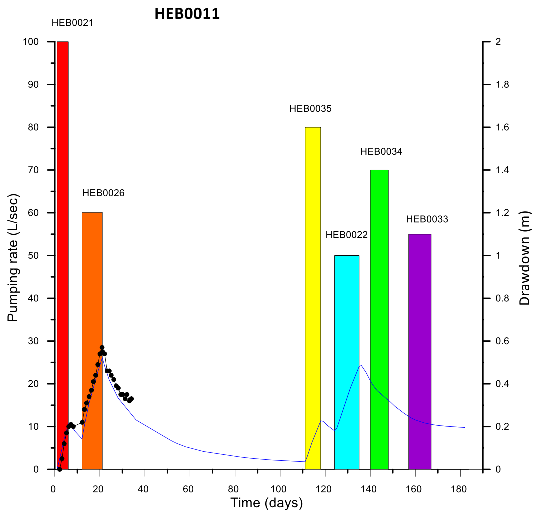

EB0178R
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)

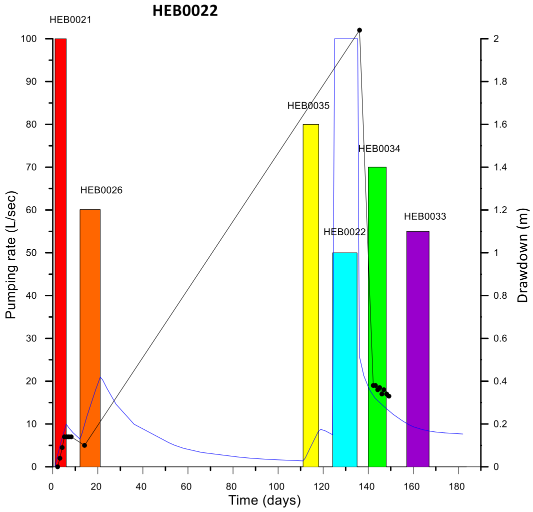

EB0072R
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)

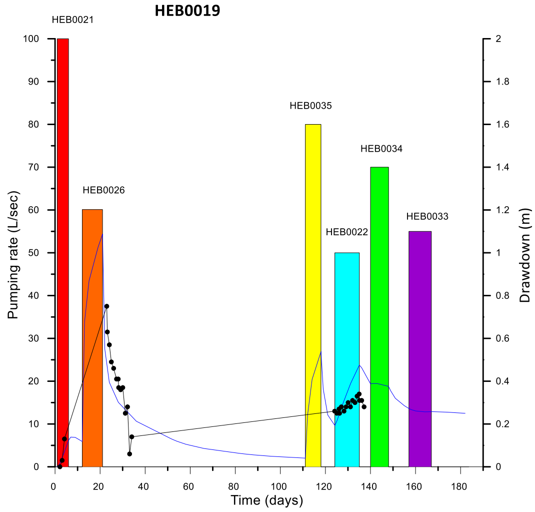

EB0057R
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)

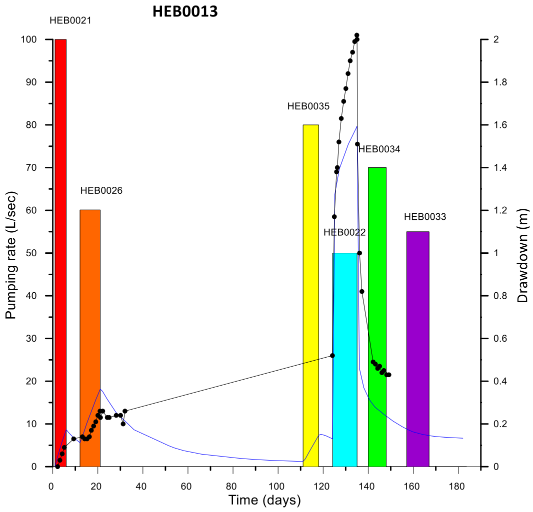

EB0052R
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)

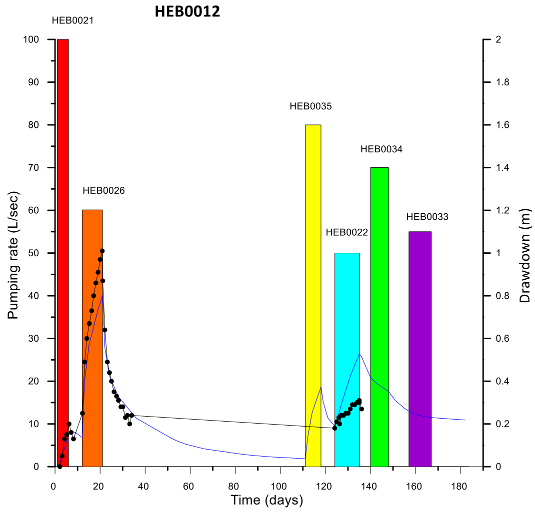

HEB0026
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)

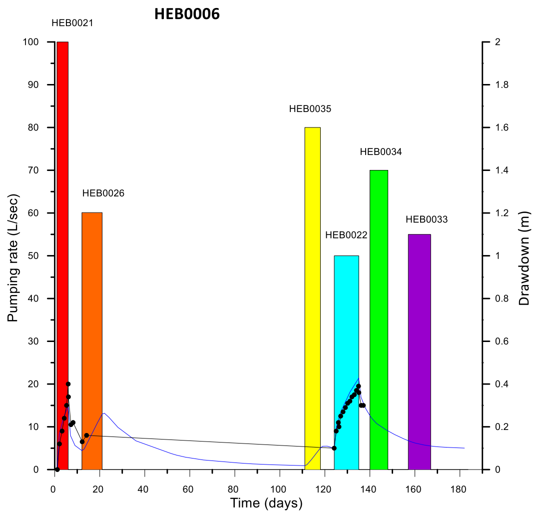

HEB0011
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)

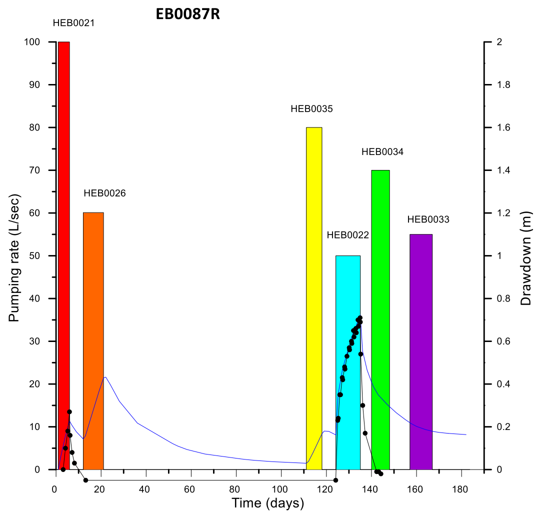

HEB0022
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)

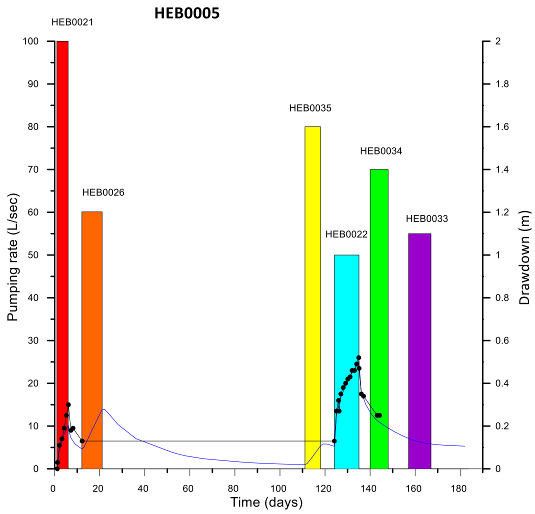

HEB0019
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)


HEB0013
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)


HEB0012
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)


HEB0006
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)


EB0087R
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)


HEB0005
HEB0021
HEB0026
HEB0035
HEB0034
HEB0022
HEB0033
Pumping rate (L/sec)
Drawdown (m)
Time (days)

gmdsi
Groundwater Modelling
Decision Support Initiative

gmdsi.com

CRICOS NO 00114A

BHP
Flinders UNIVERSITY
NATIONAL CENTRE FOR GROUNDWATER RESEARCH AND TRAINING
RioTinto