Model-Based Assessment of Coastal Aquifer Management Options

A GMDSI worked example report

by Rui Hugman, John Doherty and Kath Standen

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: Hugman, R., Doherty, J. and Standen, K., (2021). Model-Based Assessment of Coastal Aquifer Management Options. A GMDSI Worked Example Report. National Centre for Groundwater Research and Training, Flinders University, South Australia.

FUNDING ACKNOWLEDGEMENT

The co-author, Kath Standen, has received funding in the frame of the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no 814066.

ISBN:

DOI:

DOI as a weblink:

COPYRIGHT

PREFACE

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

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

The decision-support challenges that are addressed by 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 an aquifer 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 has been 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. Rui Hugman, GMDSI Project, Flinders University.

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.

Capture zone

The three-dimensional volumetric portion of a groundwater-flow field that discharges water to a well.

Connected linear network (CLN) package

This package is supported by the MODFLOW-USG simulator. Water flows through a series of one-dimensional features, each of which can be linked to another such feature, or to a cell within a two or three-dimensional groundwater model domain.

Contributing area

The two-dimensional areal extent of that portion of a capture zone that intersects the water table and surface water features where water entering the groundwater flow system is discharged by a well. (This is also referred to as the *area contributing recharge.*)

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 rate, 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-USG

A version of MODFLOW which employs an unstructured grid. This was developed by Sorab Panday in conjunction with 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 promote 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.

River (RIV) package

A MODFLOW package which provides basic simulation of the interaction between groundwater and a surface water body. Flow between the two regimes is driven by the head difference between them. Through definition of the elevation of the bottom of the river, the driving head difference can be limited.

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 of a Jacobian matrix.

Specific storage

The amount of water that is stored elastically in a cubic metre of 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.

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.

Vector

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

Well (WEL) package

A MODFLOW package that simulates withdrawal of water from a groundwater system.

EXECUTIVE SUMMARY

Coastal groundwater management must be handled with care. Excessive extraction of water from a coastal aquifer can incur sea water intrusion. This can damage the water resource.

Numerical simulation can assist in management of coastal aquifers. However its use in a decision-support context is far from straightforward. Modern modelling technology allows us to simulate details of fresh and saline groundwater flow through heterogeneous porous media. However models which can do this are complex and slow-running. Because of this, it is difficult for them to assimilate water quality and piezometric data that record the historical behaviour of a managed coastal system. The uncertainties of decision-critical model predictions may therefore remain high. This can thwart decision-making.

This GMDSI report describes a model that was built to explore options for management of a coastal aquifer in southern Portugal. The aquifer is representative of many around the world; if extraction continues at its present rate, it is only a matter of time before it suffers a serious degradation in quality. Extraction must therefore be reduced. Alternatively, or as well, recharge must be enhanced.

Enough data has been gathered over the last 20 years to support estimation of aquifer properties and inflows. These estimates are enabled by history-matching; however they are cloaked in uncertainty.

Data assimilation is implemented using a single-density, MODFLOW 6 model in conjunction with PEST_HP and PESTPP-IES. Use of a single-density model rests on the premise that water quality within the coastal aquifer is not yet seriously degraded. Meanwhile, proximity of saline waters is acknowledged by deploying a two-dimensional, sectional, density-dependent model to assist in stochastic parameterization of the coastal boundary condition employed by the single-density model.

History-matching yields 200 parameter fields, all of which allow the model to reproduce historical measurements of aquifer head under the seasonally-variable pumping regime that has prevailed over the last twenty years. Water is extracted for irrigation. A suite of history-matched, LUMPREM soil-moisture accounting models is used to calculate irrigation demand, and hence aquifer extraction.

The composite LUMPREM - MODFLOW 6, history-matched model is given two tasks. First, it is used to calculate the overall water balance of the aquifer. Elements of this water balance are used to document the state of the system, and to satisfy the terms of its short-term management according to criteria set out by the European Union Water Directive.

The model is also used to establish an upper limit on groundwater extraction subject to the condition that aquifer water quality be maintained. This requires solution of a constrained optimisation problem. This problem is solved twice – once (using PESTPP-OPT) under risk-neutral conditions, and once (using CMAES_HP) under risk-averse conditions. In the latter case, CMAES_HP is asked to maximise water extraction under the constraint that flow of water across the coastline remains seaward rather than landward for all 200 samples of the posterior parameter probability distribution.

CONTENTS

1.INTRODUCTION

1.1 General

Work that is described in this manuscript took longer to complete than we had originally intended. This is a common modelling refrain. Furthermore, if we had to start over, there are some things that we would do differently. Nevertheless, we learned a lot along the way. The GMDSI worked example reporting genre gives us a chance to describe what we did in a way that allows others to benefit from both our achievements and our mistakes.

We tried to keep the modelling simple. Nevertheless, parts of our workflow ended up being rather complicated. A benefit of model simplicity is that, with proper design, it readily supports the introduction of strategic complexity where this enhances decision-support utility. In order to maintain focus on things that matter, our description of some of these complexities is brief. Our intention is to describe our modelling approach rather than to recount all of its details.

Our study area is Vale do Lobo ("Valley of the Wolf" in English), situated on the idyllic southern coast of Portugal. A coastal aquifer faces the threat of sea water intrusion because of excessive groundwater extraction. Hence extraction must be reduced, or freshwater recharge must be enhanced. The area is being studied by personnel from Universidade do Algarve. GMDSI modelling documented herein parallels their work. Our approach recognises the centrality of data assimilation and uncertainty quantification to decision-support modelling. It also recognises the need to keep model run times short in order to enable these activities.

Modelling that is documented herein has allowed us to:

- Derive the water balance of an aquifer system that supports assessment of its status according to guidelines set out by the European Union Water Framework Directive;
- Explore options for maximisation of water extraction while ensuring system sustainability.

Before describing the Vale do Lobo site, we briefly outline some of the challenges that we faced in undertaking the work described herein. Many of these challenges are common to groundwater modelling work in general.

1.2 Challenges

1.2.1 Data Scarcity

The only acceptable response to the challenge of data scarcity is to ensure that uncertainty assessment is woven into the fabric of model design and deployment. This eliminates the need for modelling to rest on one or many untested assumptions. If a parameter, a boundary condition, or another aspect of model design, is only partially known, then its representation in a model should be stochastic rather than deterministic.

This approach to decision-support modelling is both practical and forgiving. If model simplification facilitates uncertainty analysis and reduction, while incurring errors that are small in relation to ambient uncertainties, then they are worthwhile. At the same time, they can make modelling easier. However the potential for error that model simplifications incur should be included in model predictive uncertainty intervals.

1.2.2 Salt Water

Simulation of coastal groundwater processes is numerically difficult. Models which simulate density-dependent flow must have many layers. Their run times are generally long; they operate under a

cloud of numerical instability. Model-based data assimilation and uncertainty quantification become difficult or impossible.

Problems associated with simulation of coastal groundwater flow are compounded by lack of knowledge of under-sea aquifer conditions. This is especially the case for confined aquifers from which waters may emerge tens of kilometres offshore.

Fast-running, sharp interface models can be used in place of density-dependent models. Our experience, however, is that they are not immune from numerical instability, and that they may therefore prove unreliable when deployed in data assimilation settings.

The model that is described in the present study is density-independent. Furthermore, its coastal boundary coincides with the coast itself. We employ a complementary, two-dimensional, density-dependent model to provide stochastic characterisation of this boundary. This allows us to express uncertainties in boundary parameters that arise from poorly-known but complex conditions that prevail on the seaward side of the coast, as well as those that arise from use of a simplified boundary to represent these conditions.

1.2.3 Other Lateral Boundaries

Selection of the areal extent of a groundwater model domain is often difficult. This is especially the case if the "natural" boundaries of a groundwater system are far from areas that are the focus of management interest.

Using an unstructured grid, a model domain can be extended to natural system boundaries without the incursion of an unduly high numerical cost. However this does not eliminate the need for stochastic representation of hydraulic properties, recharge, extraction and other nuances of water management in the extended model domain if these can affect conditions in a study area. Gathering of data that is necessary to support stochastic model representation of these conditions may be expensive and time-consuming.

Sometimes it may be more efficient to restrict the size of a model domain to the area of immediate management interest, and to equip model boundary properties with sufficient stochasticity to span uncertainties that are induced by conditions on the other side of the boundary. This may reduce model construction costs. However if this option is chosen, it is incumbent on a modeller to guarantee that:

- use of the boundary in place of spatially-distributed processes on its other side does not invalidate simulation integrity inside the boundary; and that
- the stochasticity of boundary parameters is sufficient to characterise predictive uncertainty arising from conditions outside the boundary, as well as from simplifications incurred by use of the boundary itself.

In the present study, one of our model boundaries coincides with a management boundary rather than with a natural boundary. Extension of the model domain would have added to the tedium of data preparation and model construction; it would have also added to the model run time. However, it could have reduced problems associated with parameterisation of the boundary.

1.2.4 Number of Model Layers

The aquifer which is the focus of our modelling attention is confined. It is separated from an overlying unconfined layer by an aquitard whose thickness varies over the model domain. Water levels within the overlying, unconfined aquifer vary seasonally. The strength of its connection with the underlying confined aquifer appears to be weak.

In the interests of speed and simplicity, the unconfined aquifer is not explicitly represented in the model that is described herein. It is replaced by a pervasive "general head" boundary with spatially varying conductance and temporally varying heads.

1.2.5 Historical Pumping

Extraction of irrigation water from the confined aquifer has resulted in unsustainable drawdown. Unfortunately, records of water extraction are incomplete. So we deploy a number of fast-running, lumped-parameter, soil moisture models to fill in temporal and spatial gaps in these records. Outputs of these models are history-matched against measured extraction rates where the latter are available. Meanwhile, stochastic parameterisation of these models ensures that uncertainties in historical extraction rates contribute to uncertainties in estimated model parameters, and thence to management-critical model predictions.

1.3 Looking Ahead

Section 2 of this report describes the Vale do Lobo site and the management problems that modelling must address. Section 3 describes simulation of irrigation demand using a series of lumped parameter soil moisture accounting models. Details of the Vale do Lobo groundwater model are provided in Section 4. Use of a companion density-dependent model to assist in stochastic parameterisation of the coastal boundary condition of the groundwater model is described in Section 5.

History-matching of the composite model (comprised of the groundwater model and the suite of soil moisture models) is described in Section 6 of this report. Use of this model to make management-pertinent predictions, and in assessing the uncertainties of these predictions, is described in Section 7. Section 8 concludes our report.

2. THE SITE

2.1 Location

Figure 2.1 depicts the domain of the groundwater model that is described in this report. Its boundaries roughly coincide with management boundaries which define the Vale do Lobo subsystem of the Campina de Faro aquifer.

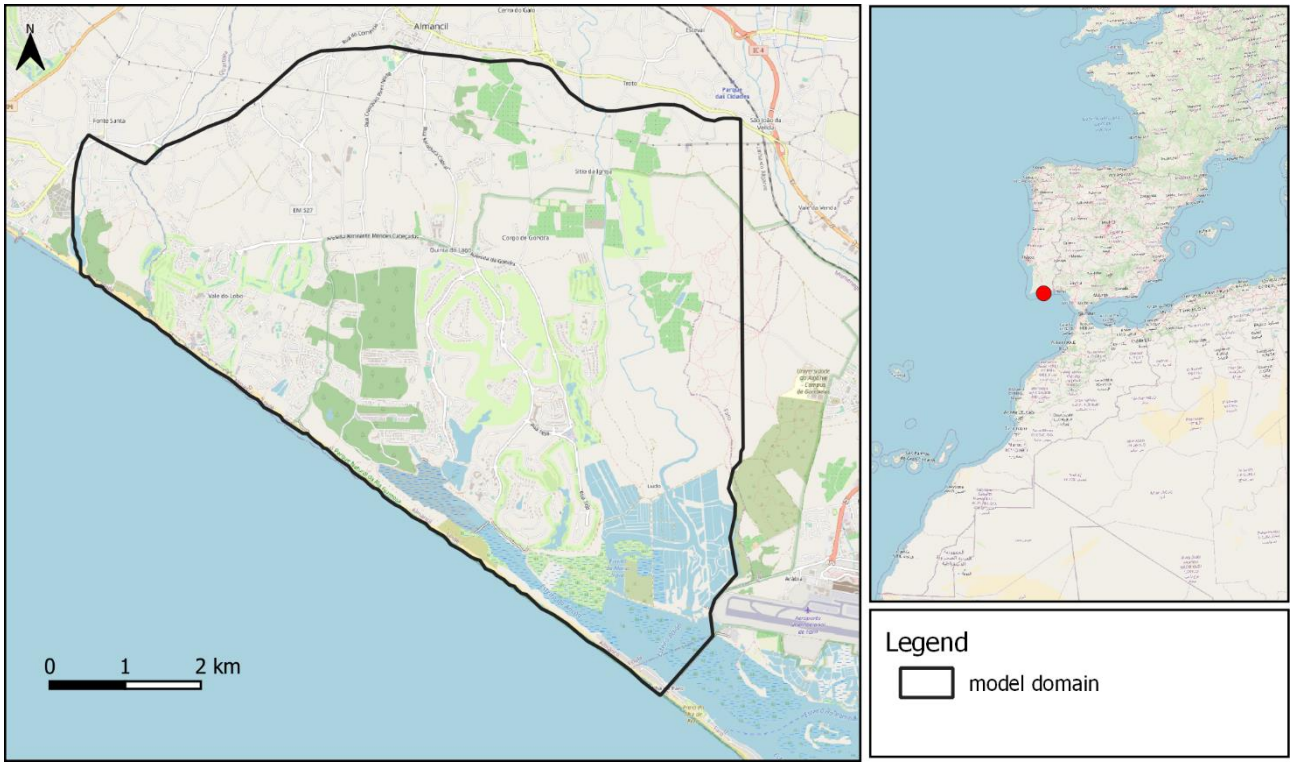


Figure 2.1 Location and domain of the Vale do Lobo model.

The study area is located to the west of Faro, capital of the Algarve province of Portugal. It is bounded to the southwest by the Atlantic Ocean. It encompasses the town of Almancil, and the two resort villages of Quinta do Lago and Vale do Lobo.

Land use is depicted in Figure 2.2. It includes seven 18-hole golf courses as well as intensive agriculture for salad crop production. The remainder of the area is covered by semi-natural sparse pine forest, as well as patches of land devoted to avocado and citrus fruit cultivation. Groundwater is used for irrigation of golf courses and agricultural land.

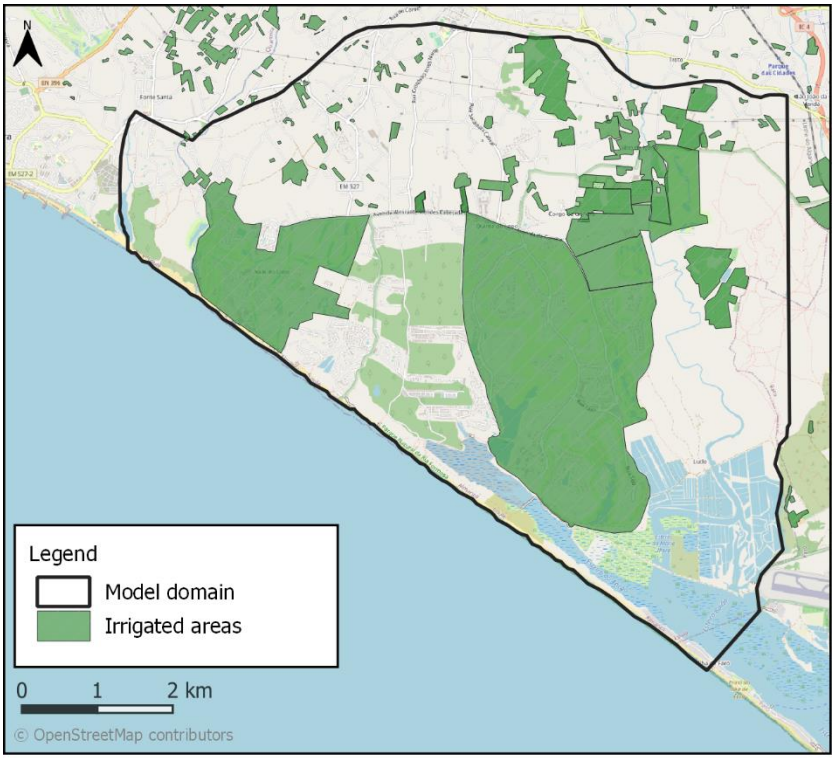


Figure 2.2 Land use map identifying irrigated areas; adapted from Reis (2007).

The long term annual average rainfall is about 600 mm/yr. Most rainfall occurs between the months of November and April. Potential evapotranspiration is around 1600 mm/yr. There is a substantial excess of potential evapotranspiration over rainfall during the summer months. Most irrigation is applied between the months of March and October.

2.2 Hydrogeology

The Vale do Lobo subsystem (hereafter referred to as the VL subsystem) of the Campina de Faro aquifer occupies an area of about 32 km^2. The boundary which separates it from the Campina de Faro subsystem to the east is a management rather than a hydrogeological boundary; it runs roughly perpendicular to topographic contours. To the northwest, the VL subsystem boundary is defined by the Carcavai fault zone, on the other side of which lies the Quarteira aquifer. Lower Cretaceous strata outcrop to the north of the northern VL subsystem boundary; further to the north, Jurassic sediments form a karstic aquifer.

The VL subsystem is part of a thick series of sediments encompassed in distinct and superimposed sedimentary basins of Mesozoic and Cenozoic age. These are underlain by Palaeozoic basement rocks. Figure 2.3 maps outcropping geology and hydrogeological features of interest.

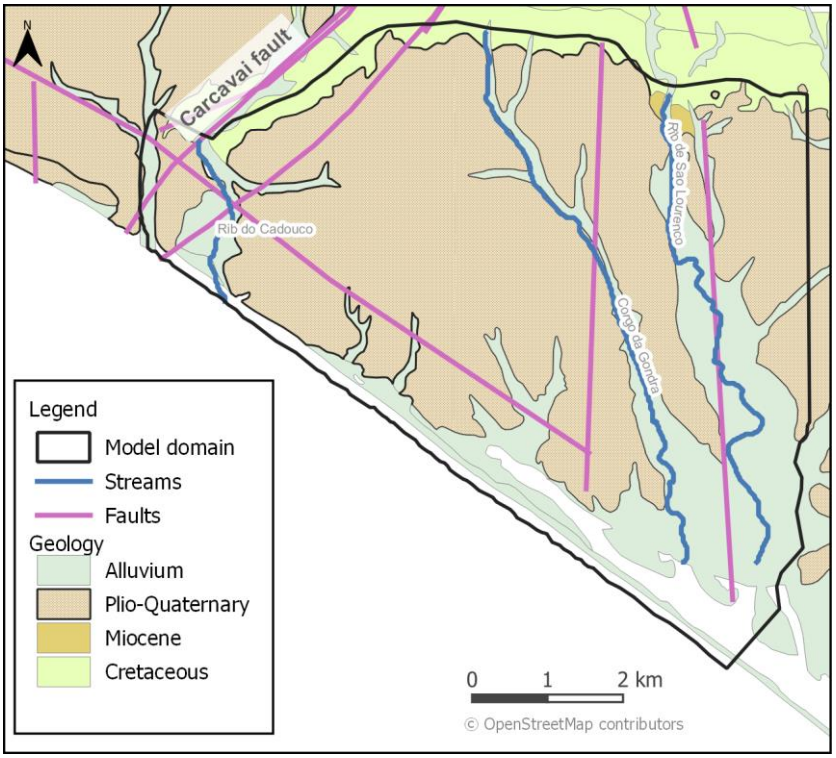


Figure 2.3 Geological outcrop map of the study area.

The VL subsystem is comprised of two aquifers. These are as follows:

- An upper, phreatic aquifer of Plio-Quaternary (PQ) age. This is composed of sands and sandy clays, with a red weathered clay typically found at the surface. This aquifer attains a maximum thickness of about 70m in the south of the study area; it thins to the north and eventually disappears where marls of Lower Cretaceous age outcrop.
- A lower, confined aquifer of Miocene (MC) age formed of calcareous sandstones and limestones. Nearby and offshore drilling suggests that the base of the MC aquifer reaches a depth of 350 metres below mean sea level in the south of the study area. It is underlain by low permeability marls of Lower Cretaceous age.

A clay aquitard, typically about 10m thick, separates the MC aquifer from the PQ aquifer. Water levels in the PQ aquifer are consistently higher than in the MC aquifer. Most water is extracted from the MC aquifer; however some boreholes may be screened across both aquifers.

The MC and PQ aquifers, together with the aquitard that separates them, dip gently towards the coast at about 4 degrees. Fresh water that flows seaward through the MC aquifer emerges into the sea at an unknown distance offshore. Presumably, it flows through the aquitard that overlies the MC aquifer in order to do so. However the offshore perseverance of this aquitard, together with all other details of offshore freshwater flow, are unknown.

It is apparent from Figure 2.3 that the study area is intersected by several faults. As stated above, the Carcavai fault zone bounds it on the west. Two NNW-SSE-oriented faults transect the eastern part of the VL area. Their dispositions are somewhat uncertain; it is possible that their alignment is closer to those of the Corgo da Gondra and Ribeira da São Lourenço streams than is depicted in Figure 2.3. A strike-slip fault parallels the coast approximately 1 km inland from it.

The Corgo da Gondra and Ribeira da São Lourenço streams flow for an average of 2 months per year. In some years they do not flow at all.

2.3 Piezometric Heads

The local regulatory authority is the Agência Portuguesa do Ambiente (APA). APA regularly measures water levels in 11 monitoring wells that lie within the study area. Three of these are screened in the PQ aquifer, while eight are screened in the MC aquifer; the latter are labelled in Figure 2.4. Only heads in the MC aquifer are simulated by the model described herein. Hence only heads in MC-screened wells are used in model history-matching.

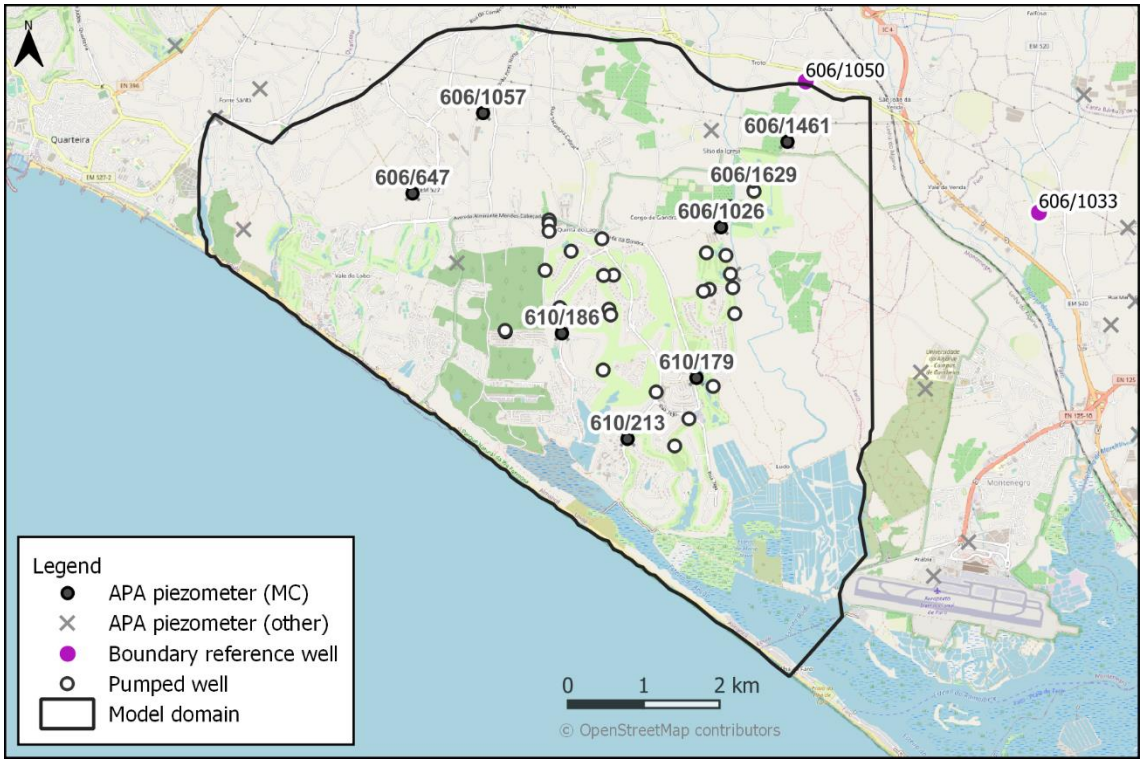


Figure 2.4 Locations of wells in which head measurements have been made.

Four major groundwater users record water levels at approximately monthly intervals in four extraction wells; these are labelled as “pumped wells” in Figure 2.4. Heads from these wells are also used for history-matching. However, because at least some of these measurements may have been made when the wells were being pumped, the history-matching process does not insist that pertinent model outputs match them exactly. Instead, history matching ensures that model-calculated heads are no lower than these measured heads.

Heads measured in bore 606/647 are graphed in Figure 2.5. This is the longest record of piezometric heads available in the VL area. It exhibits a gradual decline from the late 1970’s to the late 1990’s, at which time groundwater levels appear to reach a new equilibrium. However, seasonal fluctuations increase over this time; this probably reflects increased pumping from the MC aquifer.

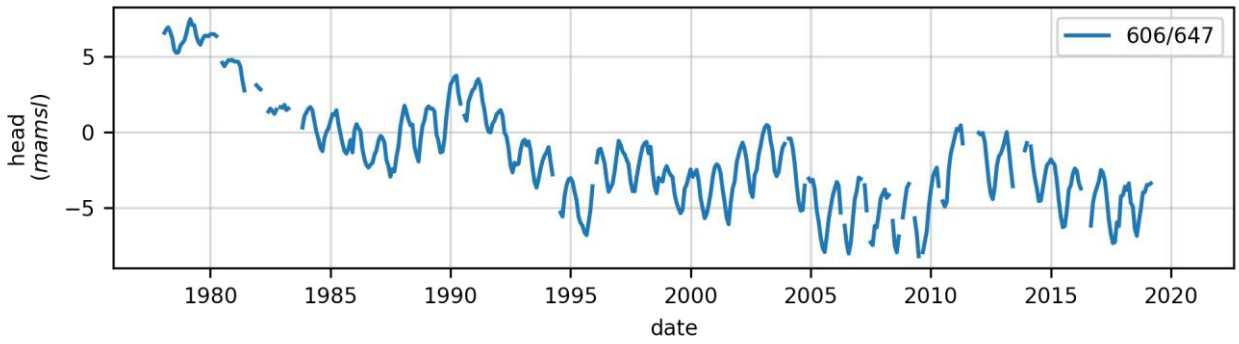


Figure 2.5. Time-series of hydraulic heads measured in piezometer 606/647. The location of this piezometer is displayed in Figure 2.4.

Figure 2.6 shows MC piezometric contours. These are based on average measured heads during 2020. Groundwater heads are, on average, below sea level in all MC piezometers. Lowest levels are in the central and north-western corner of the VL system. Groundwater appears to flow towards this area of depressed heads from all system boundaries, including from adjacent aquifers and from the coast.

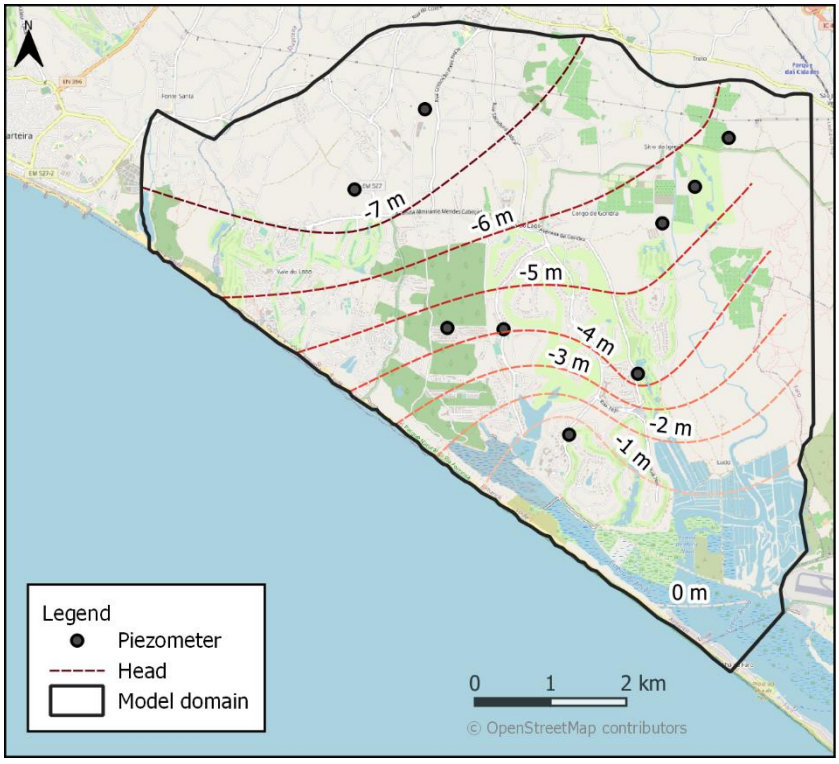


Figure 2.6 Contoured average hydraulic heads in 2020 measured in piezometers belonging to the APA monitoring network.

2.4 Recharge

Because of its impermeable surficial clay layer, diffuse recharge to the unconfined PQ aquifer is thought to be low. It is estimated to be about 3.4 mm/yr; this totals about 3.46 Mm^3/yr for the whole VL system. Ephemeral streams may provide further recharge; however the amount of this recharge is limited by short streamflow durations. Vertical recharge to the MC aquifer is likely to be smaller than that to the PQ aquifer because of the aquitard that separates these layers. The absence of a response to rainfall in MC aquifer piezometric records supports this conclusion.

Under natural conditions, the MC aquifer probably receives most of its recharge laterally through lower Cretaceous strata which abut the northern boundary of the VL system. An upper limit of 14.3 Mm^3/yr can be placed on average long-term recharge from this source. This upper limit is calculated by multiplying maximum expected diffuse recharge by the outcrop area of aquifers upgradient of the northern VL boundary. As is described below, the actual recharge through this boundary (together with its associated uncertainty) is estimated through history-matching.

2.5 Extraction

APA estimates that 6.45 Mm^3 of water was extracted from the VL system during 2019. The vast majority of this water was extracted from the MC aquifer through boreholes that are between 100 m and 200 m deep.

Figure 2.7 shows registered VL extraction wells. These are grouped into 8 major user groups and a "small users" group. Groups 1 to 5 are individual golf courses. Groups 6 and 7 are operated by companies that irrigate greenspace such as parks and private estates. Group 8 is comprised of significant horticultural irrigators. "Small users" are mostly horticultural and agricultural plots and private residences.

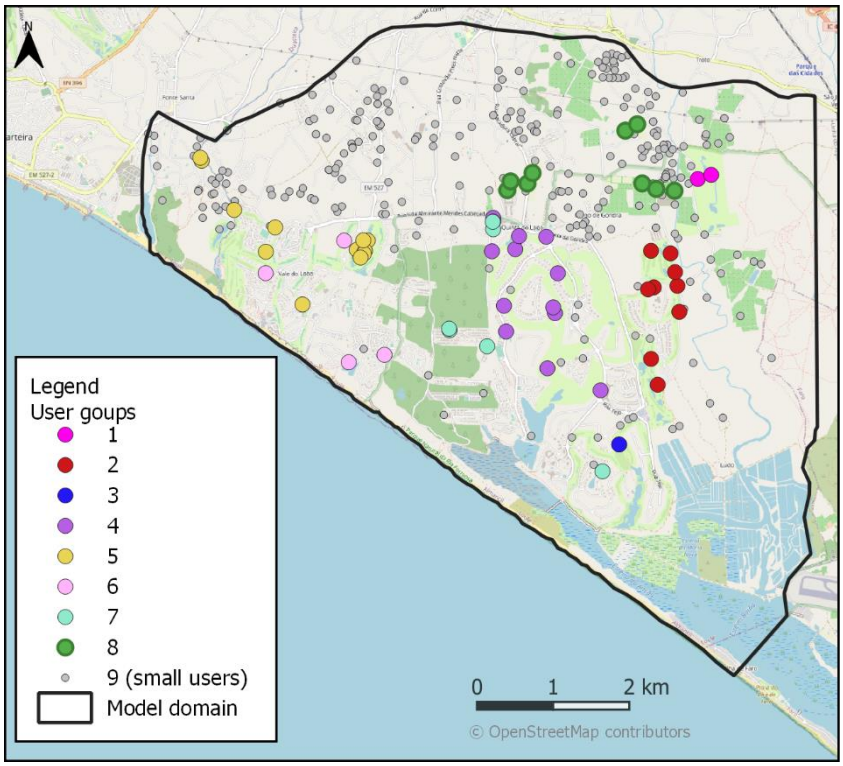


Figure 2.7 Major groundwater user groups in the Vale do Lobo area.

Since 2010, major groundwater users have been required to supply monthly records of water use to APA. These monthly water use figures comprise important model inputs. However use of these data is not without its problems. In particular:

- As is discussed below, the model history-matching period extends from October 2000 to October 2020. Measured extraction rates are not available for the first half of this time period.
- There are gaps in some water use records within the period 2010 to 2020.
- In general, water use is recorded for groups of wells belonging to a single user group, rather than for individual wells.
- Extraction rates are not available for the "small users" group.

A suite of “extraction models”, history-matched against available water use data, is used to fill in data gaps.

2.6 Groundwater Quality

Chloride concentrations measured at a limited number of observation wells within the MC aquifer are generally below 300 mg/L. However these measurements, as well as anecdotal evidence pertaining to the quality of water extracted from production wells, suggests increasing salinities over time at some locations. There is some evidence that this may be at least partly attributable to dissolution of evaporites that are disseminated through sediments which comprise the MC aquifer. It appears, therefore, that VL groundwaters have not yet suffered a serious decline in quality as a result of sea water intrusion. Nevertheless, Figure 2.6 suggests that its occurrence is inevitable.

No water bores that have been drilled into the MC aquifer have intersected a freshwater-saltwater interface. This does not preclude the existence of this interface at depths greater than have been drilled (or sampled) so far, particularly near the shoreline.

As is stated above, the base of the MC aquifer is thought to be about 350 m below sea level along at least some of the coast. The Ghyben-Herzberg principle dictates that a coastal piezometric head of 8.75 m would be required for the toe of the equilibrium freshwater-saltwater interface to lie directly beneath the coast. For piezometric heads less than this, the equilibrium position of the interface toe would be further inland. Unfortunately, records of pre-pumping heads in the MC aquifer are non-existent. The oldest recorded head in the aquifer was made in bore 606/647 (see Figure 2.4) in 1978 when extraction from the aquifer was in its infancy. This head was 7.5 m. This suggests that the equilibrium position of the toe of the freshwater-saltwater interface may be inland of the coast.

It is not known, however, whether salt water was in equilibrium with fresh water before extraction of fresh water from the MC aquifer began. It is not impossible that the interface is still out to sea in response to sea level and climatic conditions that prevailed in the past.

2.7 The Problem

Continued extraction of water from the VL system at current rates is unsustainable. In the long term, it will result in serious degradation of aquifer water quality through seawater intrusion. In the short term, it will occasion noncompliance with an EU Water Framework Directive which specifies that average ab must be lower than 90% of average annual recharge after accounting for environmental flow requirements; the Directive specifies that this condition be met by 2027.

Local authorities are looking at a number of options. Reduction of extraction through greater water use efficiency, and/or revoking of extraction licenses are two options. Managed aquifer recharge is another option. Users have already begun to significantly increase water use efficiency. However, technical and legal obstacles impede implementation of the latter two options at the present time.

2.8 Modelling

Modelling that is described herein serves a number of purposes. Firstly, it enables calculation of water extraction from time series of rainfall and evaporation. This provides estimates of historical extraction where records are unavailable. It also allows forecasting of future irrigation water requirements.

Secondly, history-matching of the model enables estimation of important components of the VL subsystem water balance, particularly lateral inflow from neighbouring aquifers. It also allows inference of aquifer hydraulic properties. Naturally, all of these inferences are accompanied by uncertainty; these uncertainties are quantified.

Thirdly, the model can be used to inquire into the design of a management strategy that maximises use of the MC water resource while mitigating degradation of water quality through seawater intrusion.

3.SIMULATION OF EXTRACTION

3.1 Introduction

Records of historical extraction from the VL system are incomplete. However they comprise the only direct measurements of any aspect of the VL water balance. Other aspects of the water balance must be inferred from the system's response to this extraction.

As is discussed elsewhere in this document, groundwater model history-matching spans the period October 2000 to October 2020. Water was extracted from the MC aquifer by all water user groups over that time; see Figure 2.7 for specification of water user groupings. The longest continuous record of monthly water extraction for any user group spans the period 2008 to 2020. Water use records for other groups are shorter than this; for some groups, records are absent for some years. The shortest record is that associated with the "small users" group; furthermore, this record comprises estimates rather than measured values of monthly water extraction. These estimates span the period 2019 to 2020.

Figure 3.1 graphs monthly extraction rates for one user group. (We purposefully omit user group identities.) This time series represents accumulated extraction over a number of wells. The same applies to water use records for most other groups. Each user group extracts water that services the irrigation needs of a specific area; this area is roughly known in all cases.

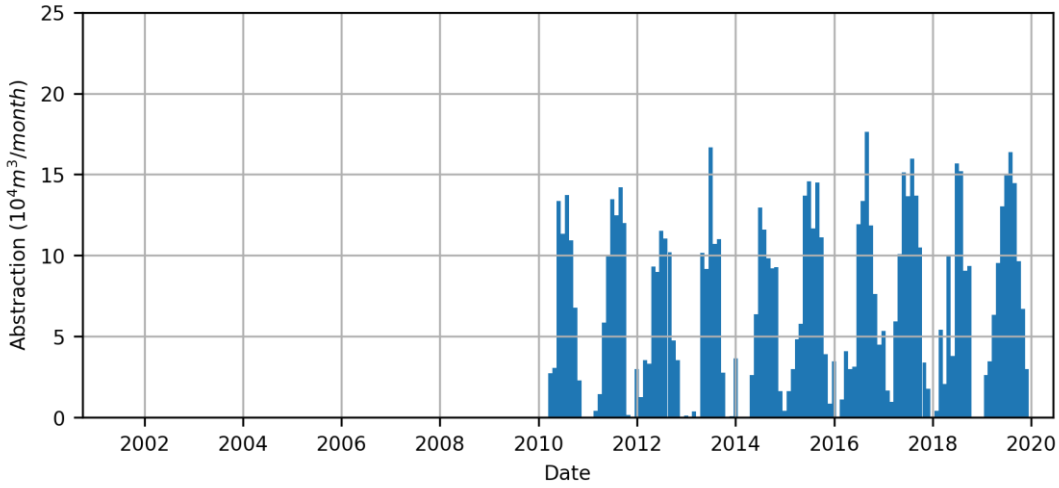


Figure 3.1 Monthly extraction volumes for a selected user group.

Because the area serviced by each water user group is known, then ideally its demand for irrigation water can be estimated from records of daily rainfall and potential evaporation.

3.2 The LUMPREM Model

3.2.1 General

LUMPREM is a lumped-parameter soil moisture accounting model. In the present worked example, it is used to calculate irrigation demand. In another GMDSI worked example, it is used to estimate recharge as well as time-varying model boundary heads.; see https://gmdsi.org/blog/water-supply-security/

LUMPREM, together with documentation, source code and examples, can be downloaded from the PEST web pages at https://www.pesthomepage.org. LUMPYREM Python support utilities can be

downloaded from http://github.com/rhugman/lumpyrem.The following description of LUMPREM is very brief. Refer to its documentation for further details.

3.2.2 Description

LUMPREM can deploy either one or two soil moisture containers. The upper container simulates process that operate in the plant root zone. The lower container simulates processes that delay drainage of sub-root-zone water towards the water table. Only the upper container is used in the present study. Processes that are operative in this container are depicted in Figure 3.2.

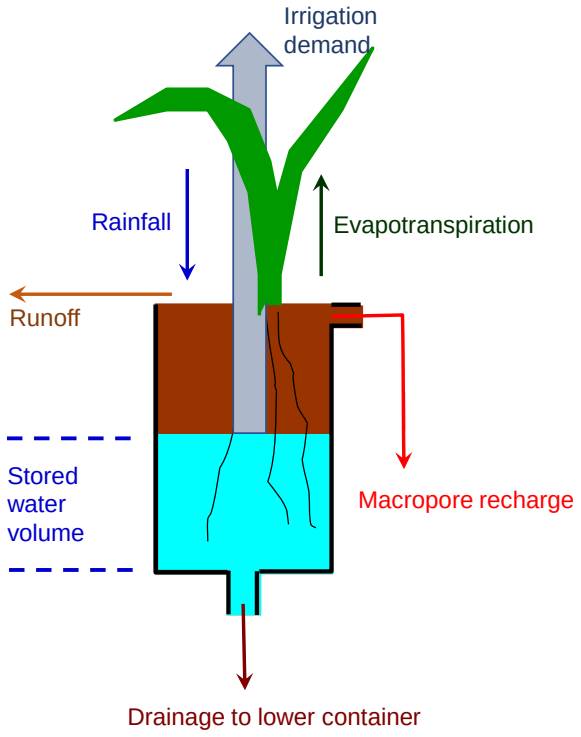


Figure 3.2 Schematic of processes that are operative in the upper container of the LUMPREM soil moisture accounting model.

LUMPREM operates on a daily time step. Water enters the upper soil moisture store as rainfall. It leaves the store as either evapotranspiration, drainage or macropore recharge. Over-topping of the store induces runoff.

Let v' denote the volume of stored moisture at any time relative to the total volume of the store. LUMPREM calculates daily evaporation E as a function of v' using the equation:

$$E = f\ E_p \frac{1-e^{-\gamma v'}}{1-2e^{-\gamma}+e^{-\gamma v'}} \tag{3.1}$$

where E_p is potential evaporation. f can be considered as a crop factor and γ as a fitting parameter.

LUMPREM calculates daily drainage R from v' using the relationship:

$$R = K_s [v']^l \left[1-\left(1-[v']^{1/m}\right)^m\right]^2 \tag{3.2}$$

where K_s represents soil saturated hydraulic conductivity, l is a pore-connectivity parameter and m is a fitting parameter.

When used to evaluate irrigation demand, LUMPREM runs in “irrigation mode”. When run in this way, *v'* is never allowed to fall to less than 0.5 during the irrigation season. This represents the average soil moisture status of all irrigators for which irrigation demand is evaluated. On any day, LUMPREM requests just enough irrigation to maintain the half-full status of the upper soil moisture store. However, should a large rainfall event occur, the store is allowed to fill. No irrigation is then required until the store drains to half full again. This mode of operation in which irrigation is requested on a daily basis facilitates LUMPREM model history-matching. It prevents the occurrence of local optima that arise if parcels of irrigation are demanded at discrete intervals of time whenever the store empties.

3.3 LUMPREM Deployment

3.3.1 Parameter Adjustment

One LUMPREM model is assigned to each water user group. Thus nine LUMPREM models are used to calculate subsets of water extraction from the MC aquifer, and hence from the groundwater model that is described in the next section of this report.

For each user group, LUMPREM parameters are adjusted by PEST in order to allow LUMPREM-calculated monthly irrigation volumes to match their user-recorded counterparts where the latter are available. Adjustable parameters are as follows:

- size of the soil moisture store;
- soil saturated hydraulic conductivity (K_s of equation 3.2);
- m of equation 3.2;
- l of equation 3.2;
- f of equation 3.1;
- γ of equation 3.1;
- the area requiring irrigation.

3.3.2 Distribution of Extraction

It is apparent from Figure 2.7 that, for most user groups, LUMPREM-calculated extraction must be distributed between a number of extraction wells. In general, distribution factors between wells that belong to a single user group are unknown. Hence they are history-matching-adjustable subject to the constraint that, for each user group, they sum to 1.0. Note that distribution factors do not affect LUMPREM outputs; they only affect heads that are calculated by the groundwater model. Hence they are adjusted during history-matching of the latter model.

3.3.3 Model Calibration

As is discussed in Section 6 of this document, model calibration seeks a unique, minimum error variance solution to an inverse problem. In the present case, the inverse problem is defined by the need for LUMPREM-calculated monthly irrigation volumes to match recorded irrigation volumes where and when these are available. As already stated, the amount of recorded irrigation data varies from user group to user group.

The nine LUMPREM models are calibrated simultaneously against their respective measurement datasets. This requires estimation of 63 LUMPREM parameters using 9 measurement datasets. The inversion process employs Tikhonov regularisation to constrain each parameter of the same type used by different LUMPREM models to adopt the same value as long as this does not compromise goodness of fit of LUMPREM outputs with respective measurement datasets. Where departures from parameter equality are necessary, the regularisation process ensures that these departures are minimal.

Simultaneous, regularised calibration of all LUMPREM models in this manner reduces the propensity for parameter nonuniqueness and for estimation of erratic parameter values. This applies particularly to LUMPREM models pertaining to user groups with small historical irrigation datasets.

Figure 3.3 shows monthly extraction volumes calculated by a calibrated LUMPREM model superimposed on the measured monthly extraction volumes depicted in Figure 3.1.

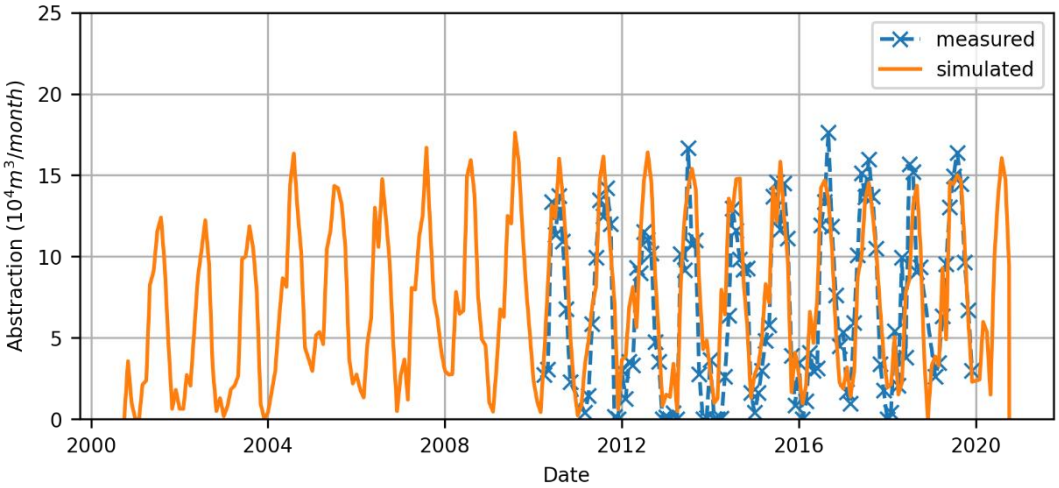


Figure 3.3 LUMPREM-calculated monthly extraction volumes for a selected user group superimposed on measured monthly extraction volumes.

During the overall VL model calibration process, LUMPREM models are calibrated separately from the groundwater model that is described in the next section. Well extraction rates calculated by calibrated LUMPREM models are provided to the groundwater model as it is itself subjected to calibration. However a different strategy is adopted when undertaking history-matching-constrained uncertainty analysis.

3.3.4 Stochastic History-Matching

Section 6 of this document describes how the PESTPP-IES ensemble smoother is used to calculate a suite of random parameter sets that allow model outputs to replicate all elements of the history-matching dataset. During this process, the nine LUMPREM models and the groundwater model are run as a single model; all parameters belonging to all of these models (including extraction distribution parameters) are subjected to simultaneous adjustment.

Figure 3.4 shows LUMPREM-calculated monthly extraction volumes superimposed on measured monthly volumes. LUMPREM outputs are calculated using an ensemble of history-match-constrained random parameter sets.

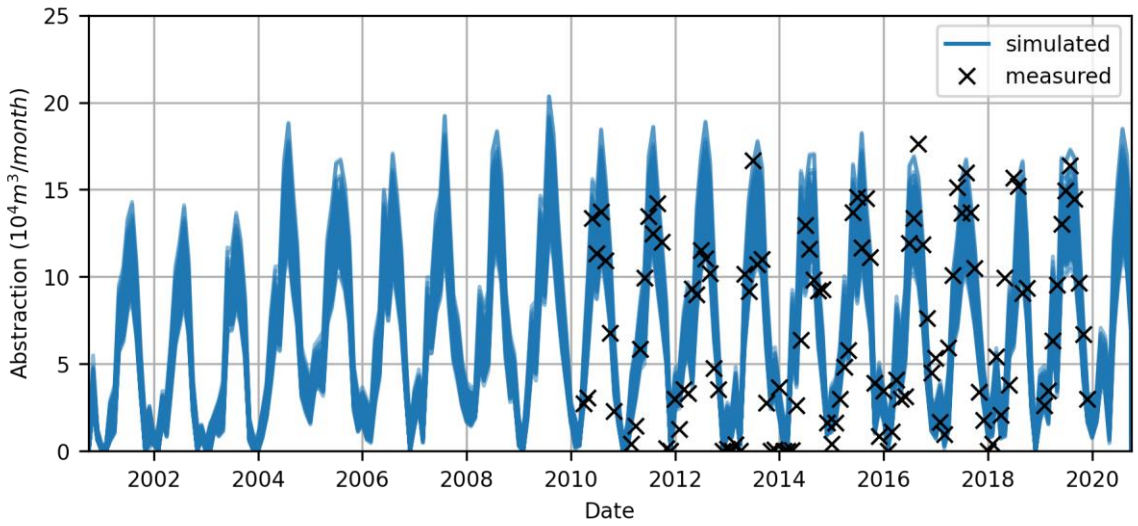


Figure 3.4 LUMPREM-calculated monthly extraction volumes for a selected user group, superimposed on measured monthly extraction volumes. LUMPREM outputs are computed using an ensemble of history-match-constrained random parameter sets.

4.THE GROUNDWATER MODEL

4.1 General Considerations

As it is for most aquifers, the spatial distribution of hydraulic properties within the MC aquifer is likely to be heterogeneous. This will affect patterns of sea water intrusion that result from excessive groundwater extraction. Groundwater modelling, when undertaken in concert with inversion packages such as PEST and PEST++, can reveal the broadscale distribution of aquifer hydraulic properties. It can also explore uncertainties in management-salient model predictions that arise from hydraulic property details that cannot be inferred through history-matching. Use of a fast-running, numerically stable model is critical to these endeavours.

The VL groundwater model possesses only a single layer. This layer represents the MC aquifer. There is no need for explicit representation of the PQ aquifer as little water is extracted from it. Furthermore, its hydraulic connection with the MC aquifer is weak. Nevertheless, the aquitard which separates the PC aquifer from the MC aquifer is represented as a pervasive boundary condition.

Salt concentration, and its effect on water density, is not simulated by the VL groundwater model. This eliminates the need for the large number of model layers that are required to simulate buoyant water movement. The decision to neglect density-dependent flow can be justified on the basis that no freshwater-saltwater interface has yet been detected within the VL system. Hence its presence can be neglected during history-matching.

The presence of off-shore saline water has some important consequences for design of the coastal boundary condition of the VL model. Lu et al (2015) outline these consequences. They show how to assign an appropriate head to the coastal boundary of a constant-density groundwater model. However their theory does not apply to models wherein current rates of groundwater extraction result in landward movement of the freshwater-saltwater interface. Hence our design of the coastal VL model boundary rests on complementary use of a density-dependent model. We describe the design of this boundary in the following section of this report.

Neglecting the presence of saline water is problematic when the model is used to explore long-term, sustainable rates of water extraction. Under these circumstances, the toe of the freshwater-saltwater interface may extend landward of the coastline. Aquifer transmissivities that are estimated under conditions of excessive extraction are inapplicable under conditions of steady-state seaward movement of fresh water above saline water that is static. For the VL model, the repercussions of neglecting the saline wedge are likely to be small; furthermore it will result in conservative, rather than exaggerated, estimation of optimal extraction rates. Nevertheless, the problem is addressed by modifying the simulator to account for near-coastal head-dependence of transmissivity that is available to freshwater flow. This is discussed in Section 7.3 of this document.

4.2 Timing

4.2.1 General

This section describes deployment of the VL model to simulate conditions which prevailed between October 2000 and October 2020. Over this time, rates of water extraction are calculated by a suite of nine LUMPREM models in the manner described in the previous section. Section 6 describes how the model is history-matched against observed water levels, this enabling stochastic inference of hydraulic and boundary condition properties. History-matching also enables stochastic inference of water balance components. As discussed above, these are required for assessing the status of the VL system according to EU Water Framework Directive management criteria.

Use of the history-matched model to investigate sustainable extraction rates and patterns is described in Section 7.3 of this document.

4.2.2 Details

The history-matching period extends for 20 years. For simulation purposes, this is subdivided into 240 monthly stress periods. Aquifer extraction rates, as well as heads assigned to most model boundaries, vary from stress period to stress period.

Initial heads for the history-matching period are calculated during a 30-year transient stress period in which time-averaged, LUMPREM-calculated pumping rates are applied to all extraction wells. Another steady-state stress period precedes this. No water is extracted from the MC aquifer during this stress period; hence it simulates pre-development conditions. Some model-generated quantities calculated during this stress period feature in the history-matching process. See Section 6.

4.3 Model Domain and Grid

Movement of groundwater within the MC aquifer of the VL system is simulated using the MODFLOW 6 simulator. A Voronoi grid comprised of 2751 cells was generated using ALGOMESH. Model cell density is greater in the vicinity of extraction and observation wells than elsewhere in the model domain. The model grid is depicted in Figure 4.1.

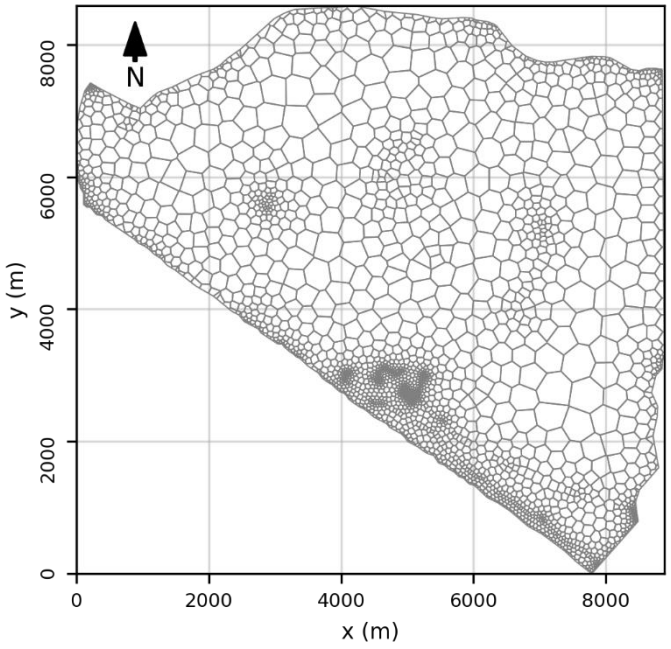


Figure 4.1 The VL model grid.

To its southwest, the model domain is bounded by the coastline. It is bounded to the north and northwest by outcropping Cretaceous limestone and the Carcavai fault, respectively. The eastern boundary of the model domain roughly coincides with the official management boundary that separates the Vale do Lobo subsystem of the Campina de Faro aquifer from the neighbouring Faro subsystem. As previously described, this boundary is a bureaucratic convenience and not based on physical properties of the system, except for rough orthogonality to topographic contours.

Top and bottom elevations of the single model layer are interpolated from onshore and offshore borehole logs and topographic elevations of outcrops. Top, bottom and thickness contours are provided in Figure 4.2a-c.

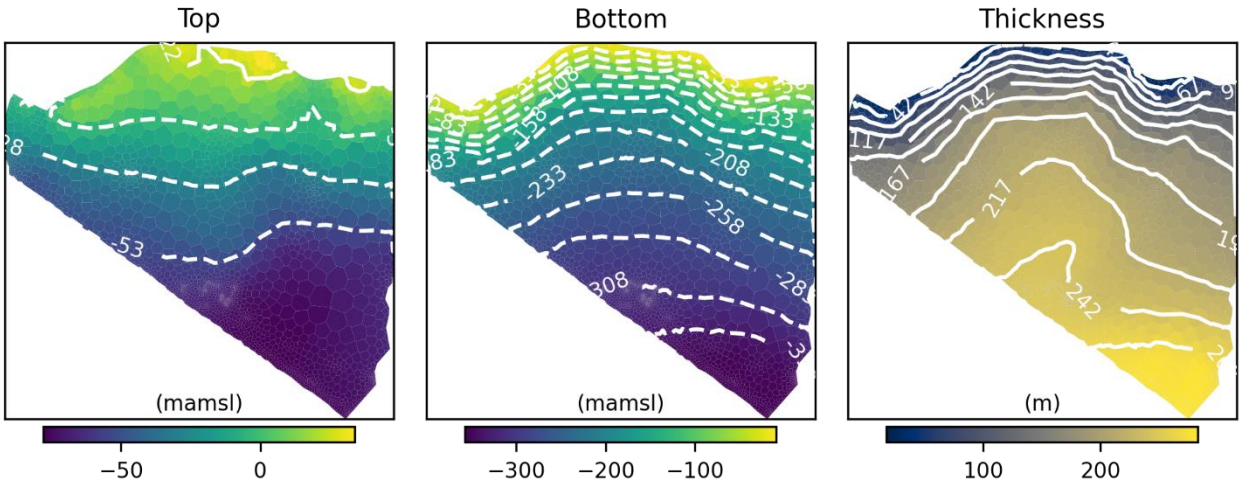


Figure 4.2 Contours of the (a) top elevation, (b) bottom elevation and (c) thickness of the single VL model layer.

Values of hydraulic conductivity and specific storage must be assigned to all model cells. Parameterisation of these hydraulic properties is affected using 565 pilot points. Hydraulic property values are assigned to these points. These values then undergo spatial interpolation to the model grid. Interpolation is implemented using the PLPROC model preprocessor supplied with the PEST suite. Pilot points are shown in Figure 4.3.

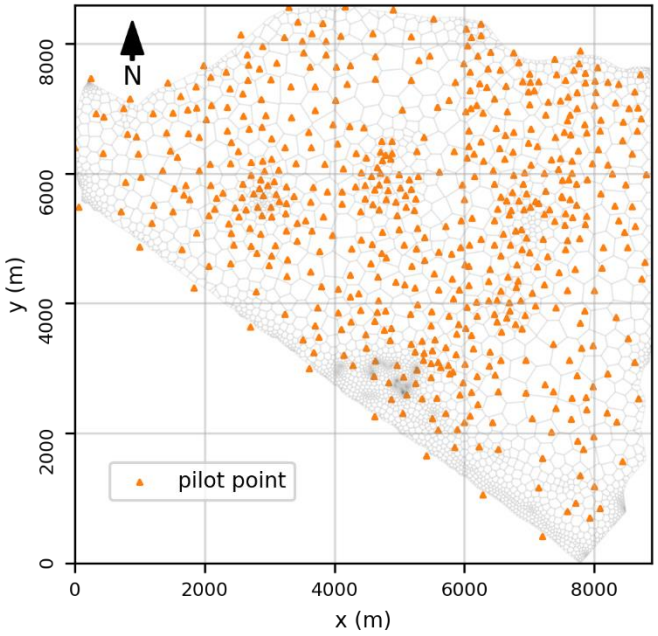


Figure 4.3 Pilot points used for parameterisation of hydraulic conductivity, specific storage and aquitard conductance.

4.4 Model Boundaries

4.4.1 Upper and Lower Boundaries

The bottom boundary of the VL groundwater model is a no-flow boundary.

A general head boundary (GHB) is assigned to every cell of the model domain to simulate its connection with the overlying PQ aquifer. The conductance of this boundary is parameterized using

the same set of pilot points as that which is used for hydraulic conductivity and specific storage; see Figure 4.3.

Heads in the PQ aquifer vary with season. Figure 4.4 shows water levels measured in piezometer 610/167 which is screened in the PQ aquifer. Head measurements are available from this well over the entirety of the simulation period. As is apparent from Figure 4.4, inter-seasonal water level variations are generally between one and two metres.

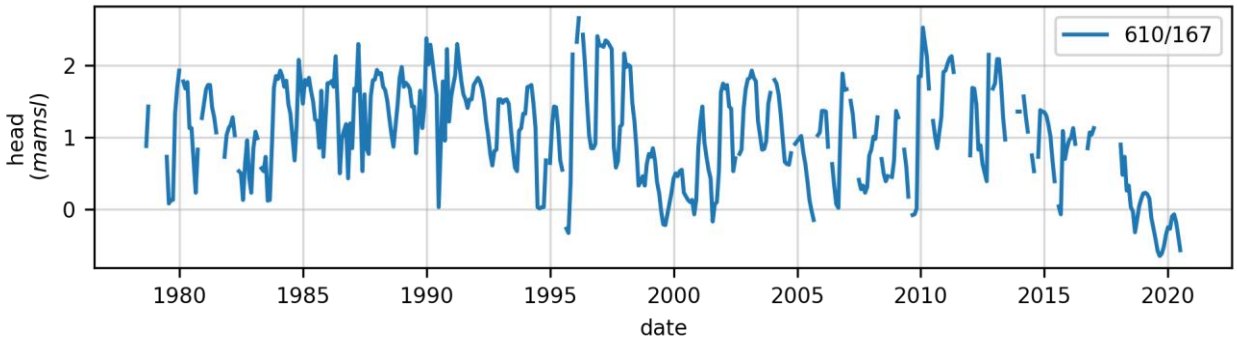


Figure 4.4 Piezometric heads measured in bore 610/167. This bore is open to the unconfined PQ aquifer.

Heads in the PQ aquifer vary with location as well as with time. However there are few observation wells in this layer; apart from 610/167, there are none with head records that span the 20 year historical model simulation time. Hence the water level record from 610/167 was used to calculate a head time series for the GHB boundary condition that occupies each model cell.

The elevation of the PQ water table at every cell-based GHB boundary is presumed to vary with time in similar fashion to water level variations in bore 610/167. The mean (over time) water table elevation at the location of each model cell is assumed to be a linear function of surface elevation. The slope and intercept of this linear function is calculated using water levels from bores which tap the PQ aquifer for which there are measured data. The mean-adjusted time series depicted in Figure 4.4 is then provided to the GHB which occupies each model cell. Meanwhile, the amplitude of water level variation is reduced where necessary in order to avoid intersection of the water table with the land surface where the latter is of diminished elevation. While this strategy is approximate, it allows representation of:

- seasonal water table variations within the PQ;
- spatial variation of these temporal variations in accordance with topography.

Approximations incurred by this GHB head assignment strategy are unlikely to introduce significant errors to the model because of the low rate of leakage from the PQ aquifer to the MC aquifer. Nor are errors incurred by representation of the PQ aquifer as a spatially continuous GHB likely to be any greater than those incurred through explicit representation of this aquifer using another model layer, for data available for history-matching is scarce in the upper aquifer. Meanwhile, gains in model solution speed incurred by omission of a thin, unconfined surficial layer are significant.

4.4.2 Lateral Boundaries – General Principles

Figure 4.6 subdivides the lateral boundaries of the VL groundwater model into three boundary segments, namely the north-western boundary, the eastern boundary and the ocean boundary. All of these boundary segments are simulated using GHB's. A head and a conductance are required for each instance of this GHB boundary condition in all model boundary cells.

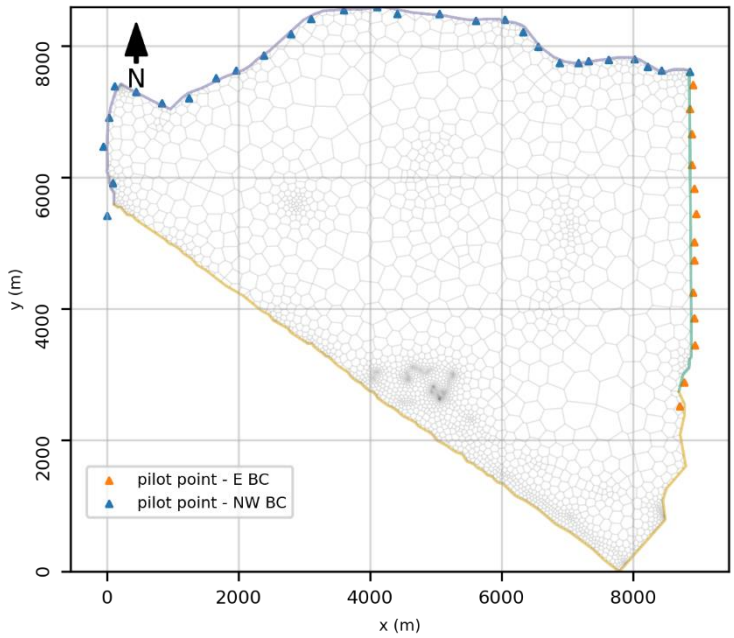


Figure 4.5 Subdivision of the lateral boundaries of the VL model into north-western, eastern and coastal segments. Pilot points used for head and conductance parameterisation of the north-western and eastern boundary segment are also shown.

For the north-western and eastern boundaries of the VL model, head varies with time while conductance is time-invariant. However, as is discussed in the next section, both head and conductance are time-invariant within each cell of the coastal boundary over the 20 year history-matching period.

Parameterisation of boundary heads and conductances is implemented using pilot points. Thus a head and a conductance is assigned to each such point. These are then interpolated along the boundary to respective model cells. Pseudolinear interpolation along model boundaries is implemented using PLPROC SEGLIST functionality.

Assignment of heads and conductances to the coastal model boundary is discussed in detail in Section 5 of this document.

4.4.3 North-western Boundary

The northern part of the north-western model boundary coincides with the outcrop of Cretaceous limestone. Inflow from this boundary is thought to be the main source of recharge to the VL system. The western segment of the north-western model boundary coincides with the Carcavai fault. The hydraulic properties of this fault are unknown; the possibility of cross-fault flow cannot therefore be ignored.

Piezometer 606/1050 is located just to the north of the model domain; see Figure 2.4. Figure 4.6 graphs heads that were measured in this borehole during the 20 year history-matching period. Seasonal fluctuations, as well as variations that span many seasons, are apparent in this time series. Water level records from other parts of the wider groundwater system that are outside the domain of the VL model are similar in character to this. (Note that only two of these bores reside near model boundaries, namely 606/1050 and 606/1033 that is discussed below.) It therefore seems appropriate to apply this head signal to all cells along the model boundary; however the signal must be altered to accommodate local topographic and other conditions.

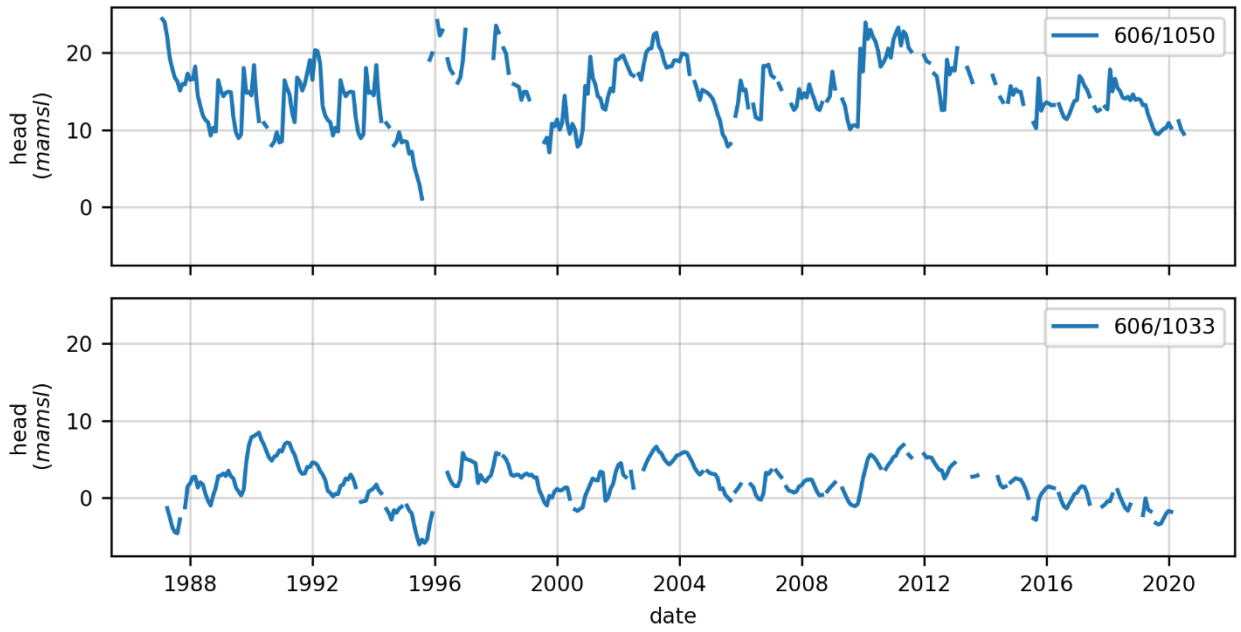


Figure 4.6 Hydraulic heads recorded in bores (top) 606/1050 and (bottom) 606/1033.

Let *h(t)* characterize heads in bore 606/1050. Time-variant heads $h_i(t)$ assigned to the *i* 'th pilot point along the north-western model boundary are calculated as:

$$h_i(t) = \alpha_i h(t) + \beta_i \qquad (4.1)$$

Values of α_i and β_i for each pilot point are history-match-adjustable. Thus they comprise part of the VL model's parameter set. The head time series ascribed to each boundary pilot point in this manner is then linearly interpolated along the model boundary using PLPROC SEGLIST functionality. Values of 0.0 and 1.0 for α and β respectively are assigned to the pilot point that is closest to bore 606/1050; these are not adjusted during history-matching.

Pilot-point-specific values of α_i and β_i are time-invariant; they apply under both steady-state, pre-development conditions, and under transient conditions that span the October 2010 to October 2020 history-matching period. However under steady state conditions, a value is estimated for the head in bore 606/1050.

4.4.4 Eastern Boundary

Head and conductance values are assigned to GHBs distributed along the eastern boundary of the VL model in an almost identical fashion to that used for assigning heads and conductances to cells which are arraigned along its north-western boundary. The reference hydrograph for the eastern boundary belongs to bore 606/1033. Its location is shown in Figure 2.4; measured heads are depicted in Figure 4.6.

5. THE COASTAL BOUNDARY

5.1 Introduction

In this section we dare to introduce some mathematics. It is not essential that a reader understand the maths to understand what we do. As far as we know, the approach to coastal boundary design that we describe in this chapter has not been implemented before. We therefore provide enough details for others to follow if they wish. If you are not interested in the mathematics, read the first few subsections of this section, and then read the summary at its end.

5.2 Considerations

5.2.1 Hydraulic Considerations

The coastline marks the seaward extent of the VL model domain. General head boundaries (i.e. GHBs) are introduced to all cells along this boundary; there are 264 coastal boundary cells in all. A value of head and conductance is required at each of these cells. Water enters or leaves a GHB in proportion to the difference between the head assigned to the boundary and that calculated by the model for the cell which the boundary occupies. The constant of proportionality is the boundary's conductance.

Coastal GHB heads and conductances are history-match-adjustable. Both are parameterised using pilot points – 29 in all. The disposition of these points is depicted in Figure 5.1. As is usual for pilot points parameterisation, values are first assigned to pilot points. These values are then spatially interpolated to pertinent cells of the groundwater model. As for other boundaries of the VL model, spatial interpolation is implemented using PLPROC SEGLIST functionality. The presence of 29 pilot points along the boundary implies the existence of 58 parameters. Half of these characterise boundary head, while the other half characterise boundary conductance.

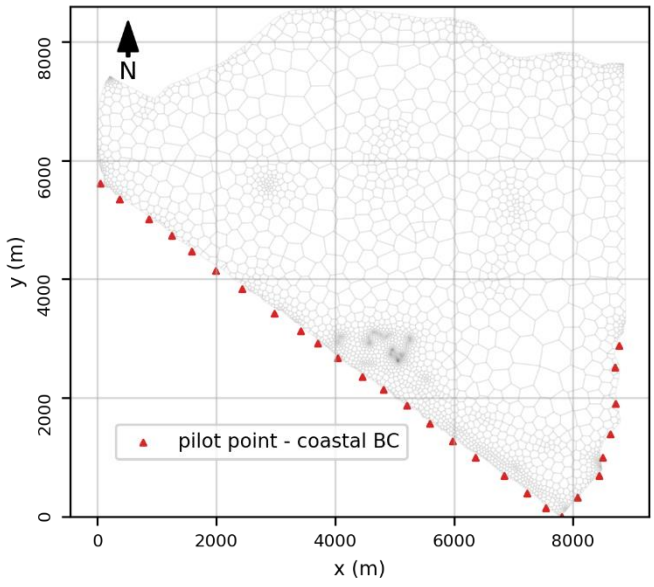


Figure 5.1 Pilot points used for parameterisation of the coastal boundary of the VL model.

The coastal boundary of the VL model provides summary numerical representation of complex processes that prevail on its seaward side. Under pre-development conditions, this boundary accommodates loss of fresh water from the confined MC aquifer. In reality, fresh water probably flows for a significant distance under the sea. Small amounts of water are lost per unit length of travel

through the aquitard which overlies the MC aquifer. Bakker (2006) and Bakker et al (2017) derive analytical equations for this condition under simplistic assumptions of aquifer geometry and properties. Meanwhile a small amount of saltwater flows landward beneath a static freshwater-saltwater interface; this sustains salt water recirculation at the interface. The situation is schematized in Figure 5.2a.

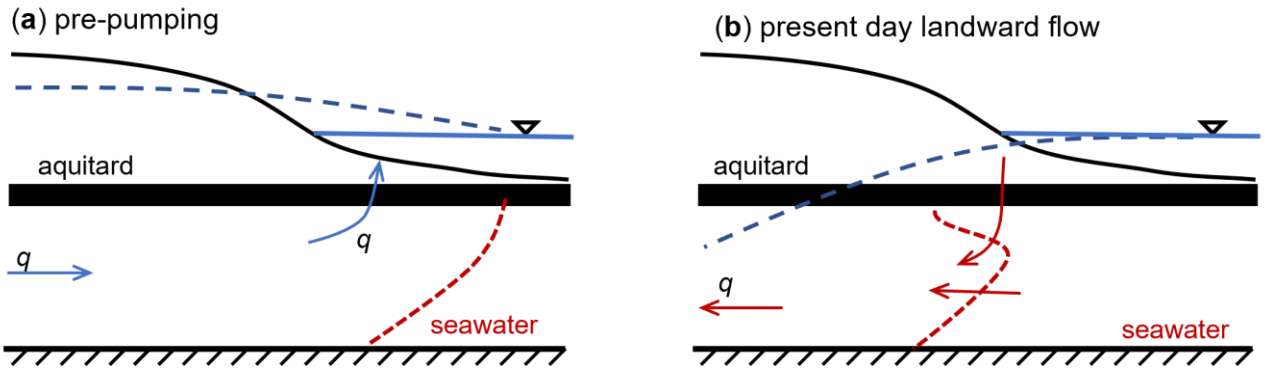


Figure 5.2 Schematic representation of (a) emergence of fresh water through the aquitard overlying the MC aquifer under pre-pumping conditions, and (b) landward flow of water under present day conditions.

Under conditions of non-sustainable pumping that prevail at the present time, both fresh and salt water flow landward across the coastline under a pumping-induced hydraulic gradient. The freshwater-saltwater interface moves inland as a consequence. This situation is schematized in Figure 5.2b.

That coastal GHB of the constant-density VL model must be capable of representing both of the conditions that are depicted in Figure 5.2.

5.2.2 Modelling Considerations

As is discussed in Section 1 of this document, model construction and deployment is based on the premise that stochastic representation be afforded to that which is poorly known. Aquifer and aquitard properties on the seaward side of the VL model coastal boundary are, indeed, poorly known, as is the means to represent them simplistically using a GHB. Nevertheless, prior probability distributions are required for head and conductance parameters ascribed to GHB coastal boundary cells. These can be refined through history-matching.

Let the vector **h** represent heads ascribed to the 29 pilot points that are used to parameterise the coastal model boundary; let the vector **c** represent pilot point conductances. Collectively, boundary parameters are therefore represented by the composite vector $\begin{bmatrix}\mathbf{h}\\ \mathbf{c}\end{bmatrix}$. Stochastic representation of boundary parameters requires that mean values be provided for these parameters at the locations of all pilot points; these mean values are represented by the vector $\begin{bmatrix}\underline{\mathbf{h}}\\ \underline{\mathbf{c}}\end{bmatrix}$. It also requires that a covariance matrix $C\left(\begin{bmatrix}\mathbf{h}\\ \mathbf{c}\end{bmatrix}\right)$ be ascribed to these parameters. This covariance matrix is used to represent spatial correlation between parameters of the same type along the boundary, as well as correlation between parameters of different types.

The prior mean vector and the prior covariance matrix of model parameters play important roles in history-matching. If undertaking model calibration in order to infer posterior mean parameter values, they can form the basis for Tikhonov regularisation through which this process achieves uniqueness. In particular, prior mean parameter values can serve as “preferred” parameter values. Meanwhile, the regularisation penalty function can feature the prior parameter covariance matrix. Alternatively, if history-matching is accomplished using an ensemble smoother in order to sample the posterior parameter probability distribution, prior parameter means and the prior parameter covariance matrix

are used to obtain samples of the prior parameter probability distribution; these samples are then history-match adjusted until they become samples of the posterior parameter probability distribution.

5.2.3 Mathematical Considerations

The prior mean vector and the prior mean covariance matrix of the $\begin{bmatrix}\mathbf{h}\\ \mathbf{c}\end{bmatrix}$ parameter set are obtained through a two-step process. The first of these steps samples a one-dimensional counterpart of $\begin{bmatrix}\mathbf{h}\\ \mathbf{c}\end{bmatrix}$. We refer to this vector as $\begin{bmatrix}h\\ c\end{bmatrix}$; it possesses just two elements (each of them random), namely a single head *h* and a single conductance *c*. Once enough samples of $\begin{bmatrix}h\\ c\end{bmatrix}$ have been obtained, its mean $\begin{bmatrix}\underline{h}\\ \underline{c}\end{bmatrix}$ and covariance matrix $C\left(\begin{bmatrix}h\\ c\end{bmatrix}\right)$ can be estimated. In the second step, the stochastic description of $\begin{bmatrix}h\\ c\end{bmatrix}$ is modified to provide a stochastic description of $\begin{bmatrix}\mathbf{h}\\ \mathbf{c}\end{bmatrix}$. That is, a mean $\begin{bmatrix}\underline{\mathbf{h}}\\ \underline{\mathbf{c}}\end{bmatrix}$ vector and a covariance matrix $C\left(\begin{bmatrix}\mathbf{h}\\ \mathbf{c}\end{bmatrix}\right)$ are determined. Once these are available, random realisations of $\begin{bmatrix}\mathbf{h}\\ \mathbf{c}\end{bmatrix}$ can be generated through standard statistical sampling.

5.3 Density-Dependent Model

5.3.1 Model Design

Figure 5.3 depicts the domain of a two-dimensional, cross-sectional, density-dependent model. This model simulates conditions which are illustrated schematically in Figure 5.2. Part of its domain lies beneath land, and part of its domain lies beneath the sea. The model simulates head-driven flow of water through a confined aquifer and its under-sea emergence through the confining aquitard. Model dimensions and hydraulic properties are representative of those that characterise the MC aquifer of the VL system.

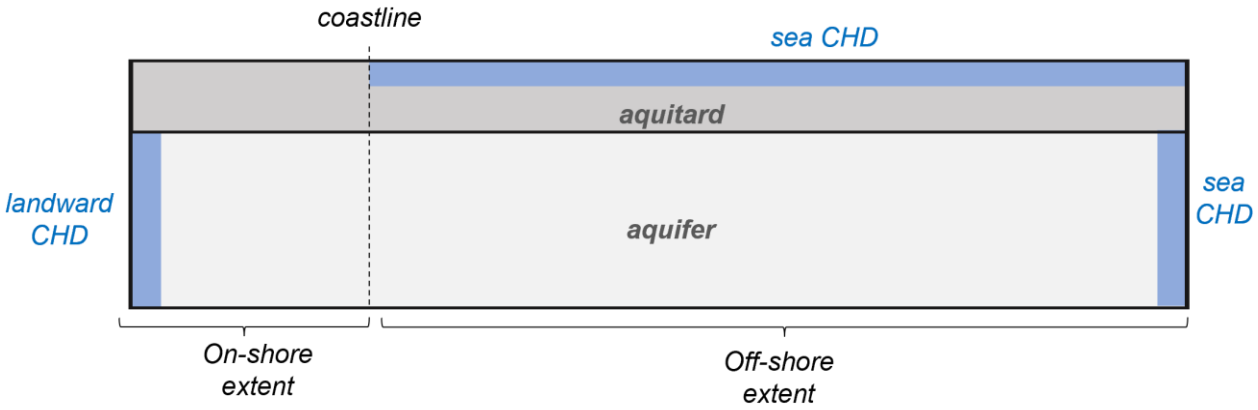


Figure 5.3 A two-dimensional, density-dependent, cross-sectional model. "CHD" signifies a constant head boundary.

SEAWAT is employed for simulation of density-dependent flow. One hundred different realisations of the model were built. Each realisation of the model is endowed with random hydraulic properties, a random landward boundary head, and randomized aspects of its geometry. Model construction is automated using FloPy.

Each column of the numerical model grid is 50 m wide. The thickness of each model layer is 5 m. The number of model layers and the number of its columns varies between model realisations. So does the elevation of the top of the aquitard.

On the seaward side of the coast, a constant-head boundary is introduced to all top-layer cells; the head is equivalent to 0 m of salt water. Constant-head boundary conditions are also introduced to cells disposed along vertical model boundaries at either end of the model domain. Along the seaward edge of the model domain, the salt water head is uniformly zero.

Cells comprising the vertical landward model boundary are assigned a uniform fresh water head. Depending on the model stress period, this is either positive (i.e. above sea level) to simulate pre-development conditions, or negative (i.e. below sea level) to simulate present-day, extractive conditions. The value of landward heads varies randomly from realisation to realisation. The distance from the coast to the landward model boundary also varies from realisation to realisation. Collectively, the range of post-development landward heads, and on-shore distances to these heads, encompass those which presently prevail in the MC aquifer at Vale do Lobo.

The distance from the coast to the seaward boundary is also realisation-dependent. The distance required for all fresh water to emerge from the aquifer under pre-development conditions depends on aquifer hydraulic properties and on the landward pre-development head. Analytical calculations based on equations derived by Bakker et al (2006; 2017) are used to set the offshore extent of the model domain for any particular realisation of model properties and boundary conditions. Limiting the length of the model domain to only that which is necessary for all fresh water to vacate the MC aquifer reduces simulation time.

5.3.2 Model Timing

The model depicted in Figure 5.3 simulates three conditions, two of which are identical. The first two conditions are those which prevailed prior to pumping. The last pertains to present day management in which the aquifer is subjected to groundwater extraction.

First the model runs for 2×10^6 years. This is long enough for steady-state conditions to be established. It is then run for two 50 year periods; over both of these time periods, outputs are requested at regular intervals so that they can be plotted. During the first of these time periods, steady state conditions are continued. During the second of these periods the model's landward head is set in accordance with present-day conditions. Fifty years was chosen for the length of these last two time periods as this corresponds to the time over which groundwater has been extracted from the VL system. The last twenty years of the last period are thus representative of conditions which prevailed during history-matching of the VL model.

Simulation times for this model vary between a minute and an hour. Solution convergence was not achieved for about 30% of model realisations; these realisations were rejected, and others used in their place.

5.3.3 Model Parameterisation

For each model realisation, aquifer and aquitard properties are homogeneous within their respective subdomains. However they are different between realisations. As for all other aspects of model design, they are representative of conditions that prevail at Vale do Lobo.

Table 5.1 lists aspects of model design which vary from realisation to realisation, together with the range of values from which these design variables are selected. A uniform or log-uniform distribution is employed in all cases. Note the following:

- the vertical hydraulic conductivity of the aquifer is equal to its horizontal hydraulic conductivity;
- the horizontal hydraulic conductivity of the aquitard is equal its vertical hydraulic conductivity;
- the porosity and specific storage of the aquitard are equal to that of the underlying aquifer;
- the thickness of the aquitard is 10m for all realisations.

Table 5.1 Aspects of the design of the density-dependent model which vary between realisations.

Design variable	Lower bound	Upper bound		Units	Distribution type
Aquifer horizontal hydraulic conductivity	1	100		m/day	log-uniform
Aquifer thickness	20	500		m	uniform
Aquifer porosity	0.1	0.3		-	
Aquifer specific storage	1×10^{-6}	1×10^{-3}		m^{-1}	log-uniform
Aquitard vertical hydraulic conductivity	1×10^{-4}	1×10^{-3}		m/day	log-uniform
Depth to top of aquitard below sea level at landward edge of model domain	0	100		m	uniform
Seaward dip of all model layers	0	0.06		degrees	uniform
Pre-development landward head	5	20		m	uniform
Distance from coast to landward model boundary	2.5	5.0		km	uniform
Post-development landward head	-10.0	0.0		m	uniform

All realisations employ a longitudinal disperssivity of 25m, a transverse disperssivity of 2.5m and a diffusion coefficient of 0 m^2/d. These result in a reasonably sharp freshwater-saltwater interface.

Figure 5.4 shows the distribution of salinity within the model domain for a single realisation. Over all realisations the location of the toe of the interface varies from 30km offshore to 500m onshore.

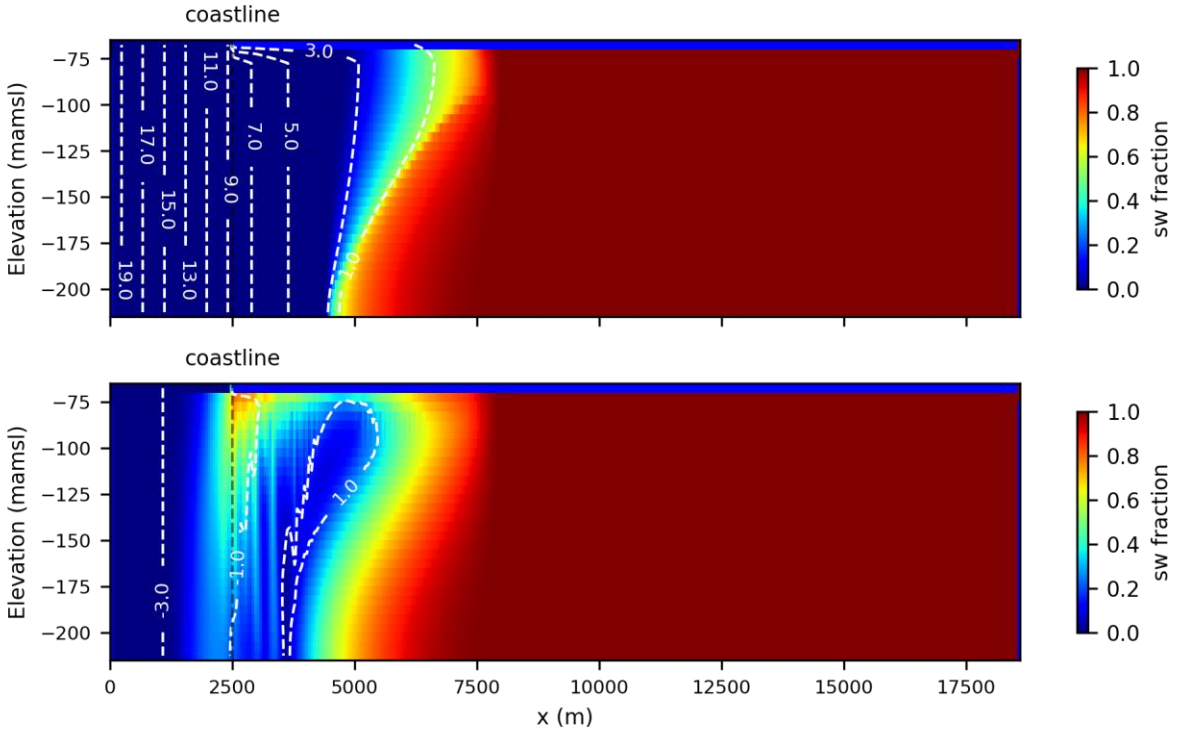


Figure 5.4 Model-calculated water salinity for a single realisation (a) prior to development and (b) post-development.

Figure 5.5 shows the fresh water head at the coast plotted against time for all 100 realisations. The fresh water head is averaged over all model cells which lie above the freshwater-saltwater interface in each case. For most realisations, a sudden change in head occurs shortly after landward extraction of water commences. The fresh water head remains reasonably constant thereafter.

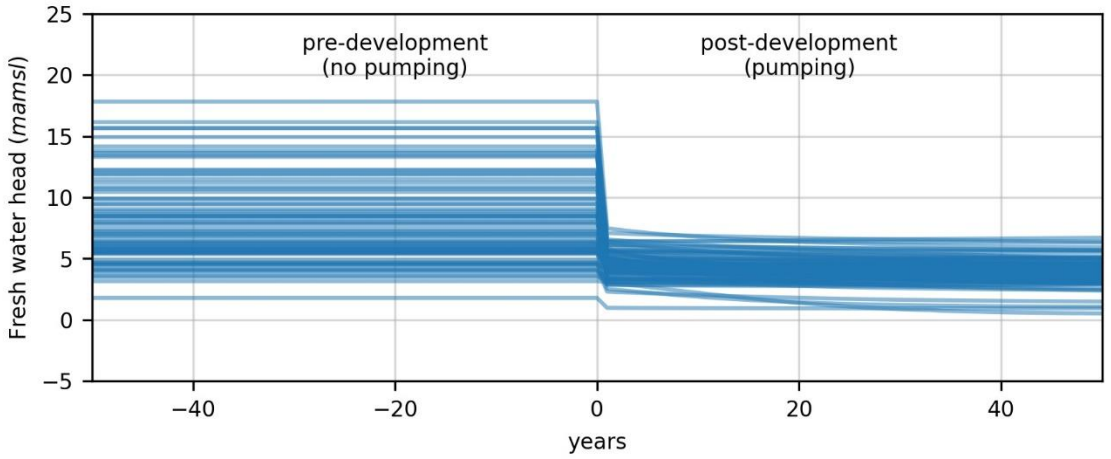


Figure 5.5 Fresh water heads at the coast under pre- and post-development conditions.

5.3.4 Calculating GHB Head and Conductance

For any one model realisation, let the fresh water head at the coastline be H_o when water flows toward the sea (i.e. under pre-development conditions), and H_i when water flows toward the land (i.e. under post-development conditions). Values of H_o and H_i are easily obtained from model outputs; see Figure 5.5. A value for H_i is established by averaging the model-calculated coastal head over the 50 year post-development stress period.

Let the flow of water (fresh and saline) under the coastline during pre- and post-development conditions be q_o and q_i respectively. In this nomenclature, the "*o*" and "*i*" subscripts stand for "out" and "in" respectively. Note also that q_i and q_o have opposite signs.

Suppose that we wish to describe flow across the coastal boundary using a GHB with head *h* and conductance *c*. Then, under outflow conditions:

$$q_o = (H_o - h)c \tag{5.1a}$$

while under inflow conditions:

$$q_i = (H_i - h)c \tag{5.1b}$$

These two equations can be solved for the two unknowns *h* and *c*. The solutions are:

$$c = \frac{q_o - q_i}{H_o - H_i} \tag{5.2a}$$

$$h = \frac{q_o H_i - q_i H_o}{q_o - q_i} \tag{5.2b}$$

5.3.5 Stochasticity of GHB Head and Conductance

By running the two-dimensional, sectional, density-dependent model many times, many different values of $\begin{bmatrix} h \\ c \end{bmatrix}$ can be obtained using equations 5.2a and 5.2b; we represent those calculated for realisation *i* as $\begin{bmatrix} h_i \\ c_i \end{bmatrix}$. After *N* model realisations have been run, the elements $\underline{h}$ and $\underline{c}$ of the mean vector $\begin{bmatrix} \underline{h} \\ \underline{c} \end{bmatrix}$ can be evaluated as:

$$\underline{h} = \frac{\sum_{i=1}^{N} h_i}{N} \tag{5.3a}$$

$$\underline{c} = \frac{\sum_{i=1}^{N} c_i}{N} \quad (5.3b)$$

The variance of *h* and *c* are calculated as:

$$\sigma_h^2 = \frac{\sum_{i=1}^{N}(h_i - \underline{h})}{N-1} \quad (5.4a)$$

$$\sigma_c^2 = \frac{\sum_{i=1}^{N}(c_i - \underline{c})}{N-1} \quad (5.4b)$$

The covariance between *h* and *c* is calculated as:

$$\sigma_{ch} = \sigma_{hc} = \frac{\sum_{i=1}^{N}(c_i - \underline{c})(h_i - \underline{h})}{N-1} \quad (5.5)$$

Putting these variances and covariances together we obtain a covariance matrix:

$$C\left(\begin{bmatrix} h \\ c \end{bmatrix}\right) = \begin{bmatrix} \sigma_h^2 & \sigma_{hc} \\ \sigma_{ch} & \sigma_c^2 \end{bmatrix} \quad (5.6)$$

Table 5.2 lists values for the statistical properties of *h* and *c* obtained in this way.

Table 5.2 Stochastic characterisation of GHB boundary parameters using a suite of 2D, density-dependent numerical models.

Symbol	Description	Units	Value
$\underline{h}$	Mean GHB head	m	5.47
$\underline{c}$	Mean $\log_{10}$ of GHB conductance	log(m²/d)	-0.19
σ^2_h	Variance of *h* (i.e. square of standard deviation of *h*)	m²	1.88
σ^2_c	Variance of $\log_{10}$ of *c* (i.e. square of standard deviation of $\log_{10}$ of *c*)	[log(m²/d)]²	0.27
σ_{ch}	Covariance of $\log_{10}$ of *c* with *h*	m - log(m²/d)	-0.35

5.3.6 Spatial Correlation

As is discussed above, 29 pilot points are used for parameterisation of the VL model coastal boundary. A value of GHB head and conductance is associated with each pilot point. Parameterisation of this boundary therefore requires that values be assigned to 58 parameters. It also requires that these 58 parameters be endowed with a joint prior probability distribution.

It cannot be assumed that head and conductance are spatially invariant along the boundary. Hence it is not appropriate to generate a random realisation of *h* and *c* based on the stochastic characterisation presented in Table 5.2, and then assign these *h* and *c* values to all 29 pilot points. Nor is it appropriate to generate 29 different pairs of values for *h* and *c* based on this stochastic characterisation and then assign these values independently to the 29 different pilot points, for parameters are expected to show some degree of spatial correlation along the boundary. Instead, parameter values should be selected from a 58-dimensional probability distribution whose mean is $\begin{bmatrix} \underline{\mathbf{h}} \\ \underline{\mathbf{c}} \end{bmatrix}$ and whose covariance matrix is $C\left(\begin{bmatrix} \mathbf{h} \\ \mathbf{c} \end{bmatrix}\right)$; see Section 5.2.3.

To do this, we first assume that the means of **h** and **c** (i.e. $\underline{\mathbf{h}}$ and $\underline{\mathbf{c}}$) are invariant along the model boundary. Each of the elements of $\underline{\mathbf{h}}$ is thus equal to $\underline{h}$, while each of the elements of $\underline{\mathbf{c}}$ is equal to $\underline{c}$.

The next task is calculation of $C\left(\begin{bmatrix} \mathbf{h} \\ \mathbf{c} \end{bmatrix}\right)$. This requires some care, as a covariance matrix must be positive definite, for this ensures that probabilities are never negative. So we first assume a specification for spatial correlation of either **h** or **c**. In our case we specify an exponential decay of **h**

correlation with distance; the spatial decay constant is 2 km. (An exponential decay of correlation with distance – that is, use of an exponential variogram - always engenders a positive definite covariance matrix.) Meanwhile the variance of **h** (i.e. the sill of the **h** variogram) is obtained from Table 5.2. In characterising stochasticity of **h** in this way, we have specified C(**h**). (Utilities such as PPCOV provided with the PEST Groundwater Utility suite allow calculation of a covariance matrix from a variogram.).

Next we assume that at any pilot point, the relationship between log(*c*) and *h* is described by the following linear equation. (We omit the log() transformation in the following equations in order to make the linear algebra clearer.)

$$c = ah + \varepsilon \tag{5.7a}$$

where *a* is a constant and ε is a random number. We turn equation 5.7a into a vector equation by applying it to all pilot points:

$$\mathbf{c} = a\mathbf{h} + \boldsymbol{\varepsilon} \tag{5.7b}$$

We assume at this stage that:

$$C(\boldsymbol{\varepsilon}) = \sigma^2_{\varepsilon}\mathbf{I} \tag{5.8}$$

where **I** is the identity matrix. Hence the value of $\boldsymbol{\varepsilon}$ at one pilot point is independent of that at another. From (5.7b) it follows that:

$$\begin{bmatrix}\mathbf{h}\\ \mathbf{c}\end{bmatrix} = \begin{bmatrix}\mathbf{I} & 0\\ a\mathbf{I} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{h}\\ \boldsymbol{\epsilon}\end{bmatrix} \tag{5.9}$$

Now we use a well-known relationship for propagation of variance. If **x** is a random vector with covariance matrix C(**x**), and if the vector **y** is calculable from **x** using the equation:

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{5.10}$$

then **y** is a random vector with covariance matrix C(**y**) given by:

$$C(\mathbf{y}) = \mathbf{A}C(\mathbf{x})\mathbf{A}^t \tag{5.11}$$

Applying this to (5.9) we obtain:

$$C\begin{bmatrix}\mathbf{h}\\ \mathbf{c}\end{bmatrix} = \begin{bmatrix}\mathbf{I} & 0\\ a\mathbf{I} & \mathbf{I}\end{bmatrix} C\left(\begin{bmatrix}\mathbf{h}\\ \boldsymbol{\varepsilon}\end{bmatrix}\right)\begin{bmatrix}\mathbf{I} & a\mathbf{I}\\ 0 & \mathbf{I}\end{bmatrix} \tag{5.12}$$

Now, because **h** and $\boldsymbol{\varepsilon}$ are independent:

$$C\begin{bmatrix}\mathbf{h}\\ \boldsymbol{\varepsilon}\end{bmatrix} = \begin{bmatrix}C(\mathbf{h}) & \mathbf{0}\\ \mathbf{0} & \sigma^2_{\varepsilon}\mathbf{I}\end{bmatrix} \tag{5.13}$$

After substituting, and then performing the necessary matrix multiplications, (5.12) becomes:

$$C\begin{bmatrix}\mathbf{h}\\ \mathbf{c}\end{bmatrix} = \begin{bmatrix}C(\mathbf{h}) & C(\mathbf{h})a\\ aC(\mathbf{h}) & a^2C(\mathbf{h}) + \sigma^2_{\varepsilon}\mathbf{I}\end{bmatrix} \tag{5.14}$$

All that remains is to determine *a* and σ_{ε}. Numerical experiments with the density-dependent sectional model have already provided the diagonal elements of C(**c**). These are all σ^2_c; see Table 5.2. It follows from (5.14) that:

$$\sigma^2_c = a^2\sigma^2_h + \sigma^2_{\varepsilon} \tag{5.15}$$

These numerical experiments have also provided the covariance between *h* and *c* at any one pilot point; see Table 5.2 again. From (5.14) it follows that:

$$\sigma_{ch} = a\sigma^2_h \quad (5.16)$$

A value for *a* can be calculated from (5.16); this value can be used in (5.15) to calculate σ^2_ε. In our case values for *a* and σ^2_ε are -0.098 and 0.23.

Calculation of $C\left(\begin{bmatrix}\mathbf{h}\\ \mathbf{c}\end{bmatrix}\right)$ is therefore complete.

5.4 Summary

Our single-density, decision-support groundwater model of the MC aquifer at Vale do Lobo features a suite of 264 GHB's arranged along its coastal boundary. This distributed boundary condition allows water to leave the VL system and flow towards the sea under conditions of low extraction, and to enter the VL system to flow towards production wells under conditions of high extraction. This coastal boundary is parameterized using 29 pilot points. Values of GHB head and conductance must be assigned to each of these points; these values are spatially interpolated to model boundary cells.

Parameterisation of GHB heads and conductances is fraught with uncertainty. Much of this uncertainty arises from the fact that little is known about hydraulic conditions offshore. Some of it arises from the use of a relatively simple boundary to represent these conditions, and from the assumption of single-density flow within the domain of the VL model.

Accommodation of uncertainty is straightforward when using modern modelling software that supports stochastic parameterisation of all aspects of a model. This software also allows constraints to be imposed on the values assigned to stochastic model parameters through history-matching. However use of these methods requires that prior probability distributions be assigned to all parameters ascribed to a model.

Where a polylinear model boundary is used to represent complex, spatially-distributed conditions that are omitted from a model, assignment of a prior probability distribution to its parameters is not straightforward. A solution to this problem is to simulate the omitted conditions, at least approximately, and to calculate the values of equivalent boundary parameters from that simulation. Evaluation of a prior probability distribution for boundary parameters requires that this process be repeated across the range of possible conditions that the boundary must represent. In order to avoid excessive numerical cost, the model that is used for this purpose in the present example simulates density-dependent flow in only the vertical plane.

A problem with using a two-dimensional sectional model to assist in parameterisation of a boundary that is perpendicular to this section is that of establishing spatial correlation of parameters along the boundary. A means of addressing this problem is described. It requires that a modeller choose an appropriate correlation length for parameters along the boundary. Outcomes of stochastic runs of the sectional model can then be used to build a covariance matrix that supports stochastic pilot points parameterisation of the boundary. This covariance matrix expresses spatial correlation of boundary head and conductance parameters, as well as correlation between the two parameter types. In so doing, it respects the need for a covariance matrix to be positive definite.

6. HISTORY-MATCHING

6.1 Concepts

6.1.1 General

As is usual in GMDSI worked example reports, we devote a few sentences to clarifying what we mean by “history-matching”.

Taken literally, “history-matching” denotes adjustment of model parameters so that model outputs are able to replicate the historical behaviour of a system. Satisfactory replication of past system behaviour comprises a necessary, but not sufficient, condition for use of a set of parameters when making predictions of future system behaviour.

6.1.2 Bayes Equation

In practice, history-matching should serve the imperatives of Bayes equation. Bayes equation is discussed extensively in GMDSI training material and webinars.

From a Bayesian point of view, the purpose of history-matching is to adjust “realistic”, random realisations of model parameters by the minimum amount necessary for them to allow the model to replicate the past. Their “realistic” status is an outcome of being sampled from the prior parameter probability distribution. Their ability to support model replication of past system behaviour also grants them membership of the posterior parameter probability distribution. Once a sufficient number of samples of the posterior parameter probability distribution has been obtained, they can be used collectively to make probabilistic predictions of future system behaviour.

Adjustment of random parameter sets in this manner can be implemented using the PESTPP-IES ensemble smoother. See White (2018) for details. PESTPP-IES works directly with ensembles of parameter sets rather than with individual parameter sets. Each member of the ensemble is called a “realisation”. The realisation adjustment process can be remarkably model-run-efficient. The number of model runs required for its implementation depends on the number of parameter realisations that comprise an ensemble, and not on the number of parameters that comprise a realisation.

6.1.3 Calibration

In the past, “model calibration” was used to denote the adjustment of a parsimonious parameter set so that model outputs provide reasonable replication of past system behaviour. However, modern-day deployment of complex parameter sets, and a requirement that model deployment be accompanied by analysis of the uncertainties of decision-critical model predictions, make the concept of “model calibration” somewhat obsolete – unless it is given a specific meaning that acknowledges the heterogeneity of natural systems, and the need to endow groundwater models with large numbers of parameters.

In modern modelling parlance, the term “calibration” refers to the process of solving a highly-parameterised inverse problem. It acknowledges that this inverse problem is nonunique. So it adopts numerical measures (referred to as “regularisation”) to render it unique. So-called “Tikhonov regularisation” promotes uniqueness by constraining parameters to respect a set of preferred values or conditions as much as possible, subject to the requirement that an acceptable fit is attained with a measurement dataset. These conditions are generally those of parameter simplicity or parameter field smoothness. Meanwhile, the use of many parameters grants the inversion process freedom to introduce heterogeneity to those places within the model domain where the calibration dataset suggests that this heterogeneity should exist. The resulting parameter set is often characterised as being of “minimized error variance” because it enables the model to make predictions that lie

somewhere near the centres of their posterior probability distributions. The potential for predictive error is therefore symmetrical with respect to these predictions.

The decision-support modelling process often benefits from postponing uncertainty analysis until a model has been calibrated according to the above interpretation of “calibration”. Modern-day model calibration allows a modeller to test the concepts on which construction of his/her model rests. If highly-parameterized, regularised inversion cannot achieve a good fit with an observation dataset, or if the simplest parameter field that enables such a fit is unexpectedly complex because parameters must adopt roles that compensate for model defects in order to attain that fit, then a modeller may decide that the conceptual basis of his/her model requires revision. Alternatively, he/she may decide that his/her stochastic characterisation of parameter variability requires revision.

Calibration can also teach a modeller much about what should comprise the contents of a history-matching dataset. It may suggest ways in which observations and corresponding model outputs should be processed before being matched in order to ensure that the model is able to replicate important aspects of system behaviour. It may also expose the need to supplement a calibration dataset with “soft” observations that discourage the inversion process from estimating parameters that result in model-calculated states and fluxes that a modeller suspects are incorrect.

Unfortunately, attainment of a unique, minimum-error-variance solution to an inverse problem can incur greater numerical costs than ensemble-based history-matching. This is because it requires calculation of a Jacobian matrix. The cost of filling this matrix is at least one model run per parameter per iteration of the inversion process.

This cost accrues benefits, however. Better fits with field measurements can generally be attained through regularised adjustment of a single parameter set than through Bayesian adjustment of an ensemble of parameter sets. Also, the Jacobian matrix that is calculated as a by-product of regularised inversion can form the basis for post-calibration linear analysis. This can be used to calculate a linear approximation to the posterior parameter covariance matrix. Hence an ensuing PESTPP-IES run can sample a linear approximation to the posterior parameter probability distribution at the commencement of its parameter adjustment process. This can enhance its numerical performance, particularly where the relationship between model outputs and parameters is nonlinear. (The inclusion of nine LUMPREM irrigation-demand models in the overall VL model renders the relationship between VL model outputs and some of its parameters highly nonlinear.)

In light of the above considerations, history-matching of the VL model comprises a two-step process. The model is first calibrated using PEST_HP. Subsequently, PESTPP-IES is employed to sample the posterior parameter probability distribution.

6.2 Parameters

6.2.1 General

The model that forms the focus of this report is a composite model. A MODFLOW 6 groundwater model simulates movement of fresh water within the unconfined MC aquifer; this model calculates hydraulic heads at the nodes of an unstructured finite-difference grid. Nine LUMPREM soil moisture accounting models simulate processes that are operative in plant root zones; these models calculate monthly extraction rates that are used by the groundwater model. All of these models possess history-matching-adjustable parameters. A total of 2241 parameters require adjustment.

6.2.2 LUMPREM

As discussed in Section 3.3.1, each of the nine LUMPREM models is endowed with seven adjustable parameters. The values of these parameters influence LUMPREM-calculated groundwater extraction. This, in turn, influences heads calculated by the groundwater model.

LUMPREM history-matching is described in Section 3 of this report. During the calibration phase of history-matching, the nine LUMPREM models undergo simultaneous parameter adjustment. Meanwhile, regularisation constraints impose a condition of preferential equality of values of parameters of the same type employed by different LUMPREM model instances; departures from parameter type equality in order to attain a good fit with the calibration dataset are minimized. The targets of the history-matching process are measured extraction rates.

During Bayesian history-matching, the nine LUMPREM models and the groundwater model are run as a composite model. During this phase of the history-matching process, random realisations of LUMPREM parameters together with random realisations of MODFLOW 6 parameters are adjusted against a history-matching dataset comprised of measured extraction rates, observed borehole heads, and some other "soft" observations that are described below. Outputs of one stochastically-history-matched LUMPREM model are compared with observed extraction rates in Figure 3.4.

6.2.3 MODFLOW 6

Three parameters are associated with each of the 565 pilot points whose locations are shown in Figure 4.3. These are:

- hydraulic conductivity of the MC aquifer;
- specific storage of the MC aquifer;
- conductance of the aquitard overlying the MC aquifer.

During calibration, Tikhonov regularisation constraints are imposed on each of these parameter types. These are of the "preferred value" type. Preferred values are uniformly 5 m/day for hydraulic conductivity, 1×10^{-5} for specific storage and 1×10^{-3} m^2/day for aquitard conductance. Each set of 565 constraints (expressed through prior information equations in the PEST control file) is accompanied by a covariance matrix whose task is to distribute emergent heterogeneity spatially, and therefore to prevent its concentration at the locations of individual pilot points. These covariance matrices are constructed using the PPCOV_SVA utility supplied with the PEST Groundwater Utility suite. PPCOV_SVA allows spatial parameter correlation to vary with pilot point spatial density. For VL model pilot point parameters, the spatial correlation length is about three times the local pilot point spacing.

Figure 5.1 depicts pilot points ascribed to the coastal boundary of the VL model. Figure 4.5 depicts pilot points ascribed to its north-western and eastern boundaries.

Along the coastal boundary, two parameters are associated with each pilot point. These are GHB head and GHB conductance. During model calibration, Tikhonov regularisation constraints for these parameters are of the preferred value type. Preferred values for head and conductance parameters are obtained using a companion density-dependent flow model as described in Section 5. Meanwhile, the composite head-conductance covariance matrix whose construction is described in Section 5.3.6 of this document is used to distribute emergent parameter heterogeneity along the boundary.

A conductance parameter is associated with each of the pilot points ascribed to the north-western and eastern boundaries of the VL model. Tikhonov regularisation specifies a preferred value of 10 m^2/day for each of these parameters. A covariance matrix for each of these boundary parameter families specifies a spatial correlation length of about 1 km.

At each pilot point along these boundaries, time-varying heads are calculated from one of the borehole-measured time series depicted in Figure 4.6 – that pertaining to bore 606/1050 for the north-western model boundary, and that pertaining to bore 606/1033 for the eastern model boundary. Calculation of each of these pilot-point-specific heads requires assignment of a scale and offset parameter (that is α_i and β_i of equation 4.1) to each pilot point. During model calibration, Tikhonov regularisation imposes a minimum-difference constraint between neighbouring parameters of the same type along each boundary. During the steady-state, pre-development stress period, heads at

bores 606/1050 and 606/1033 are also declared to be adjustable; water level records for these bores do not extend back to pre-development times.

Factors that distribute LUMPREM-calculated extraction rates among production wells that belong to each irrigation user group comprise the final set of parameters adjusted during VL model history-matching. There are a 211 of these factors. Tikhonov-preferred values for these factors distribute extraction equally over wells belonging to each group. Their estimation is further constrained by the necessity for these factors to sum to 1.0 for each user group.

6.2.4 Summary: Model Calibration

Calibration of the VL model is a two-step process. The first step comprises simultaneous calibration of nine LUMPREM soil moisture accounting models. Monthly extraction rates calculated by these calibrated models are then used by the MODFLOW 6 groundwater flow model as it undergoes calibration itself.

Calibration attains parameter uniqueness using Tikhonov regularisation. Depending on the parameter type, this is either of the preferred value or preferred difference type. In the former case preferred value constraints are accompanied by covariance matrices that subdue the emergence of “spotty” heterogeneity on a parameter-by-parameter basis.

6.2.5 Stochastic History-Matching

When undertaking stochastic history-matching, covariance matrices play a different role from that which they play during model calibration. As stated above, stochastic history-matching enables Bayesian uncertainty analysis. In this type of analysis, emergent heterogeneity in history-matched parameter fields is encouraged rather than suppressed. Ideally, emergent heterogeneity should be constrained to respect the prior parameter probability distribution. Covariance matrices are used to define this distribution.

Often, the same covariance matrices that are used to impose Tikhonov constraints can be used to define prior parameter probability distributions. Meanwhile, Tikhonov preferred values become prior parameter means. Realisations comprising the initial ensemble of the PESTPP-IES history-matching process can be sampled from prior parameter probability distributions that are defined in this way.

Unfortunately, PESTPP-IES can sometimes experience performance difficulties when the initial parameter ensemble is generated in this manner. This is particularly the case where an inverse problem is highly nonlinear, as it is in the present case with nine incidences of LUMPREM included in the composite model for which stochastic history-matching is required. So instead of sampling the prior parameter probability distribution, samples are taken of a linear approximation to the posterior parameter probability distribution when initiating the IES history-matching process.

The workflow is as follows:

1. Calibrate the model using regularised inversion in the manner described above.
2. Remove regularisation from the PEST control file.
3. Combine the nine LUMPREM models and the MODFLOW 6 model so that they run as a single composite model.
4. Use PEST to generate a Jacobian matrix containing sensitivities of all observations comprising the history-matching dataset to all parameters.
5. Use the PREDUNC7 utility supplied with PEST to build a linear approximation to the posterior parameter covariance matrix.
6. Supply calibrated parameters to PESTPP-IES as its “base” parameter set; inform it that the “prior” parameter covariance matrix is the linearised posterior parameter covariance matrix obtained in the manner described above.

PESTPP-IES performed well when initiated in this manner. In history-matching the VL model, our ensemble is comprised of 200 parameter realisations. Within four iterations, adjustment of all of these realisations achieved objective functions that are less than 1.5 times that achieved through model calibration.

6.3 Observations

6.3.1 General

The history-matching dataset that is used to constrain VL model parameters can be subdivided into three broad categories. These are:

- measurements of groundwater extraction;
- measurements of groundwater heads;
- “soft” observations expressed as penalty functions.

These are now discussed in turn.

6.3.2 Groundwater Extraction

The use of groundwater extraction data in LUMPREM model history-matching has already been discussed; see Section 3.3.

It is worth noting that groundwater extraction data appear twice in history-matching datasets provided to PEST_HP and PESTPP-IES; in so doing, they comprise two “observation groups”. The first of these groups feature raw extraction data. The second group features temporal differences of measured monthly extraction rates. Appropriate weighting ensures visibility of each of these groups in the overall objective function that PEST_HP and PESTPP-IES are asked to reduce. This strategy encourages the history-matching process to replicate seasonal variations of pumping rates, as well as pumping rates themselves.

6.3.3 Groundwater Heads

Time series of groundwater heads are available from boreholes whose locations are depicted as closed circles in Figure 2.4. Water level records are of varying length and quality. Measurements made in wells 606/647, 606/1026 and 610/179 are shown in Figure 6.1.

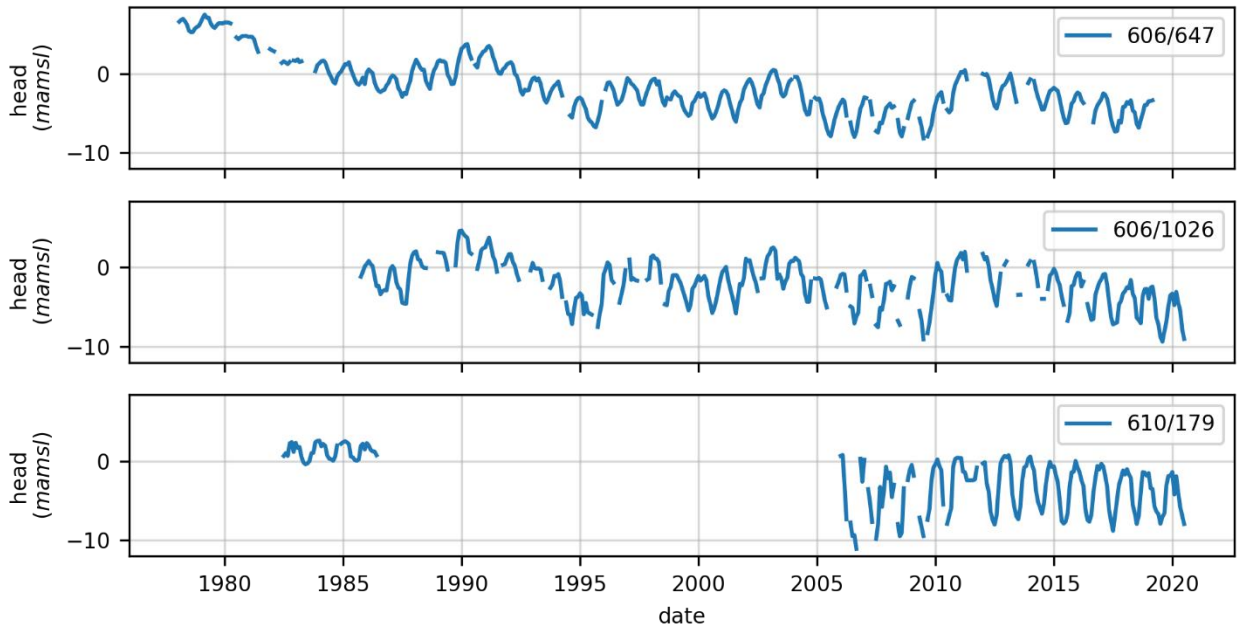


Figure 6.1 Time series of heads in three observation wells. See Figure 2.4 for the locations of these wells.

A similar strategy is adopted for matching of head data as is adopted for matching of extraction data. That is, PEST_HP and PESTPP-IES are asked to match successive differences between heads, as well as heads themselves. This renders the parameters to which these differences are sensitive more visible to the history-matching process. These parameters are principally MODFLOW 6 storage parameters and parameters belonging to LUMPREM soil moisture accounting models.

As mentioned in Section 2.3, the head measurement dataset is supplemented with measurements of water levels in extraction wells. These wells are shown as open circles in Figure 2.4. The VL model is not asked to replicate these measurements because:

- The amount of pumping (if any) at the time of water level measurement is unknown;
- Model cell areas are considerably greater than those of pumping wells; hence cell-calculated drawdowns underestimate drawdowns measured in pumping wells.

Nevertheless, these data should not be ignored. Hence they are used in formulation of a series of penalty functions. If a model-calculated head for the cell containing the production well in which such a measurement was made is lower than the corresponding measured head, then a model-to-measurement misfit is deemed to occur. However if the model-calculated head is greater than the corresponding borehole-measured head, the misfit penalty is zero.

6.3.4 Soft Data

As discussed in Section 2 of this report, the north-western boundary of the VL model allows water to enter the model domain. The amount of water that enters the VL system through this boundary is unknown. Indeed, assessment of this inflow is one of the reasons for development of the model. However, an upper limit can be calculated for this inflow based on the area of the north-western groundwater catchment, and the maximum likely recharge over this area. If inflow through the north-western boundary is calculated to exceed this amount during history-matching, a misfit penalty is deemed to occur.

As well as ensuring reasonableness of north-western inflow, enforcement of this penalty serves a secondary purpose. As is discussed in Section 2, difficult choices accompany specification of the eastern boundary of the VL model domain. It coincides with a management boundary rather than with a natural system boundary. The natural system boundary is far to the east; much uncertainty is

associated with hydraulic and extractive conditions that exist between the management boundary and the natural system boundary. Hence, it was decided to transfer this uncertainty to a much closer model boundary. The stochasticity of properties (heads and conductances) that are ascribed to this boundary must therefore reflect uncertainties associated with partly-known processes and properties to its east, as well as the abstract nature of the boundary itself. Once formulated, the uncertainties of these boundary properties can be reduced through history-matching. This process is assisted by penalizing parameter sets that give rise to unrealistic simulated hydraulic behaviour at and near the boundary. As well as respecting expert knowledge, penalising north-western boundary inflow discourages water that flows into this boundary from flowing straight out of the northern portion of the eastern model boundary.

Two further measures are introduced to encourage reasonable hydraulic behaviour near the eastern model boundary. It is desirable that the model reproduces the following conditions.

- During the transient component of the history-matching period, water flows predominantly into the model domain through its eastern boundary rather than out of it.
- During simulation of pre-development, steady-state conditions, heads along the eastern model boundary descend monotonically towards the coast. (The boundary is roughly perpendicular to topographic contours.)

Penalties are included in the history-matching dataset that ensure estimation of parameters that respect these conditions.

6.4 Results

6.4.1 Calibration

Representative modelled and measured extraction rates are plotted together in Figure 3.3.

Representative modelled and measured heads are graphed in Figure 6.2. Fits are reasonable, but also a little disappointing in the case of 610/179; the model is unable to reproduce large dry season drawdowns experienced by this well. We are unsure of the reason for this. Both 610/179 and its neighbouring well 610/613 (for which model-to-measurement misfit is similar) are close to pumping wells. Perhaps high summer drawdowns in these wells arise from local anomalies in hydraulic properties that the model is unable to reproduce. Alternatively, it is also possible that their drawdowns reflect proximity to unregistered extraction wells.

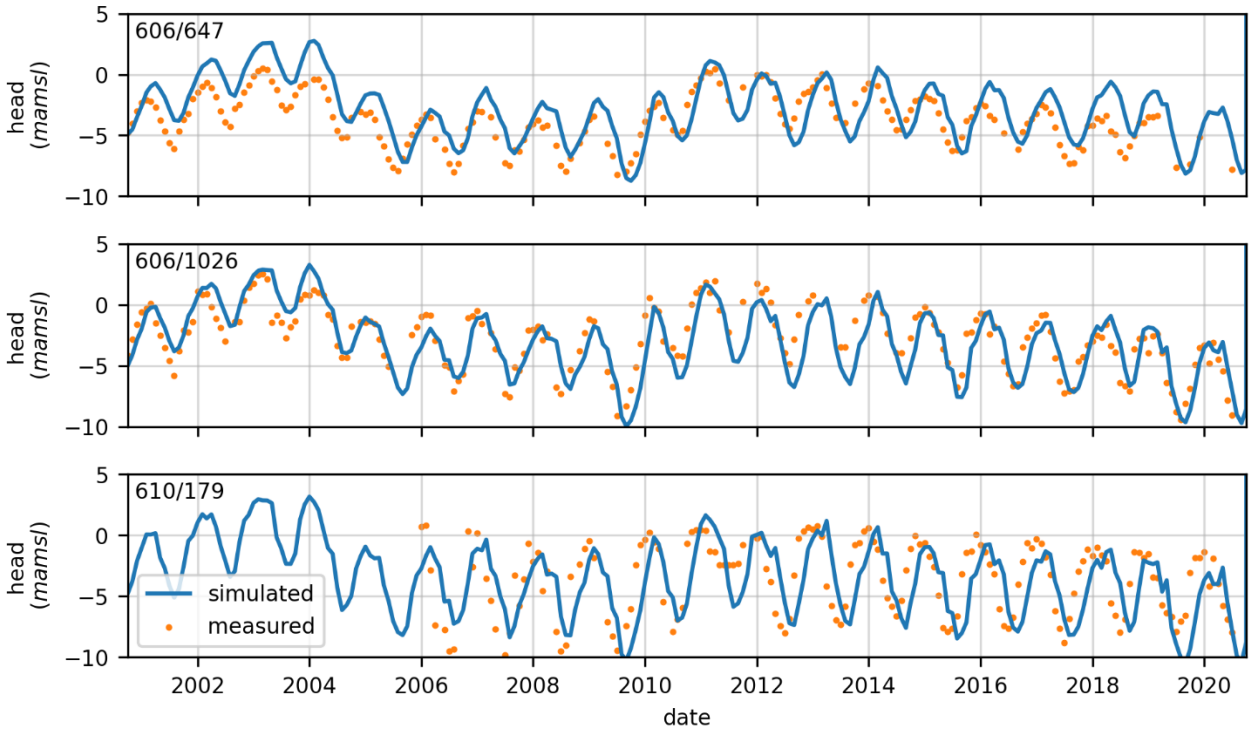


Figure 6.2 Measured and modelled heads in three observation wells. See Figure 2.4 for the locations of these wells.

Maps of calibrated hydraulic conductivity and specific storage are provided in Figure 6.3a and in Figure 6.3b respectively. Being outcomes of a calibration process, these parameter fields exhibit only as much heterogeneity as is required to fit the history-matching dataset. Nevertheless, patterns of heterogeneity are clearly visible.

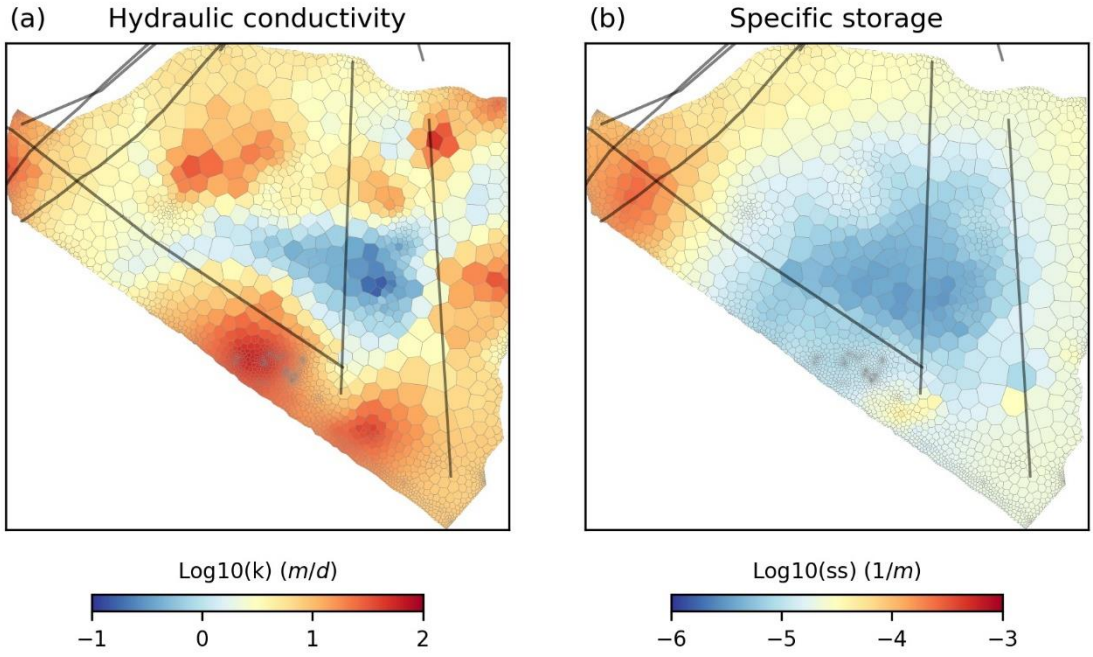


Figure 6.3 Spatial disposition of (a) hydraulic conductivity and (b) specific storage estimated through model calibration. Black lines delineate mapped faults.

Areas of anomalously high hydraulic conductivity may be linked to faults which transect the model domain. It is also possible that anomalously high and low hydraulic conductivities that are estimated

close to the eastern boundary of the VL model compensate for simplifications that are embodied in this boundary. This matter is discussed in Section 7.

High values of estimated specific storage in the western corner of the model domain suggest the possibility of a connection with the overlying PQ aquifer, possibly induced by faulting. (It is of interest to note the existence of a nearby spring.)

Figure 6.4 shows heads throughout the model domain calculated by the model (a) under steady-state, pre-development conditions and (b) at the end of the history-matching period.

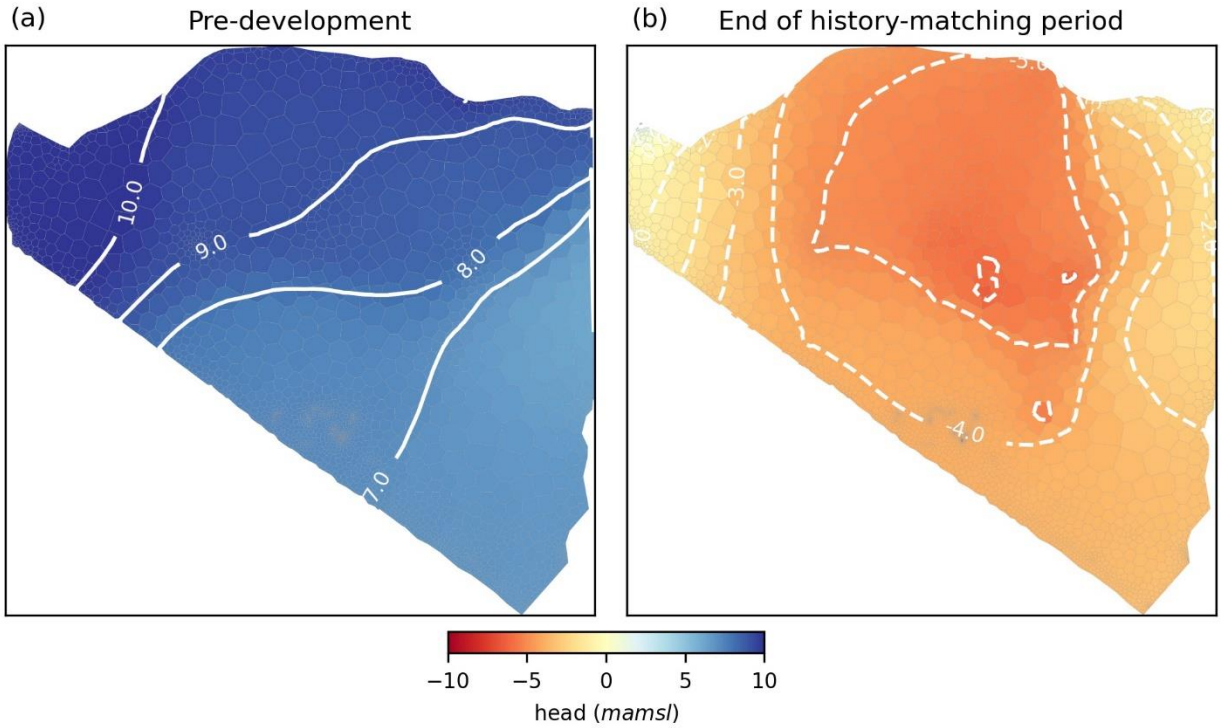


Figure 6.4 Heads calculated by the calibrated model under (a) steady state, pre-development conditions and (b) at the end of the history-matching period.

6.4.2 Stochastic History-Matching

Measured pumping rates are compared with those calculated using 200 PESTPP-IES-derived samples of the posterior parameter probability distribution in Figure 3.4. Measured borehole heads are compared with those calculated using these samples in Figure 6.5.

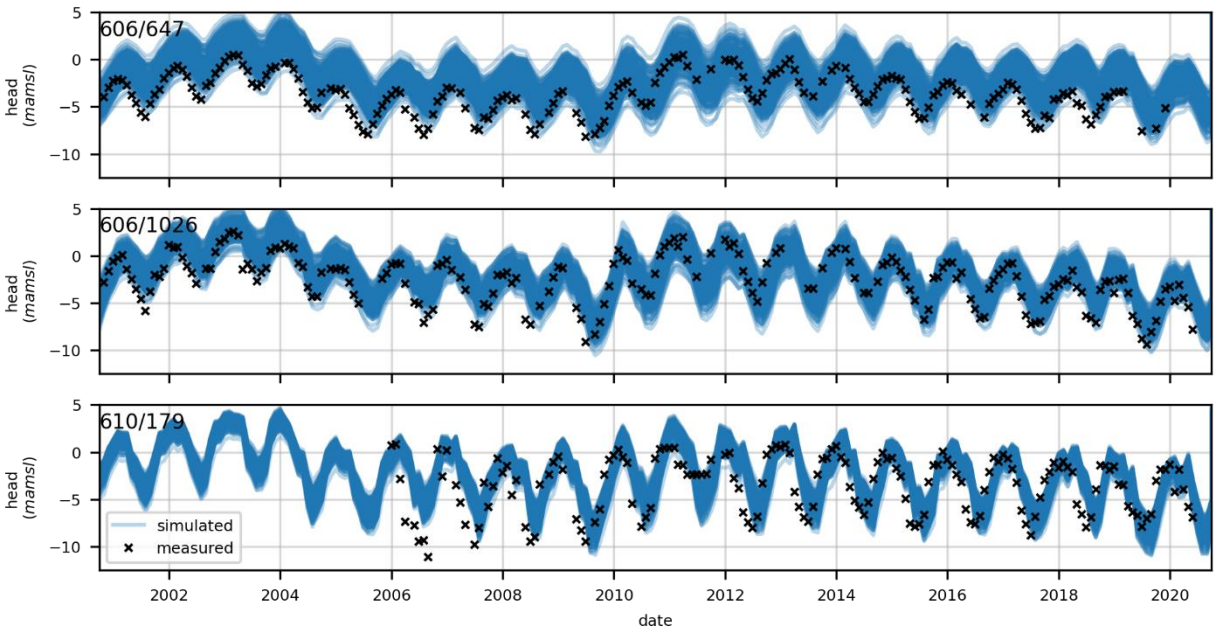


Figure 6.5 Measured and modelled heads in three observation wells. Heads are calculated using samples of the posterior parameter probability distribution derived by PESTPP-IES. See Figure 2.4 for the locations of these wells.

Figure 6.6 shows eight hydraulic conductivity fields calculated by PESTPP-IES while Figure 6.7 shows corresponding specific storage parameter fields. A comparison of these parameter fields with those provided in Figure 6.3 reveals that parameter fields emerging from stochastic history-matching show much greater hydraulic parameter heterogeneity than those emerging from calibration

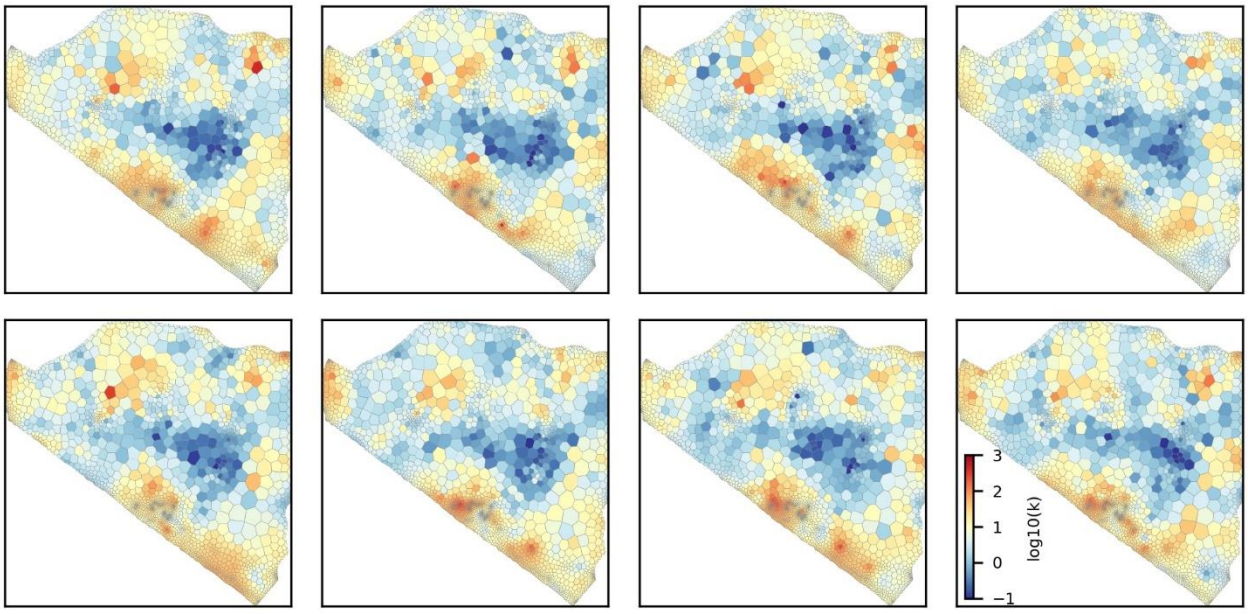


Figure 6.6 Eight hydraulic conductivity fields calculated by PESTPP-IES.

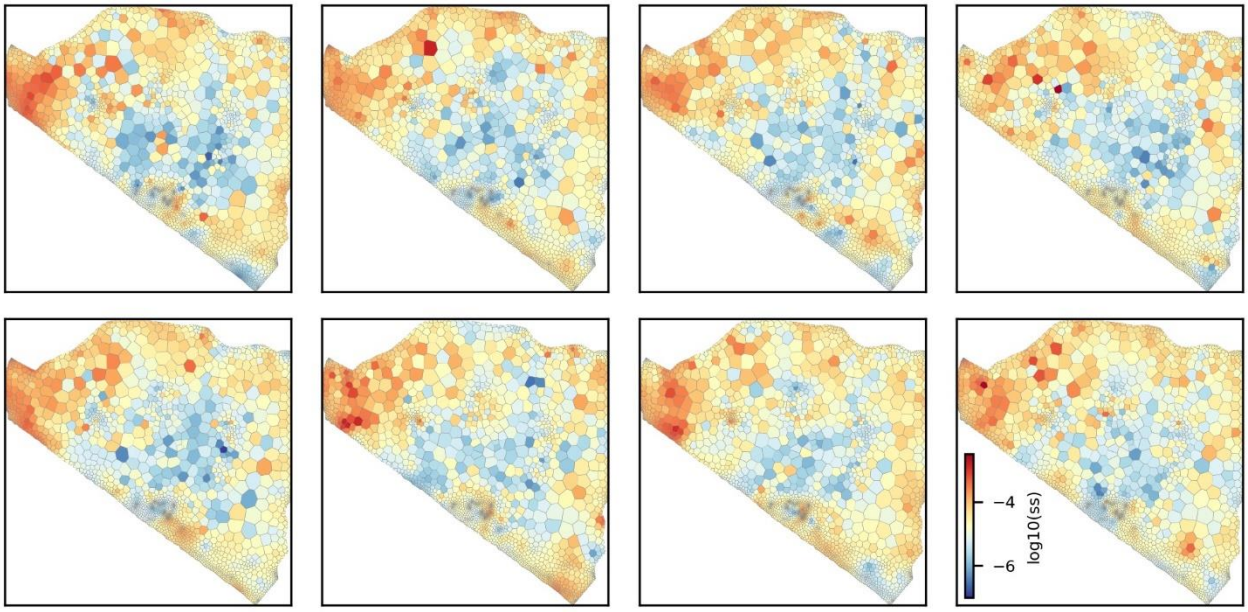


Figure 6.7 Eight specific storage fields calculated by PESTPP-IES.

7. MODEL PREDICTIONS

7.1 General

7.1.1 Predictions

In this section we use the history-matched VL model to make two predictions that are salient to management of the VL system. First we examine water balance of the MC aquifer. Then we calculate a sustainable rate of groundwater extraction from the MC aquifer. In doing so, we explore how pumping should be distributed among existing users in order to maximize this extraction while, at the same time, guaranteeing its sustainability.

7.1.2 Predictive Bias

As has been discussed, the position of the eastern boundary of the VL model is somewhat problematical. It coincides roughly with the VL management boundary. This has the advantage that model-based calculations of water balance are relatively straightforward. Another advantage is that it excludes from the model domain parts of the aquifer where hydraulic conditions and pumping rates are uncertain. Model run times and parameterisation complexity are also reduced. Meanwhile the boundary's parameterisation is designed to bear the weight of uncertainty arising from conditions further to the east.

This strategy is not without risk. White et al (2014) and Doherty (2015) point out that history-matching of a structurally simple model may induce bias in some predictions, despite the fact that model outputs fit historical data well. However the propensity for history-matching-induced bias is prediction-specific. For data-driven predictions, model imperfections are "calibrated out"; hence while some model parameters to which predictions of this type are sensitive may incur bias, these data-driven predictions do not. In contrast, other model predictions may inherit bias from history-match-biased parameters to which they are sensitive.

It follows that a model which is built solely to make data-driven predictions can have ease of history-matching as its sole design criterion. However more caution must be exercised when building a model which must make predictions which are partly informed by history-matching and partly informed by expert understanding of a system.

The previous section noted that history-matched hydraulic conductivity parameters exhibit some signs of artefacts near the model's eastern boundary. It does not follow that predictions which are discussed in this section are therefore biased or, if they are biased, that their propensity for bias is not included in quantified predictive uncertainty intervals. For the VL system, the latter are already large because of data scarcity. Nevertheless, we address this issue as we discuss these predictions.

7.2 Water Balance

Compliance with EU Water Framework Directives requires that components of the VL system water balance be evaluated. This can be done using the history-matched VL model.

Figure 7.1(a-e) shows five time series. These are inflows into the VL model domain over its history-matching period from the three families of GHBs that encircle it, and from the spatially distributed GHB which overlies it. LUMPREM-calculated water extraction rates are also shown. These components of the water balance are computed using 200 samples of the posterior parameter probability distribution calculated by PESTPP-IES. Those calculated using the parameter set achieved through model calibration are also shown.

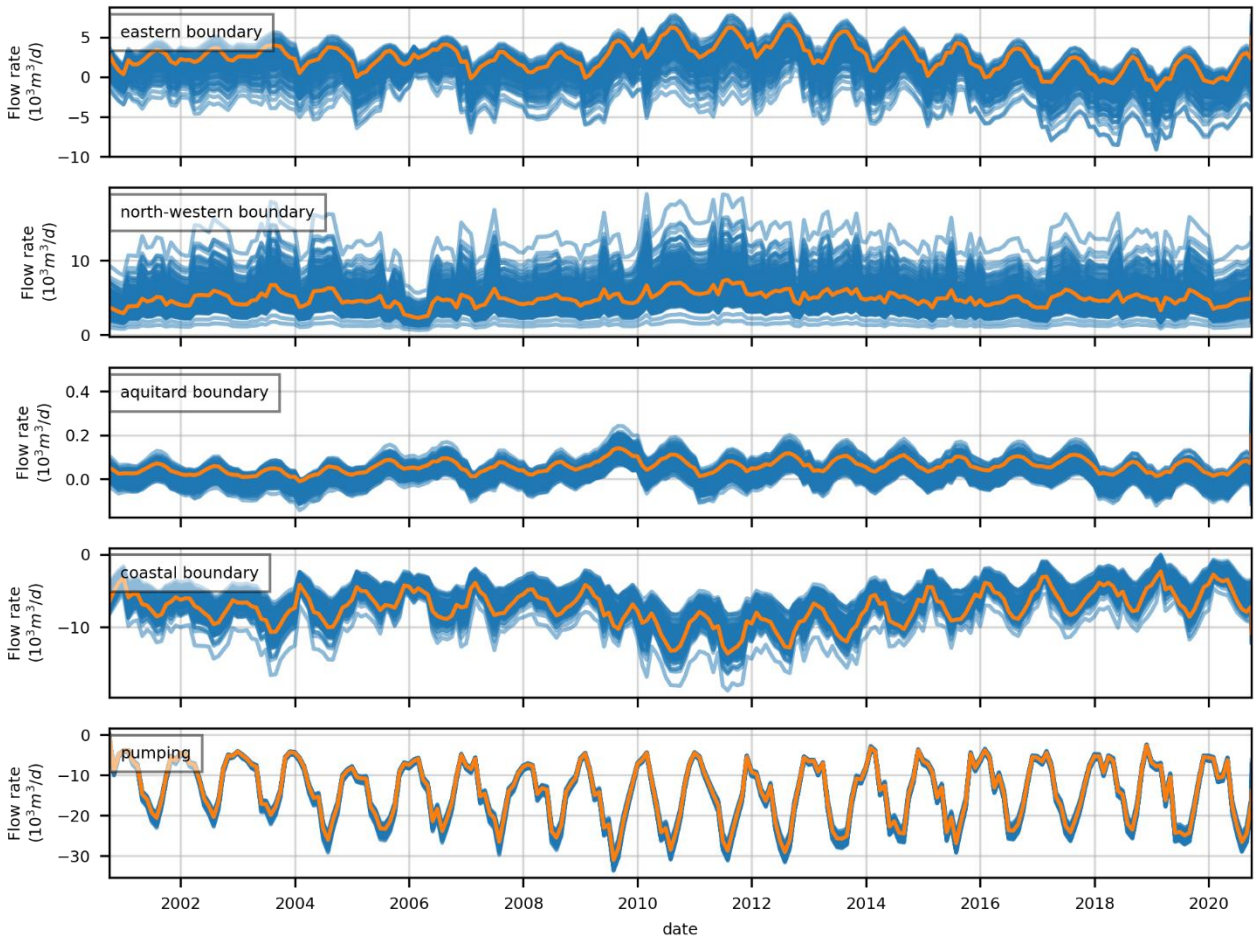


Figure 7.1 Components of the VL water balance over the 20 year history-matching period. Orange-coloured time series were calculated using the parameter set achieved through model calibration.

Time-averaged components of the water balance, together with respective uncertainty standard deviations, are listed in Table 7.1. These statistics are calculated by time-averaging each realisation of each flow component, and then computing the mean and standard deviation over all realisations using standard formulas.

Table 7.1 Water balance of VL aquifer over the history-matching period.

Inflow from	Mean (Mm^3/yr)	Standard deviation (Mm^3/yr)
Eastern boundary	0.23	0.59
North-western boundary	2.30	0.89
Aquitard	1.4×10^{-3}	0.01
Coastal boundary	2.64	0.42
Pumping	-5.17	0.17

Features of interest in Table 7.1 include the following:

- Inflow to the MC aquifer from the overlying PQ aquifer comprises only a small component of the overall water balance.
- There is a considerable influx of water to the system from the seaward side of the coastal boundary.
- Inflow to the system from the north-western boundary is large.
- Uncertainties associated with all boundary inflows are large.

The EU Water Framework Directive specifies that abstraction be less than 90% of average annual recharge in managed groundwater systems such as Vale do Lobo by 2027. This criterion for "good management" is simplistic. For the VL system it is difficult to apply, as extraction from the system induces the recharge that it requires. Nevertheless, if the above table is viewed as a static ledger, it can be established that 2.9 ± 1.07 Mm/yr of extra recharge is required if all non-coastal inflow is to exceed pumping. As stated above, managed aquifer recharge is being contemplated as a system management option.

Simplifications encapsulated in the eastern model boundary are unlikely to compromise the above calculation. Reasons are as follows:

- The posterior uncertainty ascribed to eastern boundary inflow is high; this is likely to exceed any history-matching-induced bias that is associated with this inflow.
- The water balance component of management interest is inflow over the entire boundary; local structural and parametric errors are likely to "cancel out" when calculating total inflow.
- Extraction rates form an important component of the VL model history-matching dataset; this creates a data-driven link between the prediction and field measurements.

7.3 Sustainable Extraction

7.3.1 Problem Definition

Extraction of water from the MC aquifer is threatened by salt water intrusion. Protection of the resource requires that extraction be limited, and that the spatial distribution of extraction be such as to preclude landward flow of saline water. Ideally, the spatial disposition of extraction should be such as to maximize it subject to the constraint that the system is protected. This comprises an optimisation problem. The optimal disposition of extraction will depend heavily on patterns of hydraulic conductivity that prevail within the MC aquifer. History-matching of the VL model has provided some insights into these patterns; see Figure 6.3 and Figure 6.6.

In order to demonstrate how this optimisation problem may be solved, we run the VL model under steady state conditions. This allows us to explore long-term sustainability. We designate rates of water extraction ascribed to the nine user groups discussed in Section 2.5 as our decision variables. Using software described below, we maximize these while maintaining the spatial disposition of pumping within each user group that is inherited from the history-matching process. In practice, these distribution factors should also be decision-variables.

A single constraint is imposed on the extraction maximisation problem. This is that there be zero net flow of water landward from the VL coastal boundary into the model domain. Ideally, we should insist that inflow be zero at all places along this boundary. (In fact, even this tighter constraint does not necessarily ensure the avoidance of salt water contamination of irrigation water; this requires imposition of the constraint that hydraulic heads at all pumping wells are large enough to prevent upconing of saline water.) However, we found that, with omission of spatial distribution factors from decision variables, the imposition of constraints which are too tight prevents attainment of a feasible solution to this optimisation problem when parameter uncertainty is taken into account; see below.

7.3.2 Difficulties

A number of difficulties beset solution of this optimisation problem. These are now briefly discussed.

The high conductances ascribed through history-matching to eastern GHB boundary conditions make it clear that the VL system cannot be managed in isolation. Pumping of the MC aquifer to the east of this boundary lowers groundwater levels within the VL system itself. This predisposes the system to saltwater intrusion when it is pumped.

To accommodate this issue we solve the above-defined optimisation problem with the head at the eastern boundary set at five different levels, this simulating five different intensities of water extraction from the neighbouring system. This is accomplished by assigning five different values to the head in well 606/1033, and then using calibrated α_i and β_i values (see equation 4.1) to ascribe heads to cell-specific GHBs along the boundary. The highest of these five 606/1033 head values yields GHB boundary heads that are close to those calculated for pre-development conditions; the lowest of these values yields boundary heads that resemble present-day conditions. The optimisation problem is then solved five times, once for each of these boundary head distributions. Meanwhile the groundwater level in well 606/1050 that governs heads along the north-western model boundary is set to its average value over the history-matching period.

A second problem is the possible presence of a salt water wedge beneath the coast under steady state conditions. With extraction maximized, flow of fresh water towards the coast is reduced. Consequently, heads near the coast may be sufficiently low for the elevation of the freshwater-saltwater interface to exceed that of the base of the MC aquifer. If the interface is sharp, saline water below it is motionless under steady state conditions. (In reality, salt water moves slowly inland to support the existence of diffusion-driven salt water recirculation at the interface.) Above the interface, fresh water flows towards the sea. If the freshwater-saltwater interface extends into the model domain, this limits the cross-sectional area that is open to fresh water flow, and hence the transmissivity that is available to it. This raises fresh water heads, which lowers the interface.

In the present study, this phenomenon is accommodated by using a slightly altered version of MODFLOW 6. Alterations to the code allow model-calculated transmissivity to reflect the thickness of fresh water above the freshwater-saltwater interface. During every iteration of the heads' solution process, the Ghyben-Herzberg relationship is used to calculate the elevation of the freshwater-saltwater interface from that of fresh water head; transmissivity is adjusted accordingly.

The third problem is that of parameter uncertainty. To accommodate this difficulty, we solve the optimisation problem defined above in two different ways. First we endow the VL model with a single parameter set, namely the parameter set of minimum error variance emerging from model calibration. This optimisation problem is solved using PESTPP-OPT (supplied with the PEST++ suite). Then we solve the optimisation problem using the ensemble of parameter sets that were calculated by the PESTPP-IES ensemble smoother. In the latter case, definition of the optimisation constraint varies slightly. We now maximize groundwater extraction under the constraint that inland flow from the coastal boundary is zero or less (i.e. flow is outward) for all 200 parameter fields. This simulates the operation of a very risk-averse management strategy. This optimisation-under-uncertainty problem is solved using the CMAES_HP global optimiser (supplied with PEST_HP).

A fourth difficulty was alluded to in Section 7.1.2. This is the problem of using the eastern GHB as a surrogate for processes that operate on the other side of it. We explore this difficulty using linear analysis.

The back row of Figure 7.2 shows pre-history-matching contributions to the uncertainty of model-predicted coastal inflow by different types of parameters employed by the VL model; parameter type names are listed in Table 7.2. This prediction is salient to the optimisation problem posed above as it imposes constraints on pumping rates. Note the logarithmic scale in this figure. Note also that variance is the square of standard deviation.

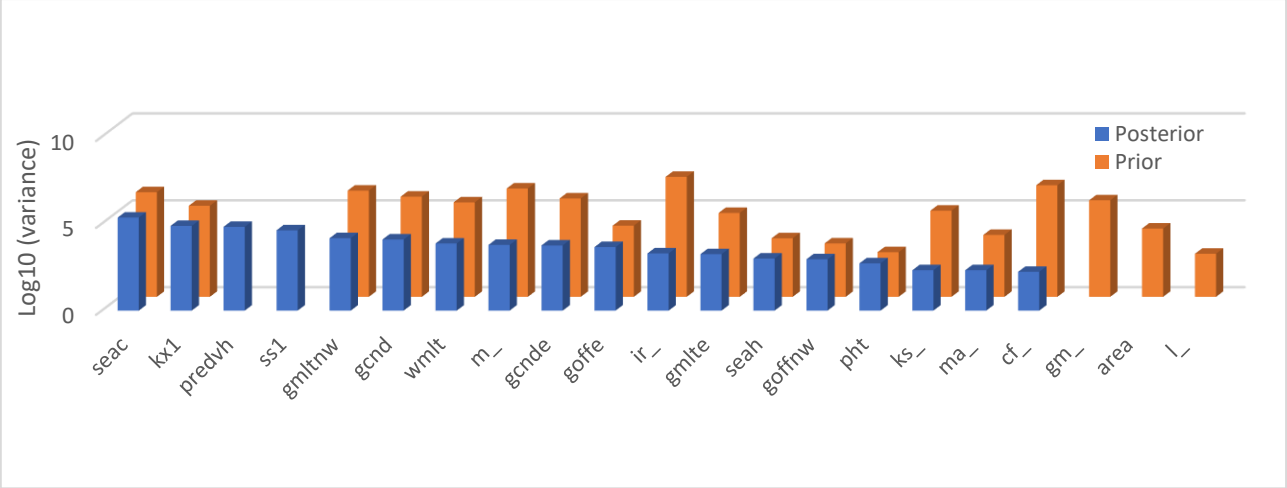


Figure 7.2 Contributions to the uncertainty of model-predicted coastal boundary inflow by different parameter types employed by the VL model. The back row provides pre-history-matching contributions while the front row provides post-history-matching contributions. Parameter group names are explained in Table 7.2.

Table 7.2 Description of parameter type names shown in Figure 7.2.

Parameter type	**Name in Figure 7.2**
LUMPREM parameters	Names end with an underscore
Irrigation areas	area
Coastal GHB conductance	seac
Coastal GHB head	seah
North-western boundary conductance	gcndnw
North-western boundary α_i (used to calculate boundary heads)	goffnw
North-western boundary β_i (used to calculate boundary heads)	gmltnw
Eastern boundary conductance	gcnde
Eastern boundary α_i (used to calculate boundary heads)	gmlte
Eastern boundary β_i (used to calculate boundary heads)	goffe
Hydraulic conductivity	kx1
Specific storage	ss1
Pre-development heads at E and NW boundaries	predvh
Intra-user group borehole fractions	wmlt
Aquitard boundary conductance	pht

Figure 7.2 demonstrates that some parameters can have large contributions to posterior predictive uncertainty while possessing small or zero contributions to prior predictive uncertainty. This is an outcome of the definition of "contribution to uncertainty by a parameter type". It is defined as "the reduction in uncertainty of the prediction that is accrued through acquisition of perfect knowledge of values of pertinent parameters". In the present case, the prediction of interest is steady state inflow from the coastal model boundary. Obviously, this is insensitive to parameters such as specific storage. Hence the prior contribution of this parameter type to the uncertainty of this prediction is zero. However, if one were to acquire perfect knowledge of specific storage throughout the model domain, this would reduce uncertainties in history-match-estimated values of hydraulic conductivity as the two parameter types are highly correlated when calibrated against a small measurement dataset. This, in turn, would reduce the posterior uncertainty of predicted coastal inflow.

Prior contributions to the uncertainties of coastal inflow are dominated by LUMPREM model parameters; these determine extraction rates. Posterior contributions to coastal inflow uncertainty are dominated by coastal boundary conductances and by hydraulic conductivities within the MC aquifer. Pre-and post-history-matching contributions by eastern boundary conductances are significant but non-dominant. Those made by parameters that determine eastern boundary heads are of much less importance.

Because eastern boundary conductances are surrogates for connections between the VL system and the groundwater system to its east, their importance is unsurprising. Nor is the reduction in their uncertainties accrued through history-matching. Any history-matching-induced bias that these parameters transfer to a prediction of coastal inflow is likely to be expressed as an unknown component of the difference between their prior and posterior contributions to the uncertainty of this prediction. It is also likely to be expressed through contributions to uncertainty, and history-matching-accrued reductions thereof, made by parameters that determine eastern boundary heads. The latter are relatively small compared to other sources of coastal inflow predictive uncertainty.

This brief analysis suggests that solution of the optimisation problem that is described herein may indeed be susceptible to history-matching-induced biases that accompany use of GHBs along the eastern model boundary in place of explicit representation of a wider system. However these biases do not appear to be sufficient to invalidate the analysis if they are viewed from a perspective that acknowledges the considerable uncertainties that arise from data insufficiency.

7.3.3 Results

The blue dots in Figure 7.3 depict solutions to the five constrained optimisation problems discussed above. As stated, these problems are solved using PESTPP-OPT. The VL model is endowed with the minimum error variance parameter field derived through model calibration; see Figure 6.3. Because use of this parameter field (ideally) yields model predictions which are at the centres of their posterior probability distributions, these are collectively referred to as “risk neutral” solutions to the extraction optimisation problem. The vertical axis of Figure 7.3 represents total extraction as a fraction of current extraction.

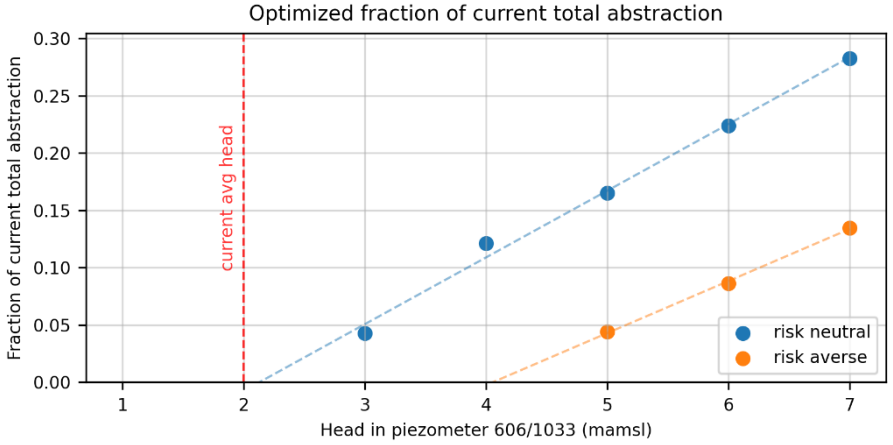


Figure 7.3 Solutions to constrained extraction optimisation problems. Blue dots are risk neutral solutions while orange dots are risk averse solutions.

The highest of the five eastern boundary heads for which the constrained optimisation problem is solved corresponds to that which would be expected under pre-development conditions, while the lowest corresponds to heads which are slightly above those which are being experienced at present. It is apparent that even under the former (benign) conditions, long-term extraction rates should be reduced to about 30% of what they are now in order to protect the MC aquifer. Alternatively, if water use on the other side of the eastern boundary continues at its present rate, any extraction from the VL aquifer is unsustainable, for eastern groundwater extraction already induces flow of water from the VL system coastal boundary and through the VL system itself.

Figure 7.4a shows fractional extraction rates by different user groups for these five solutions. Users are labelled “A” to “I” rather than 1 to 9 as we do not wish to disclose who they are.

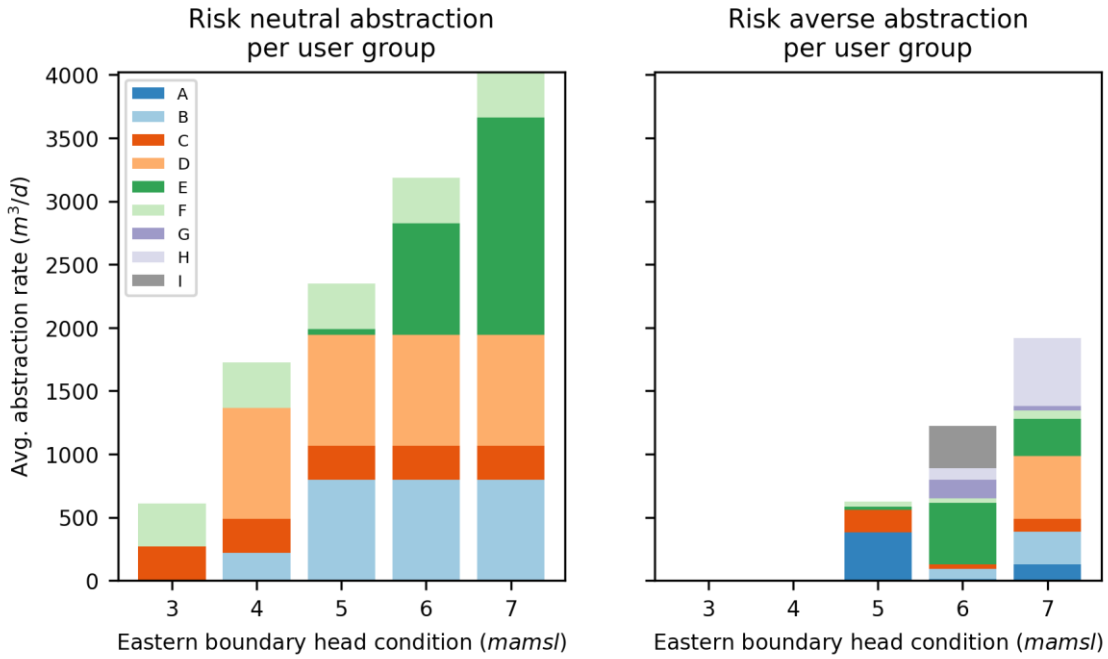


Figure 7.4 Relative extraction rates for different user groups for (a) the five solutions of the risk neutral optimisation problem depicted in Figure 7.3 and (b) the three solutions of the risk averse optimisation problem.

The orange dots in Figure 7.3 show risk averse solutions to the constrained extraction optimisation problem. As stated above, these are calculated using CMAES_HP. The VL model employs all 200 samples of the posterior parameter probability distribution calculated by PESTPP-IES; meanwhile, CMAES_HP ensures that none of these fields precipitate landward flow of groundwater from the coastal boundary. A solution to this optimisation problem is not feasible for low eastern boundary heads. This is because some parameter fields allow landward flow from the coast even if there is no pumping of the MC aquifer with these low heads assigned to the eastern model boundary.

User group extraction rates for the three viable solutions of the risk averse constrained optimisation problem are shown in Figure 7.4b.

8. CONCLUSIONS

Groundwater modelling that is the subject of this report was far from straightforward. It required the making of many design decisions. As is typical of groundwater modelling, with the benefit of hindsight some of them would have been made differently if the work were to be repeated.

In this report we describe, among other things:

- Construction of a groundwater model for a coastal aquifer system;
- Employment of a methodology for estimating monthly demand for water extracted from that system;
- Use of the model to assimilate data that can elucidate elements of the system's water balance while estimating some of its hydraulic properties;
- Quantification of the uncertainties of all such estimates;
- Optimisation of extraction under both risk-neutral and risk-averse management approaches.

To our knowledge, there are very few incidences of coastal modelling work in which the last three of the above tasks have been successfully completed. Modern simulation technology allows detailed numerical replication of complex processes that prevail at a freshwater-saltwater interfaces. However it is unable to exploit the information content of data that records the history of groundwater behaviour at a particular site. Such data are often both plentiful and information-rich.

Aquifer hydraulic properties, and the spatial disposition of heterogeneity associated therewith, exert a strong influence on a coastal system's response to human management. Movement of saline and non-saline waters is particularly sensitive to horizontal and vertical contrasts in hydraulic conductivity. Because simulation of density-dependent flow requires the use of many model layers, the parameterisation burden of density-dependent simulation is as heavy as its numerical burden if a density-dependent model is to give full stochastic expression to the potential variability of subsurface hydraulic properties. Unfortunately, density-dependent models take too long to run, and are too numerically unstable, to undergo satisfactory history-matching; nor can they be used in history-match-constrained predictive uncertainty analysis.

Compromises must therefore be made if numerical simulation is to support management of coastal groundwater systems. The present report documents some workable compromises which may be worthy of consideration at other sites.

Fundamental to the modelling approach that is documented herein is the premise that assessment of current groundwater conditions, and predictions of future groundwater behaviour, are uncertain. While acceptance of the presence of uncertainty should trigger recognition of the burden of having to quantify uncertainty, it also renders the task of decision-support modelling somewhat forgiving – especially in areas such as Vale do Lobo where data are relatively scarce. This is because simulation errors that may be incurred by numerical compromises that expedite computational speed and simulator stability may be small in comparison to uncertainties incurred by data insufficiency. If a relatively simple model can provide access to important information through its ability to assimilate site data, then its use may engender a reduction in the uncertainties of at least some decision-critical model predictions. It may also support quantification of these uncertainties. Naturally, errors accrued through use of the simple model for this purpose should be included in the uncertainty intervals that it calculates.

Model simplifications that are discussed in the present GMDSI worked example report include the following:

1. Restrictions in the area of the model domain, this requiring that conditions on model boundaries reflect properties and processes that are operative on the other side of these boundaries;
2. Use of a single-layer, single-density groundwater flow model.

The second simplification is justified by the fact that although water quality is deteriorating in the VL system, measured salt concentrations are still relatively low. However, under pre-extraction conditions, fresh water encounters saline water as it flows under the coast and eventually to the sea. Conversely, under extractive conditions saline water is induced to move inland. The presence of saline water on the seaward side of the model's coastal boundary must be reflected in its design. To the extent that hydraulic conditions on the other side of the boundary are unknown, this must be reflected in stochasticity of parameters with which that boundary is endowed.

In order to accomplish this, parameterisation of the coastal boundary is informed by stochastic runs of a density-dependent flow model that roughly simulates offshore conditions at Vale do Lobo. Processing of the outcomes of these runs allows construction of a prior probability density function from which boundary heads and conductances can be sampled. Uncertainties in these boundary parameters are then reduced through history-matching.

In this report, we describe how the history-matched model can be used to explore two issues that are salient to management of the Vale do Lobo groundwater system. The first of these issues is the system's overall water balance. Model-based data assimilation demonstrates conclusively that use of water exceeds supply. Simpler calculations, performed without the assistance of a model, would probably have led to the same conclusion. However, given the importance of water balance evaluation to compliance with EU Water Framework Directives, model-based quantification of its components, and assessment of their uncertainties, raises the credibility of these calculations.

We also demonstrate how the model can be used to manage groundwater extraction in a way that supports its sustainability. This is accomplished by using the model in conjunction with partnered software to solve a constrained optimisation-under-uncertainty problem. Formulation of this problem can account for risks that management is prepared to accept.

9. REFERENCES

Bakker, M., 2006. Analytic solutions for interface flow in combined confined and semi-confined, coastal aquifers. *Adv. Water Resour.* 29 (3), 417–425. http://dx.doi.org/10.1016/j.advwatres.2005.05.009.

Bakker, M., Miller, A. D., Morgan, L. K., and Werner, A. D., 2017. Evaluation of analytic solutions for steady interface flow where the aquifer extends below the sea. *Journal of Hydrology*, *551*, 660–664. https://doi.org/10.1016/j.jhydrol.2017.04.009

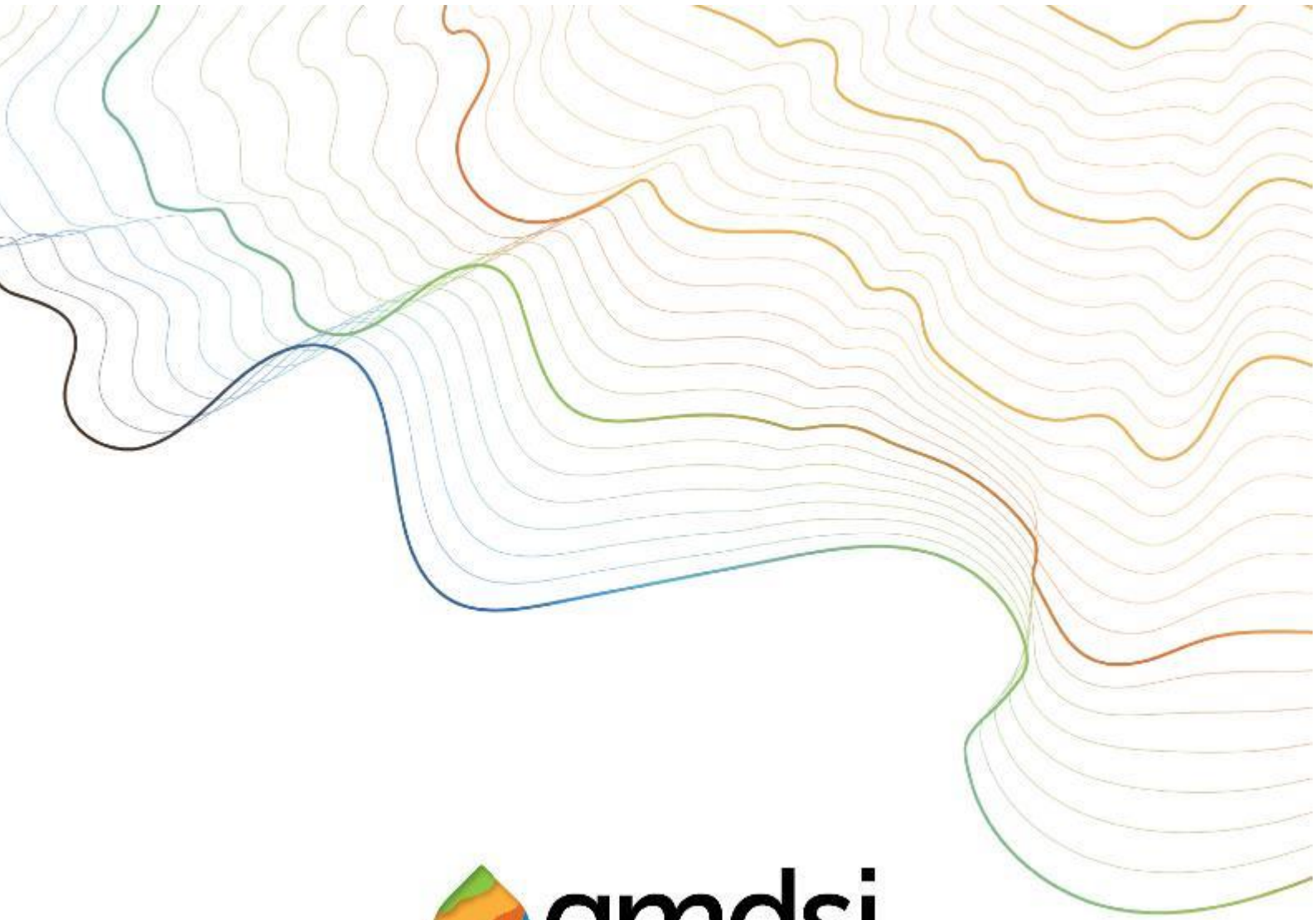
Doherty, J., 2015. Calibration and uncertainty analysis for complex environmental models. Published by Watermark Numerical Computing, Brisbane, Australia. 227pp. ISBN: 978-0-9943786-0-6. Downloadable from www.pesthomepage.org.

Lu, C., Werner, A. D., Simmons, C. T., and Luo, J., 2015. A Correction on Coastal Heads for Groundwater Flow Models. *Groundwater*, *53*(1), 164–170. https://doi.org/10.1111/gwat.12172

Reis, E., 2007. Ortho-photo assessment of landuse in the Algarve, internal report, Agência Portuguesa do Ambiente, Faro, Portugal.

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., 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.



gmdsi

Groundwater Modelling
Decision Support Initiative

gmdsi.org

BHP

Flinders UNIVERSITY

NATIONAL CENTRE FOR GROUNDWATER RESEARCH AND TRAINING

RioTinto