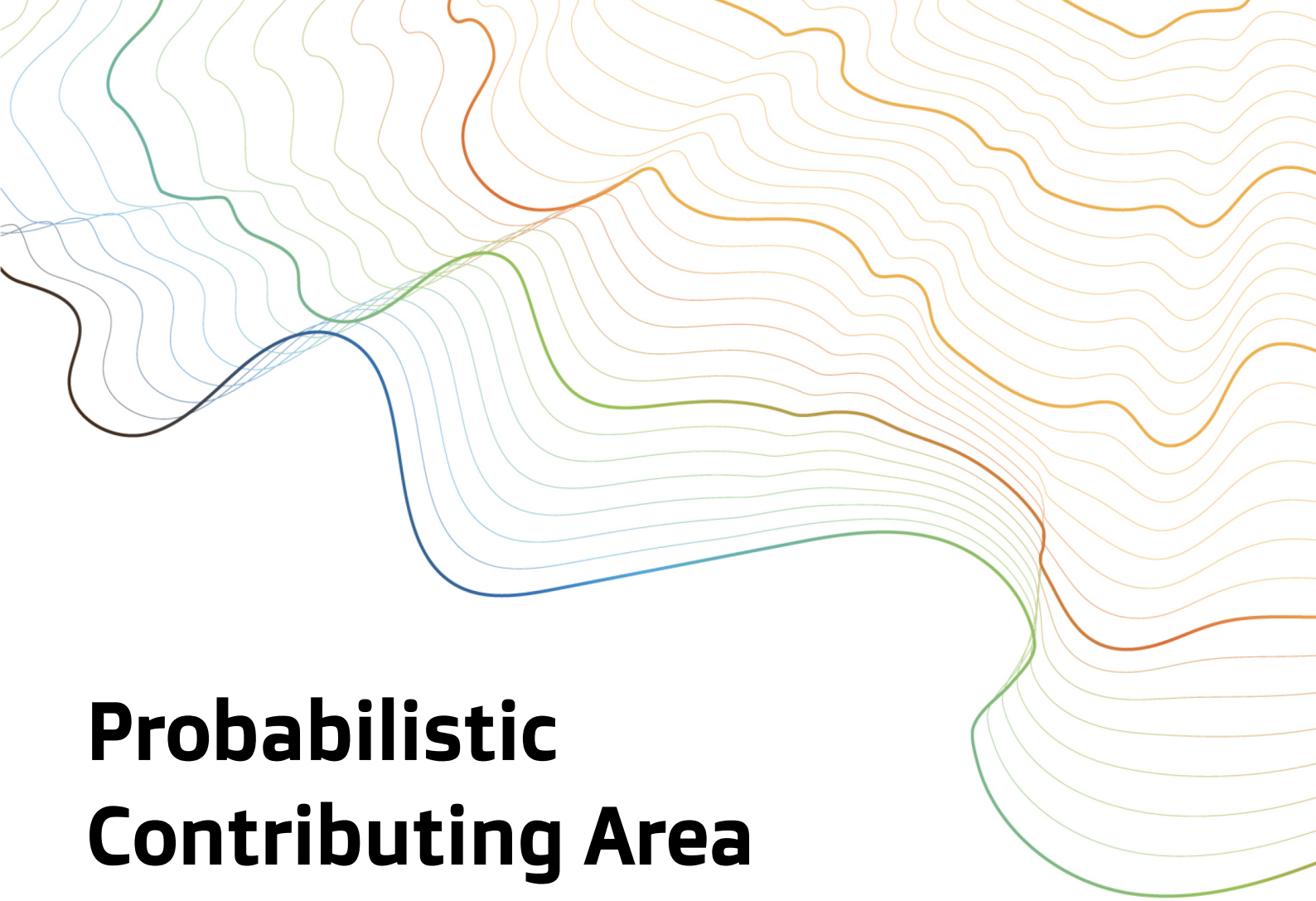

Probabilistic Contributing Area Analysis

A GMDSI worked example report

by John Doherty, Jim Rumbaugh, and Chris Muffles

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

CITATION

For bibliographic purposes this report may be cited as: Doherty, J., Rumbaugh, J. and Muffels, C., (2021). Probabilistic Contributing Area Analysis. A GMDSI Worked Example Report. National Centre for Groundwater Research and Training, Flinders University, South Australia.

ISBN: 978-1-925562-57-6

DOI: 10.25957/m47c-ra29

DOI as a weblink: https://doi.org/10.25957/m47c-ra29

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 aquifers threatened by salt water intrusion.

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

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

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

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

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 flow 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 promulgate a better fit between model outputs and corresponding field measurements. Often, but not always, these inputs represent physical or chemical properties of the system that a model

simulates. However, there is no reason why they cannot also represent water or contaminant source strengths and locations.

Phreatic surface

The water table.

Pilot point

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

Prior probability

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

Posterior probability

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

Probability density function

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

Probability distribution

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

Quadtree mesh refinement

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

Realisation

A random set of parameters.

Regularisation

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

Residual

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

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

This GMDSI worked example describes model-based evaluation of the land surface area that contributes water to a production well. Modelling to address this issue is not new. In fact, we take as our starting point a USGS model that was developed for this purpose in 1992. The model simulates movement of groundwater in a small aquifer in south-western New Hampshire. The aquifer, composed of unconsolidated glacial sediments, occupies a small valley that is drained by the Souhegan river and its tributaries. We rebuild and re-calibrate the original model in order to demonstrate some innovations in contributing area analysis that are enabled by new technologies.

We employ MODFLOW-USG for numerical simulation, together with mod-PATH3DU for particle tracking. The model is calibrated using PEST_HP. The PESTPP-IES ensemble smoother is then used to derive a suite of random parameter fields, all of which allow the model to fit the calibration dataset just as well as the parameter field achieved using PEST_HP. By evaluating contributing areas using all of these parameter fields, we develop contribution probability maps for two public water supply wells.

This study raises a number of interesting issues.

The first issue is that of enabling other modellers to do the same kind of analysis themselves. For this GMDSI worked example we provide model and PEST input datasets, as well as instructions on how to build them and deploy them. The model will also be the subject of a Groundwater Vistas tutorial. A related issue is that of data transfer to platforms such as QGIS and PARAVIEW wherein model outputs can be displayed and modified in a spatial setting against a backdrop of other data. This too is demonstrated in the accompanying set of files.

Then there are issues raised by the modelling itself. These concern structural complexity and history-matching (which are related matters).

Where an extraction well can draw some of its water from a stream, its contributing area is very sensitive to aquifer hydraulic properties in the immediate vicinity of the stream, and to the hydraulic properties of material which comprises the bed of the stream. These properties are often only vaguely known. It may be possible to infer them (albeit with uncertainty) through history-matching of near-stream head measurements if these are available. However, caution must be exercised. If these measurements are not of high quality, or if a model's structure and/or parameterisation are such that history-matching forces parameters to adopt values that compensate for model inadequacies, then estimates of near-stream hydraulic properties may be corrupted. Predictions of contributing area may then be biased. In the present study, the quality of near-stream head measurements is possibly degraded by their failure to represent average conditions; the latter is required by the steady state assumption that underpins model calibration. In other modelling contexts, failure of a model to properly represent local aquifer layering, or perforation of this layering by long well screens, may also impact the integrity of these measurements, and of local hydraulic property inferences that are drawn from them.

Where near-stream data assimilation is dangerous, near-stream head measurements should be given low weights or omitted from the calibration dataset altogether. The study that is documented herein suggests that, under these circumstances, a single layer model may be just as effective in probabilistic contributing area analysis as a multi-layer model. This eases some of the burdens of model construction and of history-matching.

CONTENTS

1. INTRODUCTION

1.1 Background

This GMDSI report differs from previous GMDSI reports in a number of significant ways.

First, we provide all modelling files, together with instructions on what they mean, how to build them, and how to use them. So, if you wish to reproduce part or all of what we have done, then you can do so. The model itself was built using the Groundwater Vistas (GV) graphical user interface. However, calibration and uncertainty analysis were performed independently of GV. Eventually, at least some of the functionality that was used to produce outcomes that are discussed herein will be available through GV itself.

Second, this study is based on a remake of an old model. The original model was built by the U.S. Geological Survey (USGS) in 1992. It simulates conditions in a small aquifer in New Hampshire whose groundwaters were contaminated by volatile organic compounds. Modelling was undertaken in order to determine areas which contribute water to two public supply wells. Calibration was undertaken through manual adjustment of zone-based parameters. Rudimentary post calibration sensitivity analysis was used to explore model predictive uncertainty.

The modelling that is described in this report is based on data that was supplied in the original USGS report, supplemented by a modern-day, publicly available digital elevation model. Our intention in building the original model was not to challenge or revise the report's conclusions. It was simply to demonstrate how new technology can improve model-based processing of the same data. We confess openly that we did not build our new model with the same attention to detail with which the original modellers built theirs. Nor is our knowledge of the area anything like theirs. We hope that this does not detract from the outcomes of our modelling work. These are to demonstrate new ways to address old (and persistent) problems.

1.2 Preview

Modelling and model-based data analysis methodologies that are illustrated in this report include the following:

- use of pilot points to modify an existing spatial parameterisation scheme;
- use of pilot points to parameterize both spatial and polylinear features;
- highly-parameterized, regularized inversion conducted on a MODFLOW-USG model using PEST_HP;
- use of penalty functions in regularised inversion;
- evaluation of contributing area using the mod-PATH3DU particle tracking package;
- creation of a suite of complex, calibration-constrained parameter fields using the PESTPP-IES ensemble smoother;
- probabilistic contributing area analysis;
- use of the QGIS and PARAVIEW packages for display and analysis of model results.

We try to keep GMDSI worked example reports short and easy to read. So the above methodologies are not discussed in detail. We make liberal use of pictures to paint thousands of boring words. Nevertheless, we hope that this report provides some interesting insights into the above methodologies in particular, and to contributing area analysis in general.

1.3 Definitions

There is some confusion in terminologies that pertain to the subject matter of this report. We adopt definitions provided by Barlow et al (2018).

Capture zone: the three-dimensional volumetric portion of a groundwater flow field that discharges water to a well.

Contributing area (or *area contributing recharge*): 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.

1.4 Organisation of this Report

Chapter 2 of this report discusses the original model, and the issues that it was built to address. Chapter 3 provides construction details of our new model. Chapter 4 describes how this model was calibrated, and how posterior parameter uncertainty analysis yielded a suite of random but calibration-constrained parameter fields on which posterior uncertainty analysis was based. Chapter 5 demonstrates probabilistic contributing area analysis; it also introduces a simpler model that is deployed for these same analyses. Outcomes of the two analyses are compared. Finally, conclusions are drawn in Chapter 6.

Some readers will be aghast at our use of miles, feet and inches instead of kilometres and metres in this report. This seemingly retrograde step maintains uniformity with the original report, thereby facilitating comparisons between our work and that of the original modellers. Conversion factors are as follows.

1 ft = 0.3048 m

I mile = 5280 ft = 1.609 km

2. BACKGROUND

2.1 Introduction

This section summarizes information recorded in Harte and Mack (1992), the USGS report that forms the basis of our own modelling work. Our site description is brief. The original report, from which the interested reader can obtain further details can be downloaded from https://pubs.er.usgs.gov/publication/wri914177.

The town of Milford is situated in south-western New Hampshire. Water for public and industrial use is drawn from the Milford-Souhegan aquifer. This small aquifer is situated in a pre-Pleistocene valley; its area is about 3 square miles. The aquifer is comprised of up to 114 ft of unconsolidated glacial sediments, predominantly stratified sand and gravel with some till. It is overlain in places by Holocene alluvium. Deposits are generally finer in eastern parts of the aquifer than in its western parts.

Good hydraulic connections exist between groundwater and the Souhegan River and its tributaries which flow through the valley. In western parts of the study area the Souhegan River recharges the aquifer; in eastern parts of the study area groundwater flow is towards the river.

In 1983 elevated concentrations of chlorinated hydrocarbons were detected in samples of water gathered from two public water supply wells. These are referred to as the “Savage” and “Keyes” wells in this report and in the USGS report. Contaminants included tetrachloroethylene (PCE); 1,1,1-trichloroethane; 1,2 trans-dichloroethane; trichloroethane (TCE) and 1,1-dichloroethane. The wells were shut down. The U.S. Environmental Protection Agency (USEPA), the New Hampshire Department of Environmental Services and the New Hampshire Division of Public Health Services then conducted investigations into potential pollutant sources. A model was built to evaluate recharge areas that contribute water to the Savage and Keyes wells under normal operation of these wells.

2.2 The Original Model

The original model employed MODFLOW 88 (MacDonald and Harbaugh, 1988) as its simulator. Figure 2.1a shows its boundary while Figure 2.1b depicts its gridding. Extraction wells are shown in both of these figures. The public supply wells which are the focus of this study are shown in red; the Savage well is to the west while the Keyes well is to the east. Harte and Mack’s model had 5 layers, each of which was approximately 20 ft thick. The number of active cells in each layer decreased with depth, so that the fifth layer had active cells in only the deepest parts of the aquifer.

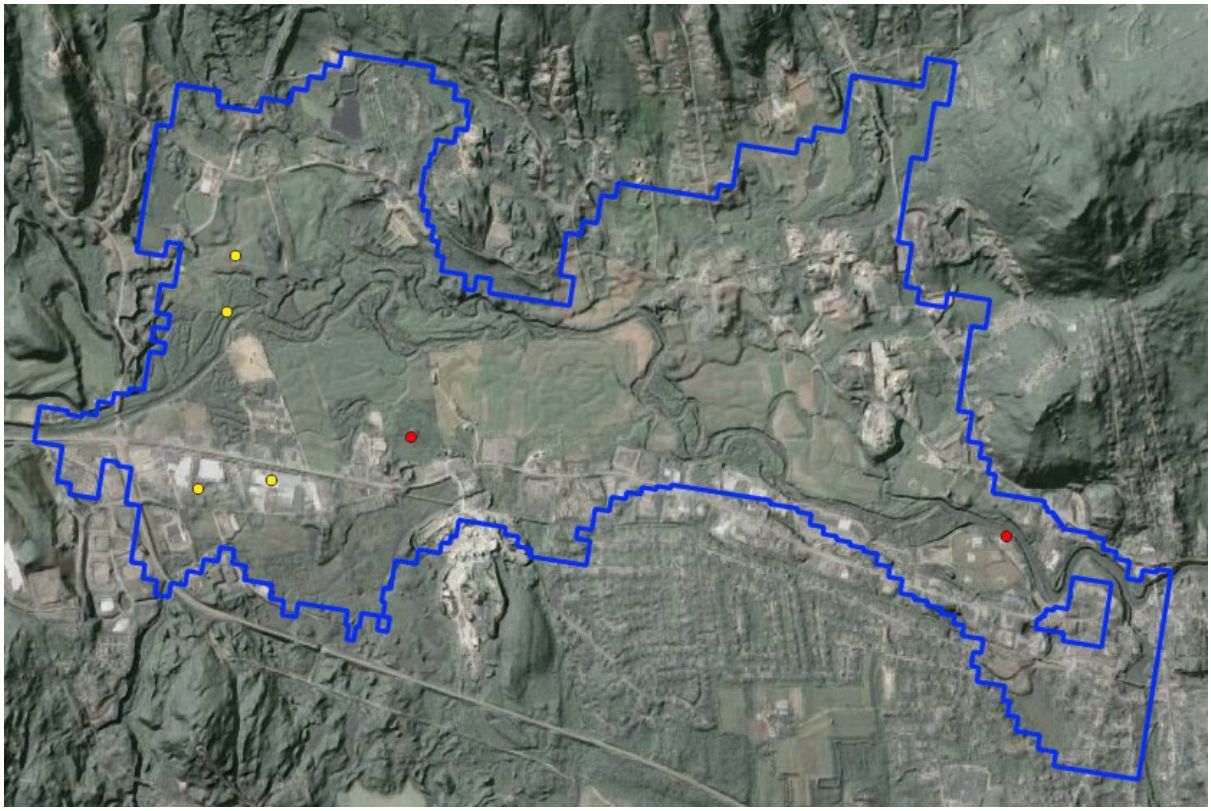

Figure 2.1a. Boundary of the original USGS model. Industrial extraction wells are shown in yellow. The Savage (west) and Keyes (east) wells are shown in red. The west-to-east extent of the model domain is about 3 miles.

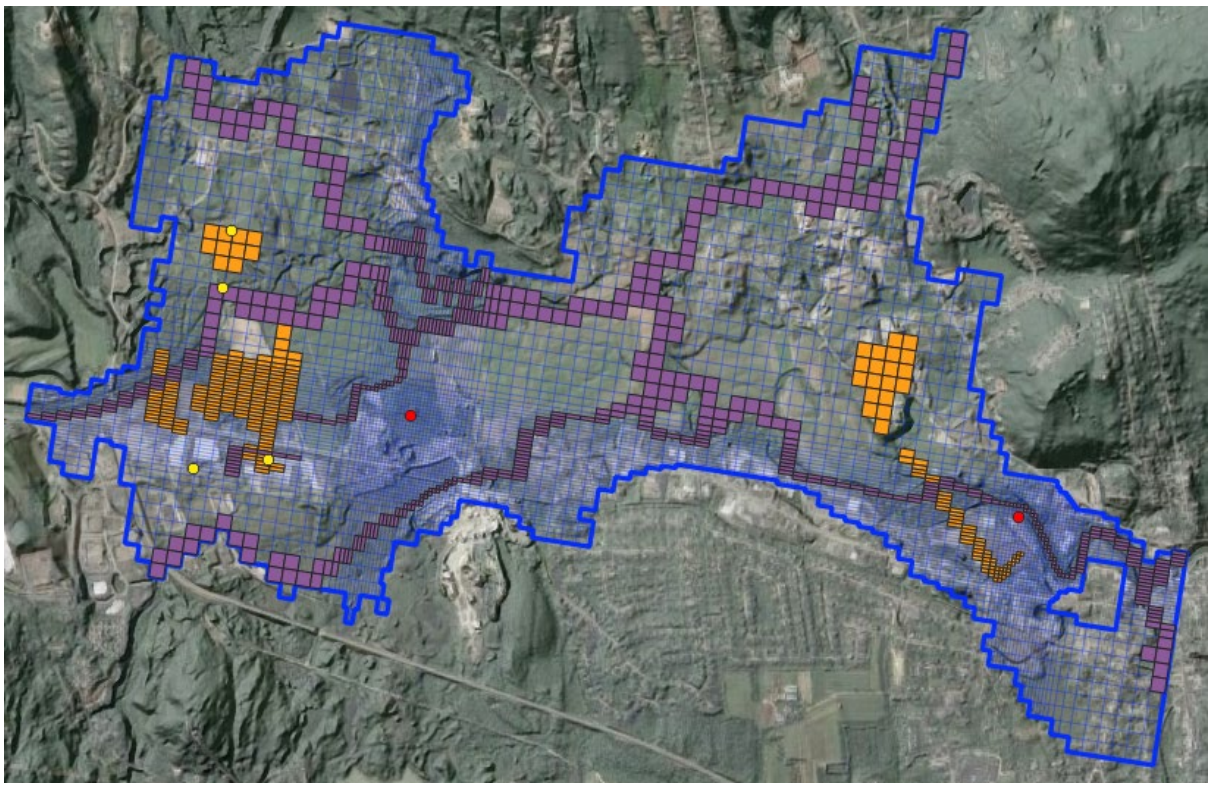

Figure 2.1b. Layer 1 of the original USGS model. The lengths of cell sides range from 50 ft to 200 ft. Cells in which a MODFLOW RIV (i.e., river) boundary condition are emplaced are shaded purple. Active cells in layer 5 are shown in orange.

The USGS model was calibrated and deployed under steady state conditions. A spatially uniform recharge rate of 3.75×10^{-3} ft/day (16.4 in/yr) is applied to all model cells. This is 37% of total average precipitation (44 in/yr); total recharge to the model domain through this mechanism is therefore 2.73×10^{5} ft^3/day. Lateral inflow (about 2.07×10^{5} ft^3/day) enters the model domain through its northern and southern boundaries; this water originates in groundwater recharge that takes place outside the model domain. The aquifer experiences further recharge through some reaches of the Souhegan river and its tributaries, particularly in the western part of the model domain. This matter is further discussed below.

The four industrial wells that are coloured yellow in Figure 2.1 extract 3.8×10^{5} ft^3/day of water from the Milford-Souhegan aquifer. No water is extracted from the red-coloured public supply wells during calibration of the model, as they had been shut down. However, extraction of water from these wells is simulated when calculating contributing areas.

2.3 Calibration

2.3.1 Calibration Dataset

The Milford model was calibrated under steady state conditions against a dataset comprised of head measurements in observation wells, and gains/losses in river/stream reaches. Figure 2.2 shows the locations of observations wells (in blue), and the locations of river gauging stations from which groundwater inflows and outflows to river/stream reaches were inferred.

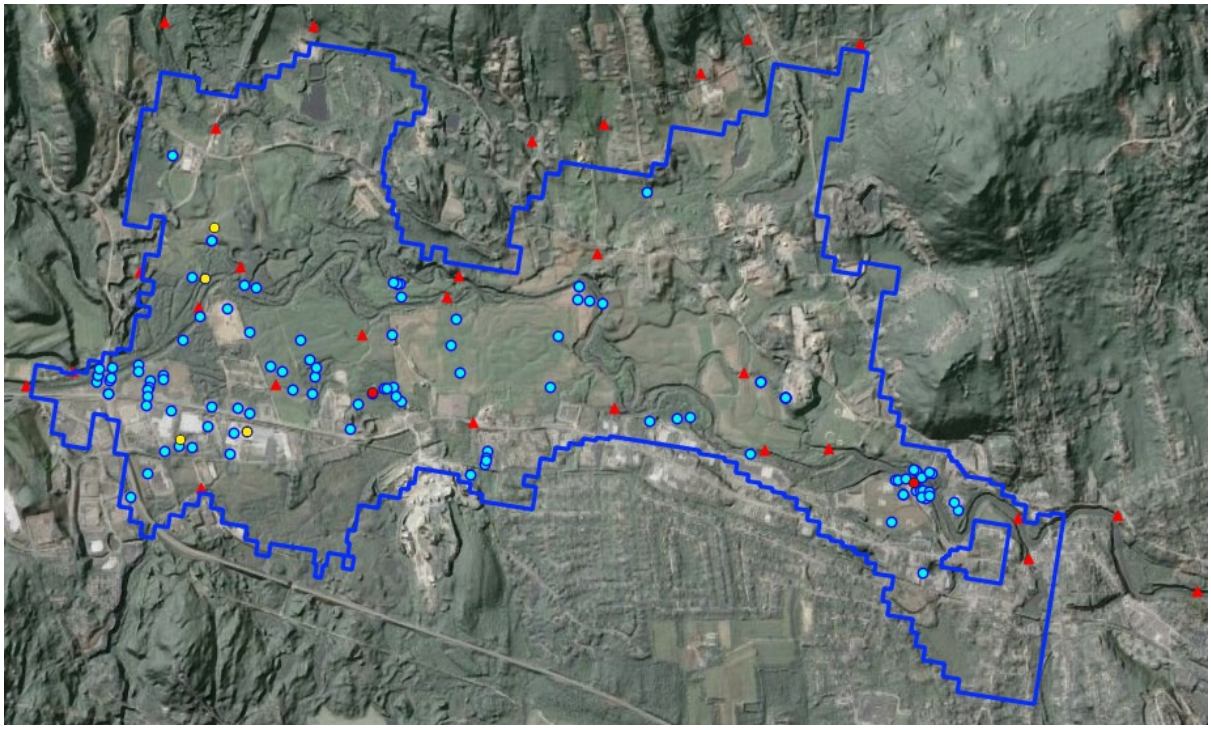

Figure 2.2. Observation wells (blue) and river/creek gauging stations (red triangles).

2.3.2 Parameterisation

Reasonable fits between model outputs and field measurements were attained by adjusting horizontal hydraulic conductivities in different parts of the aquifer, and vertical hydraulic conductivities of different segments of stream beds; the latter are used to calculate conductances that are attributed to MODFLOW RIV (i.e., river) boundary conditions. Harte and Mack report that they were prepared to adjust recharge, but that this did not prove

necessary. Meanwhile, vertical anisotropy of aquifer materials was maintained at a spatially uniform value of 10.

Spatial parameterisation of horizontal hydraulic conductivity was based on zones of piecewise constancy. The geometry of these zones was different between model layers. Zone design was based on inferences of hydraulic conductivity drawn from geological logging and pumping test analyses.

2.3.3 Calibration Outcomes

Parameter adjustment was manual. Misfits between measured heads and those calculated by the calibrated model were generally less than 3 ft. Misfits greater than this were attributed to contravention of the steady state assumption by pertinent measurements. However, some difficulties were encountered in matching groundwater exchanges with streams. Nevertheless, the authors declared these fits to be “satisfactory”. They noted that reaches that were measured as losing were generally modelled as losing, while those that were measured as gaining were generally modelled as gaining.

2.4 Sensitivity Analysis

Post calibration sensitivity analysis was implemented by subjecting the values of individual parameters to variations of 50%, and then monitoring the effects of these changes on model outputs used in the calibration process. The stated reason for this type of analysis was to understand the effect of these variations on the calibrated state of the model, and to verify that the manual calibration process would not benefit from further parameter alterations.

2.5 Model Deployment

Harte and Mack used their calibrated model to evaluate contributing areas to the Savage and Keys water supply wells; they used the MODPATH package (Pollock; 1988 and 1989) to undertake particle tracking. Uncertainties in estimates of contributing area were qualitatively evaluated by varying pumping rates, recharge, horizontal and vertical hydraulic conductivities, and streambed conductances and then recalculating contributing areas. No attempt was made to vary recharge and hydraulic property parameters in ways that maintained the model in a calibrated state.

2.6 Some Comments

In our opinion, Harte and Mack did an outstanding job, given the technology that was available at the time. Their report reveals a thorough analysis of existing data, and careful construction of a groundwater model that gives voice to these data.

The authors of the report that you are reading now devoted little time to analysis of datasets that are tabulated in the original report. Instead, we relied heavily on conclusions drawn by its authors. Model construction was also rather hurried. Our purpose in building a new model was to demonstrate that careful analyses such as were conducted by Harte and Mack can now be complemented and enhanced by modelling software and methodologies that were not available to these authors. Of particular relevance are the following:

- use of spatial parameterisation devices that allow representation of geological heterogeneity at a finer scale than that of zones of piecewise constancy;
- automatic adjustment of these and other parameters through regularized inversion;
- exploration of post-calibration predictive uncertainty using stochastic parameter fields that maintain the model in a calibrated state;

- probabilistic analysis of contributing areas.

The superior ability to analyse posterior uncertainty is of particular note. Opposing forces are at work here. On the one hand, advanced spatial parameterisation (pilot points instead of zones) supports better representation of spatial aquifer heterogeneity. This provides more opportunities for contaminants to find their way to points of water extraction. Uncertainties in contributing areas are thereby raised. On the other hand, any parameter set that is used to evaluate contributing area must maintain the model in a calibrated state. Uncertainties in evaluated contributing areas are thereby reduced.

3. MODEL RECONSTRUCTION

3.1 Model Structure

3.1.1 General

Construction of the model that is described herein was undertaken using Groundwater Vistas (GV). Display and processing of model features and outcomes was undertaken using the public domain QGIS package. Most of the figures that appear in this report are QGIS screenshots. Utility programs that facilitate data transfer between a model and QGIS are listed in Appendix A of this report.

3.1.2 Simulators

Flow of groundwater in the Milford-Souhegan aquifer is simulated using version 1.6.1 of MODFLOW-USG (Panday, 2020; Panday et al, 2013). This version of MODFLOW-USG can also simulate groundwater transport; however, its transport capabilities are not employed in the present study.

Particle-tracking is undertaken using version 2.1.5 of mod-PATH3DU (Muffels et al, 2020). Mod-PATH3DU uses the “Waterloo method” to track particles. This method can be readily employed in conjunction with an unstructured grid model whose cells are of arbitrary shape. Though not supported by the version of mod-PATH3DU that is used in the present study, the Waterloo method supports tracking of particles to extraction wells that are not necessarily at cell centres. See Rhamadhan (2015) for details.

3.1.3 Grid Design

Our model employs three layers; recall that Harte and Mack used five layers in their model. The bottom of our model domain follows that of the Harte and Mack model, while the top of the model domain is informed by a publicly available, lidar-generated digital elevation model.

Layer subdivision is proportional; at any point in the model domain the top layer of the model occupies half of the distance between the land surface and the base of the aquifer. The remaining depth is partitioned between the lower two layers. The lowest layer of the model has the same number of active cells as the highest layer. This layering strategy facilitates construction of the model; at the same time, the thick upper layer prevents drying of cells.

Despite our use of MODFLOW-USG, the model grid is structured. Nevertheless, the DISU convention of indexing cells by node rather than by layer, row and column is adopted. All cells are square, with dimensions of 50 ft × 50 ft. The model grid is shown in Figure 3.1.

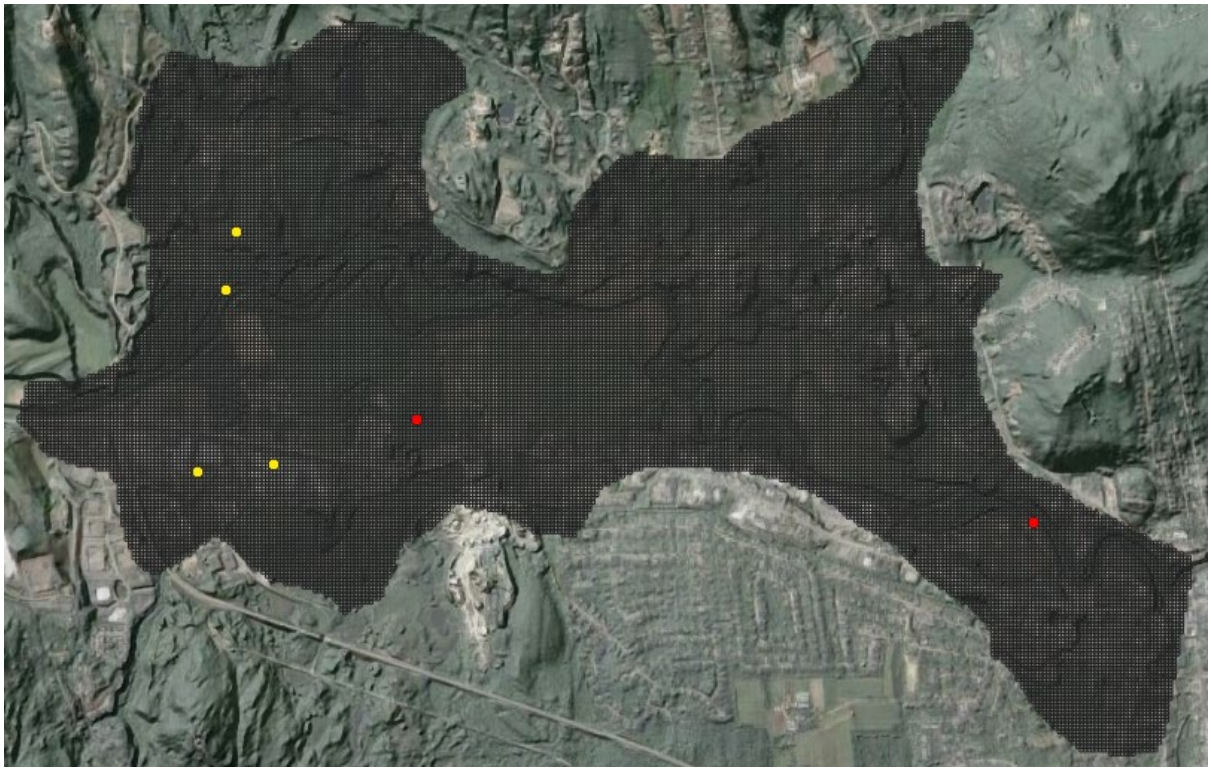

Figure 3.1. The model grid. Extraction wells (including the Savage and Keyes wells) are also shown.

3.2 Stresses and Boundary Conditions

3.2.1 Recharge

Percolation recharge is applied to the top layer of the model at a rate of 3.75×10^{-3} ft/day (16.4 in/yr), the same rate as that employed by Harte and Mack. However, as is described below, parameterisation of percolation recharge supports introduction of spatial heterogeneity to it during history-matching.

Lateral recharge is assigned to the northern and southern boundaries of the revisited model at a uniform rate of 1.1 ft^3/day/ft. Groundwater Vistas applies this rate to boundary cells in proportion to the length of the recharge front that intersects each such cell. Water is introduced to the model's third layer using the MODFLOW-USG WEL (i.e., well) package. Figure 3.2a depicts the northern and southern recharge fronts, while Figure 3.2b displays the cells to which recharge is applied.

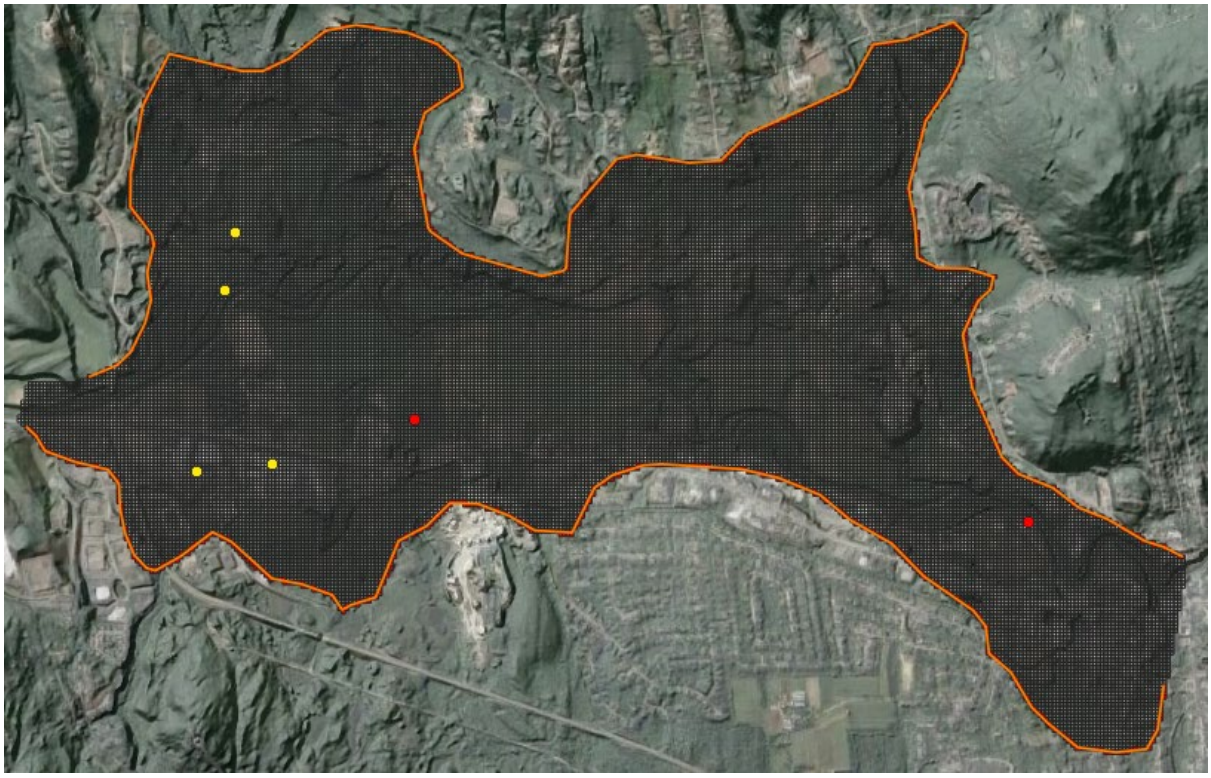

Figure 3.2a. Northern and southern recharge fronts.

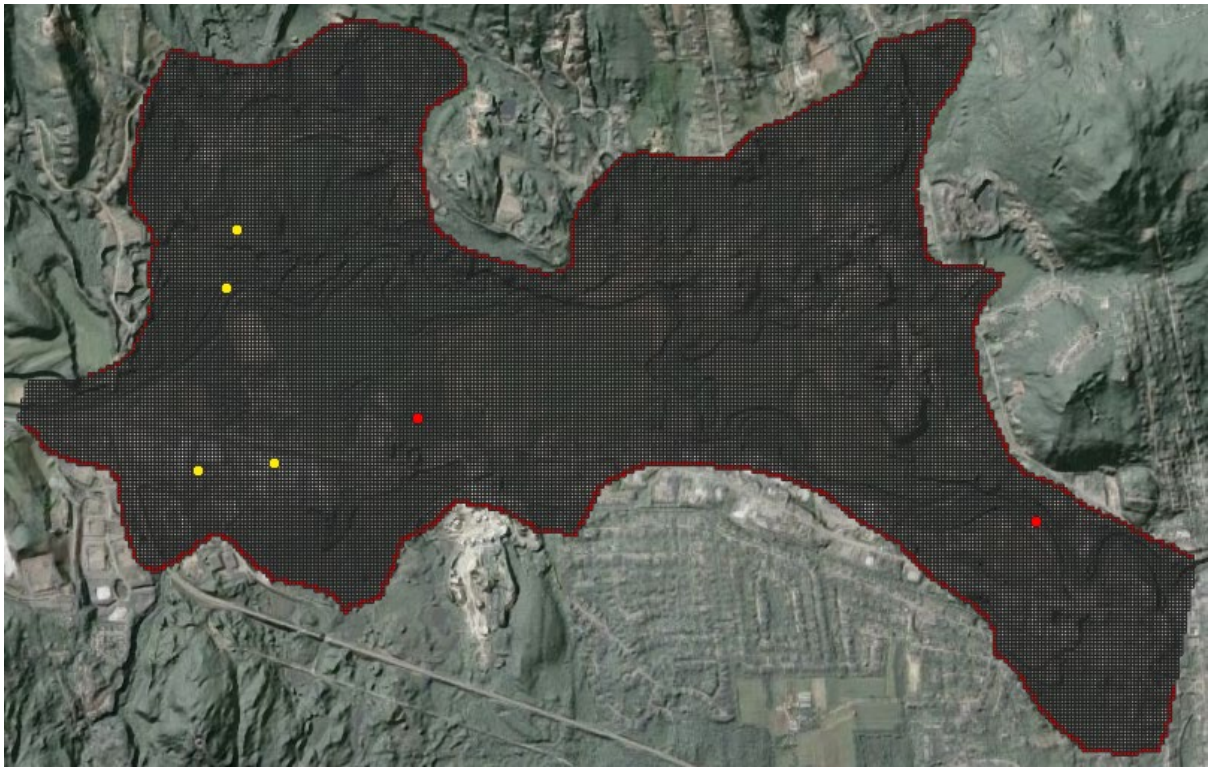

Figure 3.2b. Cells in model layer 3 to which lateral recharge is introduced using the MODFLOW-USG WEL package.

3.2.2 Rivers

MODFLOW-USG RIV (i.e., river) boundary conditions were introduced to the upper model layer to simulate the action of rivers and streams that transect the model domain. A shapefile

extracted from the New Hampshire Hydrography Dataset was supplemented with values of stream width, as well as thickness and vertical hydraulic conductivity of riverbed material; the latter were informed by the original USGS model. River/stream polylines were intersected with the model grid in Groundwater Vistas; a conductance was calculated for each RIV-affected cell based on the intersection length of the river/stream polyline with that cell. River/stream polylines are illustrated in Figure 3.3a together with river/stream gauging stations. Model cells to which a RIV boundary condition is applied are illustrated in Figure 3.3b; these are all in layer 1 of the model (i.e., the model's uppermost layer). Colouration of river cells in Figure 3.3b reflects partitioning of the stream network into reaches. A subset of these reaches was used in model history-matching.

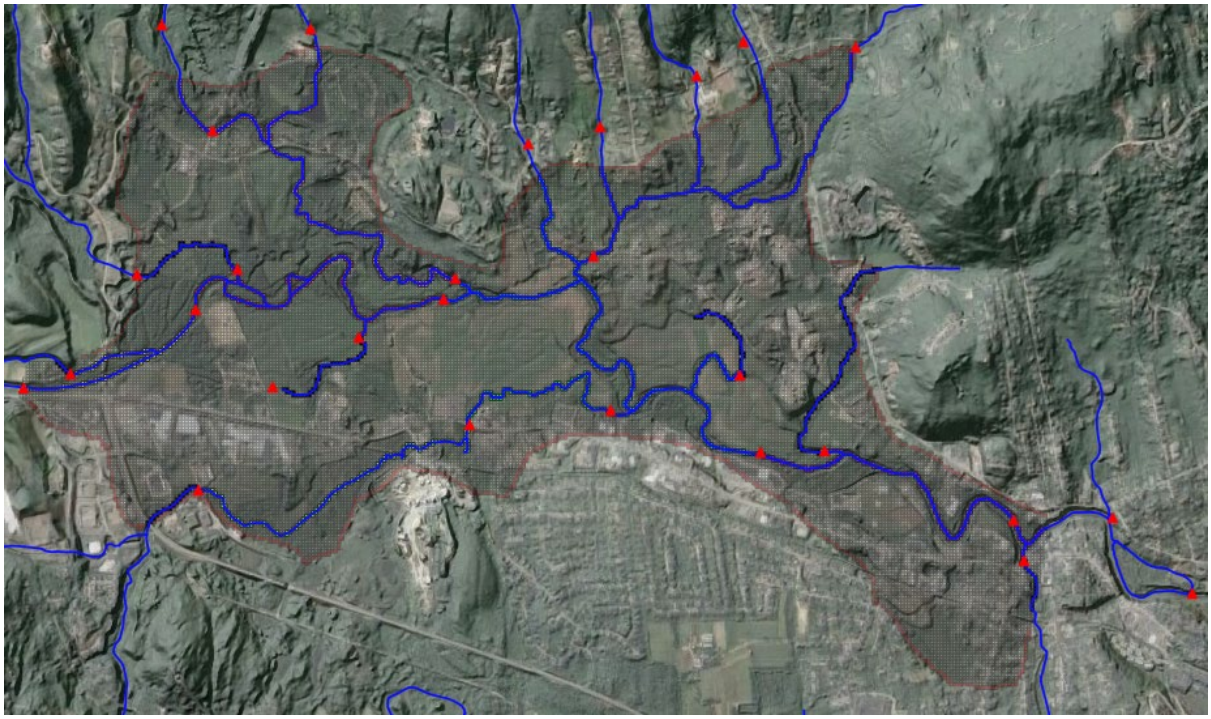

Figure 3.3a. River and stream features extracted from the New Hampshire Hydrography Dataset. Gauging stations are shown as red triangles.

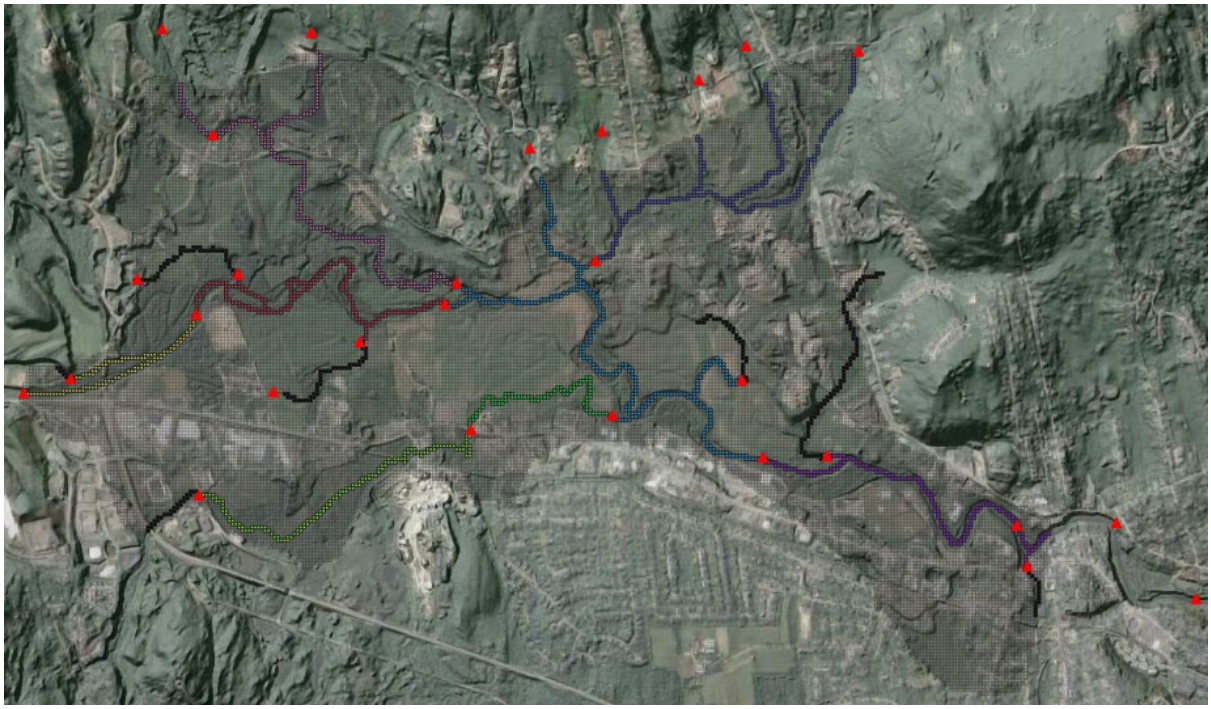

Figure 3.3b. Models cells to which RIV boundary conditions are assigned. All are in layer 1 of the model. Cells are coloured by reach; a subset of these reaches is used in history-matching.

River head elevations were calculated in Groundwater Vistas by subtracting 2 ft from the digital elevation model at pertinent cells. This displacement was subject to adjustment through model calibration.

3.2.3 Evapotranspiration and Surface Drainage

In places, the water table in the Milford-Souhegan aquifer is reasonably shallow. At these places, its elevation is probably controlled by the land surface rather than by the transmissivity of the aquifer. If mechanisms for surficial control of the water table are not introduced to a model such as this, then history-matching will introduce high transmissivities to those parts of the subsurface that underlie flat lowland topography as a surrogate for topographic control. These will distort evaluation of production well contributing areas.

Two boundary condition packages were introduced to the revisited Milford model to forestall this possibility, namely the EVT and DRN packages. Neither of these packages were employed in the original Milford model.

The EVT package withdraws water from a model if the water table is close to the land surface. For the Milford model, it is configured to remove water at a rate of 24.8 in/yr if the water table coincides with the land surface. This rate is reduced linearly to zero at a water table depth of 2 ft.

Use of the EVT package alone does not prevent the water table from rising to the surface at a small number of places within the model domain, these coinciding with breaks of slope and with gullies. In real-world groundwater systems, water escapes from the subsurface at places such as these; the attendant relief of pressure ensures that these points of water escape are limited in number and in area. In order to replicate this aspect of groundwater behaviour, every cell within the domain of the revisited Milford model is endowed with a DRN (i.e., drain) boundary condition. Drain elevation is set to surface elevation; drain conductance is assigned a high value in order to present minimal impediment to loss of water.

While locations of drain outflow are few in number, the history-matching process was instructed to penalize excessive loss of groundwater through this mechanism. This is discussed below.

3.2.4 Water Extraction

MODFLOW-USG supports “connected linear networks” (CLN’s). These can be configured as vertical, high-conductance water conduits that are connected to multiple model nodes. As such, they can simulate the action of extraction wells that penetrate multiple model layers.

The revisited Milford model simulates extraction of water from 6 wells. These are coloured yellow and red in preceding figures. The red wells are not operative during history-matching; these represent the Savage and Keyes wells which were shut down on detection of chlorinated hydrocarbons. The other wells operate under history-matching conditions, and under conditions in which contributing areas are calculated for the Savage and Keyes wells.

4. HISTORY-MATCHING

4.1 Strategy

As has already been discussed, the original Milford model was calibrated manually. This was followed by “sensitivity analysis” whose purpose was to explore, in a qualitative way, whether the calibration process could be trusted. Calculation of contributing areas to public water supply wells was then undertaken. These calculations were repeated a number of times. On each occasion, different types of parameters were varied as a block. This process comprised a rudimentary form of predictive uncertainty analysis.

We adopt a different approach in our revisitation of Milford modelling. Fundamental to this approach is acknowledgement of the fact that the Milford calibration dataset can be fit by many parameter fields; so this is what we do. Contributing areas are then calculated using all of these fields. The outcome of this process is not therefore:

- a single contributing area that is endowed with special status because it is calculated using the “calibrated” parameter field; together with
- a small number of other contributing areas that are calculated using significantly different parameter fields that do not allow the model to replicate system behaviour but are nevertheless labelled as being indicative of contributing area uncertainty.

Instead, we provide a map of contributing area probability. If the probability ascribed to a certain point within the model domain is 1.0, then that point is certainly within the area that contributes water to a public supply well. Probabilities can vary down to zero. The contributing area probability map is calculated using a suite of random parameter fields that all allow the model to replicate measured system behaviour. This prevents inflation of uncertainty arising from failure to respect constraints that are imposed on parameters by the necessity for the model to respect field measurements of system behaviour.

Before deriving a suite of history-match-constrained random parameter fields, the revisited Milford model is first “calibrated”. By “calibration” we mean pursuit of a parameter field that attains uniqueness through being the “simplest” way to fit the measurement dataset. It is important to point out that “simplest” does not mean “based on zones” – or any other simplistic parameterisation device for that matter. In fact the parameter field that is achieved through this process is continuous in space because it is based on hundreds of parameters rather than just a few. Simplicity is achieved through numerical regularisation.

In calibration of the revisited Milford model, regularisation constrains parameters to exhibit minimum departure from a preferred condition. This condition is that of broad similarity of the calibrated parameter field to that of the original Milford model. This strategy recognizes the attention that Harte and Mack paid to values of hydraulic conductivity inferred from pumping tests and grain size analyses. Meanwhile, simultaneous adjustment of a large number of parameters that define a continuous parameter field ensures that a good fit can be attainment with the calibration dataset with parameter patterns that are neither pre-defined nor artificially discontinuous.

Following model calibration, we seek a suite of parameter fields that fit the calibration dataset just as well as the “calibrated” parameter field, whilst also expressing heterogeneity that can occur in real-world aquifers. In those parts of the model domain where historical head measurements are spatially dense, differences between these parameter fields are not great. Conversely, in those parts of the model domain where historical head measurements are

sparse, these parameter fields can be significantly different from each other while still allowing the model to respect the measured behaviour of the system.

4.2 Measurements

4.2.1 Heads

Measurements of “steady-state” head are available at the wells depicted in Figure 4.1.

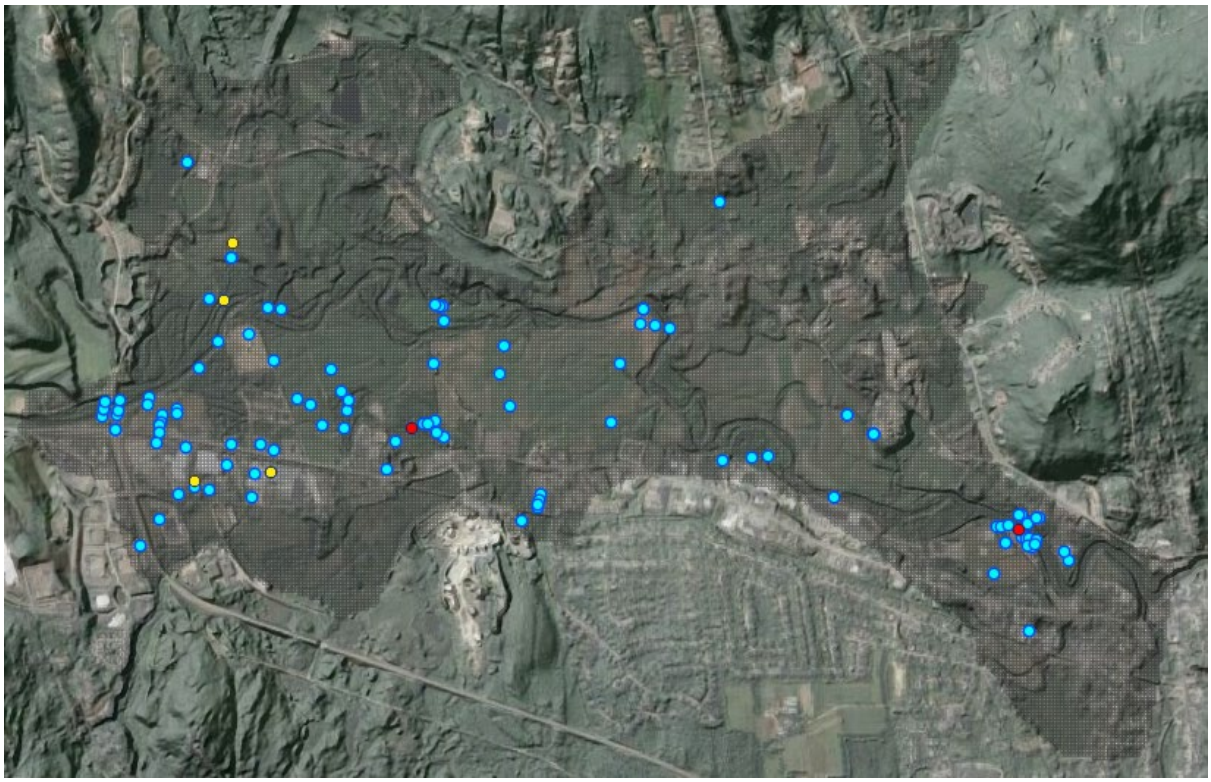

Figure 4.1. Wells at which measurements of steady-state head are available for calibration of the Milford model.

As Harte and Mack point out, the extent to which these measurements do, in fact, represent average conditions is open to question. The same applies to any “steady-state” dataset. Nevertheless, steady state calibration has its merits. The alternative is transient calibration. While a transient dataset is generally more populous than a steady-state dataset, history-matching against such a dataset cannot be guaranteed to provide better estimates of hydraulic properties to which contributing areas are most sensitive. (By “better” we do not mean “correct”, for solution of a nonunique inverse problem does not yield “correct”. We mean “unbiased”, and “with reduced uncertainty”.) The information content of a transient dataset may indeed be higher than that of a steady state dataset. However, this information must inform more parameters. Furthermore, these parameters pertain to complex processes such as recharge whose simulation rests on assumptions that are also prone to error.

4.2.2 Water Exchange with Rivers and Streams

Part of the Milford model calibration dataset is comprised of inferred values of average rates of water exchange between groundwater on the one hand, and different reaches of the stream network on the other hand. Harte and Mack remark that replication of these data by their model was not altogether satisfactory. In a number of attempts to calibrate the revisited Milford model, we encountered similar problems. Attempts to fit stream inflows and outflows that are tabulated in Table 4 of Harte and Mack’s report either proved fruitless or resulted in unrealistic

parameter fields. (One of the advantages of automated, highly-parameterized inversion, is its ability to support quick identification of model and/or data inadequacies. If a modeller uses too few parameters, or attempts to calibrate a model manually, he/she is always left wondering whether a revised parameterisation scheme, or a few more model runs, would overcome a data misfit problem. Through automated calibration of the revisited Milford model, it soon became apparent that history-matching satisfaction is unattainable.)

Rather than abandon stream inflow/outflow data altogether, we used a subset of it. Streamflow measurement sites are numbered according to the system used by Harte and Mack in Figure 4.2. Also shown in this figure are 5 stream reaches that are numbered according to our own scheme.

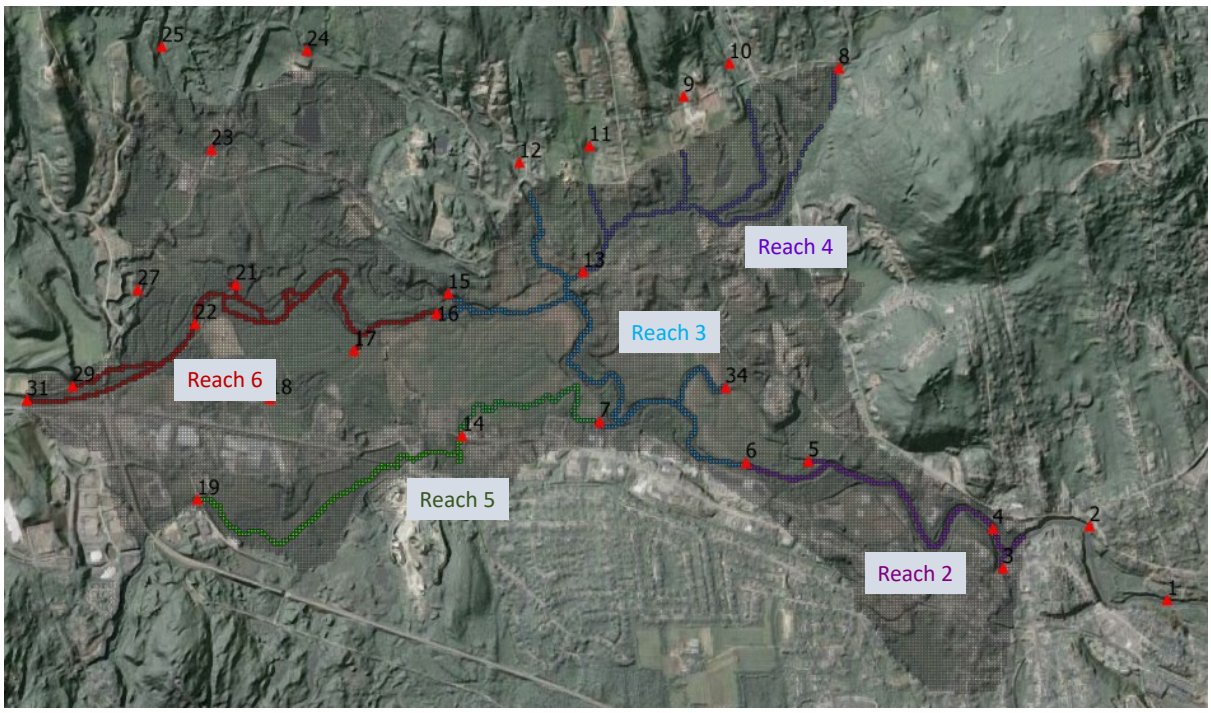

Figure 4.2. Streamflow measurement stations (red triangles and black labels), and stream reaches used in history matching of the revisited Milford model.

Stream inflows/outflows that feature in our history-matching dataset are listed in Table 4.1.

Stream reach	Exchange rate (ft^3/day)	Gain/loss to stream
2	258336	Gain
3	231552	Loss
4	3974	Gain
5	40608	Loss
6	18144	Gain

Table 4.1. Stream gains and losses employed in model history-matching.

Gains experienced by reach 2 deserve special attention. The streamflow measurement point that defines the lower end of this reach is gauge 1. (Measurements in gauge 2 are unreliable, gauge 3 is on a tributary stream, and no measurements are available from gauge 4.) This gauging station is beyond the downstream edge of the model domain. The rate of groundwater inflow into reach 2 that is listed in the top row of Table 4.1 is thus an upper limit on that which prevails within the model domain. It was treated this way in the history-matching process.

Thus, a misfit between model-calculated stream inflow and measured stream inflow is deemed to occur only if model-calculated inflow exceeds 258,336 ft^3/day. Model-calculated stream inflows that are less than this incur no model-to-measurement misfit penalty.

4.2.3 Pumping Rates

Harte and Mack provide steady-state pumping rates for the four production wells that are coloured yellow in preceding figures. For two of these wells, extraction rates are considerable, namely about 154000 ft^3/day (i.e., 80 gal/min).

MODFLOW-USG automatically reduces a well extraction rate if model-calculated drawdown at the point of extraction is excessive. In the revisited Milford model, each production well is simulated using a three-segment connected linear network (CLN); nodes of this network are connected to the three model layers to which well screens are open. MODFLOW-USG calculates heads in network segments comprising each extraction well. If the drawdown in the bottom segment approaches the bottom of the well, MODFLOW-USG automatically reduces the extraction rate in order to prevent the segment from going dry. This can occur if hydraulic conductivities ascribed to near-well model cells are too low.

While this functionality arrests model instability, model-simulated pumping rates can end up being less than actual pumping rates if this model behaviour is not discouraged as parameters are varied during model history-matching. Under-representation of extraction will result in underestimation of near-well hydraulic conductivities. Distortion of estimated aquifer properties can then distort key model predictions. To prevent this occurrence, well extraction rates are included in the measurement dataset. Differences between user-requested and simulated extraction rates thereby incur a history-matching penalty.

4.2.4 Drain Outflows

As is described above, extensive use of the MODFLOW-USG DRN boundary condition ensures that model-calculated groundwater levels do not “flood” surficial model cells. While this simulates mechanisms for groundwater pressure release that prevail in nature, loss of groundwater from the Milford-Souhegan aquifer at locations other than streams is likely to be small. Model replication of this can be ensured by including this observation as a history-matching constraint.

Figure 4.3 shows a patchwork of rectangular “drain reaches”. On each occasion that the revisited Milford model runs, the amount of water that emerges from each of these reaches is calculated. During history-matching, each such outflow is compared with an “observed value” of zero. This simple measure prevents the ubiquitous installation of MODFLOW-USG drains from having any more than a minor effect on groundwater flow.

Figure 4.3. Patchwork of drain "reaches". Observed DRN outflow through each reach is zero.

4.3 Parameters

4.3.1 General

Parameters are model inputs that are adjusted through history-matching. In general, they represent system hydraulic properties, but this does not always have to be the case. Where parameters represent spatially-variable properties such as hydraulic conductivity, it is best to adopt a parameterisation device that can be distributed liberally throughout a model domain. The history-matching process is therefore free to introduce heterogeneity to those parts of the model domain where this is required for attainment of a good fit between model outputs and field measurements, and to suppress the introduction of heterogeneity where it is not required.

Where history-matching responds to information contained in a calibration dataset by introducing hydraulic property heterogeneity to part of a model domain, it must do this in geologically meaningful ways. In many history-matching circumstances, geological reasonableness of history-match-emergent parameter fields can be assisted by associating one or more covariance matrices with these fields. Where history-matching seeks uniqueness through model calibration (normally by suppressing any heterogeneity that is not supported by the calibration dataset), covariance matrices are used in regularisation. As such, they impose penalties on patterns of emergent heterogeneity that are not in accordance with the statistical underpinnings of these matrices. (This generally occurs if these patterns are too "bumpy and lumpy".) Where history-matching seeks a multiplicity of random, history-match-constrained parameter fields, covariance matrices are used to specify prior parameter probability distributions.

Parameterisation devices used in the revisited Milford model are now described.

4.3.2 Hydraulic Conductivity

The revisited Milford model has three layers. All of these layers must be endowed with spatially distributed values of horizontal and vertical hydraulic conductivity. Ideally, these values should be:

- influenced by field-measured or field-inferred values of hydraulic conductivity where these are available;
- capable of adjustment in order for model outputs to match field measurements of system behaviour (i.e., the history-matching dataset); and
- constrained to adopt geologically reasonable patterns as they are adjusted.

Parameterisation of hydraulic conductivity within the revisited Milford model relies on a multi-tiered system of pilot points. Calculations required by their use are undertaken by the PLPROC model pre-processor supplied with PEST.

First, hydraulic conductivities are assigned to the *ppset1* family of pilot points whose locations are depicted in Figure 4.4.

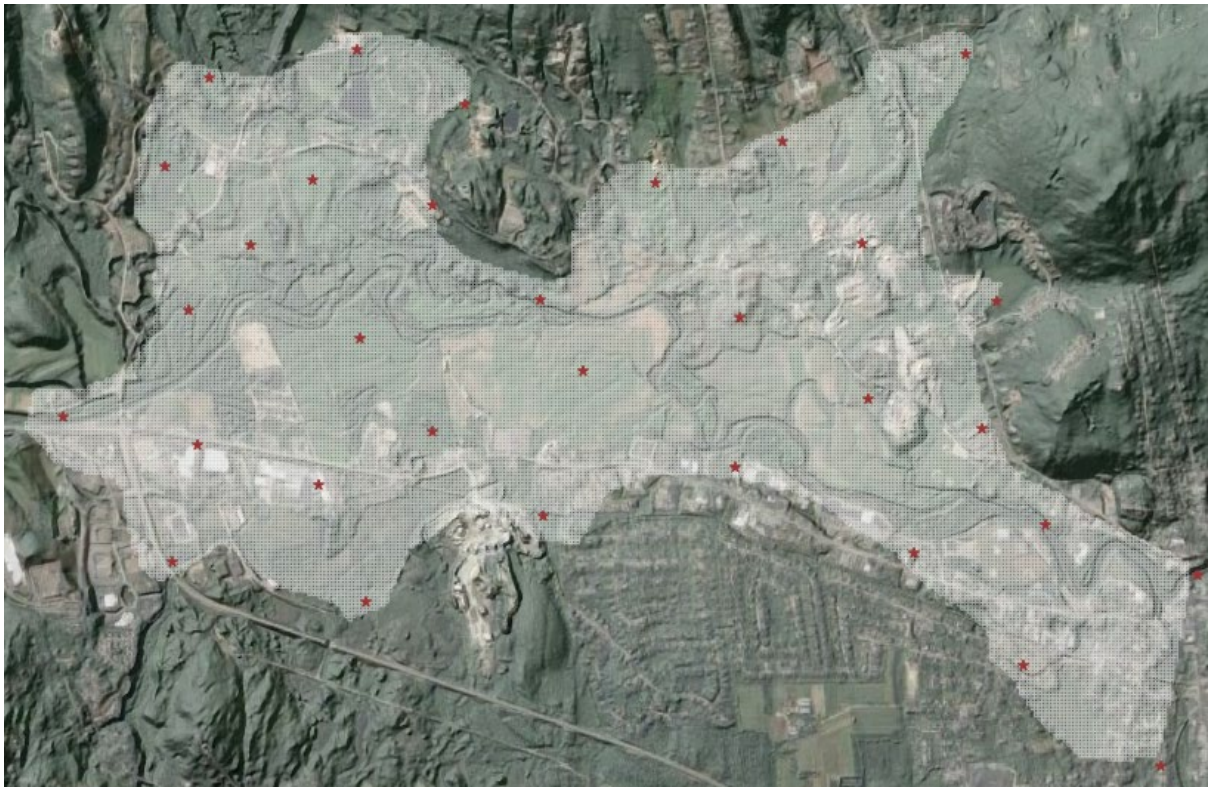

Figure 4.4. Locations of the *ppset1* set of pilot points.

As is apparent from Figure 4.4, the 32 pilot points that comprise *ppset1* are distributed relatively sparsely. Some of their locations coincide with places where hydraulic conductivity was measured or inferred through pumping test or grain size analysis. Initial/preferred hydraulic conductivity values awarded to these points are taken from the same locations in layer 3 of the original Milford model.

Spatial interpolation from *ppset1* pilot points to layer 2 of the revisited Milford model comprises the first step in hydraulic conductivity parameterisation of the entire model. Interpolation uses ordinary kriging based on a variogram whose range varies with local pilot point density. The model-grid-based hydraulic conductivity field that is obtained in this way is then subjected to multiplication by a parameter field that is interpolated from a second set of pilot points denoted

as *ppset2*. These pilot points are depicted in Figure 4.5, together with head observation wells. Initial/preferred values for multiplier parameters associated with *ppset2* pilot points are 1.0. Arrangement of *ppset2* pilot points is spatially more dense than that of *ppset1* pilot points, especially near observation wells. This allows them to represent emergent hydraulic property heterogeneity – both stochastically, and in response to information that is resident in observed heads.

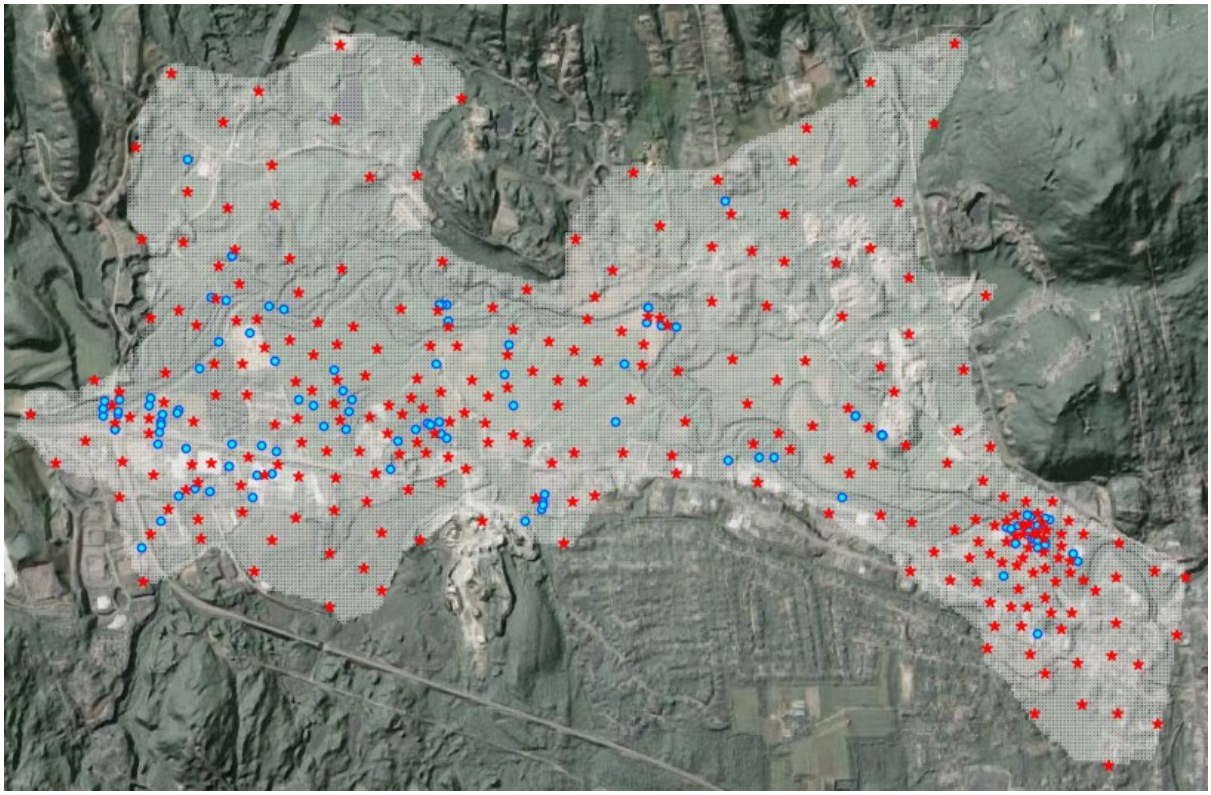

Figure 4.5. Red stars denote pilot points belonging to the *ppset2* family. Observation wells are shown as blue circles.

A third set of pilot points (*ppset3*) with intermediate spatial density is shown in Figure 4.6.

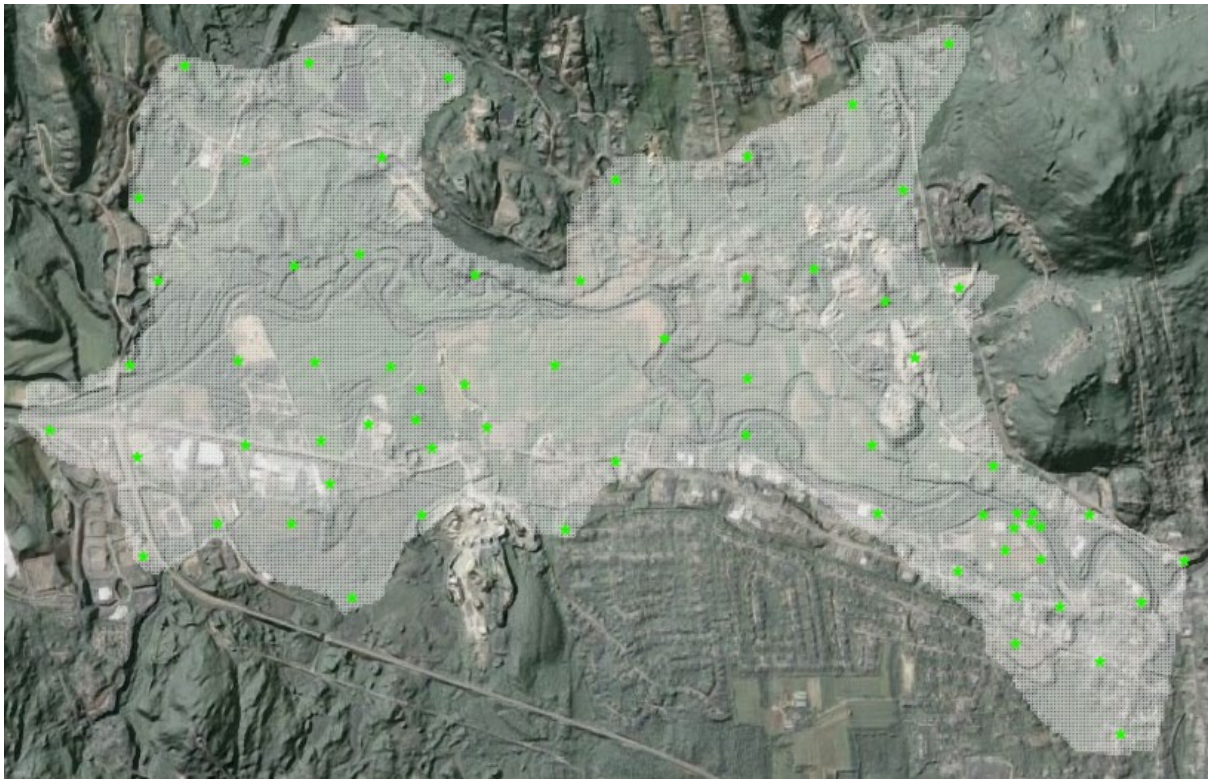

Figure 4.6. The *ppset3* family of pilot points.

The *ppset3* family of pilot points hosts multiple parameter sets. We denote the first of these sets as “K layer gradient” parameters. These parameters are endowed with initial/preferred values of 1.0. After spatial interpolation of these parameters to the model grid, the hydraulic conductivity assigned to any cell in layer 1 of the model domain is obtained by multiplying the horizontal hydraulic conductivity of the underlying layer 2 cell by the interpolated value. Meanwhile, the horizontal hydraulic conductivity assigned to a layer 3 cell is obtained by dividing the hydraulic conductivity of the immediately overlying layer 2 cell by the interpolated value of this parameter.

ppset3 pilot points also host vertical anisotropy parameters. After spatial interpolation to the model grid, the vertical hydraulic conductivity of any cell in any model layer is obtained by dividing the horizontal hydraulic conductivity of that cell by the interpolated value of vertical anisotropy. The initial/preferred value of all vertical anisotropy parameters is 10, this being recommended by Harte and Mack.

4.3.3 Recharge

Harte and Mack assigned a spatially uniform recharge rate of 3.75×10^{-3} ft/day (16.4 in/yr) to all surficial cells of their model. We do the same in our model. However, the history-matching process is provided with the opportunity to modify recharge; it is parameterized using the *ppset3* family of pilot points. The initial/preferred value of all pilot point recharge parameters is 3.75×10^{-3} ft/day.

4.3.4 Riverbed Conductance

Implementation of the RIV boundary condition requires that a conductance be assigned to every cell to which this boundary type is introduced – a total of 1783 cells in all. Conductances were initially calculated by Groundwater Vistas (GV). Each such calculation employed the width of the river, the intersection length of the river with the pertinent cell, and the thickness

and vertical hydraulic conductivity of riverbed sediments at the location of the cell. The latter were based on the original Milford model.

In the revisited Milford model, pilot point parameters are associated with RIV conductances so that the latter can be adjusted through history-matching; see Figure 4.7. These parameters act as multipliers on GV-calculated conductances. Multipliers are interpolated from pilot points to RIV-affected model cells using inverse squared distance interpolation (undertaken by PLPROC). Interpolation is conducted separately for different groups of pilot points (see Figure 4.7); grouping is based on stream segment. Group-based interpolation prevents pilot point parameters ascribed to one stream segment from informing properties of another stream segments. Initial/preferred multiplier values are all 1.0.

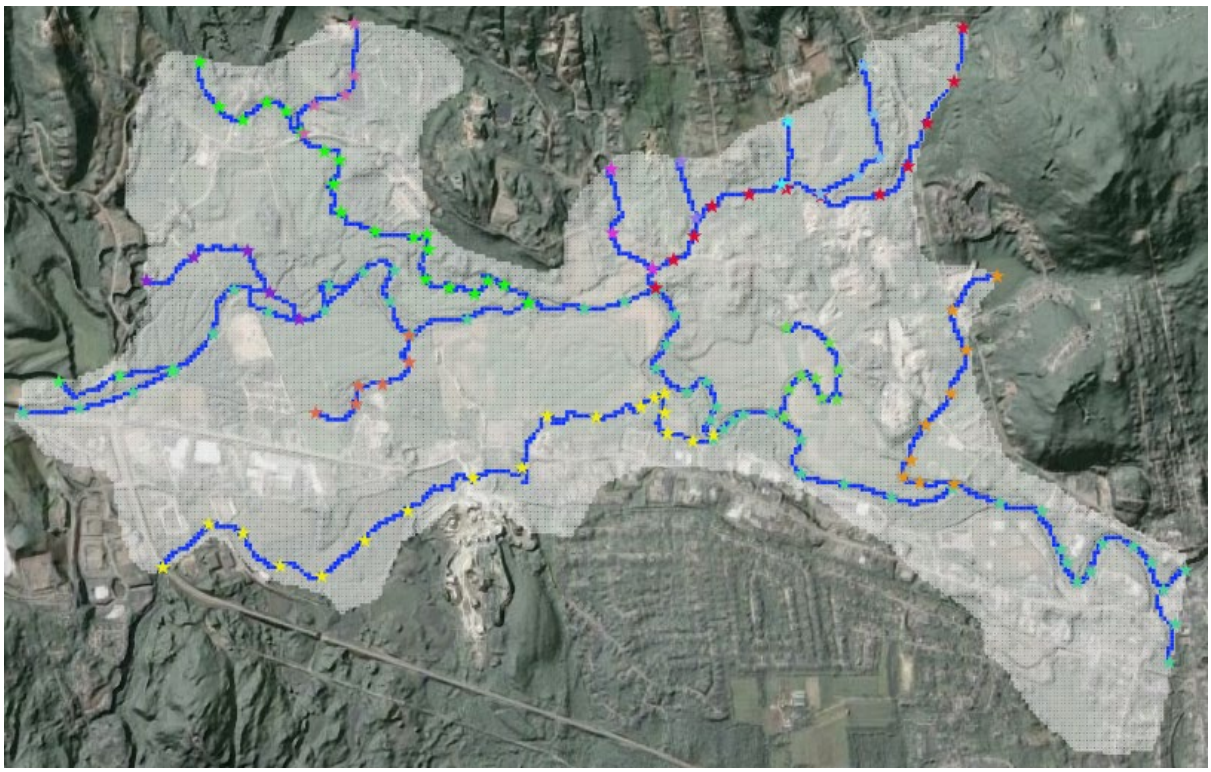

Figure 4.7. Pilot points associated with riverbed conductance coloured according to stream segment. This set of pilot points set is referred to as the *pp_riv* set.

4.3.5 River Elevation

The set of pilot points depicted in Figure 4.7 hosts another RIV-pertinent set of parameters. These are able to add or subtract altitude to the elevation of a stream. During history-matching of the revisited Milford model all of these parameters were linked together so that they effectively acted as a single parameter. The initial/preferred value of this parameter is 0.0. During initial history-matching runs, PEST_HP showed little interest in adjusting this parameter. So it was fixed at zero for later history-matching runs.

4.3.6 Lateral Recharge

Introduction of lateral inflow to the domain of the revisited Milford model is described in a previous section. These inflows are subject to history-matching adjustment using another set of pilot point multiplier parameters, the initial/preferred values of which are 1.0; see Figure 4.8. Interpolation from pilot points to lateral boundary cells (implemented using PLPROC) is through inverse squared distance. This *pp_wel* family of pilot points is subdivided into two groups, one of which pertains to the northern model boundary and the other of which pertains to the southern model boundary. As for riverbed conductance, group-specific interpolation

prevents multipliers assigned to one boundary from affecting inflows into the opposite boundary.

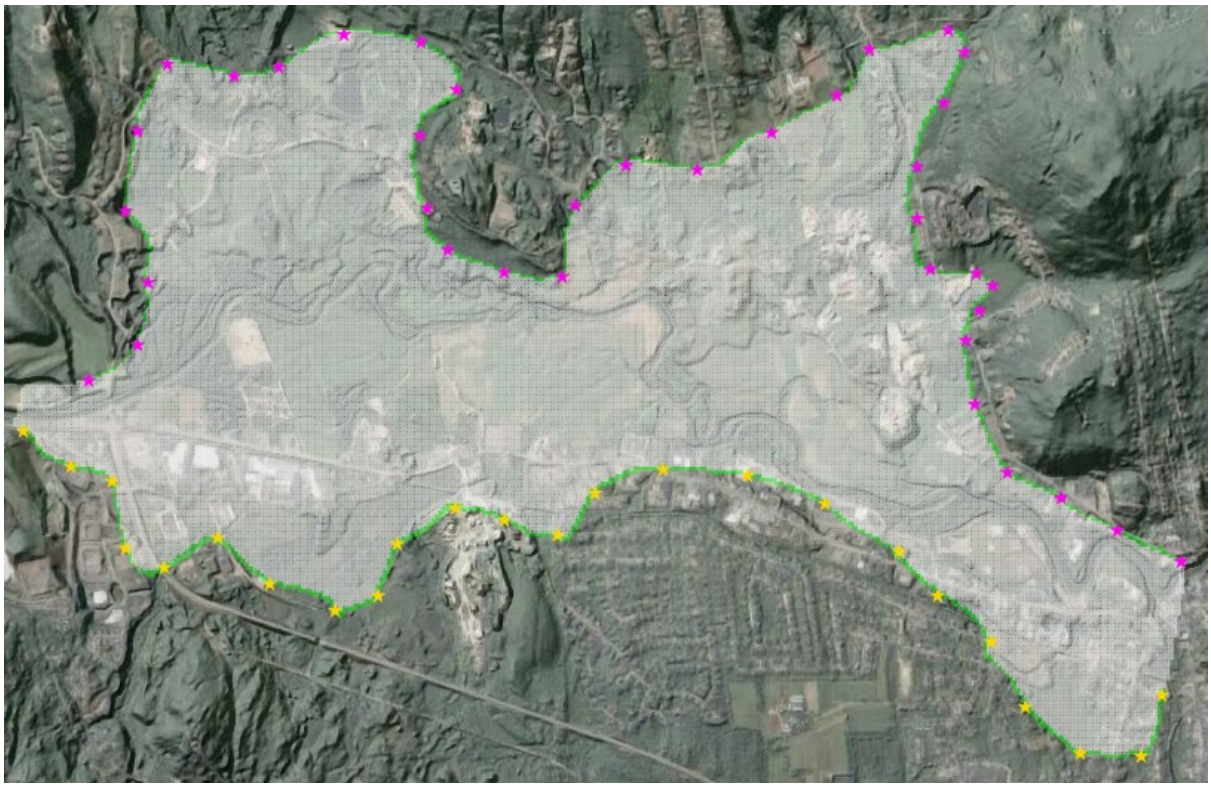

Figure 4.8. Pilot points (coloured according to group) used for parameterisation of lateral inflow. This set of pilot points set is referred to as the *pp_wel* set.

4.4 Procedure

4.4.1 General

History-matching of the revisited Milford model should not be a difficult procedure. After all, it is a steady state model with only three layers. It runs fast. It is endowed with a mere 865 parameters.

Nevertheless, some difficulties were encountered – in fitting groundwater exchanges with streams, and in fitting some measured heads. In both cases, it was possible to reduce model-to-measurement misfit to very low values. However, this came at a cost in parameter credibility. For reasons which are discussed in Chapter 5 of this report, parameter credibility can be an important issue when evaluating contributing areas to extraction wells.

History matching was implemented in two stages. First the model was calibrated using PEST_HP. As is discussed above, calibration seeks a single solution to an ill-posed inverse problem. Uniqueness is attained through seeking minimal departure of estimated parameter values from those which define a “preferred parameter state”. Initial/preferred values for parameters that are reported in previous paragraphs of this report define this state for the revisited Milford model. Meanwhile, to the extent that departures from this state are required in order for the model to fit the calibration dataset, a penalty is imposed if emergent patterns of heterogeneity become unrealistic. “Unrealistic” is measured by associating a covariance matrix with each set of parameters. The primary purpose of this matrix is to deter parameters from responding on a point-by-point basis to the demands of history-matching; instead, they

are encouraged to vary in neighbouring groups so that heterogeneity is “spread out” rather than “lumpy”.

Once a suitable fit had been attained with the calibration dataset using PEST_HP, the PESTPP-IES ensemble smoother was used to obtain a suite of parameter fields which, on the one hand, express potential hydraulic property heterogeneity while, on the other hand, fit the calibration dataset just as well as the parameter field that was previously attained using PEST_HP.

Some details are now provided.

4.4.2 Covariance Matrices and Parameter Limits

At the risk of being boring, we now say a few words about the covariance matrices that accompanied history-matching of the revisited Milford model.

As was discussed above, parameterisation of the revisited Milford model relies exclusively on pilot points. These host either point values of model inputs or point values of multipliers on model inputs. Covariance matrices are associated with groups of these pilot point parameters.

Covariance matrices express two desired characteristics of a collection of spatial parameters. Each diagonal element of a covariance matrix expresses the variance of an individual parameter. Variance is the square of standard deviation – a measure of how much a parameter can deviate from its mean value (i.e. its “expected value” in statistical parlance) while still retaining statistical credibility. Meanwhile, each off-diagonal element of a covariance matrix expresses the statistical correlation between a pair of individual parameters. This is a measure of the propensity for one member of the pair to have a higher-than-average value if another member of the pair has a higher (or lower) than average value. For spatial parameters such as pilot points, it is a measure of the tendency for heterogeneity to be dispersed rather than concentrated at a point. Generally, when constructing a covariance matrix, it is assumed that correlation between a pair of spatial parameters fades with increasing distance between them. Hence parameters that are close together are denoted as having higher correlation with each other than those that are far apart; furthermore, this correlation is positive.

Variances (and hence standard deviations) of revisited Milford model parameters were set in accordance with hydrogeological expectations. With the exception of river elevation, they pertain to the log (to base 10) of parameter values rather than directly to the values of parameters themselves. Parameter standard deviations are summarized in Table 4.2.

Parameter type	Pilot point set	Initial/preferred value	Standard deviation
Hydraulic conductivity of layer 2 (kh2)	*ppset1*	Taken from original Milford model	0.25
kh2 multipliers	*ppset2*	1.0	0.5
Layer K gradient	*ppset3*	1.0	0.023
Vertical hydraulic conductivity anisotropy	*ppset3*	10.0	0.023
RIV conductance multiplier	*pp_riv*	1.0	0.25
Precipitation recharge	*ppset3*	3.75×10^{-3} ft/day	0.023
Lateral recharge multiplier	*pp_wel*	1.0	0.023
Addition/subtraction from river elevation	*pp_riv*	0.0	0.25 ft

Table 4.2. Standard deviations of parameter types used in construction of covariance matrices.

For all parameters, correlation length is spatially variable. At any point in the model domain it is set to about four times the locally averaged distance between pilot points. Thus, in areas of high pilot point spatial density, parameter correlation spans a smaller distance than in areas of low pilot point spatial density. This does not reflect our expectations of patterns of geological heterogeneity. Rather, it reflects the fact that pilot points are distributed more densely in some places than in others because of greater data density in those places, or because of the need to ensure that predictive uncertainty is not understated in those places. At these locations, the model may be required to express parameter variability at a smaller spatial scale than at other locations.

As has already been mentioned, covariance matrices serve two purposes.

1. They impose constraints on parameter patterns that emerge when history-matching seeks a parameter field of minimized error variance through model calibration. ("Minimized error variance" can often be interpreted as "maximum simplicity".)
2. They are used to characterize the prior parameter probability distribution. Random realisations of parameters drawn from this distribution are morphed into random samples of the posterior parameter probability distribution using the PESTPP-IES ensemble smoother.

When used in the former role, packages such as PEST_HP are given licence to suppress all heterogeneity that is not directly supported by the data, notwithstanding expectations of spatial parameter variability that are embodied in covariance matrices. These matrices are used to ensure that minimalist patterns of heterogeneity that emerge through calibration are distributed rather than spotty. In contrast, where covariance matrices support posterior uncertainty analysis, full parameter variability as expressed by these matrices is retained in model parameter fields to the extent that this does not violate the fitting of field data; only heterogeneity that endangers model-to-measurement misfit is suppressed.

4.4.3 Parameter Limits

Both PEST_HP and PESTPP-IES can limit parameter variability through the imposition of lower and upper bounds on individual parameter values. However, the relationship between the lower/upper bound of a multiplier parameter, and the lower/upper bound of the model input that it multiplies is indeterminate. Hence the parameter pre-processing workflow must be such

that bounds can also be imposed directly on model inputs. Fortunately, this is easily accomplished using the PLPROC model pre-processor supplied with PEST. As well as implementing spatial interpolation from pilot points to a model grid, PLPROC can manipulate grid-based properties prior to recording them on model input files. This manipulation can include imposition of bounds.

4.4.4 Model Calibration

4.4.4.1 Background

Although use of an ensemble smoother does not require that a model first be calibrated, pre-calibration of a model has some advantages. It can also have some disadvantages. In particular, if model run times are long and parameters are many, model calibration can be a numerically intensive procedure as it requires calculation of derivatives of model outputs with respect to parameters. Where derivatives are calculated using finite differences, this requires one model run per parameter (at least); this process must be repeated during every iteration of the inversion process. In contrast, operation of an ensemble smoother requires only as many model runs per iteration as there are parameter realisations in an ensemble. This can be a fraction of the number of parameters that comprise a realisation.

Fortunately, the numerical burden incurred by calibration of the revisited Milford model is small despite the requirement for finite-difference derivatives calculation. This is an outcome of its rapid run time and its relatively small number of parameters.

Initial attempts to calibrate the revisited Milford model resulted in unrealistic parameter fields. These are an outcome of:

- measurement and conceptual errors associated with members of the calibration dataset; and
- model structural imperfections.

A significant source of conceptual error is the steady state assumption. This seems to apply particularly to stream gains and losses.

Meanwhile model imperfections abound, as they do for any numerical groundwater model. Imperfections of the revisited Milford model that may induce parameter value compensation during history-matching include the following:

- the use of only three model layers;
- the assumption that vertical heterogeneity is uniform in any vertical column of the model;
- approximations and errors associated with stream locations, elevations and widths;
- simplifications in representation of percolation-based recharge and lateral recharge.

Some of these problems could be rectified with greater attention to detail in reconstruction of the Milford model. Others are endemic to groundwater modelling itself; they were encountered by Harte and Mack when they calibrated the original Milford model.

It is tempting to suggest that problems such as the above can be overcome by increasing the complexity of a model, for example by adding more layers. It must be recalled however, that increases in model structural complexity require complementary increases in model parametric complexity. With increased parametric complexity comes the need for additional statistical constraints such as covariance matrices that maintain credibility of the more complex parameter field as it is adjusted to fit a calibration dataset. At the same time, increased model structural complexity increases the likelihood of model numerical misbehaviour.

Alternatively, a looser fit can be sought with a calibration dataset while maintaining the same level of model structural complexity (or even less structural complexity, as we explore in the next chapter). Estimation of a more reasonable parameter field generally follows this reducing the chances of predictive bias. With looser calibration constraints, parameters can express greater variability when exploring posterior uncertainty; the latter is therefore less likely to be understated. Where measurement data are compromised (by the steady state assumption in the present case), this may be a more practical course of action.

4.4.4.2 Calibration Strategy

To forestall the emergence of unrealistic recharge patterns during model calibration, recharge parameters were fixed at their initial values. This reduced the model's ability to match stream gains and losses. However, it was considered to be a reasonable course of action for the following reasons.

1. Measurements of "steady-state" stream gains and losses appear to be somewhat unreliable.
2. Model-calculated contributing areas to extraction wells are sensitive to local recharge. This is because the limits of a contributing area are set by the recharge that is needed to support extraction. As for any other model prediction, it is better that uncertainties associated with predictions of contributing area be over-estimated because of loose imposition of calibration constraints, than biased because of tight imposition of possibly inappropriate constraints.

As a further defence against calibration-induced bias, the PEST_HP inversion process was halted after only three iterations. At this stage, the fit between model outcomes and field measurements was as good as that attained through calibration of the original Milford model. With the exception of some horizontal hydraulic conductivities, parameters were endowed with reasonable values.

4.4.4.3 Calibration Outcomes

Some head calibration statistics are listed in Table 4.3. Measured and modelled stream gains and losses are listed in Table 4.4.

Statistic	**Value**
Average heads residual	-0.038 ft
Standard deviation of heads residuals	2.45 ft
Average absolute heads residual	1.57 ft

Table 4.3. Some head residual statistics.

Stream reach	**Measured gain (ft^3/day)**	**Modelled gain (ft^3/day)**
2	258336	296137.7
3	-231552	-190894.0
4	3974	10520.66
5	-40608	-27716.62
6	18144	-75196.02

Table 4.4. Stream gains and losses.

Model misfit of measured heads was dominated by two wells. These measurements are suspicious. They were removed from the history-matching dataset prior to PESTPP-IES-based posterior uncertainty analysis.

Table 4.4 demonstrates that model replication of stream gains and losses is far from exceptional. Some undocumented PEST_HP runs attained far better fits than those that are listed in this table. However, the parameter fields that achieved these fits strained hydrogeological credibility. Just as was done by Harte and Mack, we therefore accepted diminished fits with field-inferred stream gains and losses.

Figure 4.9 shows estimated hydraulic conductivities in layer 1. As stated above, some of these are too high to be credible. Limits were imposed in subsequent history-matching.

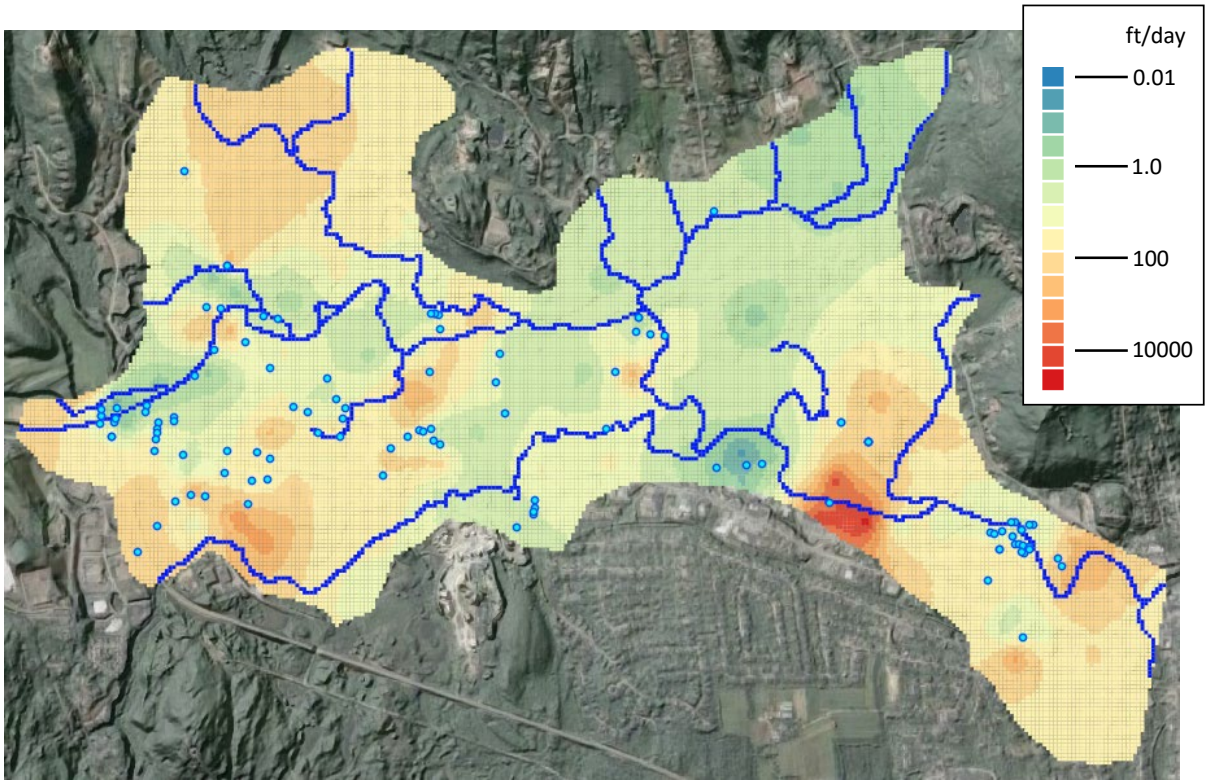

Figure 4.9 Estimated hydraulic conductivities in layer 1 together with locations of head observations.

Figure 4.10a depicts head residuals in the western part of the model domain, while Figure 4.10b shows head residuals in the eastern part of the model domain.

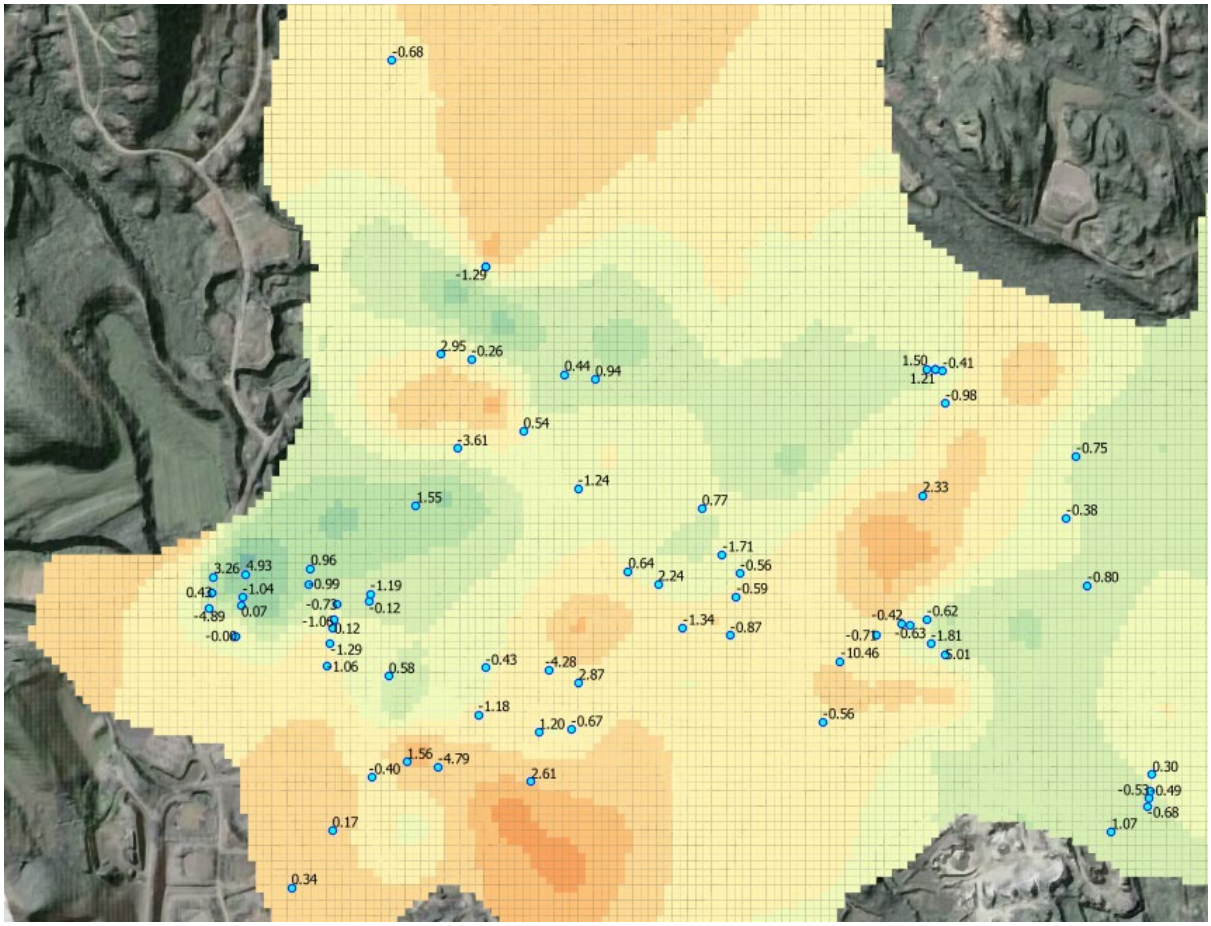

Figure 4.10a. Head residuals (in feet) in the western part of the model domain. Note that the head corresponding to the residual of -10.46 ft was withdrawn from further history-matching.

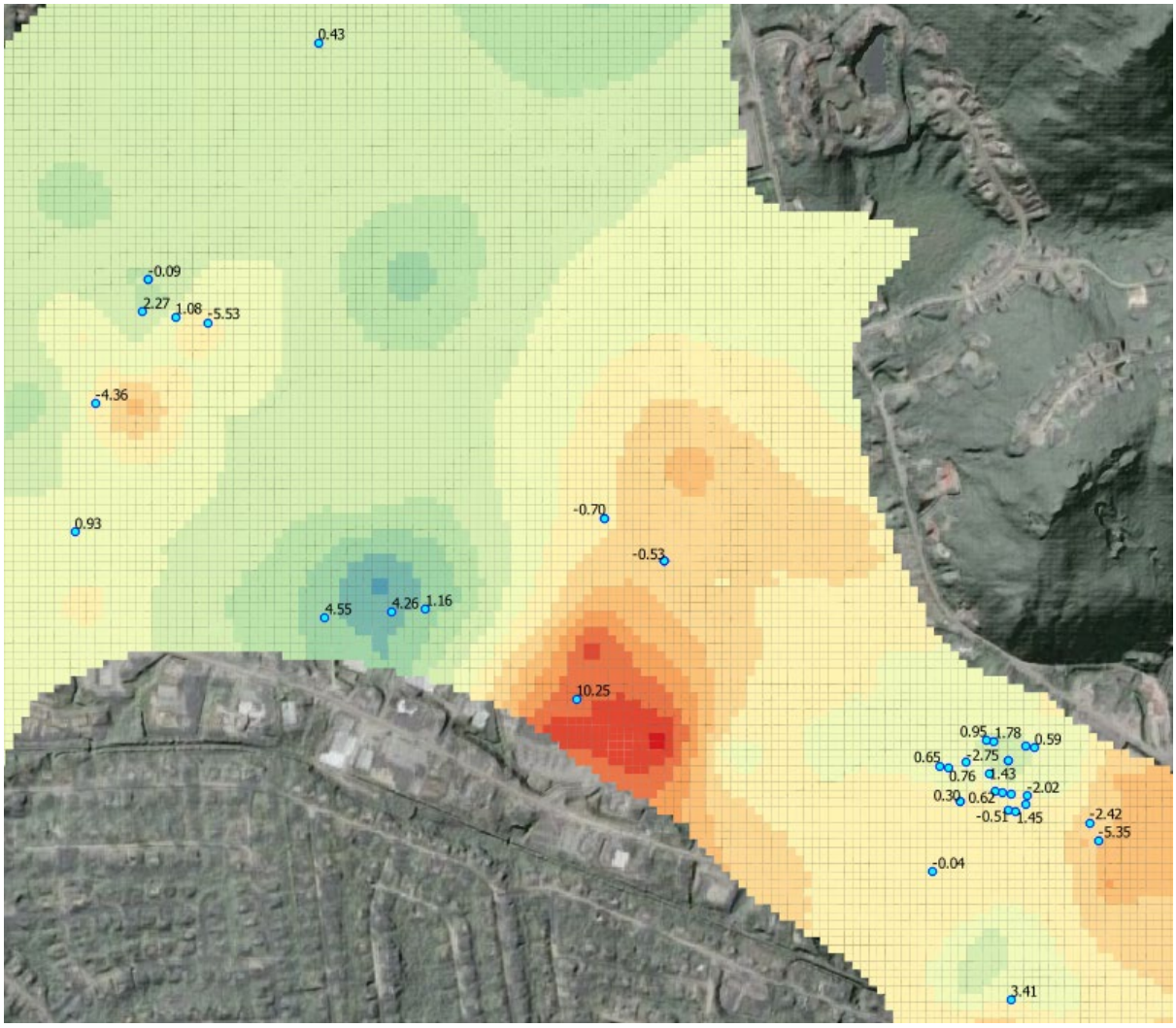

Figure 4.10b. Head residuals (in feet) in the eastern part of the model domain. (Note that the head pertaining to the residual of 10.25 ft was withdrawn from further history-matching.)

Model-calculated heads in layer 1 are shown in Figure 4.11.

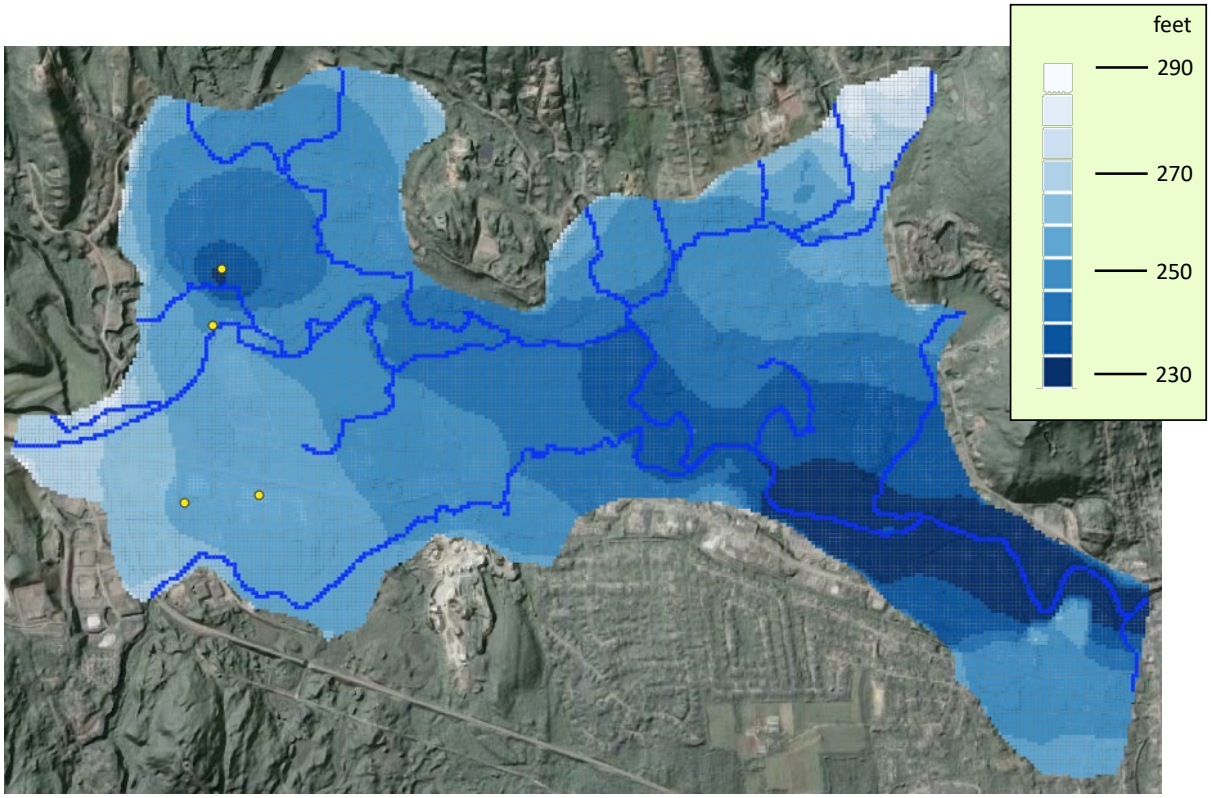

Figure 4.11. Model-calculated heads in layer 1. Extraction wells are shown in yellow; river cells are shown in dark blue.

4.4.4.4 Discussion

We pause here to discuss what calibration of the revisited Milford model achieved.

Calibration of the Milford model did not provide us with a parameter field that was used for contributing area analysis.

Actually, this is not quite true. Contributing areas for the Savage and Keyes wells were, in fact, evaluated using the parameter field depicted above. This was done in order to satisfy ourselves that they make sense. This puts the calibration process into perspective. In the present study (and in many other studies), its main purpose should be to ensure compatibility of a model with field measurements of what is being modelled. Once this has been done, the real analysis can begin.

As already discussed, calibration seeks the simplest way to fit a field dataset. The pursuit of this fit exposes model and/or data inadequacies. If parameters must adopt roles that compensate for model and/or data inadequacies, this becomes readily apparent when model-to-measurement fit is pursued under constraints of minimized emergent heterogeneity. If problems are encountered, then history-matching compromises that mitigate these problems can be sought. These compromises can then be retained during the ensuing stage of history-matching in which posterior parameter and predictive uncertainty are explored.

4.4.5 Posterior Uncertainty Analysis

4.4.5.1 Background

Generation of a suite of history-match-constrained random parameter fields was undertaken using the PESTPP-IES ensemble smoother; see White et al (2018) and Chen and Oliver (2013) for algorithmic details and supporting theory.

Ideally, a modeller supplies PESTPP-IES with an ensemble of parameter realisations which are sampled from the prior parameter probability distribution. As such, these samples are centred on prior parameter mean values. Meanwhile their variability about the mean is governed by covariance matrices that pertain to different parameter subgroups. Over a number of successive iterations, these parameter fields are morphed into samples of the posterior parameter probability distribution as they are progressively constrained by the necessity for model outputs to match field measurements.

For the Milford model (and, in the authors' experience, for many other models) this process works better if samples of the "prior" parameter probability distribution are centred on the parameter field that was attained during a previous calibration exercise. Convergence to the posterior is then rapid, while better fits are attained with the measurement dataset.

4.4.5.2 Posterior Sampling Strategy

Based on lessons learned from model calibration, the following strategy was adopted when using PESTPP-IES for history-matching.

1. Parameters pertaining to precipitation-sourced recharge were decreed to be adjustable. These parameters are hosted by the *ppset3* set of pilot points. Recall that they were held fixed during model calibration. Lower and upper bounds of 7 in/year and 26 in/year were imposed on these parameters.
2. Heads in two wells that were characterized by unusually large residuals during model calibration were removed from the history-matching dataset.
3. An upper limit of 450 ft/day was imposed on hydraulic conductivity.
4. The measurement dataset was supplemented with realisations of "measurement noise" that reflect model-to-measurement misfit attained during model calibration.

As is described in PESTPP-IES documentation, realisations of "measurement noise" can be added to members of a history-matching dataset prior to ensemble adjustment. The same number of measurement noise realisations are employed as there are parameter realisations which comprise an ensemble. This strategy ensures that parameter uncertainty not only reflects information gaps in the history-matching dataset; it also reflects potential errors in field measurements. In the present case these "errors" include their representativeness of steady state conditions.

The PWTADJ2 utility supplied with PEST allows a modeller to determine a set of measurement weights that are informed by calibration outcomes. Ideally, PWTADJ2-calculated weights approximate the inverse of the standard deviation of whatever source of error is responsible for model-to-measurement misfit. These weights can then be used to generate realisations of observation noise for the use of PESTPP-IES.

The parameter ensemble adjusted by PESTPP-IES was comprised of 300 parameter sets (i.e. 300 realisations). An important consideration in choosing the number of realisations is that this number should exceed the dimensionality of the so-called "solution space" of the history-matching inverse problem. The observation dataset is comprised of 112 head measurements, as well as 6 stream gain/loss measurements for which a relatively loose fit is sought. The dimensionality of the solution space is therefore much smaller than the number of parameter realisations. (The dimensionality of the solution space cannot exceed the number of observations comprising a calibration dataset.)

4.4.5.3 Outcomes

With recharge parameters released for adjustment, PESTPP-IES was able to achieve a better fit with field measurements than that which was attained by PEST_HP; this required about 6

iterations. However, these tight-fitting parameter fields were not used for calculating contributing areas. Instead, parameter fields from the third PESTPP-IES iteration were employed for this purpose. At this iteration, model-to-measurement fits for 275 out of the 300 parameter fields comprising the ensemble are commensurate with those attained by PEST_HP. Furthermore, they are all hydrogeologically reasonable. These 275 parameter fields were used for calculation of contributing areas.

To avoid tedium, an extensive report of PESTPP-IES outcomes is not presented herein. However, some parameter fields are presented together with calculated contributing areas in the next section.

5. CONTRIBUTING AREAS

5.1 General

Contributing areas to each of the Savage and Keyes wells were computing using 275 samples of the posterior parameter probability distribution. Probability of contribution was then evaluated using all of these contributing areas. To calculate contributing areas, the revisited Milford model is run under steady state conditions. As was done by Harte and Mack, extraction rates of 27907 ft^3/day (145 gal/min) and 19267 ft^3/day (100 gal/min) are ascribed to the Savage and Keyes wells respectively.

Contributing areas are calculated using mod-PATH3DU. To calculate a contributing area for a particular parameter realisation, a particle is introduced to the centre top of every cell in layer 1 of the model. A cell is part of an extraction well's contributing area if the particle that is introduced to it terminates at the well. After repeating this process for all realisations, a cell's contributing area probability can be evaluated by dividing the number of times that the cell is part of a contributing area by the total number of parameter realisations.

Some outcomes of this process are now shown. Utility programs which postprocess mod-PATH3DU output files in order to assign cells to the contributing area of a particular well, evaluate contributing area probabilities, and facilitate production of the plots that are shown below are listed in Appendix A of this report. All of these programs are available through the PEST Groundwater Utility suite.

5.2 Outcomes for Individual Realisations

Appendix B provides maps of contributing areas to the Savage and Keyes wells for 6 posterior parameter realisations. For each realisation, these contributing areas are plotted twice. In the first of each of these plots the contributing area is superimposed on a map of layer 1 hydraulic conductivity. In the second plot, the contributing area is superimposed on a map of recharge. The contributing area itself is coloured according to the travel time of water from the surface to the well. A spatially uniform porosity of 0.05 is used to calculate travel time.

It is evident from these figures that patterns of hydraulic conductivity and recharge bear some similarities between realisations. However, these patterns are different enough for contributing areas to be different between them. Of course, differences in vertical hydraulic conductivity and in riverbed conductance also promote differences in calculated contributing area. This is particularly the case for the Keyes well (i.e., the more eastern of the two public supply wells). While less water is extracted from this well than from the Savage well, its contributing area is smaller than that of the Savage well by an amount that is smaller than their extraction ratios. This implies lower precipitation-based recharge in the vicinity of this well, a higher rate of water withdrawal from the Souhegan River, or both. The figures in Appendix B suggest that recharge is similar in the vicinity of the Savage and Keyes wells; it follows that river extraction is important for the Keyes well.

5.3 Contribution Probability

In Figure 5.1 contributing areas calculated for all parameter fields are combined into a single probability map. Probabilities range from 0.0 to 1.0, with red indicating 1.0. A contributing area probability of 1.0 signifies that a particle placed at that point in the model domain is captured by an extraction well for all parameter realisations.

Any cell which contributes to well extraction for at least one parameter realisation receives a colour in Figure 5.1. Those which never contribute do not.

In Figure 5.1, cells containing RIV boundaries are coloured dark blue (these are in model layer 1) while those receiving lateral recharge (and hence contain WEL boundaries) are coloured green (these are in model layer 3).

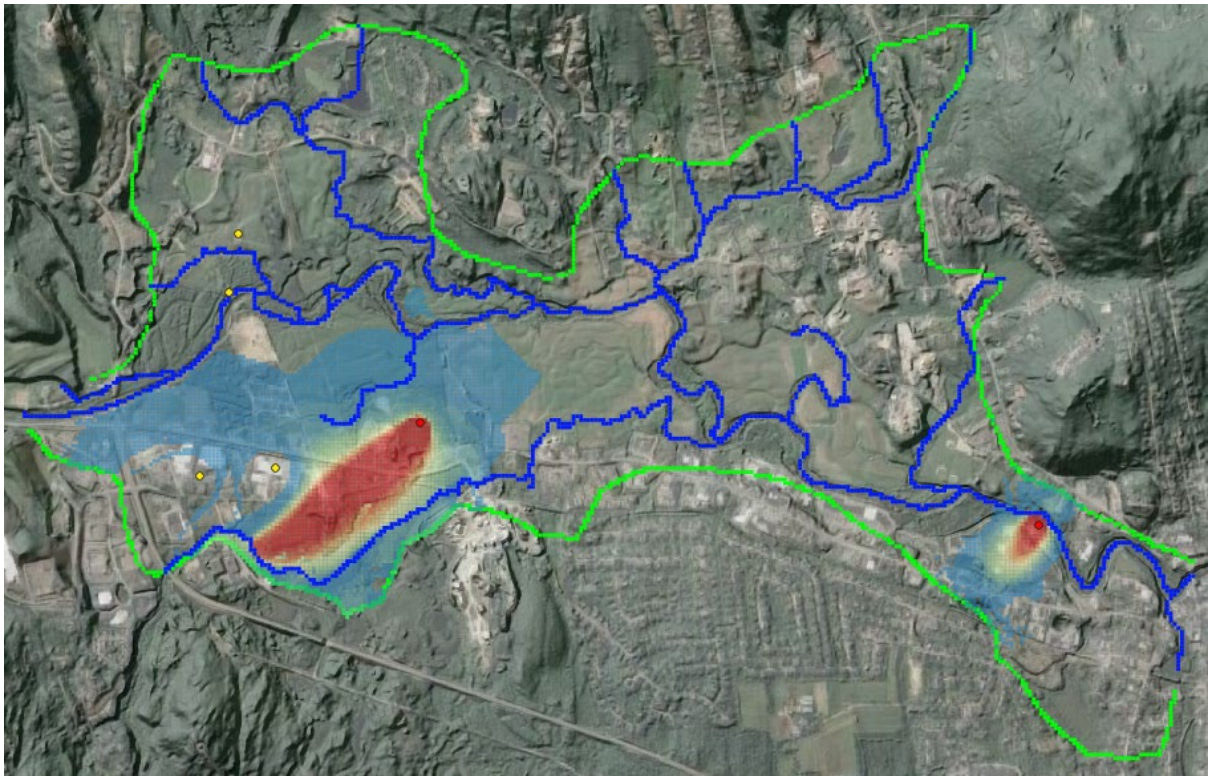

Figure 5.1a Coloured shading depicts the probability that a cell contributes to water extracted from the Savage (west) or Keyes (east) extraction wells; the colour scale is shown in Figure 5.1b. The Savage and Keyes wells are plotted in red, while other extraction wells are plotted in yellow.

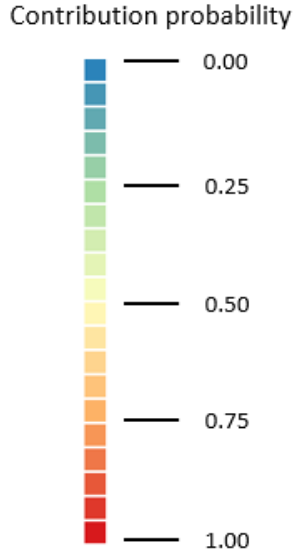

Figure 5.1b. Colour scale used in Figure 5.1a.

Obviously, the area of non-zero contribution probability for the Savage well is much larger than that for the Keyes well, further suggesting a substantial contribution to the latter from the Souhegan River.

What is interesting from a management point of view is the possibility that water that recharges the Milford-Souhegan aquifer far to the west, and even somewhat to the north, of the Savage well has a finite probability of being discharged from this well. For both the Savage and Keyes wells, possible contributing areas lie on either side of nearby rivers and streams, and can even extend to the boundary of the model domain to capture more distant recharge waters.

Also noteworthy is the fact that the two production wells which lie to the southwest of the Savage well limit its contributing area to some extent because of their own contributing areas.

The probabilistic contributing areas depicted in Figure 5.1 are broadly aligned with contributing areas reported by Harte and Mack. However, Figure 5.1 reveals the possibility that the contributing area for the Savage well may extend further to the west than was predicted by Harte and Mack. This is an interesting, but perhaps unsurprising, outcome of the modelling approach documented herein that views uncertainty analysis as its core requirement. Of concern, however, is that the area of nonzero contributing probability to Keyes extraction depicted in Figure 5.1 is smaller than the Keyes well contributing area calculated by Harte and Mack. This matter is further discussed below.

5.4 One Layer Model

How important is it that a Milford-type model which is built to undertake contributing area analysis be multi-layered? The original Milford model possessed five layers; the revisited Milford model possesses three layers. However, flow of water in the Milford-Souhegan aquifer is predominantly horizontal; Harte and Mack report that vertical head gradients are minimal. This suggests that a one layer model may suffice for contributing area analysis. Use of a one layer model instead of a three layer model offers the possibility of faster construction, simpler parameterisation, and higher execution speed. If, at this or any other site, uncertainties in calculated contributing areas dominate predictive biases arising from model simplification, then this justifies a simpler modelling approach over a more complex approach (provided the ability to quantify contributing area uncertainty is not compromised by use of a simpler model). A simpler approach may be useful where, for example, a local authority is required to undertake contributing area analysis at many locations to assess the vulnerability of community water supply wells to pathogens.

The case for model simplicity is strengthened by noting that, under steady state conditions, a well's contributing area is equal to that for which total recharge is equal to well extraction; the task of modelling then becomes that of defining where this area is actually situated with respect to the extraction well. Analytical and semi-analytical methods for contributing area analysis have been available for a long time; see, for example, Paradis and Martel (2007), Haitjema (2002) and USEPA (1994).

Groundwater flows horizontally according to the Dupuit Forchheimer assumption that underpins use of a single layer model. In spite of this, particle movement has a vertical component, as older particles are displaced downwards by younger particles along a flow line. This improves their ability to handle heterogeneities in hydraulic conductivity and in recharge in the vicinity of a well. On the other hand, a single layer model may suffer from overly simplistic representation of conditions near a stream, perhaps artificially enhancing pumping-induced loss of water from that stream because of its reduced ability to simulate pumping-induced flow of water under the stream.

To explore these issues, the Milford model was re-constructed as a single layer model. Most specifications of the single layer model are identical to those of the three layer model described above. Lateral grid dimensions, and the active area of the single layer model are identical to those of the three layer model. Pilot point parameterisation of precipitation-based recharge, lateral recharge and riverbed conductance are also identical; however lateral recharge is supplied to layer 1 instead of layer 3. Horizontal hydraulic conductivity of the single layer model is parameterized using the *ppset1* and *ppset2* families of pilot points in the same manner as that described above for layer 2 of the three layer Milford model.

History-matching of the single layer model followed the same course as that of the three layer model. First the model was calibrated using PEST_HP; then the PESTPP-IES ensemble smoother was employed to derive an ensemble of 300 calibration-constrained parameter fields. Model-to-measurement misfit was defined in the same way for both models, including the use of penalty functions for drain flow and diminution of well extraction. The same level of fit with the history-matching dataset was attained. Patterns of hydraulic conductivity and recharge that emerged from history matching are similar. These, along with accompanying contributing areas for the Savage and Keyes wells, are shown for 6 parameter fields in Appendix C. (Note that while these parameter fields have the same indices as those depicted in Appendix B, they are not directly comparable because of randomness in random parameter field generation.)

Contributing area probabilities for the Savage and Keyes wells, calculated using 300 samples of the posterior parameter probability distribution, are presented in Figure 5.2.

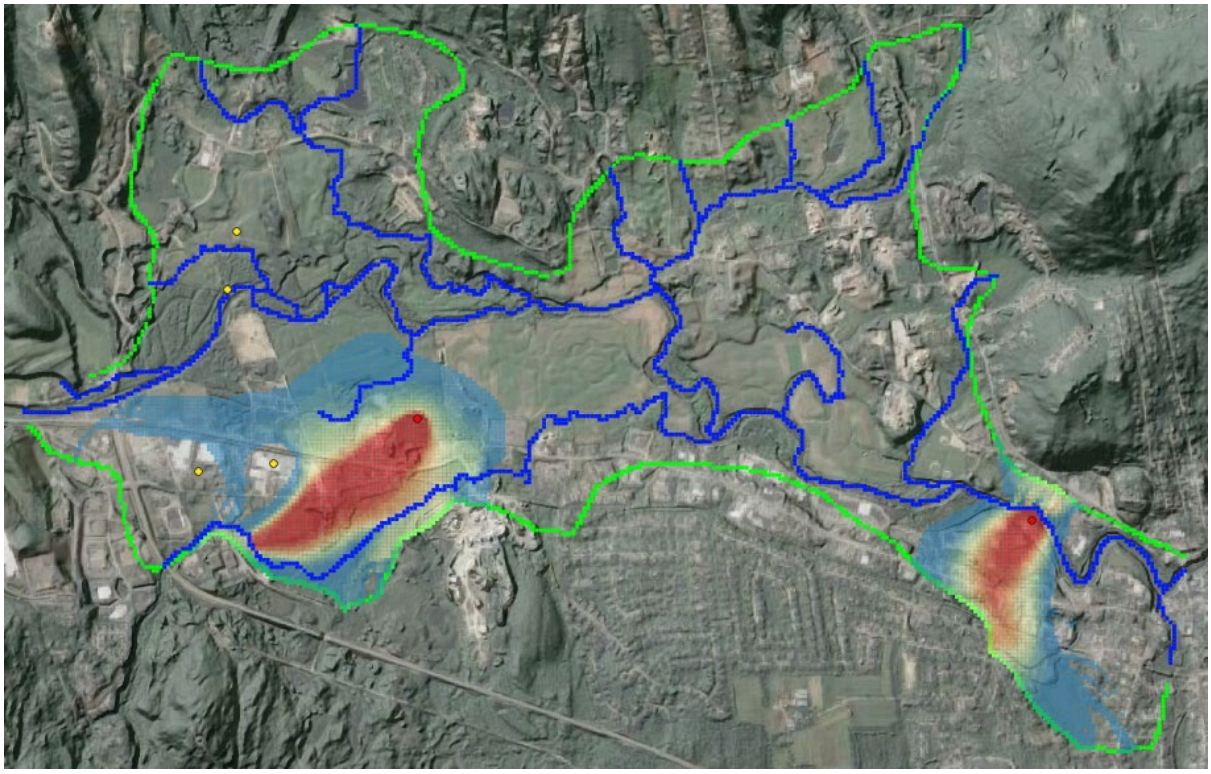

Figure 5.2. Contributing area probabilities calculated using a single layer model. See Figure 5.1b for the probability colour scale.

Contributing area probabilities for the Savage (i.e., western) well are similar in Figures 5.2 and 5.1. However, areas of moderate contributing area probability depicted in Figure 5.2 are slightly more extensive than those in Figure 5.1. Furthermore, they cross the stream that lies to the south-east of the Savage well. This appears to refute the suggestion that use of a single

layer model results in over-estimation of contributions to contributing area from rivers and streams.

Contributing areas for the Keyes well are dissimilar between Figures 5.2 and 5.1. The area of non-zero contribution is much larger for the single layer model than it is for the three layer model. Furthermore, this area extends along the south-eastern boundary of the model domain for the single layer model. This is in accordance with Harte and Mark's characterisation of the Keyes well contributing area.

Figure 5.3 shows a close up view of that part of Figure 5.2 which lies in close proximity to the Keyes well. It also shows head observation wells (in blue) together with pilot points that support parameterisation of riverbed conductance. (These are coloured green for the river reach near the Keyes well). An inspection of PESTPP-IES numerical outputs reveals that riverbed conductances for the single layer model are lower across all parameter realisations in the vicinity of the Keyes well than they are for the three layer model. This restricts extraction of water from the river; the area that contributes to Keyes well extraction must thereby expand in order to gain access to precipitation and boundary recharge.

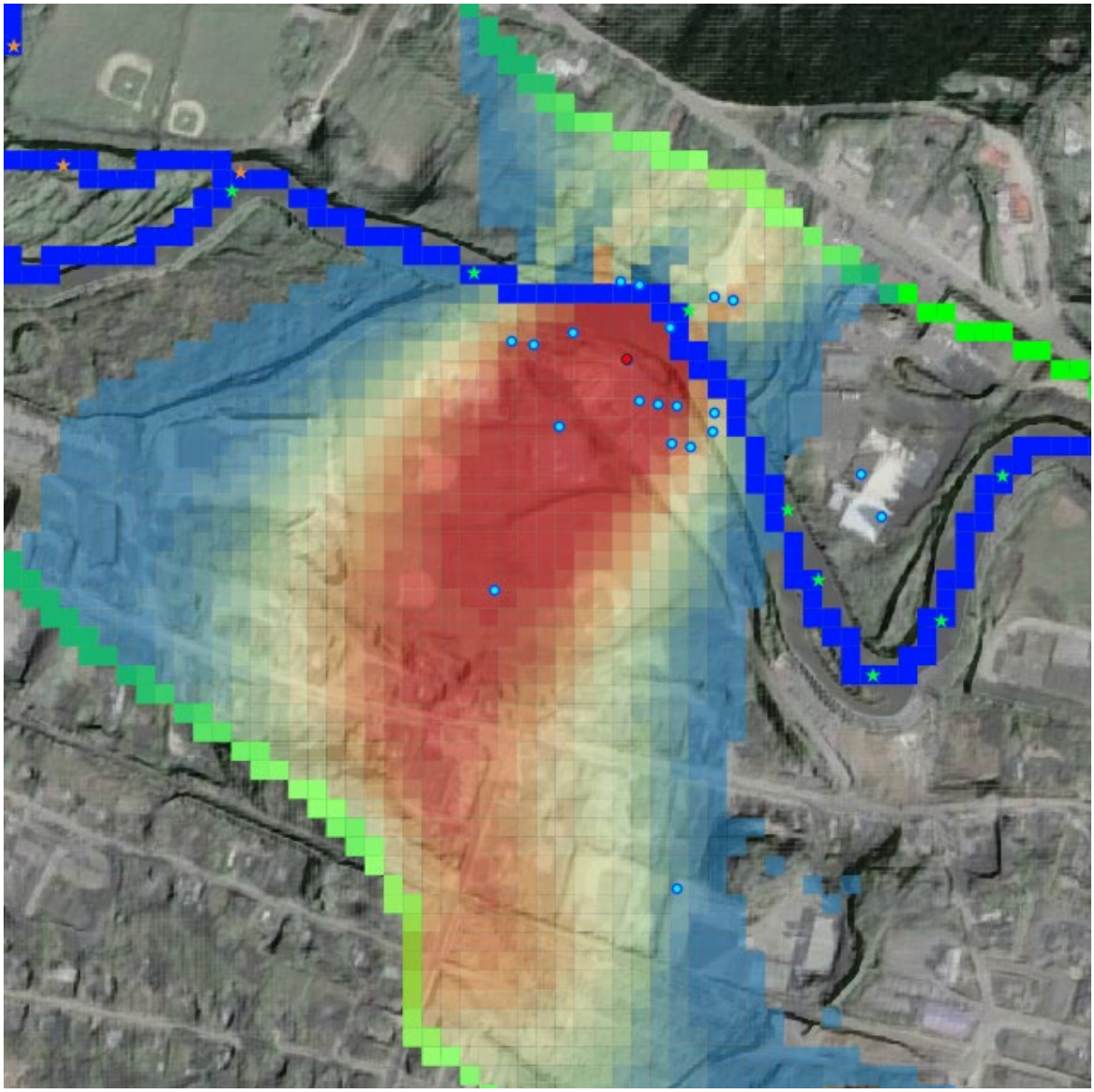

Figure 5.3. Magnified view of part of Figure 5.2 with observation wells (blue circles) and riverbed pilot points (stars) added.

The observation wells depicted in Figure 5.3 are predominantly in layers 1 and 3 of the revisited three layer Milford model. History-matching of the three layer model is able to replicate differences in heads in these layers by altering local values of vertical hydraulic conductivity. At the same time, adjustment of riverbed conductance and layer 1 hydraulic conductivities supports the fitting of observed heads in layer 1. In contrast, where a single layer model is employed, fewer parameters are available for supporting the attainment of a good fit with all members of the local calibration dataset. Hence vertical gradients induced by inflow to the river during the calibration period when the Keyes well was not operating cannot be accommodated except through introduction of a low riverbed conductance. The same applies even if these gradients are spurious because they do not, in fact, reflect steady state conditions. In either case, there exists a potential for bias in estimation of riverbed conductance if local structural simplicity of the model requires that these parameters compensate for this simplicity during history matching, or if they must bear all of the parameter adjustment burden of responding to erroneous data.

So, which of the Keyes contributing area probability maps are correct – that displayed in Figure 5.1 or that displayed in Figure 5.2?

We do not know.

What we do know is that model structural and parameterisation detail in the vicinity of the Keyes well are important because of the proximity of this well to the Souhegan River. It would be better to build a model that is dedicated solely to calculation of its contributing area than including this calculation in a larger, multi-purpose model. If this were done, then considerable attention would need to be paid to hydrogeological conditions in the vicinity of this extraction well, these including riverbed dimensions and river elevation. The quality of local observation data, and its representativeness (or otherwise) of steady state conditions, would also require considerable attention. A judgement call would be required on which nuances of this dataset are worth fitting, and which are not. Of particular importance would be vertical head gradients. If this portion of the calibration dataset (or certain aspects of it) are construed to be worthy or respect, then local grid refinement supported by denser parameterisation may be warranted in order to ensure that history-matching of these data does not bias prediction-sensitive parameters.

Alternatively, if it were judged that some or all of these near-well measurements are of questionable integrity, then contributing area analysis would benefit from their removal, for their presence in a history matching dataset may bias important model predictions. With omission of these data, predictive bias would be replaced by predictive uncertainty. The above example suggests that a single layer model, and its associated parameterisation, is quite capable of quantifying this uncertainty while responding adequately to information contained in other data.

5.5 The Third Dimension

We conclude this section with some pictures of particle tracks in three dimensions. All of these pictures pertain to the Savage well. The viewing point is from the south-east of this well. For Figure 5.4 it is above ground while for Figure 5.5 it is below ground; see Figure 5.6. The upper part of each of these figures shows particle tracks calculated by mod-PATH3DU for the three layer model, while the lower part shows particle tracks calculated for the one layer model. The Savage well is shown in red while other extraction wells are shown in yellow. Cells that contain a RIV boundary condition are also shown; these are thinner for the three layer model than for the single layer model as the upper layer of the three layer model within which RIV boundary conditions are placed does not extend to the bottom of the aquifer.

In each of these figures particles are coloured according to residence time, with purple and blue indicating a shorter time and red indicating a longer time (up to 400 days under the assumption of a spatially uniform porosity of 0.05).

Note that particle tracks are not directly comparable between each pair of figures. They were calculated using random samples of posterior parameter probability distributions provided by PESTPP-IES for the three layer model on the one hand, and the single layer model on the other hand.

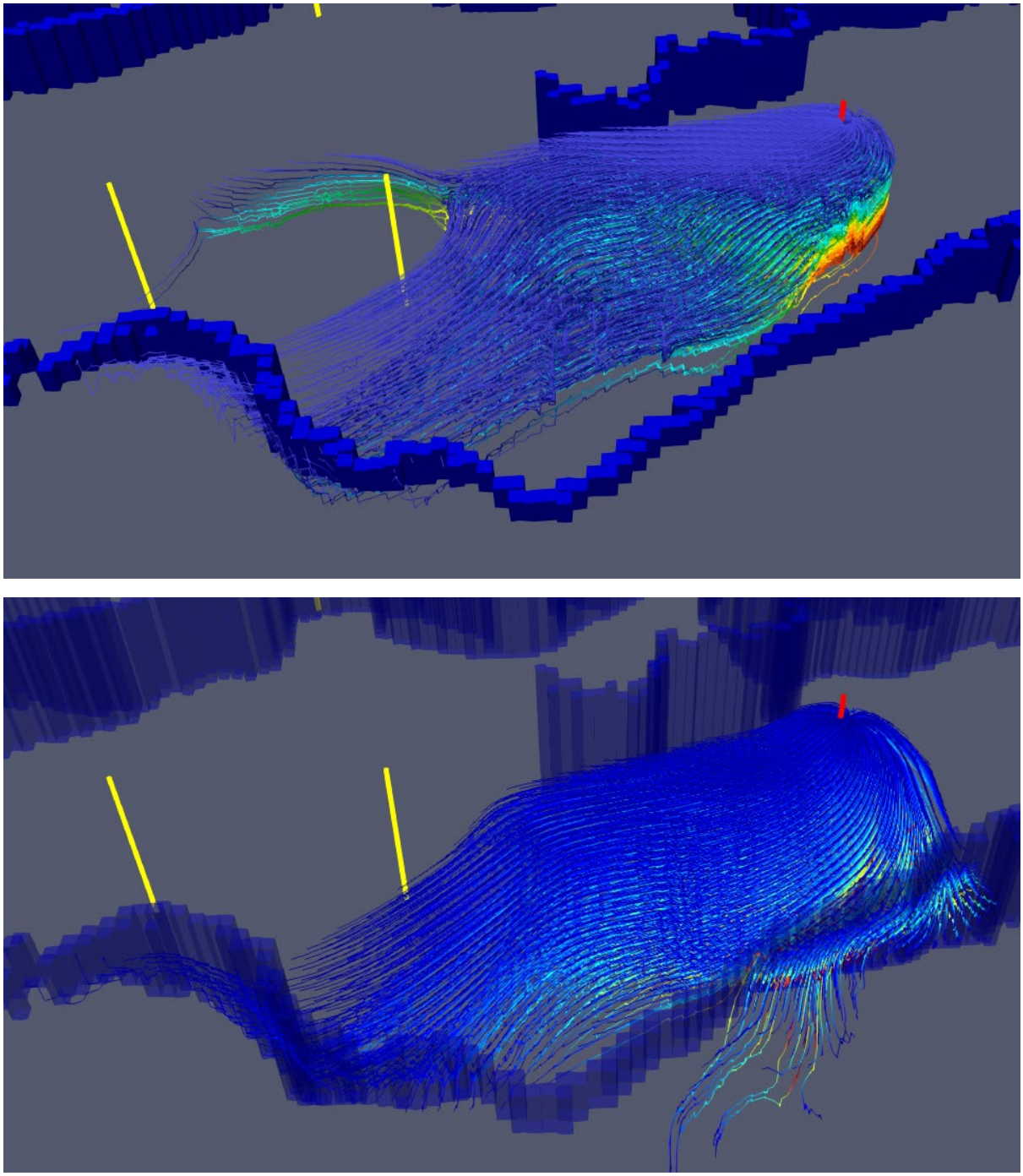

Figure 5.4. Tracks of particles captured by the Savage well. Redder colours indicate greater residence time. The upper figure pertains to the three layer model while the lower figure pertains to the single layer model.

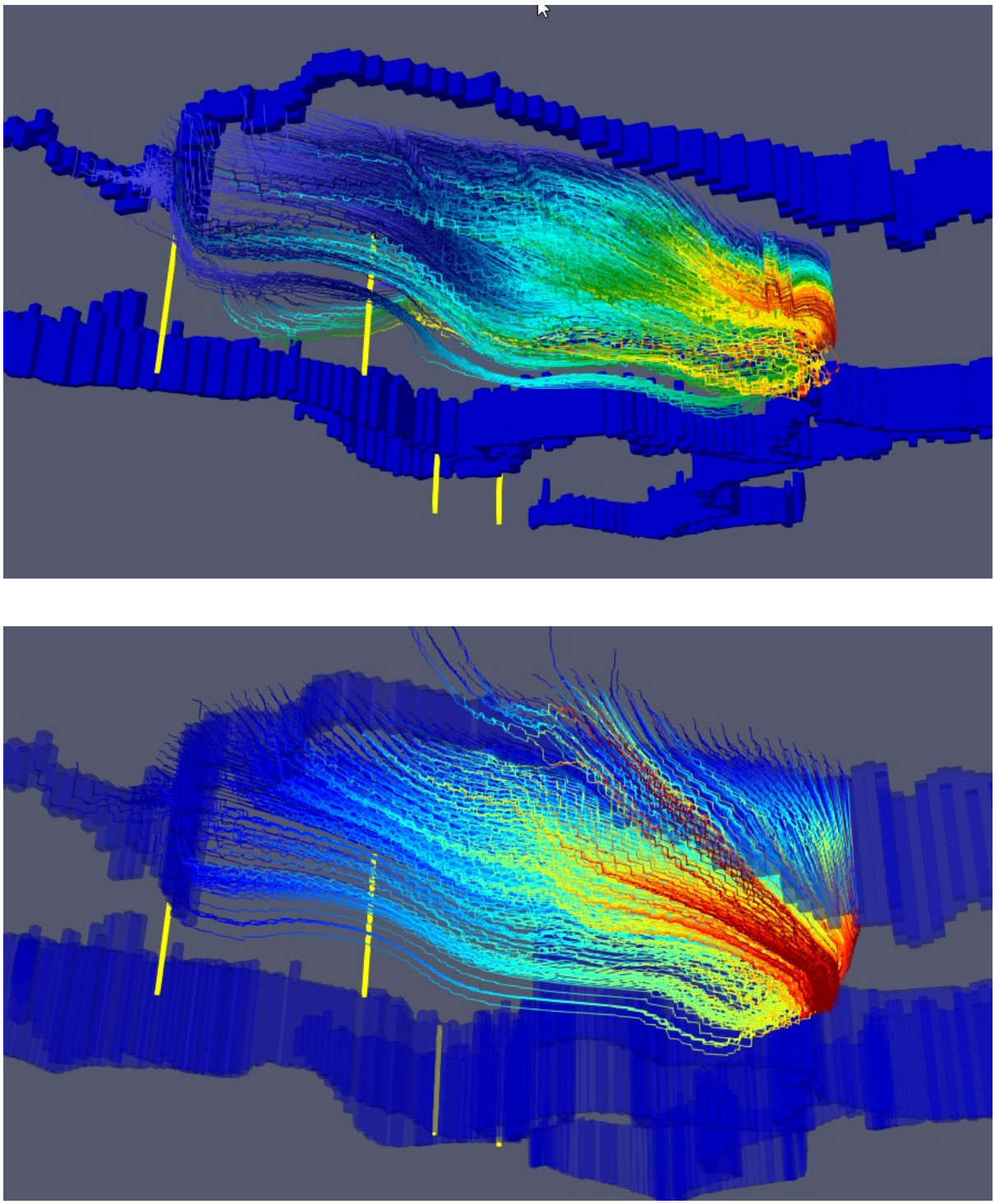

Figure 5.5. Tracks of particles captured by the Savage well. Redder colours indicate greater residence time. The upper figure pertains to the three layer model while the lower figure pertains to the single layer model.

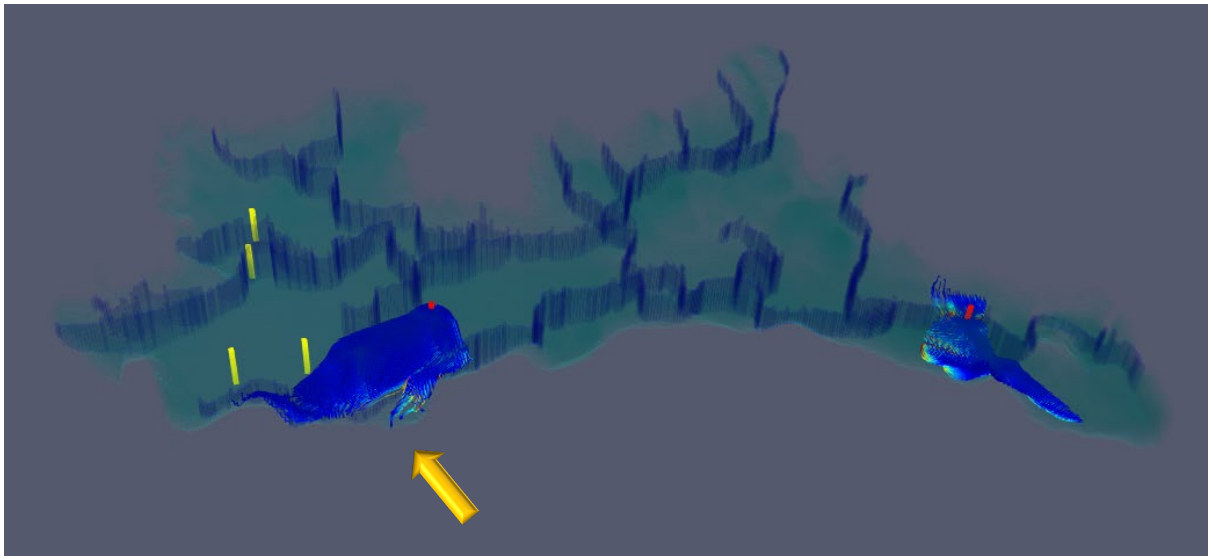

Figure 5.6. Viewing direction for previous two figures.

Immediately noticeable in Figures 5.5 and 5.6 is the “jaggedness” of each particle track. This is an outcome of the way in which particle movement is calculated. When a particle reaches the boundary of a model cell, its passage to the neighbouring cell is such that its local (i.e. cell-based) vertical coordinate is preserved. Mod-PATH3DU records in-cell and between-cell particle movement in its output files. Because of this, vertical particle displacements at cell boundaries are readily apparent. (Note that the default behaviour of the MODPATH particle tracker is to record particle locations only where they cross cell boundaries.)

Vertical between-cell particle displacements would be reduced if layer boundaries were flatter. In the revisited Milford model, layering divides the distance between the surface and the aquifer into thicknesses whose ratios are roughly the same throughout the model domain. In contrast, the thickness of each layer in the original Milford model is roughly constant. However, this layering strategy has its own costs, as lower layers must appear and disappear where bedrock topography undulates.

Figures 5.4 and 5.5 depict trajectories of particles that originate on the opposite side of a river or stream to the extraction well which captures them. In the three layer model, most of these particles travel under the cells that contain RIV boundary conditions. Conversely, in the single layer model, particles travel through the cells that contain RIV boundary conditions.

It is important to note that water is exchanged between any MODFLOW cell that contains a RIV boundary condition and the boundary condition itself. Both MODPATH and mod-PATH3DU require that a user specifies the cell face through which this exchange takes place. Identification of this face determines the distribution of vertical velocities within a groundwater model cell. This vertical velocity distribution, in turn, determines whether a particle emerges through a RIV boundary condition or passes through the cell.

6. CONCLUSIONS

Probabilistic contributing area analysis is not new. Molson and Frind (2012) showed how probability of capture can feature as a state variable in a partial differential equation that is very similar to the standard advection dispersion equation. Numerical solution of this equation is easily achieved by running a transport model backwards while supplying the correct boundary conditions. Meanwhile, local scale hydraulic property heterogeneity and mixing of waters can be taken into account through macrodispersion. See Neupauer and Wilson (2005) and Frind et al (2002) for further details and examples.

In our opinion, the methodology for probabilistic contributing area analysis that is documented in the present study is able to provide a better guarantee that uncertainties in its evaluation are not understated than solution of a single state variable based on a single parameter field. Plots of contributing area probability that are presented in Section 5 of this report acknowledge the possibility that, depending on the spatial distribution of hydraulic properties, waters that enter a groundwater system at very different locations may terminate their journeys in the same extraction well. These disparate hydraulic property distributions share compatibility with the current conceptual hydrogeological model, at the same time as they allow model outputs to replicate measured system behaviour. Identification of widely separate locations from which recharge waters may possibly reach the same extraction well is only possible when travel paths are explicitly simulated in models that embody realisations of hydraulic property heterogeneity that promulgate their capture.

Like other worked examples that feature in this GMDSI series, the importance of data assimilation through history matching, and of post-history-matching uncertainty quantification are emphasized. Ideally, the uncertainties of decision-critical model predictions can be reduced through history-matching. However, this worked example demonstrates that for some predictions caution must be exercised. This is because history-matching has the potential to inculcate bias in some predictions if deficiencies in a model's structure force some of its parameters to adopt roles that compensate for these deficiencies. This matter is discussed extensively by White et al (2014) and Doherty (2015). These authors point out that history-matching-induced parameter bias does not automatically promulgate predictive bias; this depends on the prediction. The GMDSI worked example that is documented herein demonstrates that, under certain circumstances, contributing area is one such prediction.

If an extraction well may draw some of its water from a nearby stream, then representation of this stream in a model should be of high quality. Features which require high-fidelity representation include directly-measurable properties such as elevation and width. Measurements of system state from which other stream properties may be back-calculated must also be of high quality. Erroneous estimation of streambed conductance and near-stream aquifer hydraulic properties from error-affected data may bias model-calculated contributing areas.

If it is not possible to represent stream features and hydraulic properties with integrity, then they should be represented stochastically. If measurements from which stream hydraulic properties can potentially be back-calculated are not of high quality, then they should be omitted from a history-matching dataset. Predictions of extraction well contributing area are not therefore understated; nor are they prejudiced by model structural errors or history-matching-induced parameter bias. High, but quantifiable, uncertainty is a far better modelling outcome than unquantifiable predictive bias.

The study that is reported herein suggests that there may be many circumstances where a single layer model can be used just as effectively for contributing area analysis as a multi-layer model. Construction of a single layer model is easier than that of a multi-layer model. Its faster run time and greater numerical stability can support high levels of parameterisation complexity. If steps are taken to ensure that its simple structure does not create conditions for history-matching-induced predictive bias, this strategy may enhance the ability of the modelling process to assimilate prediction-pertinent data, thereby reducing uncertainties in contributing area evaluation, at the same time as it supports quantification of these uncertainties. If near-stream measurements of system state are scarce, or if they are omitted from a history-matching dataset because they are of questionable integrity, then use of a single layer model becomes even more attractive; discarding this data denies it the opportunity to induce parameter and predictive bias during history-matching.

We close this report by reminding the reader that all of the analyses that are documented herein are supported by PEST_HP, PESTPP-IES and members of the PEST Groundwater Utility Suite. All figures that are presented in these pages were produced by either QGIS or PARAVIEW. Formatting of model and other datasets for viewing and manipulation in these packages was undertaken using members of the PEST Groundwater Utility suite.

All model input files, together with instructions and some documentation, are available from the GMDSI web site at http://www.gmdsi.org.

7. REFERENCES

Barlow, P.M., Leake, S.A. and Fienen, M.N., 2018. Capture versus captures zones: clarifying terminology related to sources of water to wells. *Groundwater*, 56 (5), 694-704.

Chen, Y. and Oliver, D.S., 2013. Levenberg–Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification. *Computational Geosciences*, 17(4): 689–703.

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

Doherty, J., 2021. Manuals for PEST and associated utility support software. Downloadable from www.pesthomepage.org.

Frind, E.O., Mohammed, D., Molson, J.W., 2002. Delineation of three-dimensional capture zones in a complex aquifer. *Ground Water* 40 (6), 586–598.

Harte, P. and Mack, T.J., 1992. Geohydrology of, and Simulation of Ground-Water Flow in, the Milford-Souhegan Glacial-Drift Aquifer, Milford, New Hampshire. USGS Water-Resources Investigation Report 91-4177. Downloaded from https://pubs.er.usgs.gov/publication/wri914177.

Haitjema, H. M., 2002. *WhAEM User's Manual*. Environmental Protection Agency (EPA), USA.

McDonald, M.G. and Harbaugh, A.W., 1988. A modular three-dimensional finite-difference ground-water-flow model. U.S. Geological Survey Techniques of Water-Resources Investigations, Book 6, Chapter A1, 586p.

Molson, J.W. and Frind, E.O., 2012. On the use of mean groundwater age, life expectancy and capture probability for defining aquifer vulnerability and time-of-travel zones for source water protection. *Journal of Contaminant Hydrology*, 127, 76-87.

Muffels, C., Scantlebury, L., Wang, X., Tonkin, M., Neville, C., Ramadhan, M., and Craig, J.R., 2020. User's Guide for mod-PATH3DU. S.S. Papadopoulos and Associates Inc and University of Waterloo.

Neupauer, R.M. and Wilson, J.L., 2005. Backward probability model using multiple observations of contaminant to identify groundwater contamination sources at the Massachusetts Military Reservation. *Water Resour. Res.* 41 (W02015), 1–14.

Panday, S., 2020. The Block-Centered Transport Process for MODFLOW-USG, *GSI Environmental*, January, 2021 http://www.gsi-net.com/en/software/free-software/USG-Transport.html

Panday, S., Langevin, C.D., Niswonger, R.G., Ibaraki, M. and Hughes, J.D., 2013. MODFLOW–USG version 1: An unstructured grid version of MODFLOW for simulating groundwater flow and tightly coupled processes using a control volume finite-difference formulation: *U.S. Geological Survey Techniques and Methods*, book 6, chap. A45, 66 p.

Paradis, D. and Martel, R., 2007. HYBRID: a wellhead protection delineation method for aquifers of limited extent. *Geological Survey of Canada*. Technical Note 1. Catalogue No. M41-10/1-2007E-PDF.

PEST++ Development Team (2020). Manual for PEST++. Downloadable from https://github.com/usgs/pestpp

Pollock, D.W., 1988. Semianalytical computation of path lines for finite-difference models: *Ground Water*, v. 26, no. 6, p. 743-750.

Pollock, D.W., 1989. Documentation of computer programs to compute and display path lines using results from the U.S. Geological Survey modular three dimensional finite-difference ground-water flow model: U.S. Geological Survey Open File Report 89-391, 188 p.

Rhamadhan, M., 2015. A Semi-Analytical Particle Tracking Algorithm for Arbitrary Unstructured Grids. Unpublished MASc. Thesis. University of Waterloo, Waterloo Ontario.

USEPA, 1994. Handbook: Ground Water and Wellhead Protection. U.S. Environmental Protection Agency. EPA/625/R-94/001.

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.

APPENDIX A. UTILITY PROGRAMS

A.1 Introduction

This appendix tabulates utility programs from the PEST and PEST Groundwater Utility suites that were used in the modelling work that is discussed in this report. They are subdivided according to function; programs which have multiple functions are listed in more than one table.

Before tabulating these programs, some of the file types that they employ are listed.

A.2 File Types

File	Description
Covariance matrix file	An ASCII file that holds a covariance matrix.
CSV file	"CSV" stands for "comma-separated variables". These files are used extensively by programs of the PEST++ suite. Or particular use in the present context are CSV files that store a sequence of random parameter sets comprising an ensemble.
MIF/MID	A pair of ASCII files that record data pertaining to points, lines and polygons. "MIF" stands for "Mapinfo Interchange Format".
Node data table file	This type of file is used by those members of the PEST Groundwater Utilities suite that support MODFLOW-USG. It is an ASCII file in which data are listed in columns. The first column must list the index of every node in a MODFLOW-USG unstructured grid.
Parameter uncertainty file	Used by programs of the PEST and PEST++ suites to specify prior or posterior parameter uncertainties.
PEST control file	Contains variables which govern the use of PEST or PEST++, as well as initial parameter values, and the values of measurements that model outputs must match.
PEST template file	Used by PEST/PEST++ to inform these programs how to record parameter values on a model input file.
PEST instruction file	Used by PEST/PEST++ to read model outcomes of interest from a model output file.
Site sample file	This is also referred to as a "bore sample file". This type of file is used extensively by members of the PEST Groundwater Utility suite. It records time series of measurements (or the model-generated counterparts to measurements), along with the dates and times at which measurements were made.
Tabular data file	This is a general file type. In the present context it refers to files in which one column contains model node numbers, while other columns contain data pertaining to these nodes. However, in contrast to a node data table file, not all nodes within a model need be cited in the first column of the file.
Unstructured grid specification file	This is used by members of the PEST Groundwater Utility suite. It lists coordinates of all vertices of an unstructured grid, and the model cells to which these vertices pertain.

A.3 GIS Interface

Program	Function
MP3DU2MIF	Reads a particle file produced by mod-PATH3DU. Records particle paths in MIF/MID file format.
MP3DUFBE	Filters a particle file written by mod-PATH3DU such that the re-written particle file includes only particles that terminate in user-specified cells. Writes another file that can be imported into a GIS to display starting points and total travel times of end-point-filtered particles.
TAB2NDTF	Reads a tabular data file in which one column lists model node numbers. Re-writes data in node data table format.
USGNDTF2MIF	Writes a MIF/MID file using data contained in a node data table file.

A.4 Paraview Interface

Program	Function
MP3DU2VTK	Reads an ASCII particle file written by mod-PATH3DU. Records particle tracks in a VTK file. These can then be visualized in packages such as PARAVIEW.
USG2VTK	Reads an unstructured grid specification file and a node data table file. Writes a VTK file that supports display of model properties or results in three dimensions.
USG2VTK1	Similar to USG2VTK but can accommodate non-rectilinear model cells.
USGMODGSF	Reads a MODFLOW-USG grid specification file. Modifies it so that the tops and bottoms of cells are flat, and at elevations provided in MODFLOW-USG input files. Display of a model grid and accompanying properties in this way can be useful for understanding particle travel paths, particularly as they cross cell boundaries.

A.5 Manipulating Mod-PATH3DU Files

Program	Function
MP3DU2MIF	Reads a particle file produced by mod-PATH3DU. Records particle paths in MIF/MID file format.
MP3DU2VTK	Reads an ASCII particle file written by mod-PATH3DU. Records particle tracks in a VTK file. These can then be visualized in packages such as PARAVIEW.
MP3DUACCUM	Used to build a histogram of particle starting cells when undertaking stochastic contributing area analysis.
MP3DUFBE	Filters an ASCII particle file written by mod-PATH3DU. Writes a particle file that features only particles that terminate in user-specified model cells. Writes another file that lists residence times of these particles.

A.6 Model Pre- and Postprocessing

Program	Function
PLPROC	General pilot point model pre-processor. Undertakes spatial interpolation to a model grid, and writes or modifies model input files accordingly.
SMP2SMP	Undertakes temporal interpolation from one site sample file to another.
USGBUD2SMP	Reads a binary budget file written by MODFLOW-USG. Records boundary-condition flows in user-specified groups of cells.
USGMOD2OBS	Reads a binary system state file written by MODFLOW-USG. Undertakes spatial interpolation to the sites of observation wells. Records outcomes in site sample file format.
USGQUADFAC	Calculates interpolation factors for the use of USGMOD2OBS.

A.7 PEST/PEST++ Input Dataset Construction

Program	Function
MKPPSTAT	Reads a pilot points file; writes a file that PPCOV_SVA can use to write a covariance matrix file.
PESTPREP1 PESTPREP2	Adds information to the “observation groups” and “observation data” sections of a PEST control file; writes instruction files to read model-generated data stored in site sample file format.
PPCOV_SVA	Writes a covariance matrix file based on a spatially-varying variogram for the use of pilot point parameters. This file can be used for regularised inversion and/or for linear/nonlinear uncertainty analysis.
PWTADJ1	Rewrites a PEST control file using weights that guarantee visibility of each component of a multi-component objective function.
PWTADJ2	Re-writes a PEST control file using “correct weights” calculated from observation-group-specific objective functions achieved during a previous PEST run.
TPL2PST	Creates a “partial PEST control file” citing parameters that are cited in one or a number of template files.

APPENDIX B. MAPS OF CONTRIBUTING AREA

This appendix provides maps of contributing area for the three layer Milford model calculated using 6 realisations from its posterior parameter probability distribution. These realisations were evaluated using the PESTPP-IES ensemble smoother.

Each contributing area is pictured twice. In the first picture, it is underlain by a map of estimated layer 1 hydraulic conductivity. In the second picture it is underlain by a map of estimated recharge. Colour scales used to display hydraulic conductivity and recharge are the same for all realisations. They are provided in Figure A.1.

Contributing areas themselves are also coloured. Coloration is according to time of travel from the recharge point to the extraction point. Steady state conditions are assumed; porosity is assigned a spatially uniform value of 0.05. Note that it would have been possible to use an expanded colour scale to represent travel time in these plots; however, the same time colour scale is used in Appendix C where some travel times are higher.

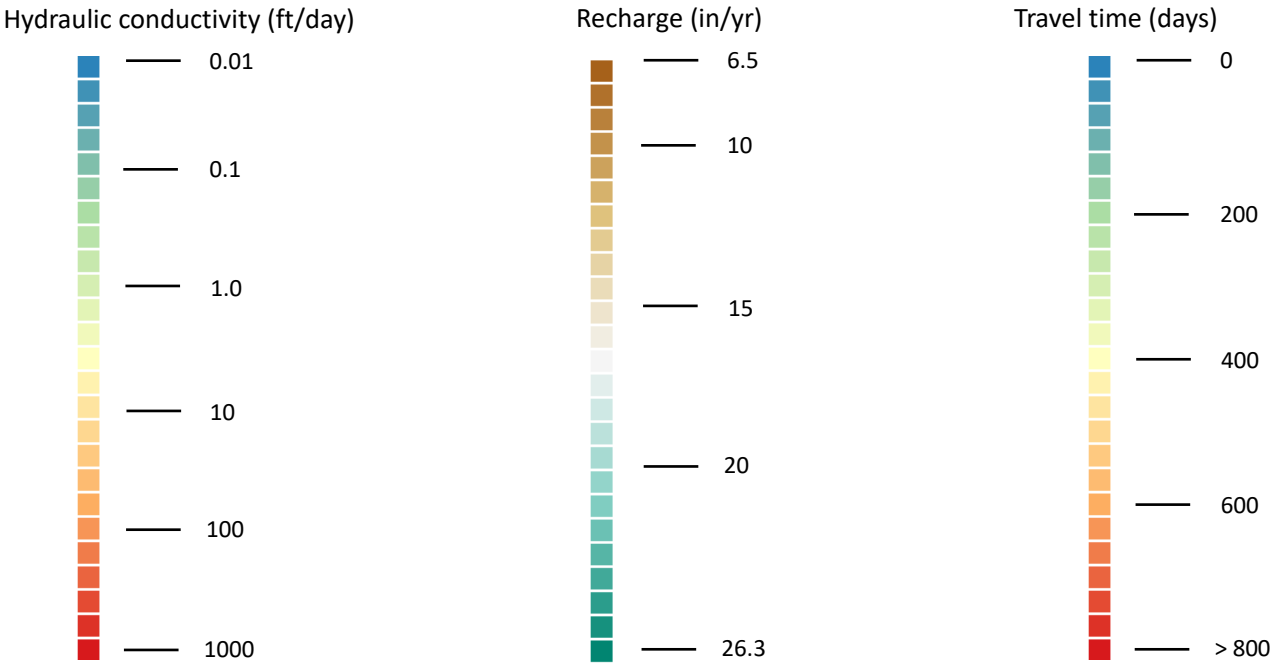

Figure B.1. Colour scales used in following pictures.

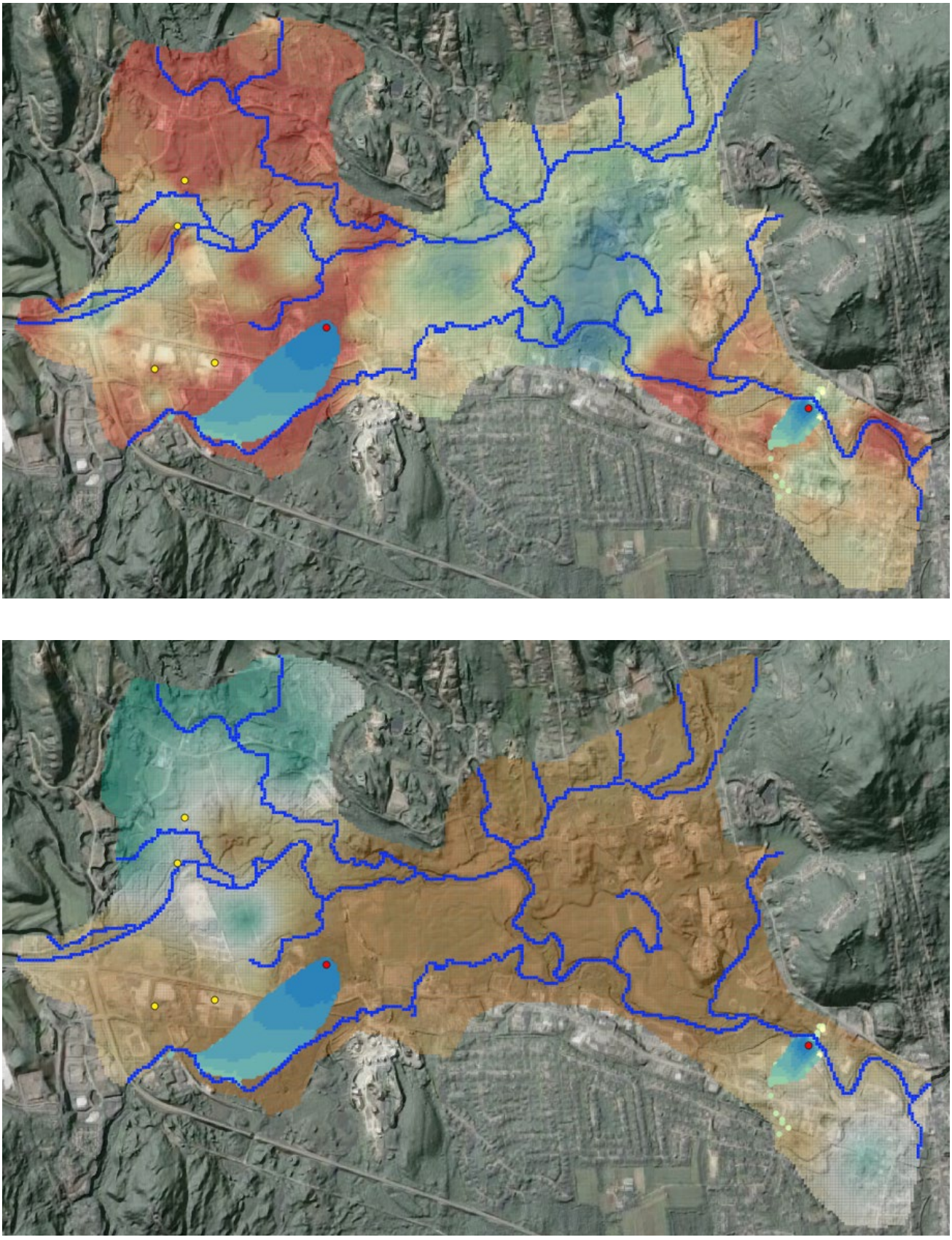

Figure B.2. Contributing area for realisation "0" superimposed on maps of layer 1 hydraulic conductivity (top) and recharge (bottom).

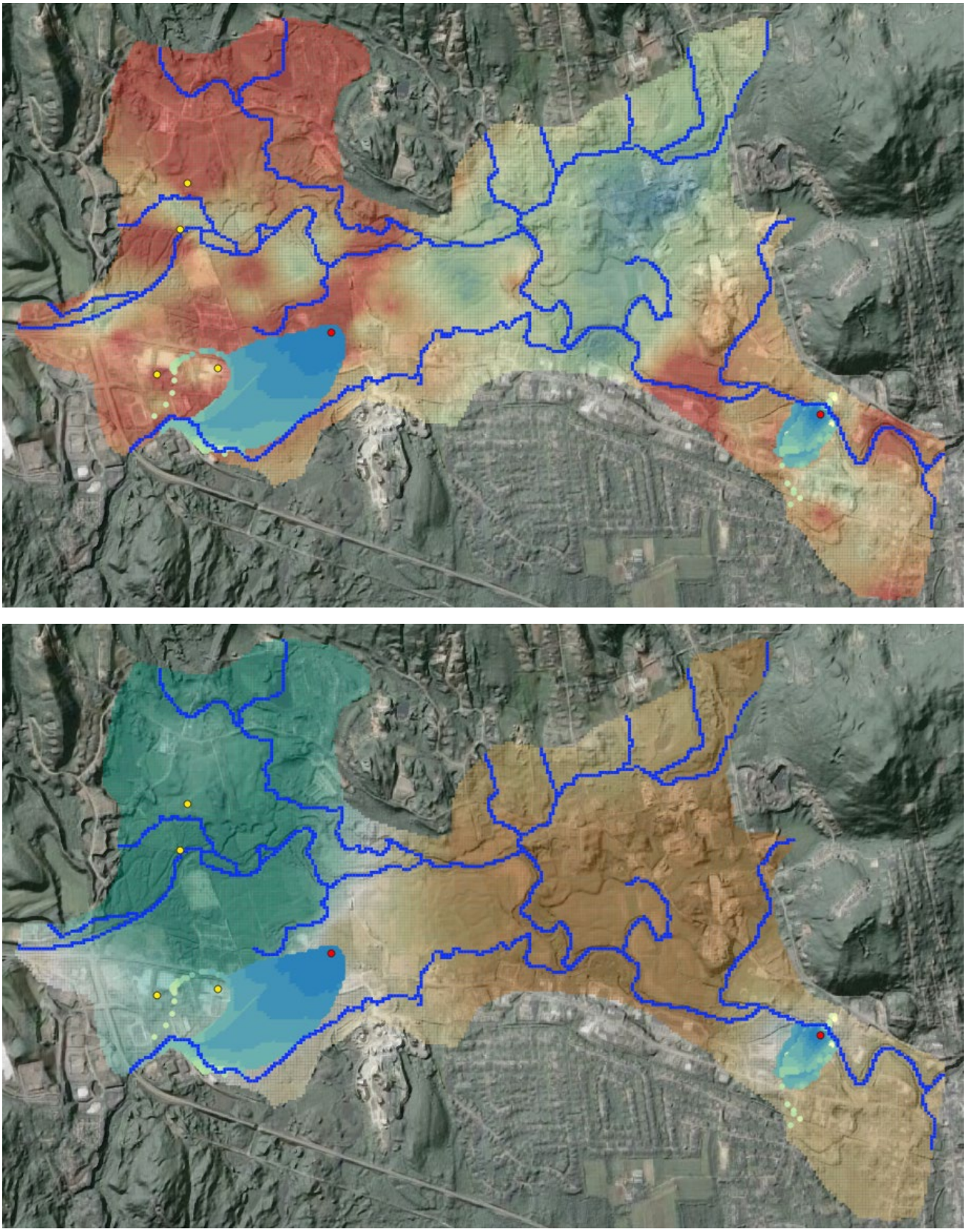

Figure B.3. Contributing area for realisation "25" superimposed on maps of layer 1 hydraulic conductivity (top) and recharge (bottom).

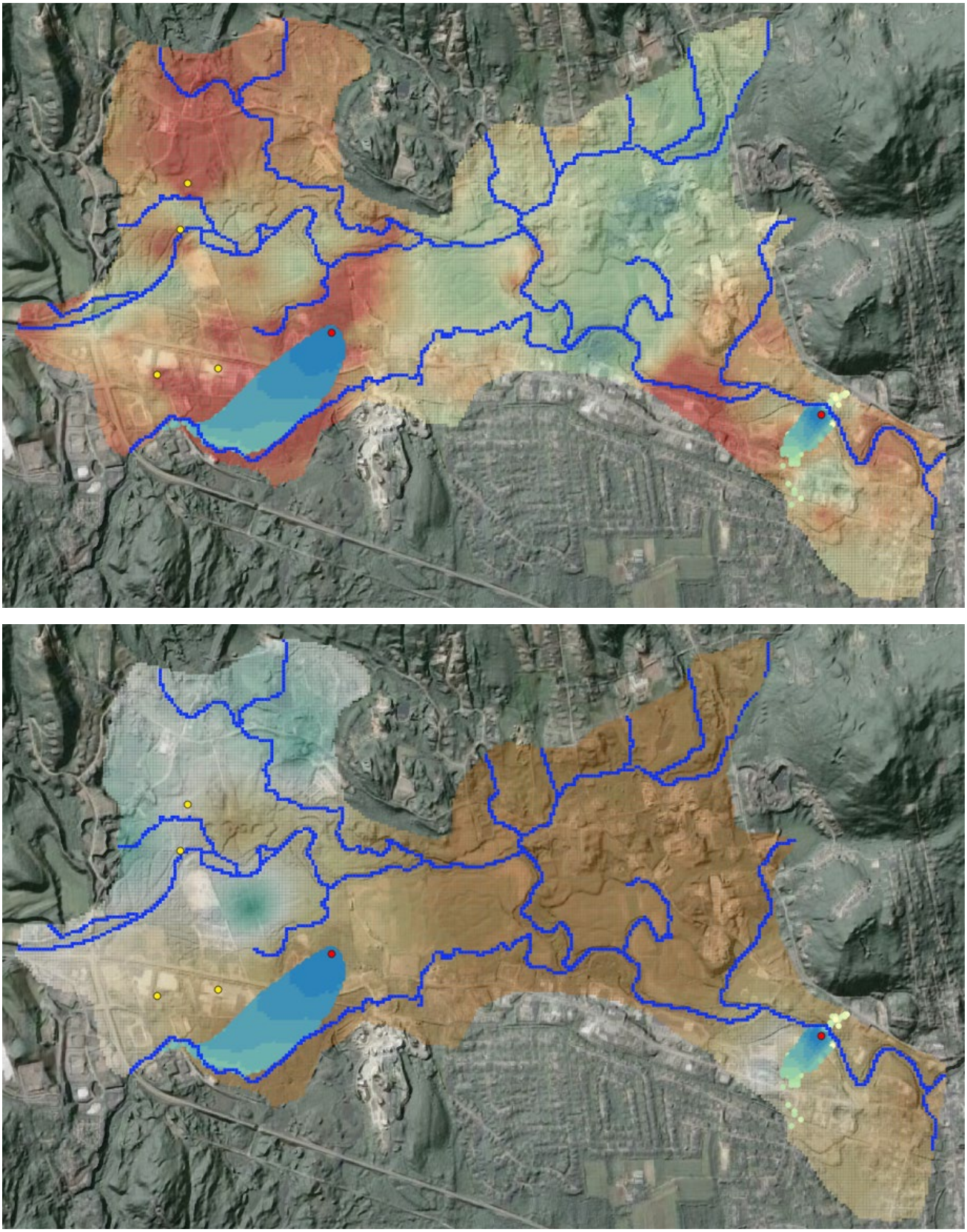

Figure B.4. Contributing area for realisation "50" superimposed on maps of layer 1 hydraulic conductivity (top) and recharge (bottom).

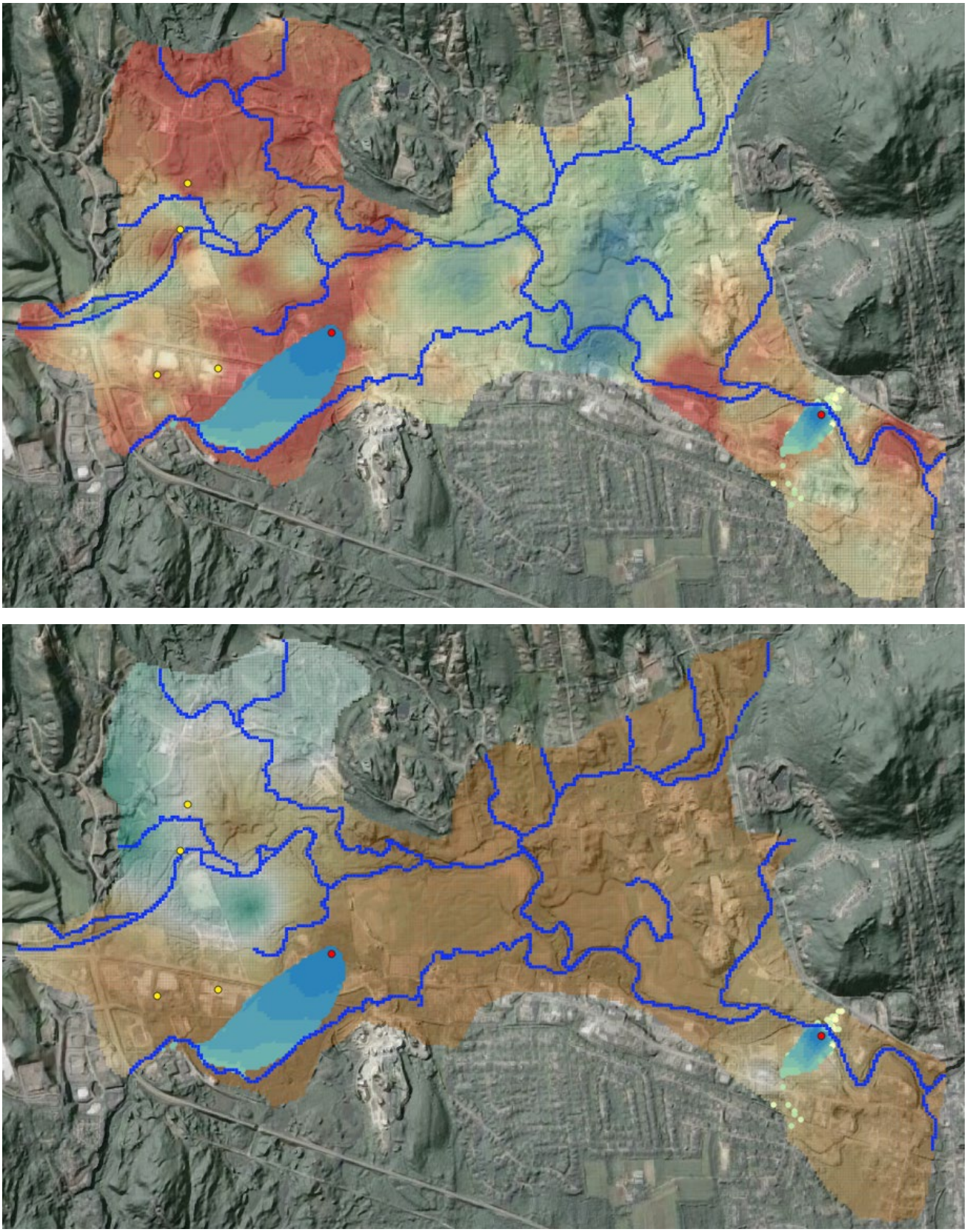

Figure B.5. Contributing area for realisation "100" superimposed on maps of layer 1 hydraulic conductivity (top) and recharge (bottom).

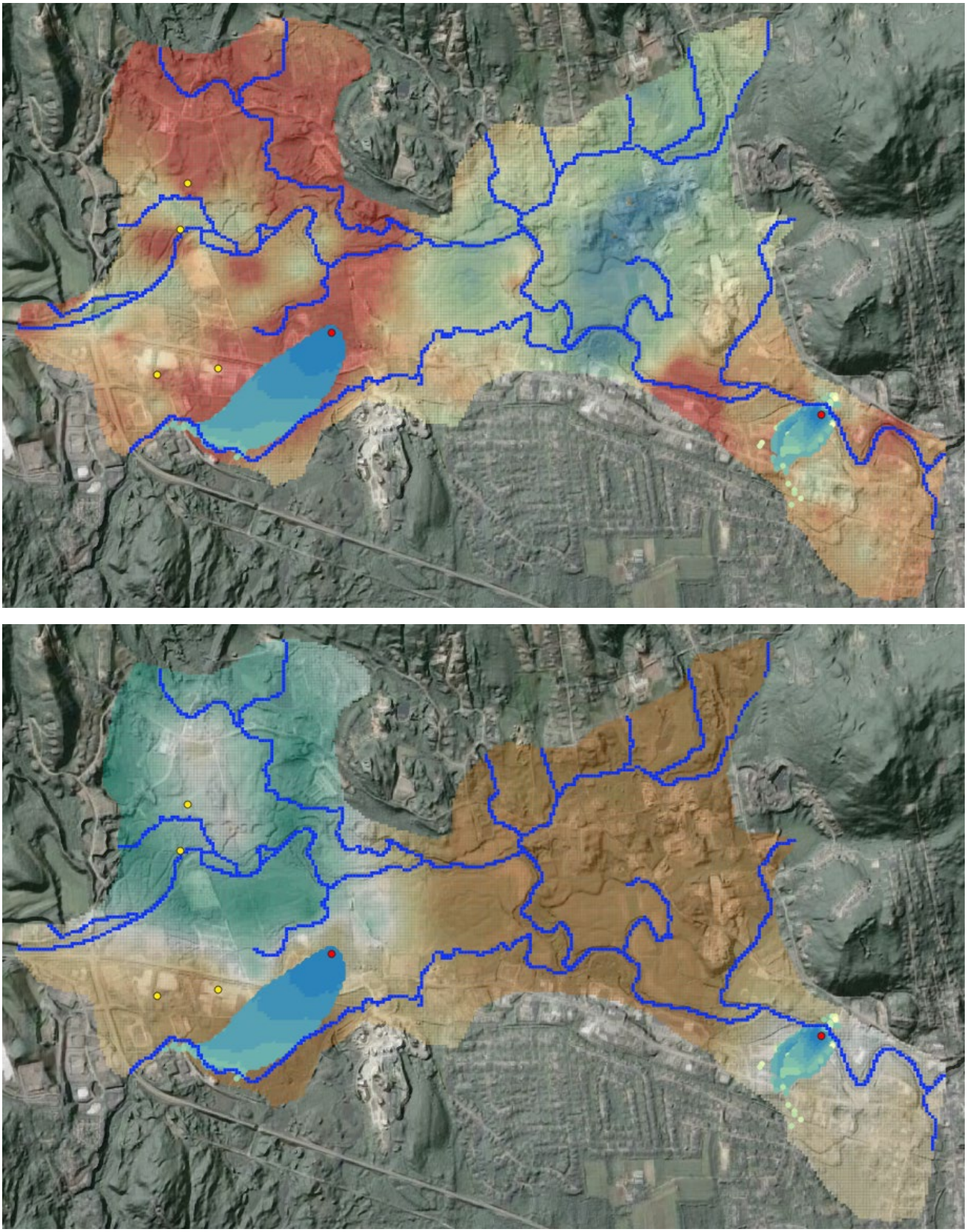

Figure B.6. Contributing area for realisation "125" superimposed on maps of layer 1 hydraulic conductivity (top) and recharge (bottom).

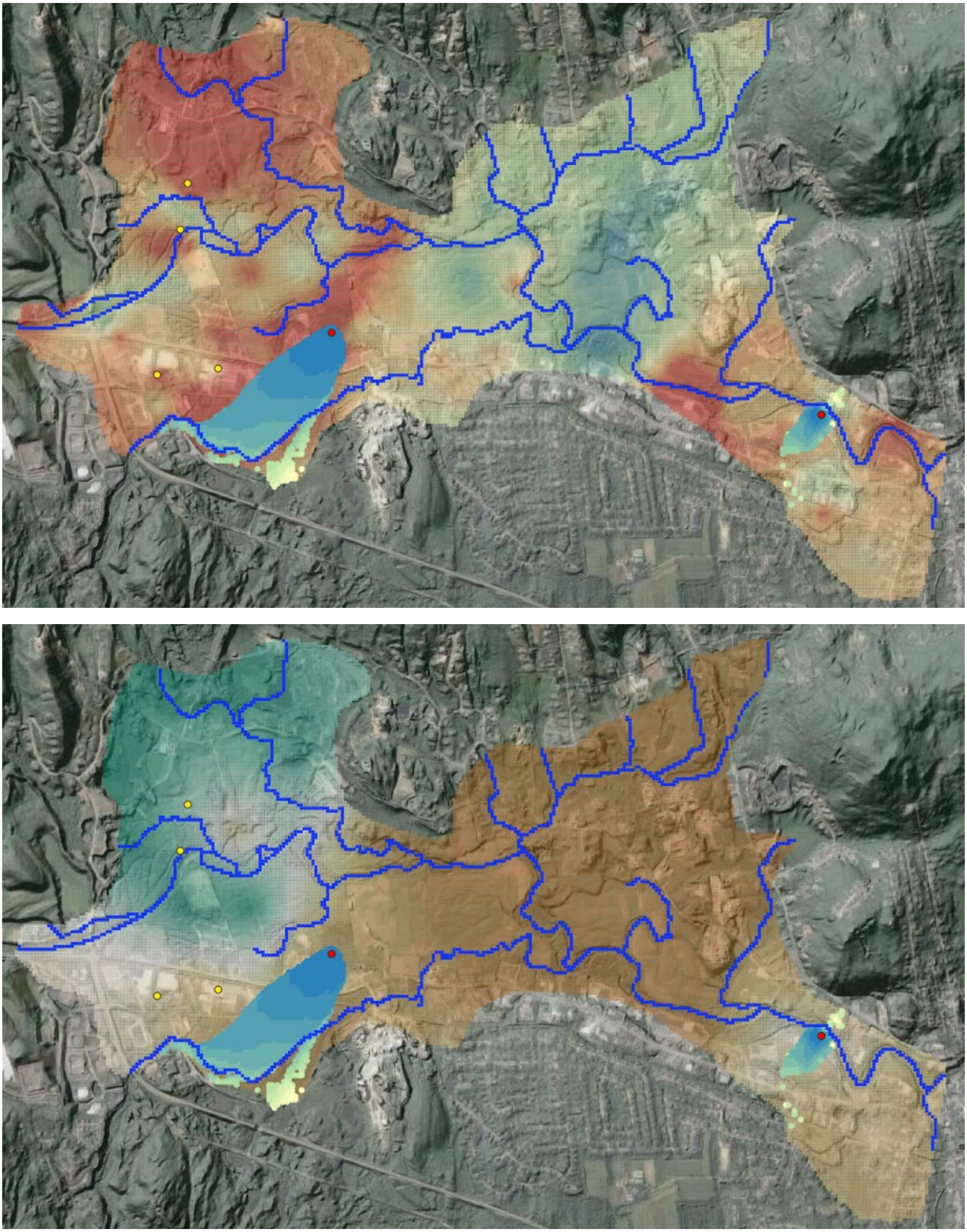

Figure B.7. Contributing area for realisation "base" superimposed on maps of layer 1 hydraulic conductivity (top) and recharge (bottom).

APPENDIX C. CONTRIBUTING AREAS FOR ONE LAYER MODEL

This appendix provides maps of contributing area for the single layer Milford model calculated using 6 realisations from its posterior parameter probability distribution. These realisations were evaluated using the PESTPP-IES ensemble smoother.

Each contributing area is pictured twice. In the first picture, it is underlain by a map of hydraulic conductivity. In the second picture it is underlain by a map of recharge. Colour scales used to display hydraulic conductivity and recharge are the same for all realisations. They are the same as were used previously for the three layer Milford model; see Figure B.1. Calculation of travel time assumes a porosity of 0.05.

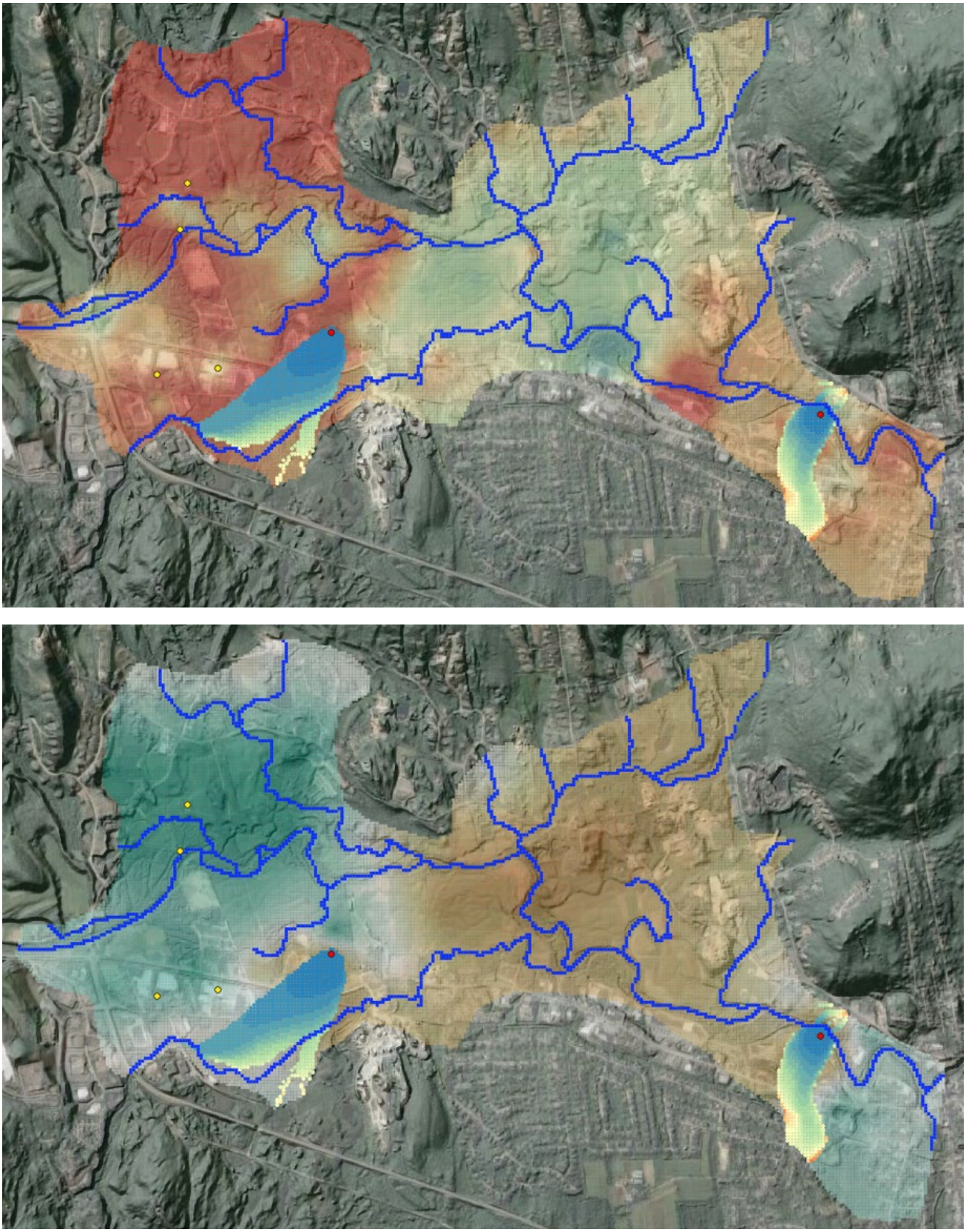

Figure C.1. Contributing area for realisation “0” superimposed on maps of hydraulic conductivity (top) and recharge (bottom).

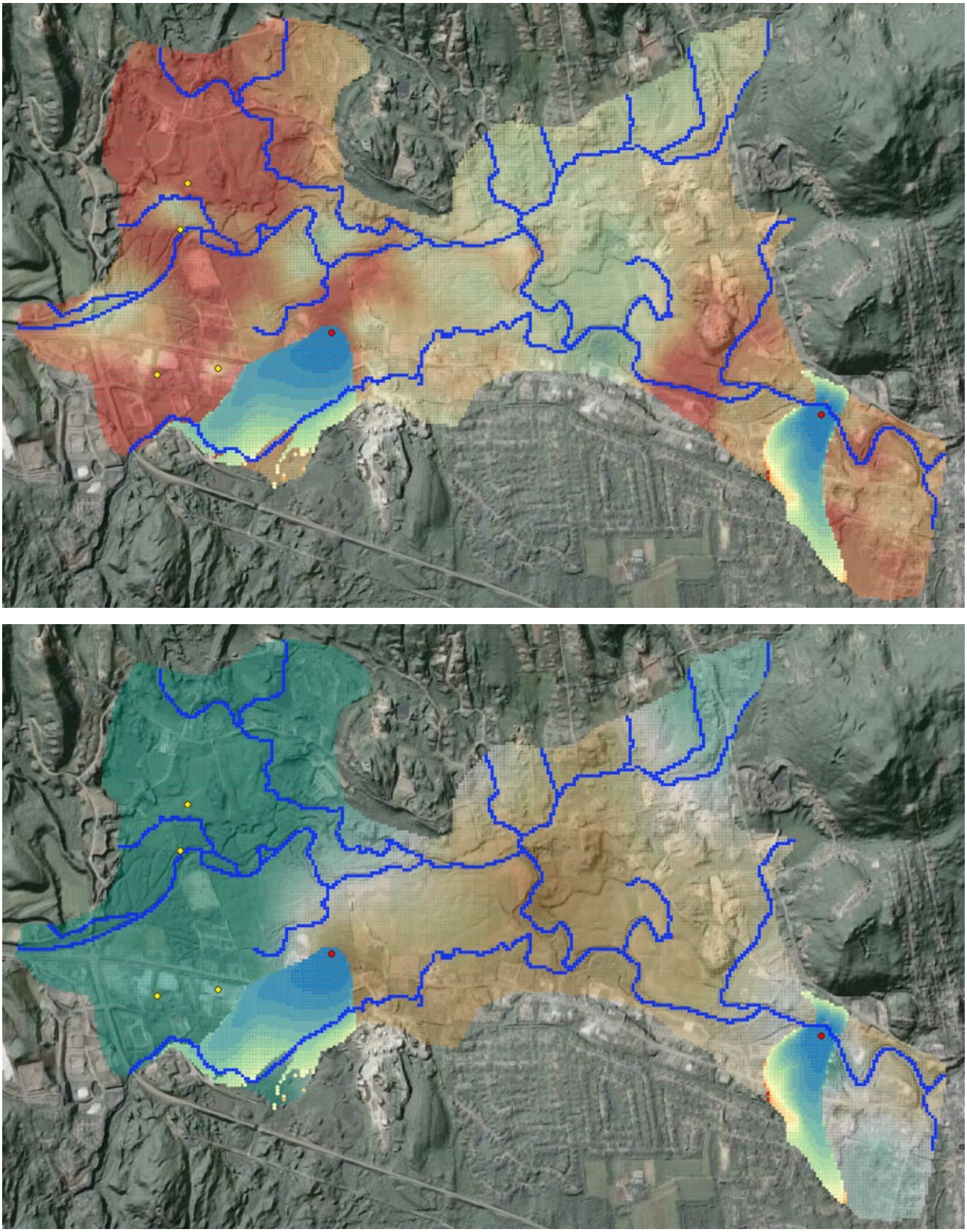

Figure C.2. Contributing area for realisation “25” superimposed on maps of hydraulic conductivity (top) and recharge (bottom).

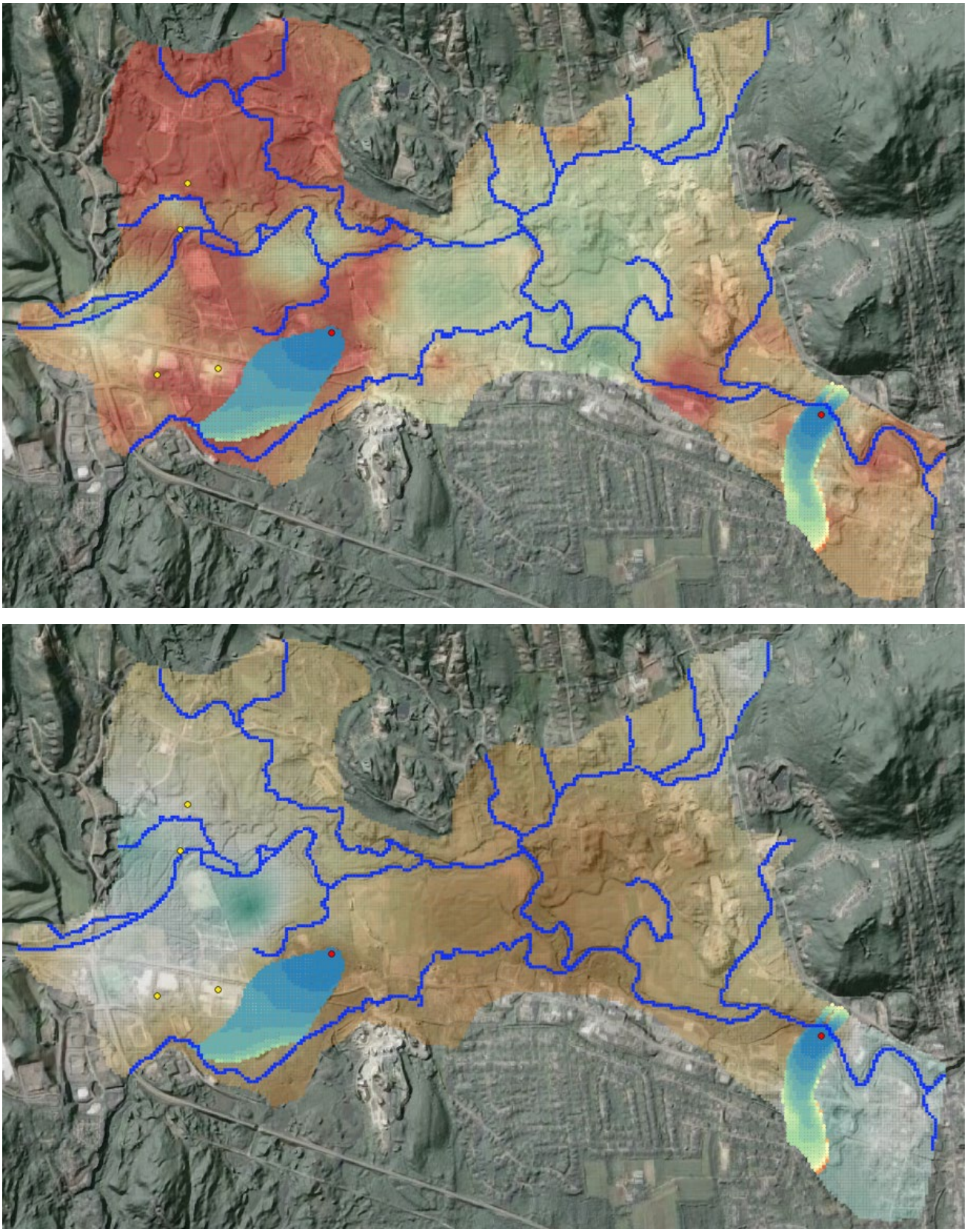

Figure C.3. Contributing area for realisation “50” superimposed on maps of hydraulic conductivity (top) and recharge (bottom).

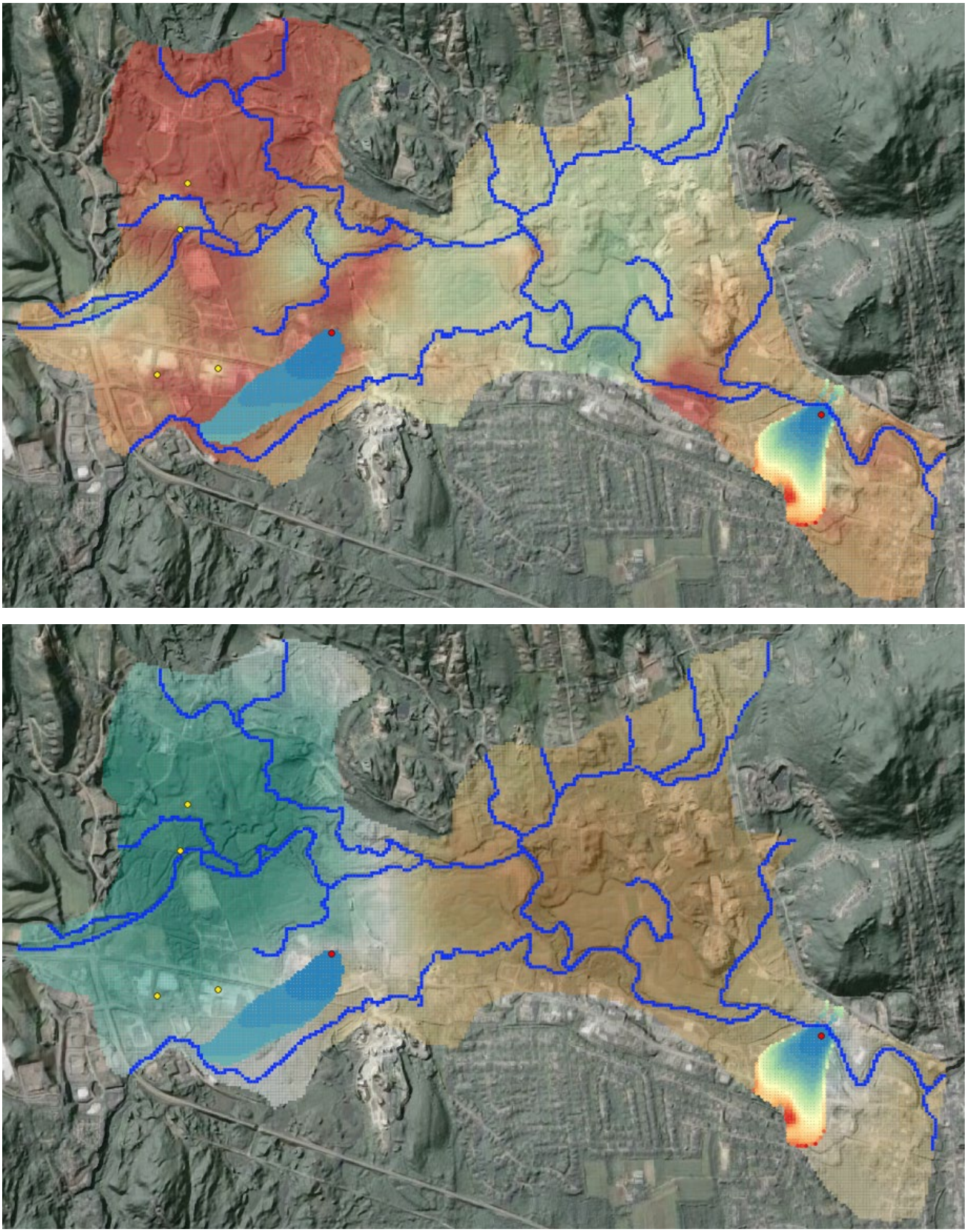

Figure C.4. Contributing area for realisation "100" superimposed on maps of hydraulic conductivity (top) and recharge (bottom).

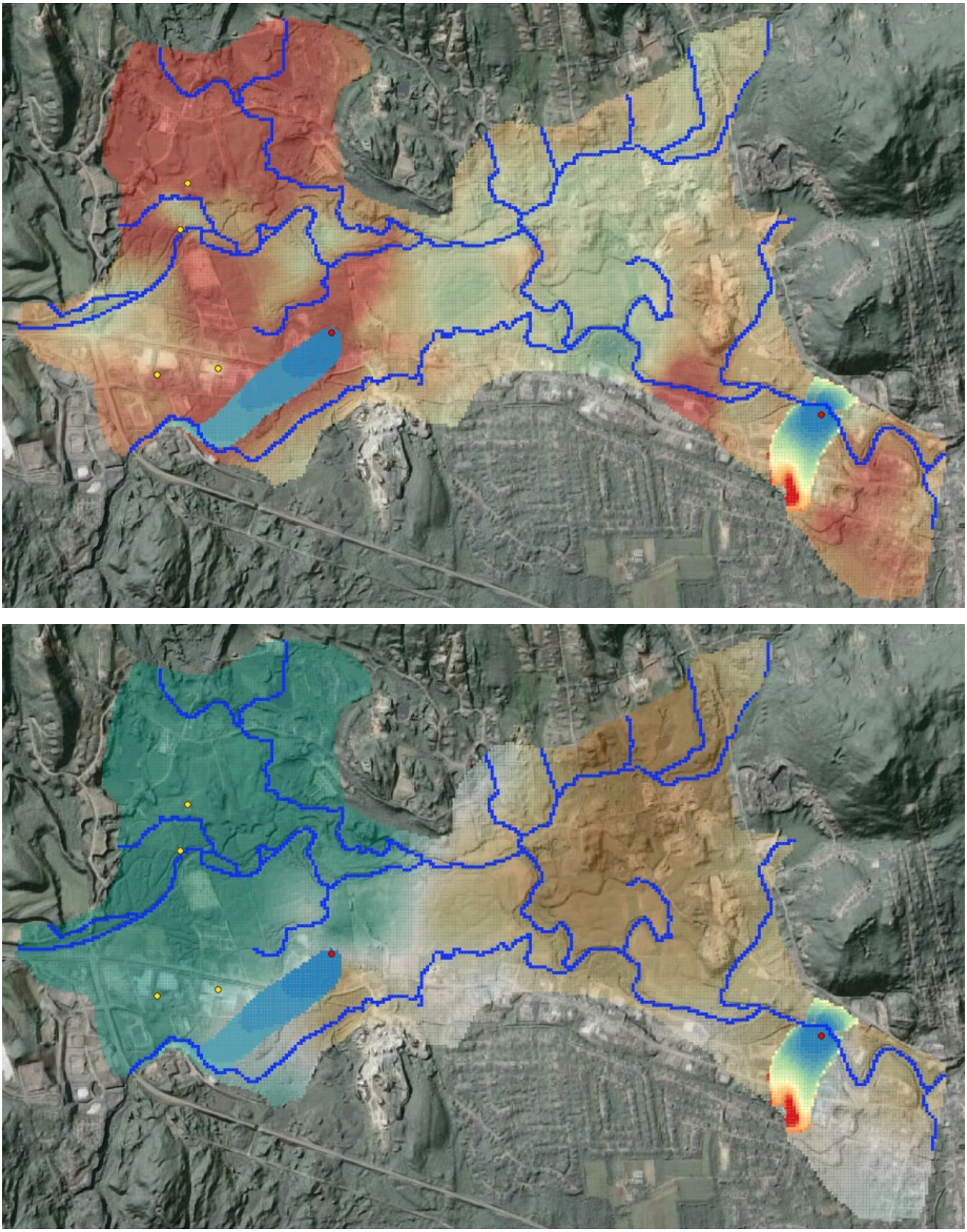

Figure C.5. Contributing area for realisation "125" superimposed on maps of hydraulic conductivity (top) and recharge (bottom).

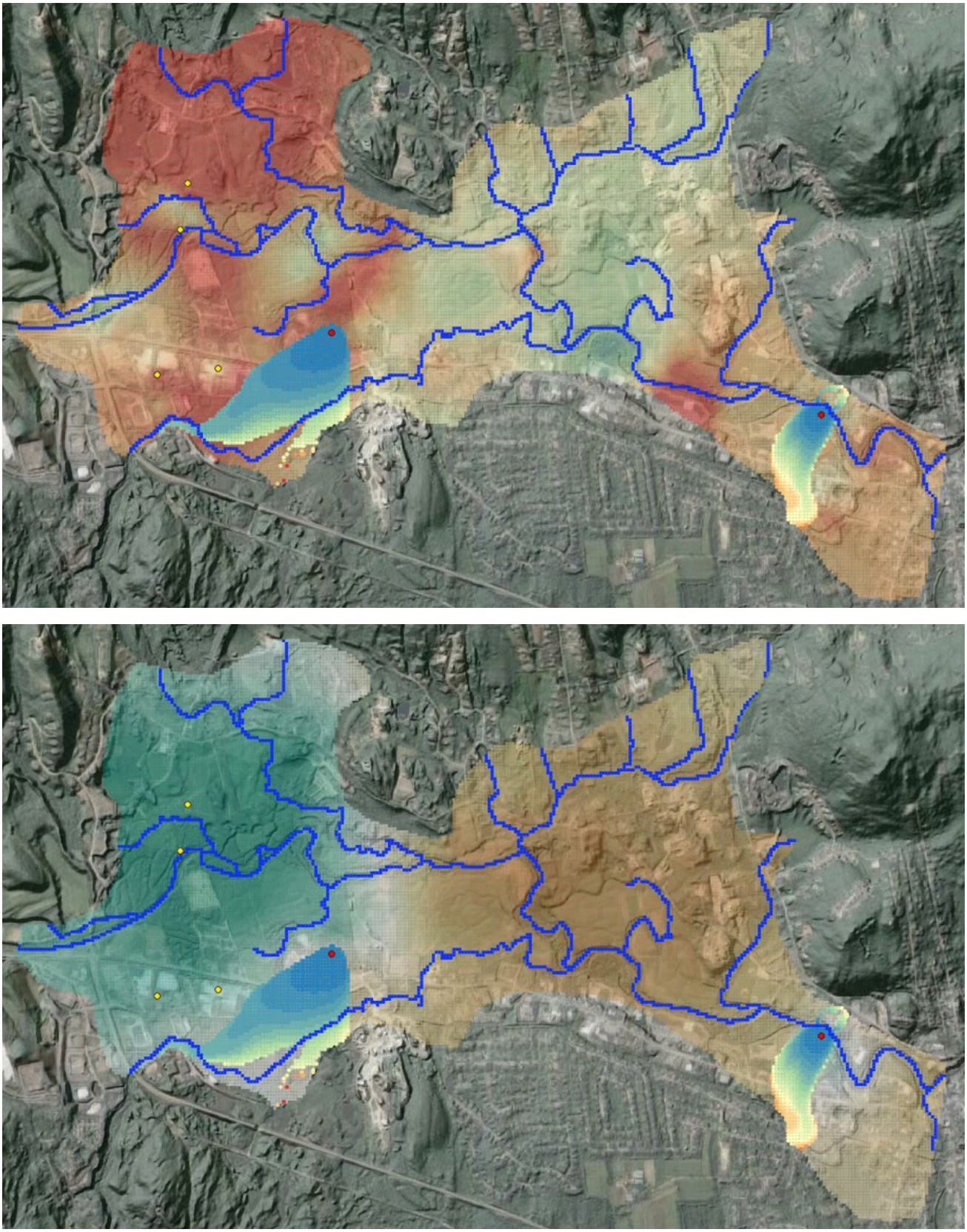

Figure C.6. Contributing area for realisation “base” superimposed on maps of hydraulic conductivity (top) and recharge (bottom).

gmdsi

Groundwater Modelling
Decision Support Initiative

gmdsi.com

CRICOS NO 00114A

BHP

Flinders UNIVERSITY

NATIONAL CENTRE FOR GROUNDWATER RESEARCH AND TRAINING

RioTinto