Eberhard Karls Universität Tübingen

Mathematisch-Naturwissenschaftliche Fakultät

Fachbereich Geowissenschaften

MSc Applied & Environmental Geoscience (AEG)

Master Thesis

Investigations of River-Groundwater exchange fluxes in the Wairau River

submitted by:	Moritz Gosses
First Supervisor:	Prof. Dr.-Ing. Olaf A. Cirpka
Second Supervisor:	Dr. Thomas Wöhling
Date of Submission:	March 2015

Affidavit

This thesis was prepared at the Department of Geosciences of the University of Tübingen as a final requirement to obtain a Master of Science degree in Applied Environmental Geoscience. The work was supervised by Prof. Dr.-Ing. Olaf A. Cirpka.

I hereby confirm that I have written this thesis independently, by exclusive reliance on the literature and tools indicated therein. This thesis has not been submitted to any other examination authority in its current or any modified version, and it has not been published.

Tübingen, Germany.

27.03.2015

Moritz Gosses

0 Abstract

The Wairau River is the main source of recharge for the highly conductive gravel aquifer underlying the Wairau Plains at Blenheim, New Zealand. Utilized both agriculturally and municipally, a thorough understanding of the aquifer system is mandatory for sustainable water allocation in the area. Therefore, we investigate the Wairau aquifer with a focus on the river-groundwater exchange fluxes.

A regional steady-state groundwater model for the low-flow summer period 2013-2014 of the system was set up in Modflow. All major incoming and outgoing water sources were implemented as boundary conditions. The Wairau River geometry was implemented in detail by twenty-two surveyed cross-sections. The influence of (braided) river geometry on the exchange fluxes was considered in the model via the Stream-Flow Routing (SFR) package of Modflow. The hydraulic conductivity field of the aquifer was treated conceptionally both in a two-zone model and via Pilot-point parameterization.

Model calibration was undertaken with PEST and AMALGAM. The model was fitted to observed groundwater-levels, low-land spring discharge, and river exchange fluxes derived from differential stream-flow gauging. Uncertainty analysis of parameters and boundary conditions for different model predictions was carried out with linear (PREDUNC) and partially non-linear (Null-Space Monte-Carlo) methods that are part of the PEST suite of utilities.

The resulting models manage to fit most of the data acceptably. The two Pilot Point models can incorporate further information about the aquifer and river into their parameterizations, compared to the simple two-zone model. Uncertainties of parameters and boundary conditions are acceptably low. Model structural problems that prevent fitting two of the observations are recognized and discussed.

Contents

List of Tables

List of Figures

1 Introduction

Groundwater is a resource that is hard to explore due to its location beneath the earth's surface - nonetheless, it is used in different ways by mankind and endangered by its activities as well. To better one's knowledge about quantity and quality of groundwater and about the processes involved, measurement, field test and various studies are undertaken world-wide.

System state models of groundwater use knowledge about the past for predictions about the present and future, further utilizing the information collected. Different models, as well as a magnitude of various calibration techniques exist to undergo this process, making choices hard for users while allowing incredible flexibility for experts. However, both models and predictions are incorrect. Therefore, it is important to incorporate the knowledge about this incompleteness into the modelling and the use of its results. Again, many tools for the analysis of uncertainties are available.

This study is based on an actual regional groundwater system: the Wairau River and aquifer near Blenheim in the Marlborough District, New Zealand. A steady-state model is set up for low-flow summer conditions and three different parameterizations are used. Furthermore, two different calibration methods are employed – PEST and AMALGAM. On basis of two calibrated models resulting out of this analysis, uncertainty analysis of parameters, observations and predictions is carried out with two different methods. In the end, the information gained from the different models, their problems and the results of the uncertainty analysis are summarized in respect to the ability to investigate the river-groundwater interaction in the system.

2 Study Area

2.1 The Wairau system

The Wairau system of Wairau River and aquifer lies near Blenheim in the Marlborough District in the South Island of New Zealand. The area is mainly under agricultural use with vineyards as the predominant type. A large distribution of wells over the Wairau plain for irrigation and municipal purposes makes in-depth understanding of the Wairau system indispensable.

The Wairau aquifer, with a catchment area of around 260 km², is defined as the subsurface formed by the Rapaura Formation gravels which are under influence by the Wairau river (Cunliffe, 1988). The aquifer's extent is governed in the north by the Wairau River, in the south by a clay-composed structure called the Renwick Terrace, and the western boundary is characterized by the intersection of Wairau River and Renwick Terrace. The eastern boundary is less precisely defined, as the aquifer probably extents into the sea. The average aquifer thickness is around 20 meters, thinning out to the east due to the confining wedge of marine clay deposits lying on top of the gravel aquifer. Unconfined in the west and middle, the aquifer is first semiconfined, then confined in the east due to these clays (see Figure 1).

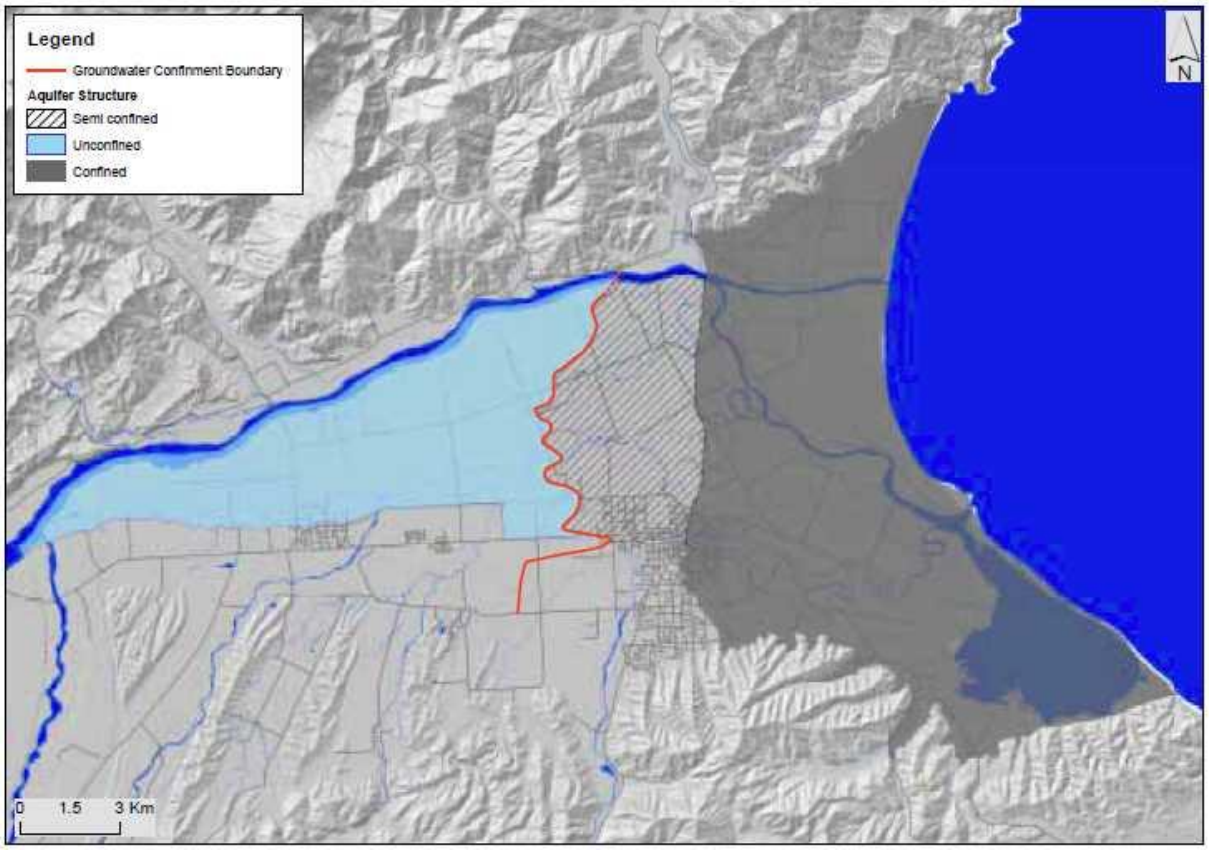

Figure 1: The Wairau Aquifer structure (Cunliffe, 1988)

The Wairau River is the main source of recharge for the Wairau Aquifer, losing around 7 m³/s over the aquifer extent at low-flow conditions. Along the unconfined aquifer, the Wairau River is mostly perched, with water levels and river bed above the surrounding groundwater heads. With the starting confinement in the east, the water levels in the aquifer rise closer to the surface, leading to water fluxes from the groundwater back to the river. Further water input

into the system comes from smaller rivers in the southern region. Water losses, aside from the reflux into the river, are several low-land springs along the confinement boundary, a small flux emerging at the coast and well pumping in the region.

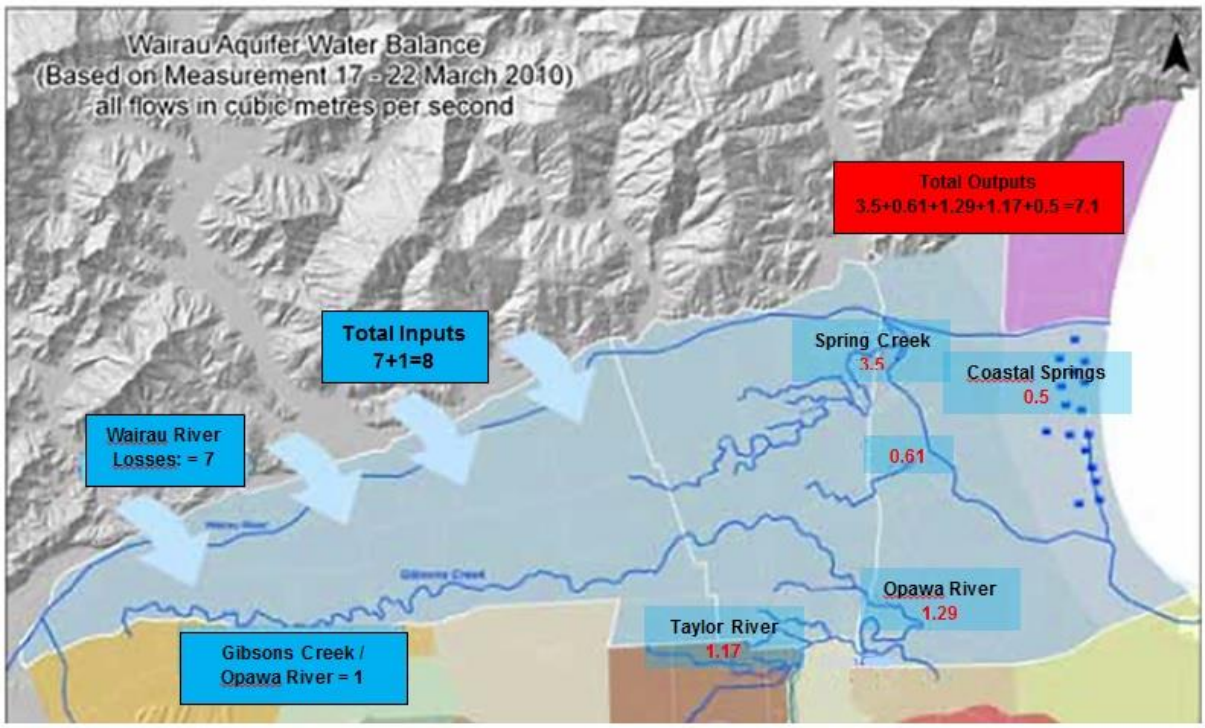

Figure 2: Wairau late-summer water balance (Cunliffe, 1988)

The above Figure 2 shows the system again, with all gains and losses through natural waterways (not including the well pumping over the area).

2.2 Data Availability

For the construction and calibration of a numerical groundwater model different data of the regional aquifer and its boundaries is required. All data described in this chapter was provided by the Marlborough District Council (Peter Davidson, MDC, personal communication).

First, the model geometry, i.e. the upper and the lower aquifer boundary were necessary.

For the upper aquifer boundary, a topological map in GIS of the region's surface was available. This data was taken from a Digital Elevation Model (DEM) created by a LIDAR survey from 23rd to 28th of February 2014 (Peter Davidson, MDC, personal communication). The DEM has a resolution of 2 meters horizontally. The vertical accuracy depends on the slope of the landscape.

Around the river, this data was supplemented with interpolations of various river bed cross-sections, as these gave more accurate information of the river depth, and therefore the aquifer top, along the river. A total of twenty-two river cross-sections were available for the model area, taken from a survey from end of 2012 / beginning of 2013. For this survey, a GPS base station was used and water levels were taken by wading across the river, resulting in a measurement accuracy of around 2 cm. To combine these two data types in GIS, first the DEM data was scaled to the model grid (100x100m cells). Then, the spatial information of the cross-sections was interpolated along the river. This data was also upscaled to the

resolution of the model grid (100x100m cells). Finally, the gridded data of the DEM was overwritten along the river with this new spatial information from the cross-section profiles. This merged spatial information was used as the basis for the model top's geometry in Modflow.

The aquifer bottom's topology, made up from the confining layer limiting the main water-bearing material, was also necessary for the model geometry. Here, a GIS map with interpolations of measurements of aquifer depth was provided, generated from data taken from Cunliffe (1988). Again, the map was interpolated over the model grid to generate the aquifer's bottom extent.

For the eastern constant flux boundary, the flux across this model border had to be estimated, since the border was chosen arbitrarily. The educated guess from MDC (Peter Davidson, MDC, personal communication), based on the overall water balance of the system, suggested that a rather small amount of water leaves the model area over this boundary compared to other outflows. This flux was assumed to be constant at 0.5 m^3/s.

The well abstraction over the model area was also needed for a correct water balance of the system. Due to the size of the area, its mostly agricultural land-use and the number of small individual wells with only sparse information on exact pumping rates and dates, it is almost impossible to simulate each individual well. Fortunately, MDC had some comprehensive information in the overall water abstraction from the groundwater in the model area: a total of 17 different zones with daily estimates of water abstraction in each zone were available as GIS and excel data (see Figure 3 below and Table 24 in the Appendix). These zones were based on the distribution of soil hydraulic properties, and the pumping demand on the corresponding soil moisture deficit (Peter Davidson, MDC, personal communication). Figure 3 only shows a representation of the zones, since a detailed figure would be too exhaustive here. Exact locations of the zones can be found in the relevant Modflow file.

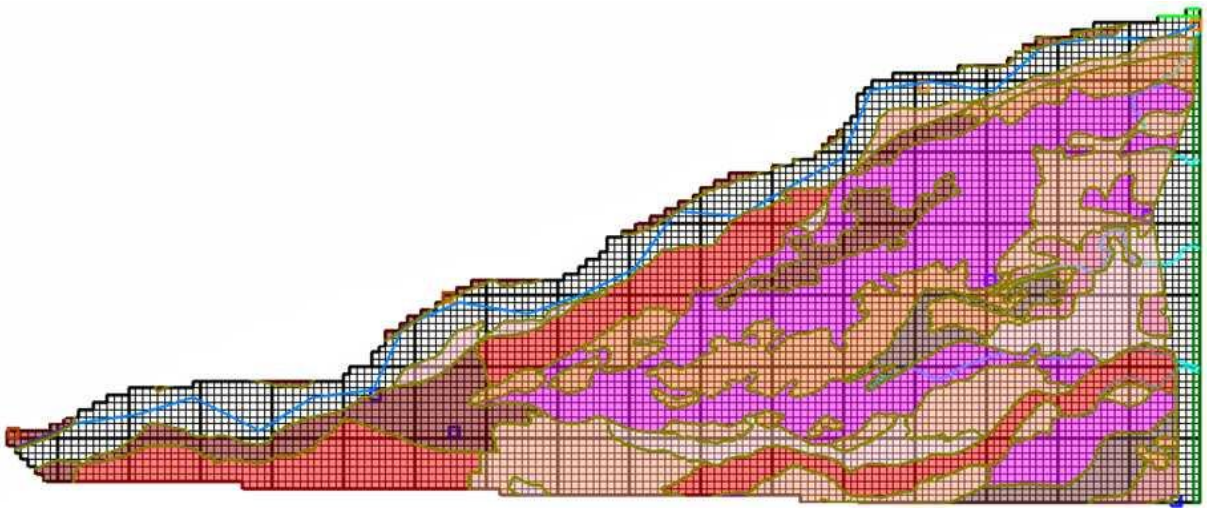

Figure 3: Zones of Well Abstraction

The Wairau River, most important boundary condition in the model, required several different types of data. First of all, the cross-sections, mentioned earlier in their use for the model geometry, were incorporated into the river via the `Stream-Flow Routing`'s flux equations (see 3.1.3.4). The spatial locations of the cross-sections were also used as points for the polyline representing the river in the model, which connected these locations linearly. Along

the river, four different gauging stations were (almost) within the model area: Rock Ferry, State Highway 6 (SH6), Wratt's Road and State Highway 1 (SH1). In reality, the gauging station at SH1 is a little downstream of the model end, but for ease of use it was shifted slightly upstream to serve as the last river point (compare Figure 4).

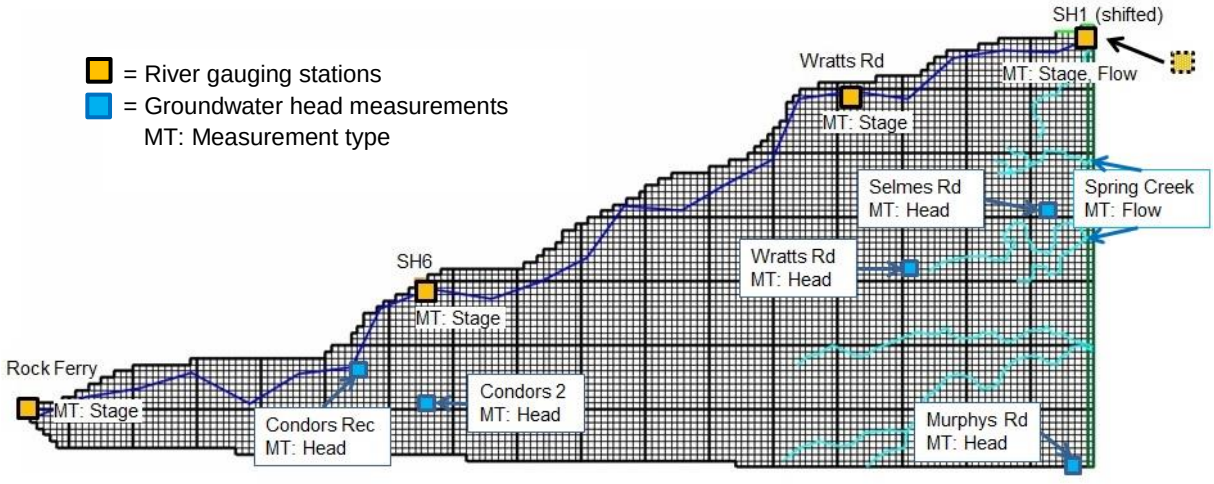

Figure 4: River, Spring & Groundwater measurement data

River stage recorders were available for all four river gauging stations, as well as a flow recorder measuring the river flux at SH1, measuring in 15-minute intervals. See Figure 5 for a plot of the data from 20.12.2013 to 10.04.2014, the time span averaged over for the steady-state model calibration.

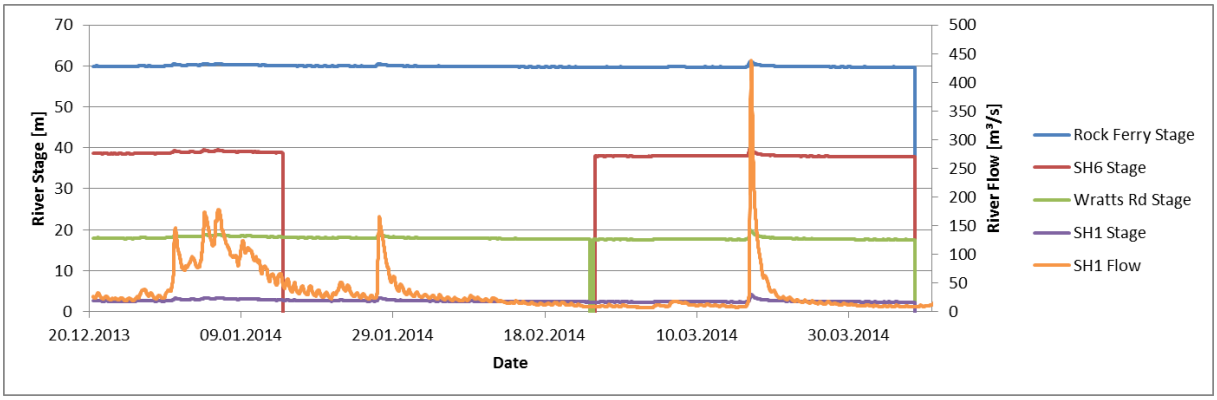

Figure 5: Wairau River data

Regarding the overall water balance of the system, low-flow estimates of the river losses in between the gauging stations were available from several one-time measurements during such conditions (see Cunliffe, 1988). These were later used as observations for the model calibration process (see 3.1.3.4 and 3.2.3 for further details).

For the low-land springs, data regarding the topographic locations was sparse, so a geo-referenced map was used as an orientation (see 3.1.3.3). Spring-flow data was available for Spring Creek (see Figure 4 and Figure 6), though at a location outside the model area. Due to the thinning of the aquifer towards the eastern boundary, with a simultaneous thickening of the aquitard, it was assumed that most water entering Spring Creek leaves the aquifer in the model area. Therefore, the Spring Creek flux data, with an average flow of 3.13 m³/s, was

used as an observation of the total Spring Creek flux in the system. Measurements at the station were taken in fifteen-minute intervals.

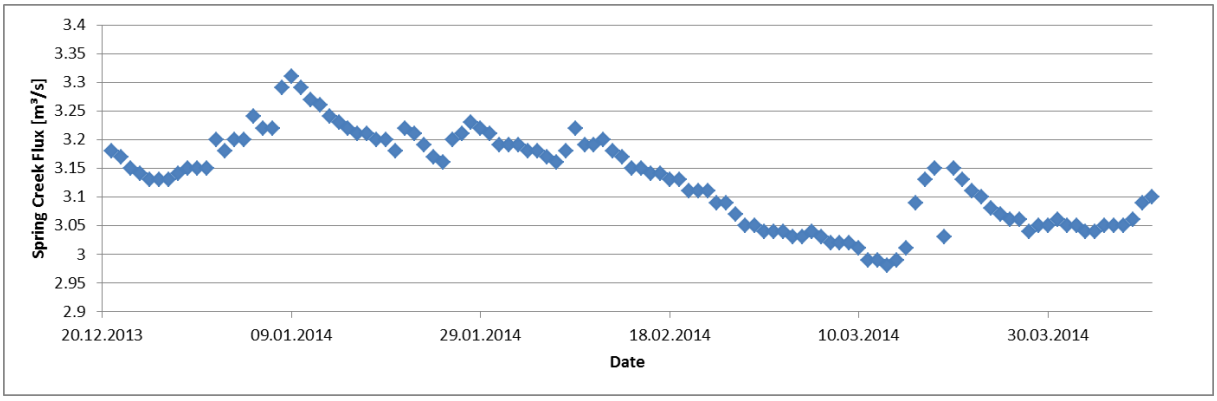

Figure 6: Spring Creek flux data (daily averages)

While covered with wells for small-scale pumping, actual water-level measurement wells in the model area were thinly spread. Data from a total of five groundwater head measurement stations was given as observations for the model calibration. Here, loggers measure the ground-water level in fifteen-minute intervals. See Figure 4 for their location and Figure 7 and Figure 8 for the measured data during the calibration interval.

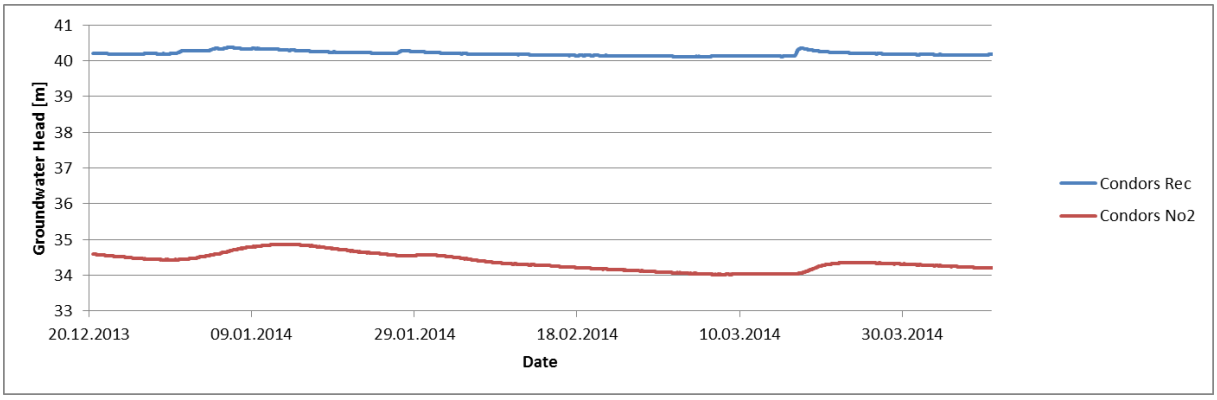

Figure 7: Groundwater head data 1: Condors Recharge and Condors No2 wells

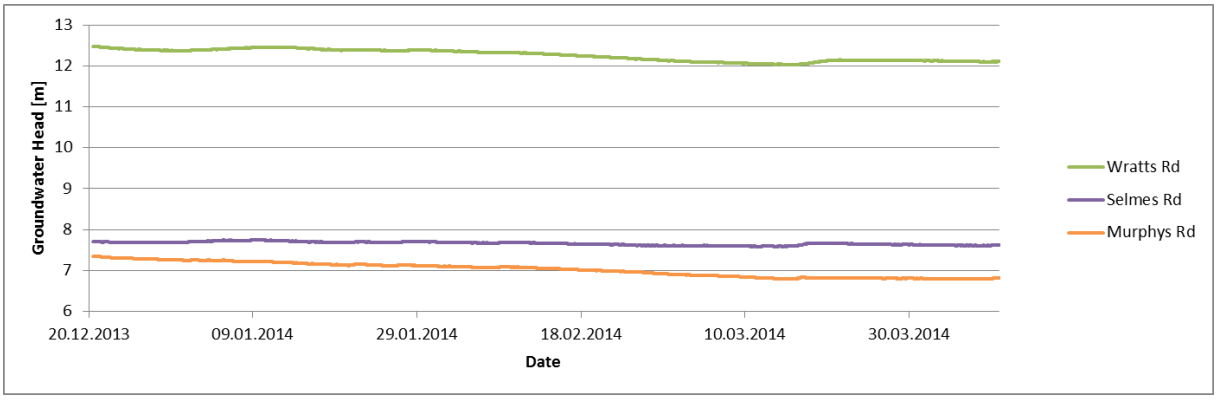

Figure 8: Groundwater head data 2: Wratts Road, Selmes Road and Murphys Road wells

Finally, several aquifer tests had been conducted, yielding approximate information on the hydraulic conductivity of the aquifer material. The aquifer tests located in the model area (see 3.2.3.1 and Table 1) were taken from a MDC data base, submitted there by a variety of different people (both MDC and private) at various times. This information was included into the model via limiting intervals for the corresponding hydraulic conductivity parameters.

Table 1: Aquifer test data

Aquifer Test	x [cells]	y [cells]	HK Estimate [m/d]
1	70	55	188
2	89	69	755
3	79	46	876
4	83	60	875
5	52	63	1000
6	52	58	767
7	101	39	1162
8	18	65	777

3 Methods and tools

This chapter of the thesis explains the different methods and tools used. In the beginning, the setup of the Modflow groundwater model is detailed: the graphical user interface and the model geometry are shortly introduced, followed by a comprehensive presentation of the different Modflow packages representing various model boundary conditions. With the setup complete, the model calibration with two different techniques is described, focusing on the PEST parameter estimation and a short introduction to the global optimization. To end the chapter, the methods for uncertainty analysis used in this thesis are portrayed.

3.1 Setup of the Modflow model

3.1.1 Graphical User Interface ModelMuse

For this thesis, ModelMuse (Winston, 2009) was used as a graphical user interface (GUI) for Modflow (Harbaugh, 2005). ModelMuse is a free GUI from the U.S. Geological Survey, available with all related programs and documentation from the website (http://water.usgs.gov/ogw/modflow/). The use of a GUI allows building a Modflow model via the data input through arrays, polygons and other aids, creating the corresponding Modflow files on the basis of this data. This greatly simplifies the task, as information exported from GIS as arrays or polygons can be directly imported into ModelMuse and linked with Modflow packages to represent boundary conditions or other model-necessary data. Detailed information on ModelMuse's mode of operation would be too exhaustive here and can be found on the website.

3.1.2 Model geometry

The outlines of the Model geometry generally follow the Wairau aquifer's extent as shown in 2.1. The model area was translated into a Modflow-usable network; specifications of the grid can be seen in Table 2.

Table 2: Modflow Grid specifications

Grid size	170 x 70 cells
Cell size	100 x 100 meters
Number of Layers	20
Layer step-size (as % of top-bottom)	2, 4, 6, 8, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 80, 90, 100

Due to the necessity for a rectangular grid, parts outside the aquifer extent are included in this area. These cells were set inactive in the model, i.e. not part of the computation (see Figure 9). The outer model boundaries along the grid's edges correspond to the Wairau aquifer's boundary: north of the Wairau River, along the southern boundary and at their intersection in the west, the model is framed in no-flow boundaries. Only the eastern boundary is open for flux (see 3.1.3.1).

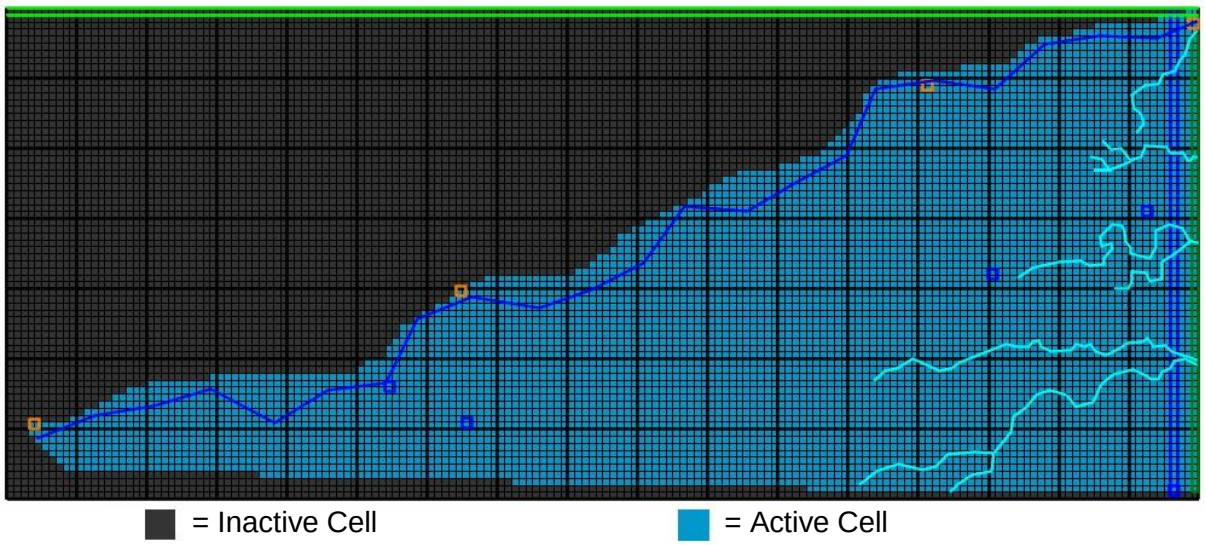

Figure 9: Active Model Area

The vertical geometry of the aquifer was taken from data provided (see 2.2). The data for top and bottom elevation was interpolated onto the regular grid used in the model, generating the upper and lower boundary of the aquifer. Then, the twenty layers were created by vertical interpolation between the upper and lower elevation. The step-size used was non-uniform (see Table 2 for ratios) to generate thinner layers at the model top and thicker layers at the bottom. This setup was used to get a higher resolution of likely water-level fluctuations in the top layers, while the bottom layers are always saturated, allowing for a lower resolution.

3.1.3 Modflow packages: boundary conditions

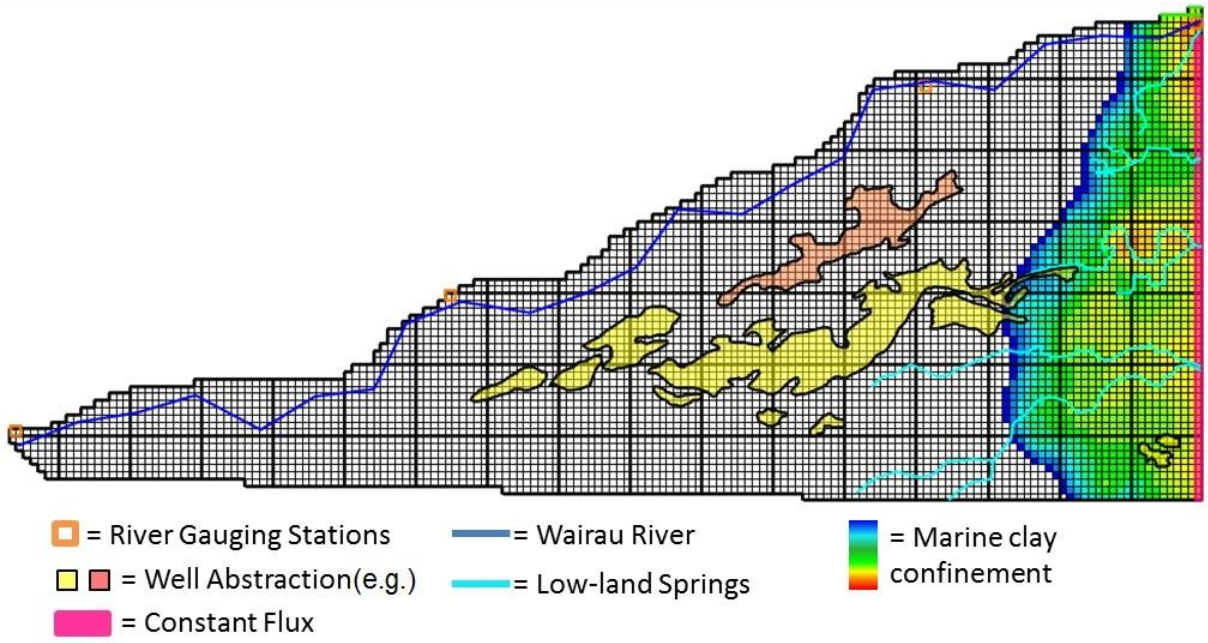

Figure 10: Model boundary conditions

3.1.3.1 Flow and Head Boundary package: constant flux boundary

The constant flux boundary along the eastern edge of the model was implemented using the `Flow and Head Boundary package` in Modflow. This package allows creating either a constant head or a constant flux boundary – in this case, the second option was chosen. A

constant flux boundary removes water with a certain specified flux for each assigned cell from the model.

The boundary goes over the whole horizontal extent of the eastern model edge from below the river to the south-eastern corner. Vertically, it stretches over all twenty layers, assuring a complete coverage of the model boundary. The total assumed flux of -0.5 m³/s was divided by the number of layers and then implemented into the package as a “Total flux per layer”. Internally, ModelMuse then divides this flux evenly over all affected cells in each layer, resulting in an equal flux in each cell along the boundary summing up to the total assumed flux.

3.1.3.2 Well package

The Modflow `Well package` was used to simulate water extraction from the aquifer over its horizontal extent, contrary to the representation of fluxes over vertical boundaries at the model edges with the `Flow and Head Boundary package`. In this package, water extraction is assigned to cells (individually or as areas) of a certain model layer. From these cells, as long as they are saturated, water is removed (i.e. “pumped”) with the specified flux.

As mentioned in 2.2, the average well abstraction rates for different zones in the model area were available. The corresponding zone polygons from GIS were directly imported into ModelMuse. To each of these zones, the average abstraction rate was then assigned in Modflow’s `Well package`. Internally, ModelMuse allocated the flux to each cell overlaid by the polygon. The well depth, characterized in the Model by the specification of the model layer from which the water is removed, was not known specifically due to the representation of the wells as abstract zones. Therefore, all wells were put into the bottom layer 20 to assure the affected cells are always saturated, resulting in the defined fluxes.

3.1.3.3 Drain package: low-land springs

In this thesis, the Modflow `Drain package` was used to model the low-land springs. Drain cells affected by this package allow water to leave the model area dependent on the groundwater level in relation to the drain bed elevation, as well as a flux-resistance factor called drain bed conductance. The flux through the drain cell is calculated by the formula

$$\mathbf{Q_{drain}} = -\mathbf{C_{bed}} * \left(\mathbf{h_{aq}} - \mathbf{h_{bed}}\right) [\mathrm{m}^3/\mathrm{s}] \quad (1)$$

where $\mathbf{h_{aq}}$ is the groundwater level [m], $\mathbf{h_{bed}}$ is the drain bed elevation [m] and drain bed conductance is

$$\mathbf{C_{bed}} = \frac{\mathbf{K_{bed}} * \mathbf{L} * \mathbf{w}}{\mathbf{m}} [\mathrm{m}^2/\mathrm{s}] \quad (2)$$

In this formula, $\mathbf{K_{bed}}$ is the drain bed conductivity [m/s], $\mathbf{L}$ is the drain length [m], $\mathbf{w}$ is the drain width [m] and $\mathbf{m}$ is the drain bed thickness [m].

The locations of the low-land springs in the model area were estimated: first, a topographic map of the model area with the springs delineated was geo-referenced onto the model area in ModelMuse, using known locations like the river gauging stations for reference. Then, polylines were drawn in ModelMuse following the spring markings in the map. Coupling these polylines with the `Drain package`, each intersected cell then represented a drain cell.

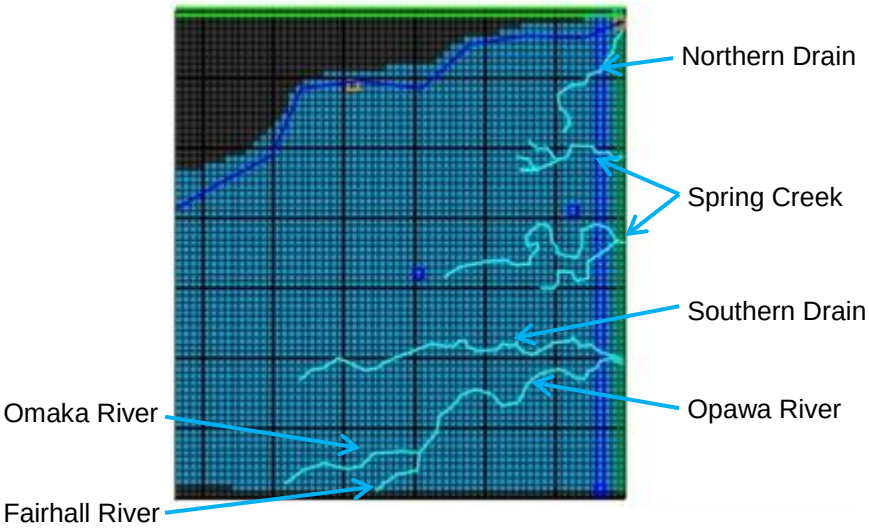

Figure 11: Spring Locations

Drain cells need vertical information assigned to them, as the drain bed elevation h_{bed} governs the gradient (and therefore amount) of the drain flux. Since the model grid with its 100x100m resolution was too coarse to reproduce the indentation of the spring beds in the real topology into the model topology, the model surface along the drain cells was deemed too high to lay the drains on top of it. Since altering the model geometry was not easy and this inaccuracy had no further negative effect on the model, it was instead chosen to simulate the relatively lower spring elevation by assigning the drain cells not to the top of layer 1 but to a deeper layer. It was assumed that the spring channels are around 1.5m below the surface in this area. This depth corresponds roughly to the bottom of layer 4 in this part of the model domain. Therefore, the drain cells were assigned to the bottom of layer 4. Inside ModelMuse, the drain bed elevation was then automatically set to the elevation of the bottom of layer 4 in each cell.

Furthermore, the drain bed conductance $\mathbf{C_{bed}}$ or its composing variables had to be determined. Since the drain length $\mathbf{L}$ in each cell could be taken from the automatic calculation of the intersecting length of the polyline through the cell, it was decided to use this information: therefore, $\mathbf{C_{bed}}$ was calculated. All three remaining variables were unknown, though. As their influence on $\mathbf{C_{bed}}$ is proportional, two of the variables ($\mathbf{w}$ and $\mathbf{m}$) were set to a rough estimate (see Table 3), while the third, drain bed conductivity, was estimated in the model calibration process.

Table 3: `Drain package` settings for low-land springs

Drain bed elevation ($\mathbf{h_{bed}}$) [m]	Layer 4 Bottom
Drain length ($\mathbf{L}$) [m]	Length of polyline intersection
Drain width ($\mathbf{w}$) [m]	5
Drain bed thickness ($\mathbf{m}$) [m]	0.2
Drain bed conductivity ($\mathbf{K_{bed}}$) [m/s]	One parameter for each spring

The available flux measurement data for Spring Creek was also implemented into the model. With a steady-state flow of 3.13 m³/s over the calibration interval, flux into Spring Creek amounts for a significant water output in the overall model budget. A tool called `bud2hyd` from the Groundwater Utility Tools available with PEST (Doherty, 2013) was used to read out

the total flow over all drain cells amounting for Spring Creek. This was then used as an observation for the model calibration (see 3.2.3.2).

3.1.3.4 Stream-Flow Routing package: Wairau River

From the selection of different Modflow packages for river simulation, the `Stream-Flow Routing package` (`SFR package`) was chosen to represent the Wairau River in the model. The `SFR package` routs the water flow through a series of continuous cells, while interaction with the groundwater (loss or gain of water) affects the flow for each cell. Similar to the `Drain package`, stream leakage is computed via

$$\mathbf{Q_{Loss}} = \mathbf{C_{riv}} * \left(\mathbf{h_{riv}} - \mathbf{h_{aq}}\right) [\mathrm{m}^3/\mathrm{s}] \tag{3}$$

where $\mathbf{h_{aq}}$ is the groundwater level [m], $\mathbf{h_{riv}}$ is the water level in the river [m] and river bed conductance is

$$\mathbf{C_{riv}} = \frac{\mathbf{K_{riv}} * \mathbf{L} * \mathbf{w}}{\mathbf{m}} [\mathrm{m}^2/\mathrm{s}] \tag{4}$$

with $\mathbf{K_{riv}}$ being the river bed conductivity [m/s], $\mathbf{L}$ the river length [m], $\mathbf{w}$ the river width [m] and $\mathbf{m}$ the river bed thickness [m].

For a water level below the river bed bottom, i.e. a perched river, $\mathbf{h_{aq}}$ in Equation (3) is substituted by $\mathbf{h_{bed}}$, the bottom elevation of the river bed. Therefore, between river bed bottom and groundwater table, unit gradient flow occurs with a leakage rate not exceeding the saturated hydraulic conductivity of the material in between. These assumptions are generally sound for narrow unsaturated intervals, which is the case in this model. Further details can be found in Niswonger and Prudic (2005).

Analogous to 3.1.3.3, the river length $\mathbf{L}$ was calculated automatically, the river bed thickness $\mathbf{m}$ was estimated and the river bed conductivity $\mathbf{K_{riv}}$ was parameterized. The value of the `SFR package` lies in the various options available for the computation of the river width $\mathbf{w}$ and the river water level $\mathbf{h_{riv}}$: they include specifying values, Manning's equation, power equations or a table of values relating flow to depth and width.

With the available data, it was decided to use power functions to relate both river head and width to flow according to

$$\mathbf{h_{riv}}(\mathbf{Q}) = \mathbf{a}\mathbf{Q}^{\mathbf{b}} \tag{5}$$

and

$$\mathbf{w}(\mathbf{Q}) = \mathbf{c}\mathbf{Q}^{\mathbf{d}} \tag{6}$$

To generate these formulas, several steps had to be undertaken. This was taken from Wilson and Wöhling (2015). First, the twenty-two cross-sections along the river were used to generate relationships of river head and width to $\mathbf{A}$, the cross-sectional wetted area. For the river gauging station at SH1, simultaneous measurements for both river water level and river flow were available. Therefore, Equation (5) could be produced directly, while Equation (6) was calculated from the cross-sectional area A:

$$\mathbf{Q\ in\ h}(\mathbf{Q}) \rightarrow \mathbf{h}; \mathbf{h\ in\ h}(\mathbf{A}) \rightarrow \mathbf{A}; \mathbf{A\ in\ w}(\mathbf{A}) \rightarrow \mathbf{w}; = \mathbf{w}(\mathbf{Q})$$

Unfortunately, no such (or a not sufficient amount of) information was given for all other cross-sections. Therefore, the $\mathbf{h}(\mathbf{Q})$ relation of SH1 had to be translated onto the other cross-sections for the generation of the respective Equations (5) and (6). On average, normal flow was presumed along the river, leading to (at each time) constant velocity. On these assumptions a relationship of the flow to the area, $\mathbf{Q}(\mathbf{A})$ could be created for SH1. The corresponding areas for the different flows were then put into the $\mathbf{w}(\mathbf{A})$ and $\mathbf{h}(\mathbf{A})$ relationships, resulting in the desired $\mathbf{w}(\mathbf{Q})$ and $\mathbf{h}(\mathbf{Q})$ relationships for each cross-section on the basis of their geometry. Figure 12 shows an example of this work: the measured cross-section with its corresponding stage is depicted in the upper-left picture. The upper-middle and upper-right pictures show the observed and fitted exponential functions gained from this information for the stage-to-area $\mathbf{h}(\mathbf{A})$ and wetted perimeter-to-area $\mathbf{w}(\mathbf{A})$. Below is the translation into the corresponding $\mathbf{w}(\mathbf{Q})$ and $\mathbf{h}(\mathbf{Q})$ relationships.

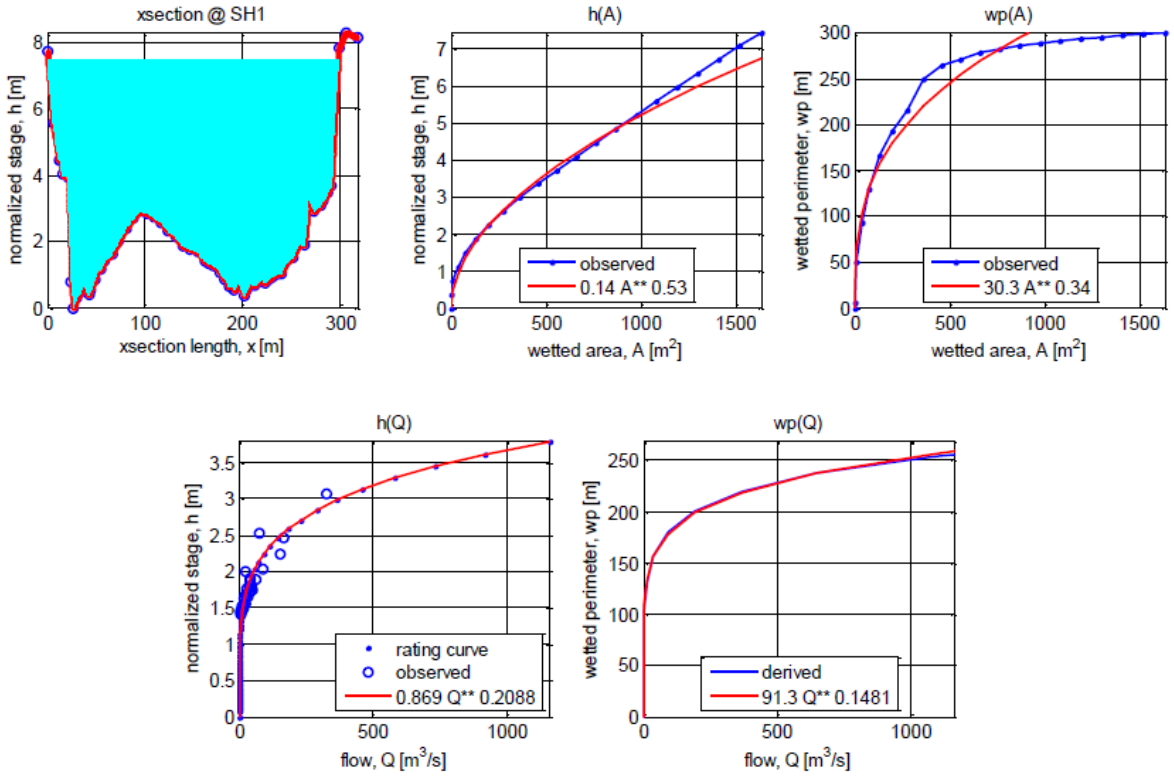

Figure 12: Relationships between stage (h), flow (Q), cross-sectional area (A), and wetted perimeter (wp) at the SH1 recorder site (Wilson and Wöhling, 2015)

With all information for the `SFR package` complete, the river was instated in ModelMuse with twenty-two straight lines connecting the twenty-two cross-sections (and one point manually put in at the model-cut-off river end). For each of the twenty-two cross-sections (river segments in *SFR*), river bed thickness, river bed elevation (at segment start and end), river bed hydraulic conductivity and $\mathbf{w}(\mathbf{Q})$ and $\mathbf{h}(\mathbf{Q})$ power functions were put in. The very first segment also needed information of the inflow into the river. Here, the average measured flux at SH1 plus 7 m³/s (the average loss over the model area) was used. All data input is summarized in the following:

Table 4: `SFR package` settings for Wairau River

Number of river segments [-]	22
River inflow ($\mathbf{Q}$) [m³/s], into segment 1	34.24
River length ($\mathbf{L}$) [m]	Length of polyline intersection
River water level ($\mathbf{h}$) [m]	$\mathbf{h_{riv}(Q) = aQ^b}$
River bed width ($\mathbf{w}$) [m]	$\mathbf{w(Q) = cQ^d}$
Bed bottom elevation ($\mathbf{h_{bed}}$) [m]	From cross-sections, defined at start end and of segment, interpolated in-between
River bed thickness ($\mathbf{m}$) [m]	1
River bed conductivity ($\mathbf{K_{riv}}$) [m/s]	11 Parameters, each one for two consecutive segments

Since river flux loss estimates were available for the low-flow period of the steady-state model, these could be implemented into the model as well. Between all four river gauging stations present in the model area, the exchange fluxes of river and groundwater were estimated (see Figure 13). The flux estimate Q_3, between the Wratts Road and SH1 gauging stations, is problematic in the model due to the shift of the SH1 station in the model area (see 2.2). Consequences of this will be explained in detail in chapter 4.

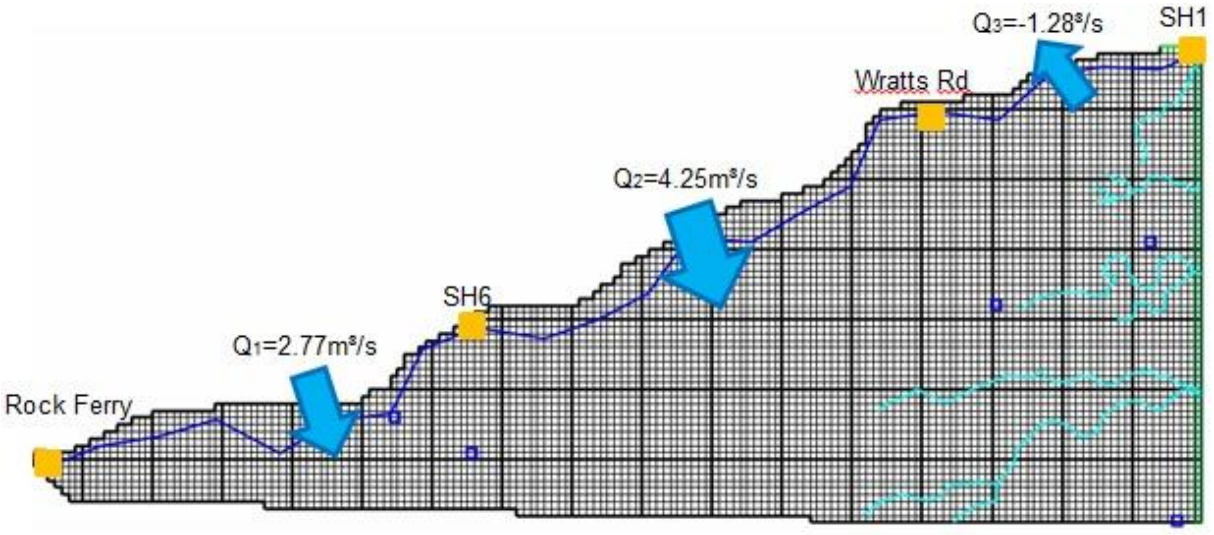

Figure 13: Wairau river exchange fluxes

In the *Wairau.lst* file, the main Modflow output file, specific budget information for each river cell is available and was used to calculate those exchange fluxes in between the gauging stations. This made it possible to use them as observations in the calibration of the model.

3.1.3.5 Head Observation package

The `Head Observation package` of Modflow is used to simplify the output of groundwater head data at specific locations in the model. Placing points into the model at the real-world observation locations and linking them with the package, Modflow automatically generates a file listing the simulated head values at these locations (and the stated observed values plus corresponding residuals, as well). Head data and locations are summed up in Table 5:

Table 5: Head observations

Observation Name	x-Location (in grid)	y-Location (in grid)	Observed Value [m]
Condors Recharge	55	54	40.199
Condors No. 2	66	60	34.370
Wratt's Road	141	38	12.258
Selmes Road	163	29	7.662
Murphy's Road	167	69	7.022

Head observations in Modflow also incorporate a vertical dimension. Similar to the wells, the observation points were built into layer 20 to assure the observation points are always saturated. Since the groundwater head is vertically uniform in the aquifer at each location, no error is introduced by this simplification.

3.2 Model calibration

3.2.1 PEST: Model-independent parameter estimation

The first calibration algorithm used in this thesis is PEST – a model-independent parameter estimator (see Doherty, 20010a). PEST is a free-to-download software package for model calibration, uncertainty analysis and further functions. It comes with a variety of tools aiding its setup, especially regarding Modflow models. In this chapter, PEST's calibration technique and the interface to the model is outlined shortly. Other parts of the software suite used are explained in the relevant chapters.

3.2.1.1 PEST - Theory

For a basic understanding of PEST mathematics, assume a linear system of

$$\mathbf{Xb} = \mathbf{c} \tag{7}$$

where $\mathbf{X}$, a $m \times n$ matrix, represents the model. The vector $\mathbf{b}$ of order n holds the system parameters. The vector $\mathbf{c}$ of size m contains the system outputs, for which corresponding measurements for the determination of the system parameters $\mathbf{b}$ are available.

Under the assumption that $m > n$, i.e. there are more observations than parameters to determine, the elements of $\mathbf{b}$ can be inferred with the elements of $\mathbf{c}$.

An objective function is defined which is the sum of squared deviations between model-generated simulations and measurements. By minimizing this objective function, the optimal parameter set is determined. This objective function is defined mathematically by following equation:

$$\mathbf{\Phi} = (\mathbf{c} - \mathbf{Xb})^{\boldsymbol{t}}(\mathbf{c} - \mathbf{Xb}) \tag{8}$$

The "t" superscript stands for a matrix transpose operation. Now, the optimal parameter set would be the vector $\mathbf{b}$ which minimizes $\mathbf{\Phi}$. This vector $\mathbf{b}$ is given by:

$$\mathbf{b} = (\mathbf{X^t X})^{-1}\mathbf{X^t c} \tag{9}$$

Therefore, the parameter estimation process aims towards obtaining this optimal parameter set $\mathbf{b}$.

This is only the basic theory, though. In practice, models are non-linear (and therefore need to be linearized), the optimization is iterative and lots of options and setups are available to aid this estimation process. Explaining those would go beyond the scope of this thesis, though, but further information can be found in Doherty (2010a).

3.2.1.2 PEST / Model Interface

As the tag "model-independent" reveals, PEST can be used with any type of model. Therefore, the PEST / Model interface is something that needs to be set up (partly at least) by hand.

The centrepiece of PEST is the *"pest control file"*. This file contains all information relevant for PEST. See Figure 14 for a shortened overview of such a file.

```
pcf
* control data
restart estimation
    58     14      5    466      5
…   …   …   …
* parameter groups
pps          relative    1.00000E-02  0.0001       switch  2.0000    outside_pts
…   …   …   …
* parameter data
hkpp1        log    factor    997.801800      170      1000      pps     1.0000    0.0000      1
…   …   …   …
* observation groups
Heads
…   …   …   …
* observation data
Condors_Rec  4.0199000000E+001  10        Heads
…   …   …   …
* model command line
 Wairau.bat
* model input/output
hk_points.tpl hk_points.dat
Heads.ins Wairau.hob_out
…   …   …   …
```

Figure 14: Shortened example of a *pest control file*

The header *pcf* states that this is a *"pest control file"*. Then, several blocks of data, all with a header starting with an asterisk, follow. The first block, **control data*, states information for the optimization settings, from number of iterations to different control variables to the number of parameters and observations. The segment **parameter groups* contains the different groups of parameters, as well as some settings regarding them. Then, the **parameter data* follows, with information on every single parameter, from starting value to permitted interval. Likewise, the **observation groups* and **observation data* names observation groups and individual observations with their measurement value and weight. Under **model command line*, a batch file containing all program run information necessary for the model run is stated. The last block, **model input/output*, names all model files that are written or read, as well as the associated *pest template file* or *pest instruction file*.

These two file types are the important part in connecting PEST with the model. As it changes parameter values during the calibration, those new values need to be given to the model. This is what the template files do. They contain information on how to rewrite a model input file with the new parameter values. The instruction files help sorting the relevant model

output. With these files, the simulated values corresponding to the observations for the calibration are read from the model output.

For further understanding of this PEST / Model interface, see Doherty (2010a) for in-depth information and the following Figure 15 for a graphical summary.

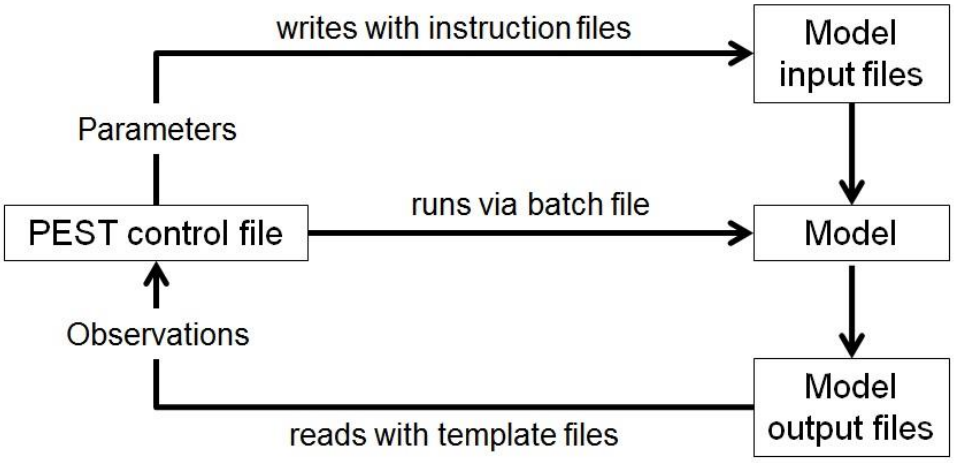

Figure 15: Diagram of PEST **/** Model interface

3.2.1.3 Singular Value Decomposition Assist (SVDA)

PEST has the option to run in a mode called Singular Value Decomposition Assist (SVDA). This method combines Tikhonov regularization (see more on regularization in 3.2.3.3) with truncated Singular Value Decomposition (SVD). Broadly speaking, SVD reduces the dimensionality of the problem by estimating only a limited number of linear combinations of the parameters. SVDA allows using prior information of parameter values (meeting the variogram with Pilot Point parameters, for example) together with the dimensionality-reduction of SVD (and therefore greatly decreasing calibration run time). Since the theory behind these individual methods (and their combination) would be too exhaustive to state here, the reader is referred to the PEST manual (Doherty, 2010a).

3.2.2 Multi-objective Optimization with AMALGAM

The second optimization scheme that was used as part of this thesis is the multi-objective genetic search algorithm named AMALGAM. It was first presented in Vrugt and Robinson (2007). Its method and the application on the Wairau system are shortly explained in the following. All information in this chapter is taken from Vrugt and Robinson (2007) and Thomas Wöhling (WESS, personal communication), who used the method and produced the AMALGAM results kindly utilized in this thesis.

AMALGAM is a multialgorithm and multi-objective genetically adaptive method. It combines the two concepts of simultaneous multimethod search and self-adaptive offspring creation. Where simple genetic methods use a single algorithm to create offsprings of populations and thereby minimizing an optimization problem, the AMALGAM method applies sever algorithms simultaneously and “favor[ing] individual algorithms that exhibit the highest reproductive success” (Vrugt and Robinson, 2007, p. 710). Multi-objective here means that several individual optimization criteria can be used. The algorithm computes a great number of parameterizations and their corresponding objective functions. From this, a Pareto front is created: along this front are all solutions where reduction of one objective function is only possible with an increase in one or more of the other objective functions. This front then

represents the trade-off between the different criteria, with each individual solution a plausible fit for the overall minimization.

In the case of the Wairau model, several steps were taken along the way of implementing AMALGAM for model calibration. First, a single aggregated optimization criterion was tried out, but proven to be unsuccessful. Therefore, it was decided to use five different objective functions for the minimization, all sums of squared errors of their respective observations. For computational purposes, twice two of the objective functions were combined to a single objective function:

- OF1: the two river flux observations and river bed conductivity regularization
- OF2: the four head measurements
- OF3: the Spring Creek flux and the hydraulic conductivity regularization

For more details on the regularization constraints and the implementation see 3.2.3.3. From all of these realizations along the Pareto front, a single model realization was used in this thesis. It was decided to use the compromise solution, which has the smallest overall trade-off between all three objective functions. Note that this is a subjective decision made on the idea of valuing all three objective functions the same. Other solutions along the Pareto front could be plausibly used as well.

3.2.3 Parameters, observations and prior information

In the following, the different parameters, observations and prior information and their implementation into the PEST / Model interface are explained. All tools used in this application are either part of the PEST software (Doherty (2010a)) or its accompanying Groundwater Support Software (Doherty (2013)).

3.2.3.1 Parameters

For both steady-state model parameterizations, the individual parameters were assigned to five different groups, outlined in Table 6.

Table 6: Parameter groups

Parameter group	Number of parameters	
	Two zones	*Pilot Points*
Horizontal hydr. conductivity	2	40
Aquitard conductivity	1	1
K_{xy}/K_z	1	1
River bed conductivities	11	11
Drain bed conductivities	5	5

As apparent from the above Table 6, the two model calibration approaches only differ in their parameterization of the hydraulic conductivity – either with two (zonal) or forty (Pilot Points) parameters. Both times, the instruction files generate the hydraulic conductivity field necessary for the Modflow model, only in different ways.

As the name suggests, the two zone approach separates the hydraulic conductivity field into two zones (vertically in the middle of the model area, resulting in an eastern and a western zone), both with one hydraulic conductivity value. While this is probably not a good representation of the real-world heterogeneity, it is a simple enough parameterization to investigate contrasting to the complicated Pilot Point methods.

The field was generated from the two parameter values by employing the PEST utility called `int2real`. With this tool, a prepared integer array of the zones was read, as well as a text file containing the hydraulic conductivity values of the two parameters corresponding to the zones. From this, the tool creates a Modflow-ready array containing the hydraulic conductivity values at the right cells.

The second method of parameterization for the steady-state model calibration assumes a medium-scale heterogenetic conductivity field. To generate a field according to this assumption, spatial variability in the hydraulic conductivity had to be resolved in the parameters. A common approach to this problem is using so-called Pilot Points: a number of points are spread over the model area, each representing a single parameter. From these points, the hydraulic conductivity field is generated by interpolation, usually using kriging methods.

In this case, tests with different setups suggested that a total number of 40 Pilot Points were sufficient of introducing heterogeneity into the conductivity field. The Pilot Points were mostly spread evenly over the model area, supported by points at aquifer test locations and several auxiliary points along the edges. See Figure 16 for the exact distribution.

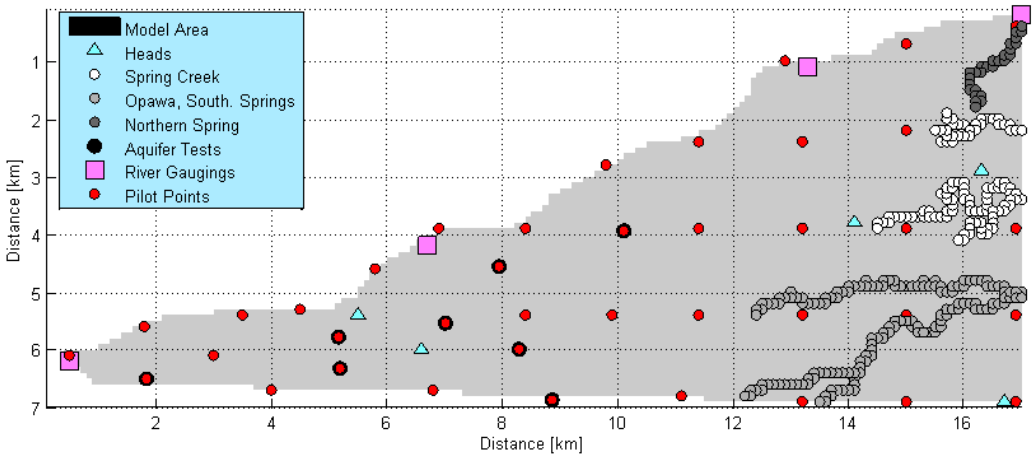

Figure 16: Location of Pilot Points (red dots)

Several of the PEST utilities were specifically designed to aid with the use of Pilot Points in a Modflow model. First, `ppk2fac` was used, a tool generating the kriging factors for the model grid from the Pilot Point locations and a governing kriging structure (see Table 7 for the setup of the kriging). This structure (and its accompanying variogram) was inferred from the aquifer tests. From the factors and the Pilot Point values, `fac2real` created the hydraulic conductivity field.

Table 7: Pilot Point kriging setup

Kriging type	ordinary kriging
Kriging search radius [m]	5000
Minimum number of points used	1
Maximum number of points used	10
Variogram type	exponential
Variogram range [m]	5000
Anisotropy	1.0

For the hydraulic conductivity of the aquitard (marine clay confinement) in the eastern part of the model, a single parameter was used, applying one uniform value to the whole wedge. The aquitard extent was available as a polygon of aquitard depths from the ground, interpolated from various measurements. This information was translated into aquitard depth in model layers, which was used to create arrays for all affected layers (1-12) containing the aquitard extent for this layer. A tool named `int2real` was then used to modify the hydraulic conductivity files for these layers, rewriting the aquitard conductivity into the related cells.

The next parameter, K_{xy}/K_z, is a value containing information of the vertical hydraulic conductivity of the aquifer. Instead of resolving this in a spatially individual way from the horizontal hydraulic conductivity, a single value representing a ratio of horizontal to vertical hydraulic conductivity was used. This parameter was only assigned to the aquifer, so again the array had to be altered with `int2real` to overwrite the value for the marine clay confinement (where a K_{xy}/K_z ratio of 1/1 was assumed).

For the river, eleven parameters of river bed conductivity were used. While the model has a total of twenty-two river segments (see 3.1.3.4), each two subsequent segments got the same parameter. This was deemed flexible enough while keeping the number of parameters down. The parameter values could be directly passed to the relevant Modflow file via a PEST instruction file.

The low-land springs, modelled as drains, were parameterized by five drain bed conductivity parameters, one for each spring. Again, giving these values to the model needed no further tools.

For all parameters explained in this section, starting values as well as minimal and maximal values had to be assigned in the calibration process. In the Pilot Point calibration, a few Pilot Points represented real-world aquifer tests. For those, the aquifer test values were used as starting values, with a possible alteration interval of +/- 25%. Only the aquifer test 1 (with a hydraulic conductivity estimate of around 200 m/d) was treated like a parameter with no information on, since early tests showed that the value is probably an outlier. With no detailed information for the other parameters, starting values were arbitrarily chosen and intervals wide enough to permit freedom of variation were used. Actual values can be found in the corresponding chapters.

To analyse the uncertainty of the assumed boundary conditions, they had to be incorporated into (a calibrated) PEST setup as parameters. This allows for the variation of their values, computing the sensitivity of the observations to these changes. The boundary conditions were inserted similar to the real parameters with accompanying template files to write their values. Table 8 has information on which boundaries were included this way.

Table 8: List of boundary condition “parameters“

Boundary condition	Number of “parameters”
Eastern constant flux	1 (total flux)
Well Abstraction	17 (one for each zone)
River Geometry	88 (22 zones with 4 geometry variables each)

3.2.3.2 Observations and predictions

Similar to the parameters, different observations (and predictions) of several groups were used in the model. Table 9 shows all observations, their grouping, the observed values and their weights.

Table 9: Observations

Observation group	Observation name	Observed value	Weight
Groundwater Heads [m]	Condors Recharge	40.199	10
	Condors No.2	34.370	10
	Wratts Road	12.258	10
	Selmes Road	7.622	10
	Murphys Road	7.022	1 / 0
Spring Flow [m³/d]	Spring Creek	-270222	0.001
River Flow Losses [m³/s]	Rock Ferry – SH6	2.768	25
	SH6 – Wratts Road	4.274	25
	Wratts Road – SH1	-1.282	0.01 / 0

Of the first group, groundwater heads, five observations were used. The measured data at those locations (see 2.2) was averaged to create a single value for the modelling interval. The observations “Murphys Road” proved impossible to fit during early tests, due to its position at the edge of two model boundaries (directly in the south-eastern corner). Therefore, a smaller weight was assigned to this observation at first, rendering it less important during the calibration process. Later, it was excluded completely from the calibration process, its weight set to zero. All head observations could be read out directly from a Modflow output file.

The only spring flow observation used in the calibration process was the flux at “Spring Creek”. Again, the average value of the measured data was used in the calibration process. Since the model used meters and days as standard units, the measured value had to be converted. Then, a small weight was assigned to ensure that expected weighted errors of this observation have the same magnitude as the other observations. This “Spring Creek” flux is no direct output of the model. Instead, the flow losses over the drain cells composing “Spring Creek” had to be read out and summed. This was done using a tool called `bud2hyd`.

Finally, the river flow losses (or gains) in between the four gauging stations were incorporated into the model as observations. Again, weighting made sure that the weighted errors of all observations have the same magnitude. Only the last river flux, between the Wratts Road and SH1 gauging stations, got a significantly smaller weight. Again, early tests showed this flux gain could not be adequately modelled with this model, due to the relocation of the SH1 gauging station (see 2.2). Therefore, at first a smaller weight was assigned to this observation, while later omitting it completely. Again, these values were no direct output of the model. `Obs2obs` was used to read out the river fluxes at the gauging stations and then calculate the flow losses (or gains) in between.

Several predictions were also implemented for the use in the uncertainty analysis. In general, a prediction is any kind of (sensitive) model output, i.e. an observation. In the PEST framework, a prediction is therefore easily integrated, usually with a weight of zero. This ensures it was not utilized in the calibration process. Table 10 shows all predictions in addition to the observations.

Table 10: Predictions

Observation group	Prediction name	Weight
River Flows [m³/s]	SH6	0
	Wratts Road River	
	SH1	
Spring Flow [m³/d]	Southern Streams	

The river flows at the gauging stations downstream of Rock Ferry (where the river flow was used as an input) were read out during the calculation of the river flow losses, so they were incorporated mostly due to their availability. The flow into the two streams in the south, the "Southern Drain" and "Opawa River" (see 3.1.3.3), was read out similar to the "Spring Creek" flux.

3.2.3.3 Prior information

Prior information is a further type of data input into a model calibration that generally works similar to observations: each prior information has a certain error, which is weighted and added to the objective function, and is therefore sought to be minimized. Unlike observations, tough, prior information is data or knowledge about either the parameters or boundary conditions of the model.

In this case, prior information was added as regularization data. In regularization mode, PEST calculates the standard objective function and a regularization objective function (pertaining to the prior information errors). This second objective function is automatically weighted and then added to the standard objective function, resulting in a combined objective function which is minimized. More information on this can be found in Doherty (2010a).

Two prior information groups were used in this calibration: one for the river bed conductivities, the other for the Pilot Points.

For the river bed conductivities, the regularization aim was to favour a homogeneous or gradually changing river bed (as expected to naturally occur), thus penalizing highly heterogeneous sequences of river bed conductivities. This was implemented by calculating the difference of each river bed conductivity to its first three neighbours and adding these differences as prior information to minimize. Differences to closer neighbours were weighted higher in the process. See Figure 17 for clarification.

Figure 17: River bed conductivity – prior information

Prior information regarding the Pilot Point parameters worked similarly: differences to neighbouring Pilot Points were calculated and weighted inversely to the distance of the points. For the Pilot Points, this was done with a tool called `ppkreg`. This tool automatically creates the regularization data on basis of the variogram. Furthermore, the hydraulic conductivity values of the Pilot Points were constrained by a mean value calculated from the variogram of the pumping tests. For each Pilot Point, a regularization observation equal to this mean value was added. Therefore, deviation from this mean was penalized.

3.3 PREDUNC: Predictive uncertainty analysis

In the following, the theoretical background of the predictive uncertainty analysis with the PREDUNC tools and its practical implications are outlined shortly. Further explanations and the full derivation of the stated formula can be found in Doherty (2010b).

Underlying all of PREDUNCs uncertainty analysis is the following formula:

$$\sigma_s^2 = \mathbf{y^t C(p) y} - \mathbf{y^t C(p) X^t [X C(p) X^t + C(\varepsilon)]^{-1} X C(p) y} \quad (10)$$

in which σ_s^2 is the variance of prediction $\mathbf{s}$, $\mathbf{y}$ states the sensitivity of the prediction to the parameters $\mathbf{p}$, $\mathbf{C(p)}$ is the covariance matrix of the parameters' variability, $\mathbf{C(\varepsilon)}$ is the covariance matrix of measurement noise and $\mathbf{X}$ is the model matrix.

This equation therefore expresses the post-calibration predictive uncertainty variance of a certain prediction $\mathbf{s}$ by combining the prediction's pre-calibration uncertainty (the first term on the right-hand side: $\mathbf{y^t C(p) y}$) and the reduction of this uncertainty through calibration (the second term).

For the practical use of this method, the four different variables on the right-hand side of the equation need to be obtained:

- The vector $\mathbf{y}$: the sensitivity of the prediction to the parameters can be gained by including the prediction into PEST as a zero-weight observation. Now, the sensitivities are included in the Jacobian matrix and can be extracted.
- The matrix $\mathbf{X}$: the matrix containing the sensitivities of observations to parameters is calculated in PEST as the Jacobian matrix and therefore readily available.

- The matrix $\mathbf{C}(\mathbf{p})$: the covariance matrix of the variability of the parameters. This matrix is user-built, requiring information on the variability in the form of standard deviations of parameters, for example. For Pilot Point (or other spatially distributed) parameters, the variability can be calculated automatically with a small PEST tool called `ppcov`.
- The matrix $\mathbf{C}(\boldsymbol{\varepsilon})$: the covariance matrix of measurement noise is calculated from the observation weights

Based on this method it is also possible to calculate not only the uncertainty of a prediction, but also the influence that single (or groups of) parameters or observations have on this uncertainty without actually knowing their real values. Therefore, knowledge of parameters or observations can be assumed and the gained uncertainty reduction (i.e. the worth of data) be computed. Furthermore, by including boundary conditions as parameters into a calibrated PEST framework, their influence on the predictions can also be calculated. This is interesting due to the fact that knowledge of boundary conditions is usually uncertain, too.

3.4 Null-Space Monte-Carlo (NSMC): Partially nonlinear uncertainty analysis

The PEST framework provides another way of uncertainty analysis with its tools: the partially non-linear (in contrast to PREDUNC's linear approach) Null-Space Monte-Carlo analysis. Again, the underlying method is only shortly described here, with further information available in Doherty (2010a).

The concept of this analysis is based on the partition of the parameter space into the solution space and the null space, as shown in Figure 18. Basically, this states that the change of a parameter value has an influence on the solution space (and therefore the calibration dataset, i.e. the measurements) and the null space (and therefore NO influence on the measurements). Since the calibration is based on the parameters influence on the measurements, the parameters variability in the solution space is constrained by their measurement error (and therefore the acceptable model-to-measurement fit, i.e. the objective function value). The parameters variability in the null space, on the other hand, is only restricted by the overall parameter restrictions (upper and lower bounds, regularization) as long as their projection onto the solution space stays acceptable.

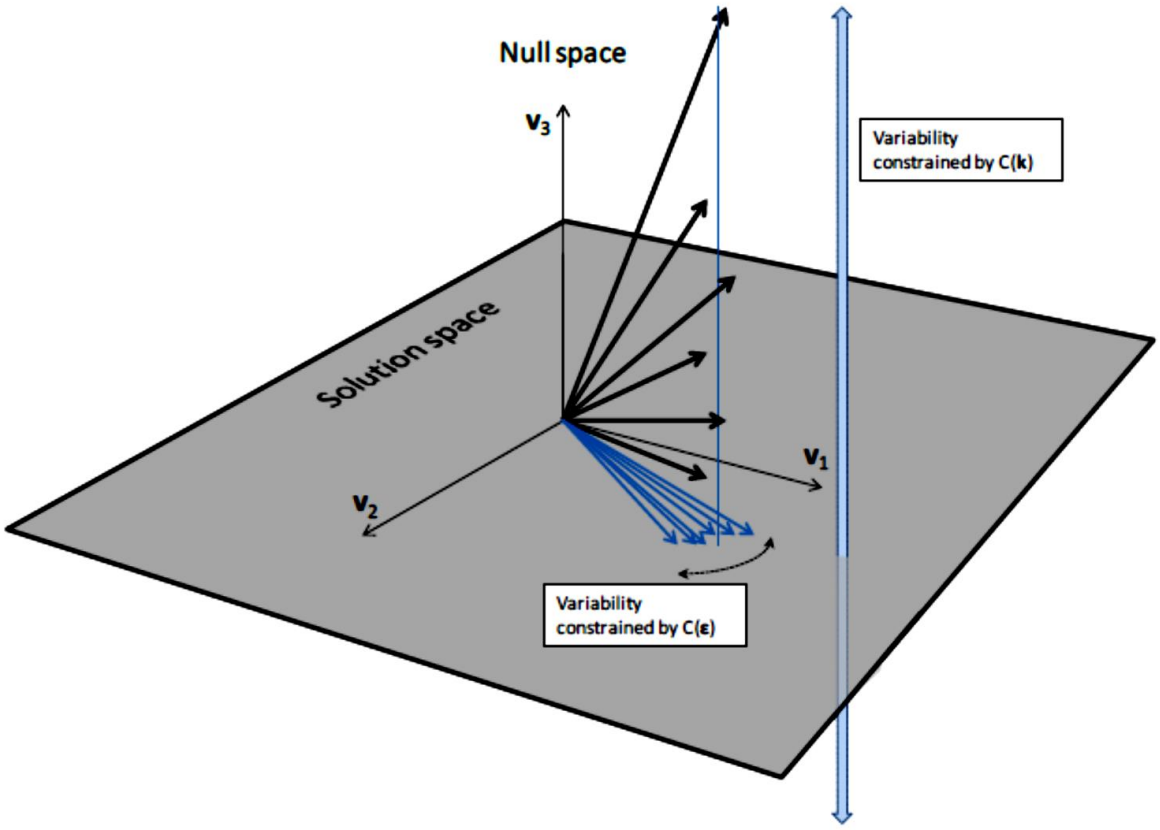

Figure 18: Post-calibration parameter variability (taken from Doherty (2010c))

Uncertainty analysis with Null-Space Monte-Carlo explores this parameter variability. The PEST tools `Randpar` and `Pnulpar` generate random parameter fields and then restricting the parameter field departure from the calibrated field to the null-space. These random parameter fields are then recalibrated, making sure the model-to-measurement fit of the random fields is at least as good as the original calibration results. This analysis yields a number of parameter fields and their corresponding observation and prediction values. The resulting variabilities can then be analysed to gain information on the uncertainty of the parameters, boundary conditions, observations and predictions.

For the two analysed models in this thesis, 200 random fields were generated for each model. After recalibration, 135 parameter fields had a suitable model-to-measurement fit (i.e. lower or equal to the respective original fit) for the PEST Pilot Point case and 71 for the AMALGAM Pilot Point case. These were then used for the uncertainty analysis in chapter 4.3.

4 Results and Discussion

In this chapter of the thesis, the calibration and uncertainty analysis results are presented. For all three different approaches of model calibration (two zones, Pilot Points with PEST, Pilot Points with AMALGAM), the estimated parameters (with focus on the conductivity field) and the model fit are examined. At the end, a synthesis of the different calibration techniques is given and discussed. The uncertainty analysis was undertaken only with the two Pilot Point calibrated models. It is divided into the two methods that were applied: first, the results of the analysis with PREDUNC are presented, and then the Null-Space Monte-Carlo analysis is portrayed. Afterwards, a complete discussion includes the results of these uncertainty analyses and summarizes the information gained on the three different calibrations.

4.1 Steady-state model calibration

4.1.1 Two zones of hydraulic conductivity

For the first model case, the hydraulic conductivity field only consists of two values for the two different zones. Calibrated values can be seen in Table 11. With both values between 1200 and 1300 m/d (~ 1.5 * 10^{-2} m/s), the values lie within the range of expected values for highly conductive gravel aquifers like the Wairau aquifer (see Wilson and Wöhling (2015), p.19). The similarity of the two values is somewhat surprising and leaves the impression that no heterogeneity at all is necessary to fit the model predictions. The river bed conductivities, on the over hand, range from 10^{-2} m/d to over 1 m/d. This variation seems to be necessary to meet both head and especially river flux observations with a homogeneous hydraulic conductivity field.

Table 11: Two zone calibration: calibrated parameter values

Parameter Name / Group	Unit	Parameter Number	Range	Calibrated value
Hydraulic Conductivity	m/d	1 (west)	100 - 2500	1222.95
		2 (east)		1286.47
Aquitard Hydraulic Conductivity	m/d	-	0.1 - 10	10.00
K_{xy}/K_z	-	-	1 - 10	5.19
River bed conductivity	m/d	1	0.05 - 10	0.87
		2		$5.51*10^{-2}$
		3		$5.00*10^{-2}$
		4		$5.84*10^{-2}$
		5		$9.39*10^{-2}$
		6		$5.00*10^{-2}$
		7		$5.30*10^{-2}$
		8		0.78
		9		0.53
		10		1.26
		11		0.13
Drain bed conductivity	m/d	1 (Spring Creek)	0.01 – 10	2.68
		2 (Southern Drain)	0.01 – 5	2.51
		3 (Opawa River)	0.01 – 1	1.00
		4 (Spring Creek)	0.01 – 10	0.89
		5 (Northern Drain)	0.01 - 1	0.11

The hydraulic conductivity of the eastern aquitard wedge is at the upper boundary of 10 m/d, ensuring sufficient water flow through the aquitard into the drains, the major outflow of the model domain. The ratio of horizontal to vertical hydraulic conductivity, K_{xy}/K_z, lies in the middle of its range with a value of around 5. Therefore, vertical flux in the aquifer is five times slower than horizontal flux. The bed conductivities of the various drains vary within their respective ranges. Only the drain bed conductivity of the Opawa River is at its upper limit of 1 m/d, which was set rather conservative, though.

The benchmark of model calibration is the difference between the real-world observations and the model calculations. The model fit is summarized in Table 12 and Figure 19.

Table 12: Two zone calibration: model fit

Observation Name	Type [Unit]	Weight	Observed Value	Calculated Value	Residual (abs)	Residual (%)
Condors Recharge	Groundwater Heads [m]	10	40.199	40.200	-0.001	< 0.00
Condors No.2		10	34.370	34.369	0.001	< 0.00
Wratts Road		10	12.258	12.255	0.003	0.02
Selmes Road		10	7.662	7.660	0.002	0.03
Murphys Road		1	7.022	9.318	2.296	32.70
Rock Ferry – SH6	River Flow Losses [m³/s]	25	2.768	2.767	0.001	0.04
SH6 – Wratts Road		25	4.247	4.245	0.002	0.05
Wratts Road – SH1		0.01	-1.282	0.257	-1.549	120.83
Spring Creek	Drain Flux [m³/s]	0.001 (in m³/d)	-3.128	-3.128	< 0.000	< 0.00

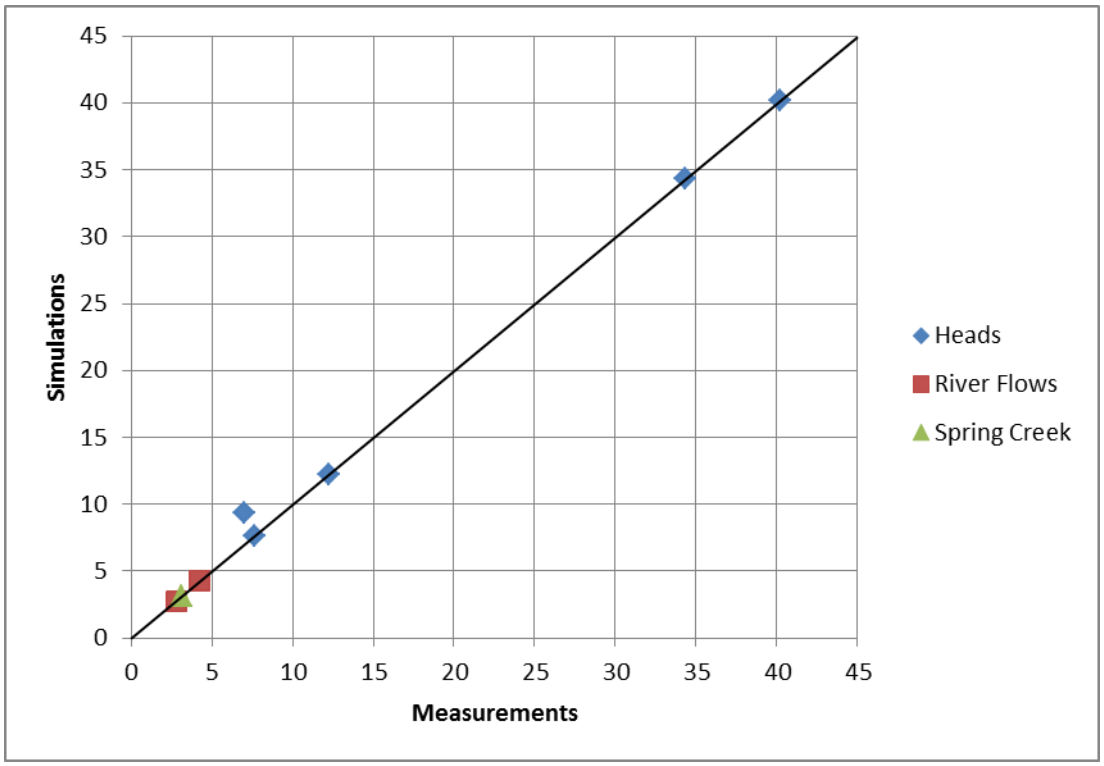

Figure 19: Two zone calibration: Measurements vs. Simulations

The calibrated model fits four of the five head measurements almost perfectly with residuals smaller than 0.1%. The fifth head, Murphys Road, is off by over 2 meters. This head measurement is in the south-eastern corner of the model domain. During earlier tests, it was found out that the model cannot reproduce this head – therefore, its weight was reduced dramatically. Speculations about this problem can be found in chapters 4.1.4 and 4.4.

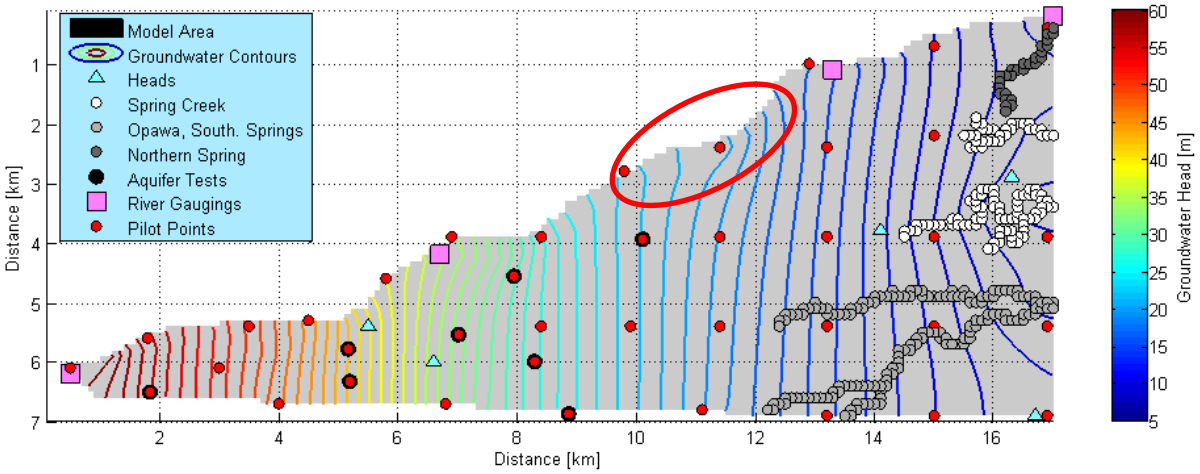

Figure 20: Two zone calibration: head field

The above Figure 20 shows the head contours for the calibrated field. The contour lines are relatively straight north-south, indicating a flow from west to east. At some parts along the river, there are strong indentations in the head contours where the water enters the aquifer (see red marking). Furthermore, the head field is inclined north-ward in the south-east corner, redirecting the flow here. This inclination depicts again how the head at Murphys Road in this corner was not met.

Next is the analysis of the simulated values for the river flow losses (and gain): the first two, between Rock Ferry and State Highway 6 and between State Highway 6 and Wratts Road gauging stations, are practically identical to the measured values, again with residuals smaller 0.1%. The third one, though, between Wratts Road and State Highway 1, differs largely from the estimated river gain of more than 1 m³/s. Instead, the modelled river loses a total of 0.26 m³/s over this area. As the State Highway 1 gauging station was shifted upstream in the model (see 2.1), this is not surprising, since the river gain increases downstream in this area. Again, this is explained in more detail in chapters 4.1.4 and 4.4. Therefore, the fit was ignored and the weight of the particular observation reduced dramatically in this model.

The last observation, drain flow into Spring Creek, was met perfectly by the simulated value in the model with the same simulated as observed value.

4.1.2 Pilot Point parameterization: PEST

The second model calibration was undertaken with a different parameterization. The hydraulic conductivity field was created by interpolation from 40 Pilot Points, whose values were used as parameters in the calibration process (see 3.2.3.1). Due to the relatively large amount of parameters, Table 13 shows only the values of the other parameters, while Figure 21 shows the hydraulic conductivity field (parameter values of the Pilot Points can be found in Table 25 in the Appendix).

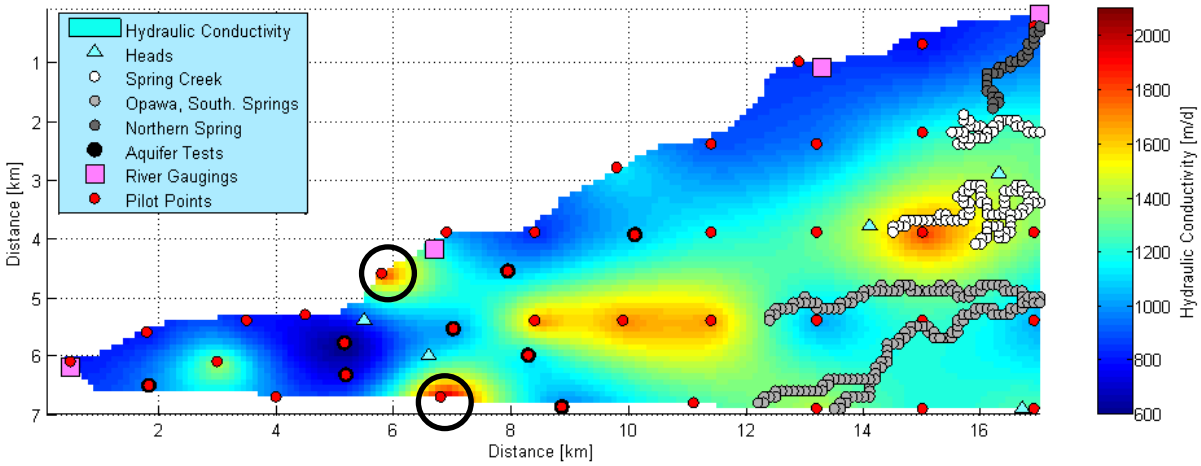

Figure 21: PEST Pilot Point calibration: hydraulic conductivity field

The hydraulic conductivity values from the Pilot Point calibration with PEST range from as low as around 600 m/d to a maximum of over 2000 m/d. Since the hydraulic conductivity estimates gained from the aquifer tests range from around 700 m/d to around 1200 m/d, some calculated Pilot Point values exceed this range. This is not concerning, though, since up to 2000 m/d hydraulic conductivities seem plausible geologically and most aquifer test information is situated in the eastern half of the model domain, leaving the western part's conductivity unknown. Looking at the structure of the field, it looks plausible in most parts of the model domain. Only the high conductivity at the northern and southern model edge (around 6 to 7 km from the western edge of the model, circled black) is surprising.

Table 13: PEST Pilot Point calibration: calibrated parameter values

Parameter Name / Group	Unit	Parameter Number	Range	Calibrated value
Aquitard Hydraulic Conductivity	m/d	-	0.1 - 10	6.19
K_{xy}/K_z	-	-	1 - 20	17.84
River bed conductivity	m/d	1	0.01 - 10	0.49
		2		1.49
		3		$1.00*10^{-2}$
		4		0.12
		5		0.12
		6		0.13
		7		0.27
		8		0.31
		9		0.35
		10		0.52
		11		0.47
Drain bed conductivity	m/d	1 (Spring Creek)	0.01 – 10	2.20
		2 (Southern Drain)	0.01 – 5	0.49
		3 (Opawa River)	0.01 – 1	0.99
		4 (Spring Creek)	0.01 – 10	2.07
		5 (Northern Drain)	0.01 - 1	$8.30*10^{-2}$

The hydraulic conductivity of the aquitard wedge is somewhat smaller in this case (6.19 m/d instead of 10 m/d), but still in the upper part of its range, allowing enough groundwater to exfiltrate into the drains. The ratio of horizontal to vertical hydraulic conductivity, K_{xy}/K_z, is almost 20 for this calibration, meaning the vertical flow is relatively slow compared to horizontal flow (i.e. anisotropy in the vertical direction). Apart from river section number 2 and 3, the values of the river bed conductivities are relatively homogeneous, ranging from 0.12 m/d to 0.52 m/d. Of the drain bed conductivities, only number 3 (Opawa River) lies close to its upper bound, again around 1 m/d.

In the following, the model fit is summarized. As can be seen from Table 14 and Figure 22, the groundwater head measurement at Murphys Road and the third river flux loss target were excluded from the model calibration this time.

Table 14: PEST Pilot Point calibration: model fit

Observation Name	Type [Unit]	Weight	Observed Value	Calculated Value	Residual (abs)	Residual (%)
Condors Recharge	Groundwater Heads [m]	10	40.199	40.320	-0.121	0.30
Condors No.2		10	34.370	34.276	0.094	0.27
Wratts Road		10	12.258	12.262	-0.004	0.03
Selmes Road		10	7.662	7.702	-0.040	0.52
Rock Ferry – SH6	River Flow Losses [m^3/s]	25	2.768	2.644	0.124	4.48
SH6 – Wratts Road		25	4.247	4.266	-0.019	0.45
Spring Creek	Drain Flux [m^3/s]	0.001 (in m^3/d)	-3.128	-3.127	0.001	0.03

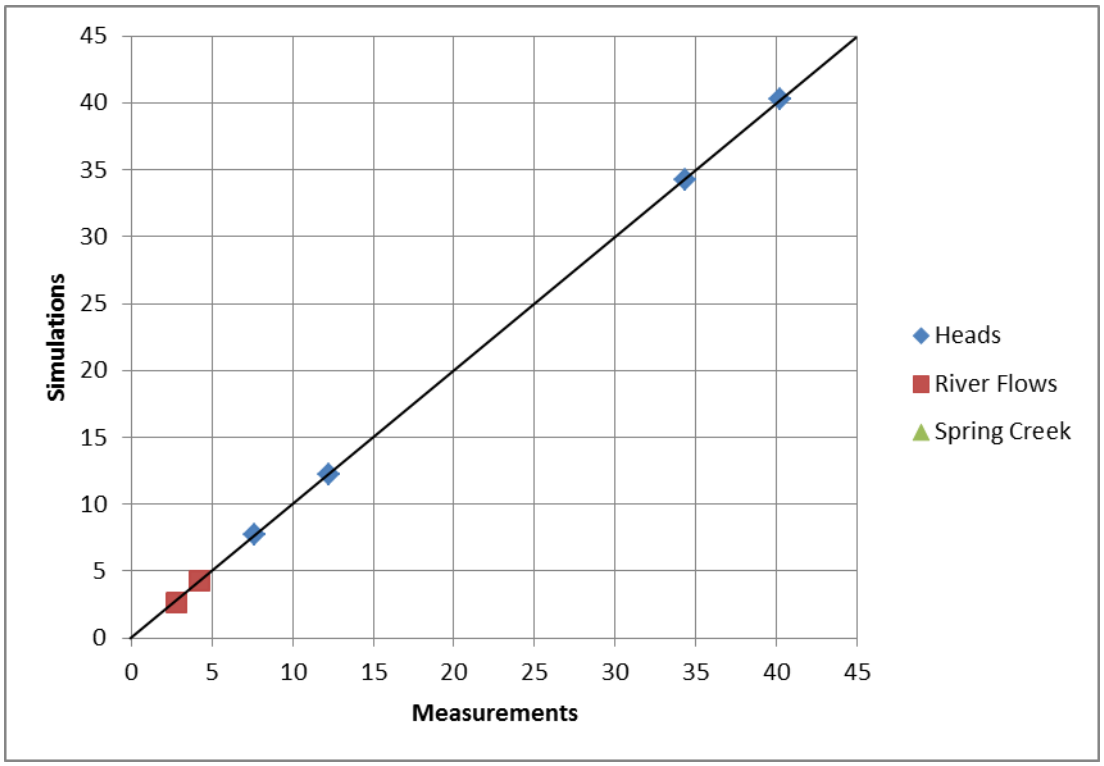

Figure 22: PEST Pilot Point calibration: Measurements vs. Simulations

The model fit for the groundwater heads is good, with relative residuals under 1% and absolute deviations between measurements and simulations in the range of centimetres, which is within the measurement error for groundwater heads. The second river flow loss is met very well, with a somewhat higher deviation (~ 5%) for the first flux between the Rock Ferry and State Highway 6 gauging stations. Since the river flow losses are only crude estimates based on an observed overall pattern, residuals in this range are acceptable. The simulated flux from the groundwater into Spring Creek is very close to the measured value with a residual of 0.5%.

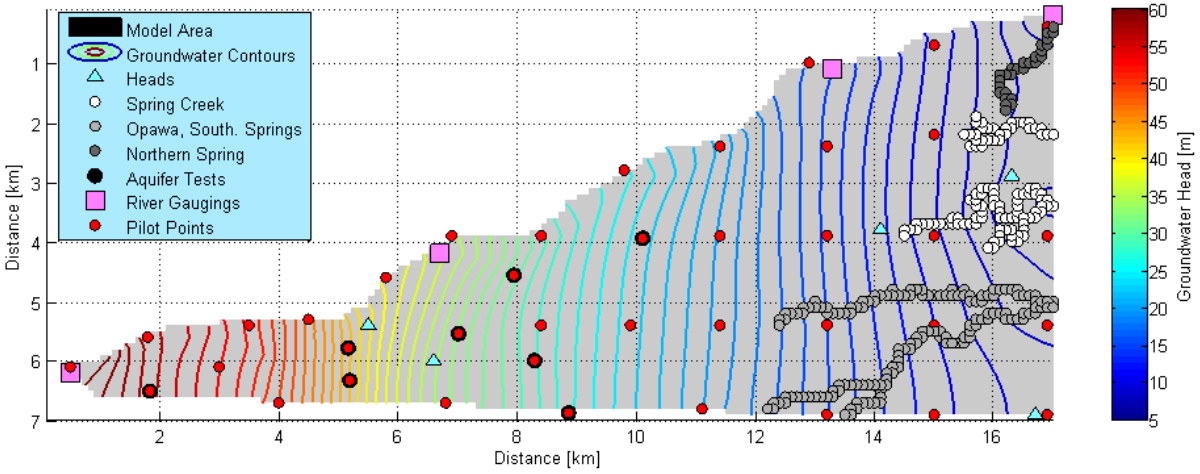

Figure 23: PEST Pilot Point calibration: head field

The head field for the Pilot Point calibration with PEST (shown in Figure 23) shows the intended flow from west to east in the aquifer. The water entering the aquifer from the river along the north is seen in the head contours, as is the flux back into the river in the north-

east corner. At the eastern boundary, the groundwater flow is shifting towards the north-east, creating a mount in the south-eastern corner. This is where the simulation of the head at Murphys Road is too high, again.

4.1.3 Pilot Point parameterization: AMALGAM

The third steady-state calibration that is presented here is a realization computed by AMALGAM. As noted before, the compromise solution (realization with the smallest trade-off between all three objective functions) was chosen from the Pareto front. Again, 40 Pilot Points were used for the interpolation of the hydraulic conductivity field. Figure 24 shows the calibrated hydraulic conductivity field, while Table 15 shows the resulting values for the other parameters. Parameter values for the Pilot Points can be found in Table 25 in the Appendix.

The hydraulic conductivity values of all Pareto solutions from the AMALGAM calibration range from as low as around 600 m/d to a maximum of around 1500 m/d. The western part of the model domain has relatively low conductivity values, though this area is somewhat constrained by the amount of aquifer tests in it. Again, two small high-conductivity areas can be found around 7 km eastward from the western model edge (circled black). The middle and eastern part is mostly a high-conductivity zone, with considerably lower conductivities in the north- and south-western corner (circled red).

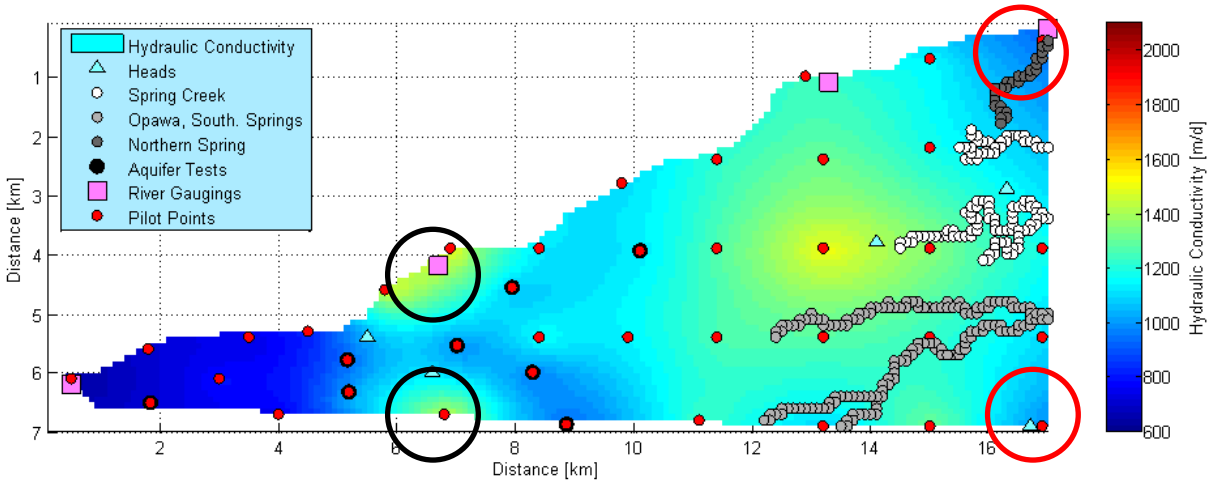

Figure 24: AMALGAM Pilot Point calibration: hydraulic conductivity field

Table 15: AMALGAM Pilot Point calibration: calibrated parameter values

Parameter Name / Group	Unit	Parameter Number	Range	Calibrated value
Aquitard Hydraulic Conductivity	m/d	-	0.1 - 10	2.24
K_{xy}/K_z	-	-	1 - 20	10.19
River bed conductivity	m/d	1	0.01 - 10	$8.22*10^{-2}$
		2		0.11
		3		$6.92*10^{-2}$
		4		$8.95*10^{-2}$
		5		$2.09*10^{-2}$
		6		0.14
		7		0.25
		8		0.26
		9		0.29
		10		0.27
		11		0.24
Drain bed conductivity	m/d	1 (Spring Creek)	0.01 – 10	4.75
		2 (Southern Drain)	0.01 – 5	1.11
		3 (Opawa River)	0.01 – 5	1.67
		4 (Spring Creek)	0.01 – 10	5.70
		5 (Northern Drain)	0.01 - 1	0.31

In the compromise solution of the AMALGAM run, the hydraulic conductivity of the aquitard has a rather small value of around 2 m/d, compared to around 6 m/d and 10 m/d for the other calibrations. In contrast, the drain bed conductivities are relatively high (1 m/d to 5 m/d, compared to 0.5 m/d to 2 m/d for the other calibrations), probably compensating for the lower speed with which the water can travel through the aquitard due to its comparably lower hydraulic conductivity. The ratio of horizontal to vertical hydraulic conductivity, K_{xy}/K_z, lies in the middle of its range at around 10. The river bed conductivities are rather low in the first part of the river (< 0.1 m/d), and very uniform in the second part (0.25 m/d to 0.29 m/d).

Again, the model fit is given as both a table and a figure in the following. Both absolute and relative residuals for all observations are satisfactorily small.

Table 16: AMALGAM Pilot Point calibration: model fit

Observation Name	Type [Unit]	Weight	Observed Value	Calculated Value	Residual (abs)	Residual (%)
Condors Recharge	Groundwater Heads [m]	10	40.199	40.216	-0.017	0.04
Condors No.2		10	34.370	34.353	0.017	0.05
Wratts Road		10	12.258	12.298	-0.040	0.33
Selmes Road		10	7.662	7.692	-0.030	0.39
Rock Ferry – SH6	River Flow Losses [m^3/s]	1	2.768	2.744	0.024	0.87
SH6 – Wratts Road		1	4.247	4.230	0.017	0.40
Spring Creek	Drain Flux [m^3/s]	1	-3.128	-3.141	-0.013	0.42

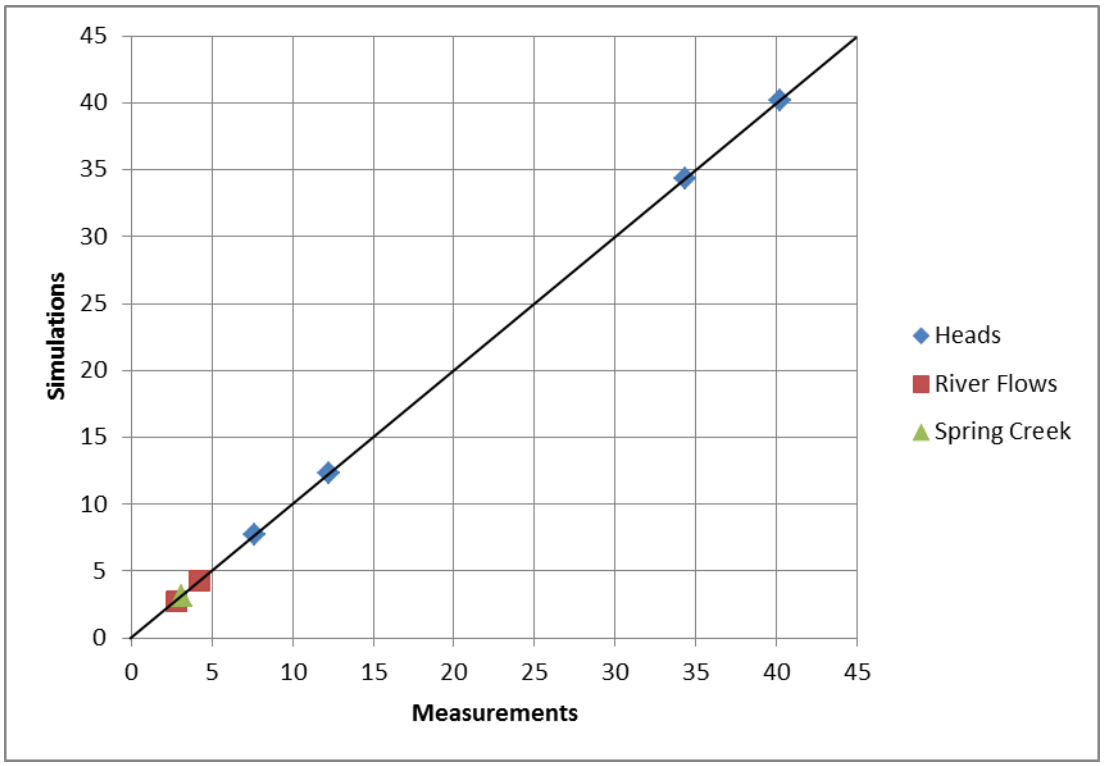

Figure 25: AMALGAM Pilot Point calibration: Measurements vs. Simulations

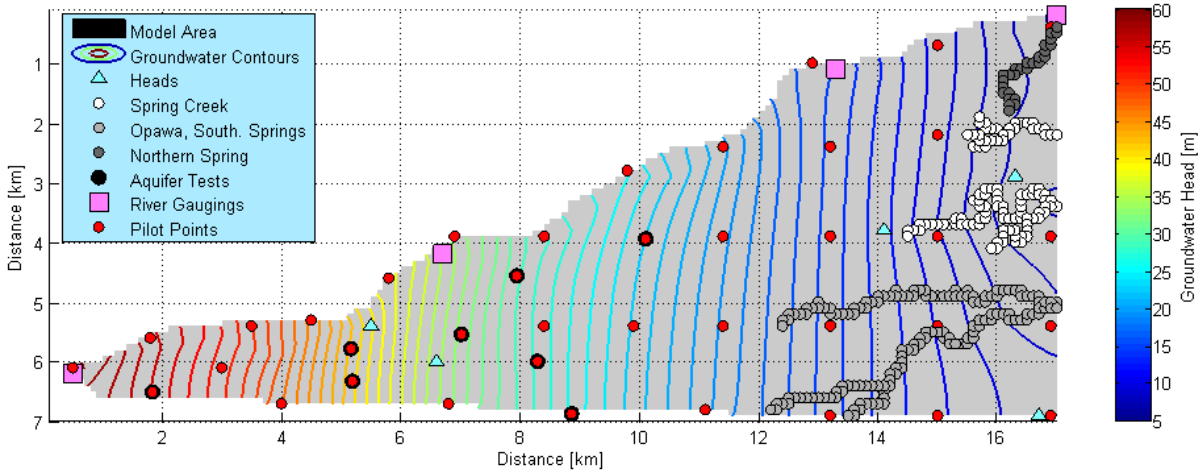

Figure 26: AMALGAM Pilot Point calibration: head field

The head field for the calibrated model is similar to the field gained with the PEST Pilot Point calibration. There are indentations in the head field along the north, revealing where river water infiltrates into the aquifer. Again, groundwater flow shifts from east to north-east at the eastern model edge, resulting in the groundwater mount at the south-eastern corner.

4.1.4 Model calibration: Synthesis

In this chapter, the three different model calibrations presented earlier are compared briefly in their parameter values, their model fit and their predictions. This comparison focusses on the models' ability to reflect real-world knowledge of the system and on conceptual differences of the models.

First, the hydraulic conductivity fields of all three calibrations are analysed: the two-zonal approach separates the aquifer into two zones of constant hydraulic conductivity, both with a value between 1200 m/d and 1300 m/d. The two different Pilot Point calibrations interpolate a field from 40 Pilot Point parameters and the same variogram. Looking at the fields (Figure 21 and Figure 24), the PEST calibration has an average hydraulic conductivity slightly higher than the two-zone model, while the AMALGAM calibration has a slightly lower one. Comparing the two Pilot Point fields, both show certain similar distinctive features:

- the eastern part of the model domain has a uniformly low conductivity (as it is somewhat constrained through the aquifer tests)
- there are small high-conductivity zones around 7 km eastward from the western model edge
- the highest conductivity is in the middle of the western model domain, around the beginning of Spring Creek, the main water outlet of the system

In contrast, there are some features differing between the two hydraulic conductivity fields:

- the field of the PEST calibration has much higher conductivities, up to over 2000 m/d compared to the maximum of around 1500 m/d of the AMALGAM calibration. This is not due to different parameter ranges, though – they are the same for both models
- the field of the AMALGAM calibration has an additional two low-conductivity zones at the north- and south-western edges of the model domain
- the field of the PEST calibration has a long area of lower conductivity underneath the river from the State Highway 6 gauging station westward

In conclusion, all three hydraulic conductivity fields seem theoretically plausible. The Pilot Point calibrations resolve some heterogeneity and fit the aquifer tests, therefore incorporating more real-world knowledge into the model than the two-zone approach.

Second, the other parameter values are of interest in this comparison, too. Since no real knowledge about the hydraulic conductivity of the aquitard and the ratio of horizontal to vertical hydraulic conductivity, K_{xy}/K_z, is available, and the values for all three calibrations seem reasonable, no real distinction in favour of one calibration can be made from these parameters. Similarly, the bed conductivities of the drains differ only slightly between all three realizations: the Spring Creek conductivities are relatively high to ensure enough water can leave the system here, while the other drains have rather low values. These are probably limited through the overall water balance, ensuring only small amounts of groundwater leave the system through these drains. Looking at the river bed conductivities, though, a difference between the two Pilot Point approaches and the two-zone calibration can be observed: while the Pilot Point calibrations have both rather similar and pretty uniform river bed conductivity values in the middle and down-stream part of the river, the two-zone calibration resulted in a more heterogeneous river bed. This is due to the lack of spatial heterogeneity in the underlying hydraulic conductivity field, which is therefore compensated in the river bed. These values are probably not realistic on the modelling scale, where a pretty uniform river bed (and therefore uniform river bed properties) should be expected. Therefore, the two Pilot Point calibrations are probably a more realistic approach to modelling the river-groundwater interactions parameter-wise.

Next, the model fit to observations is an important criterion to analyse the different model calibrations. Overall, the two-zone calibration has the best fit. This seems counter-intuitive,

since the increase of the number of parameters leads to an increase of degrees of freedom, and therefore usually to a better model fit. Since both Pilot Point parameterizations had to fit not only the measurements, but also the regularization constraints, they are actually more constraint, thus explaining the (slightly) worse fit. As was observed in the respective chapters for the different calibrations, all three fit the data reasonably well, though. Therefore, all three calibrations are acceptable for their ability to fit the measured data.

A model's worth depends not only on its fit to measured data that were used in the calibration, but also on its ability to make accurate predictions. A prediction is any sort of model output that was not used in the calibration process. For this purpose, four different predictions were examined:

- the groundwater head at Murphys Road
- the third exchange flux from river to groundwater in between gauging stations Wratts Road and State Highway 1
- the flux from groundwater into the drains for the two southern streams (Southern Drain and Opawa River)
- the distribution of exchange fluxes along the Wairau river

For the first two of these predictions, measurements were available, but not used in the calibration process due to the inability to meet them for model structural reasons. For the other two predictions, only general ideas of their real-world values and behaviour are available.

Table 17: Calibration comparison: predictions

Prediction Name [Unit]	Observed Value	2 Zones	PEST Pilot Points	AMALGAM Pilot Points
Murphys Road [m]	7.022	9.318	9.686	9.525
Wratts Road – SH1 [m^3/s]	-1.282	0.257	0.237	0.248
Southern Streams [m^3/s]	-	-0.917	-0.839	-0.807

The above Table 17 displays the simulated values for the first three predictions, as well as the observed values (where applicable). As one can see, all three calibrated models do poorly for the prediction of the groundwater head at Murphys Road as well as for the third river-groundwater flux. As was mentioned earlier, these problems result from structural deficits of the general model setup, as explained now.

The Murphys Road head lies at the south-eastern corner of the model area. Here, two boundary conditions meet: the constant flux of the eastern boundary, and the no-flow southern boundary. Both are based on general assumptions of experts about the aquifer, and are not valid in the real world, especially on a finer scale of the model area's edge. The problem at this corner can also be seen in the three head fields of the models (see Figure 20, Figure 23 and Figure 26). All show a flow field shifting from east to north-east along the model edge there, resulting in the too high heads at the south-east corner. To illustrate this further, Modpath was set up for the model calibrated with the PEST Pilot Points. Modpath is a Modflow tool for the implementation of particle tracking into the model. This is usually used

for the modelling of solute transport and the computation of travel time distributions. In this case, it is only used to clarify the change in the flow field in the south-eastern model corner. For this, three conservative particles (i.e. not retarded in the aquifer) were introduced into the model in each river cell. Figure 27 shows the resulting flow field.

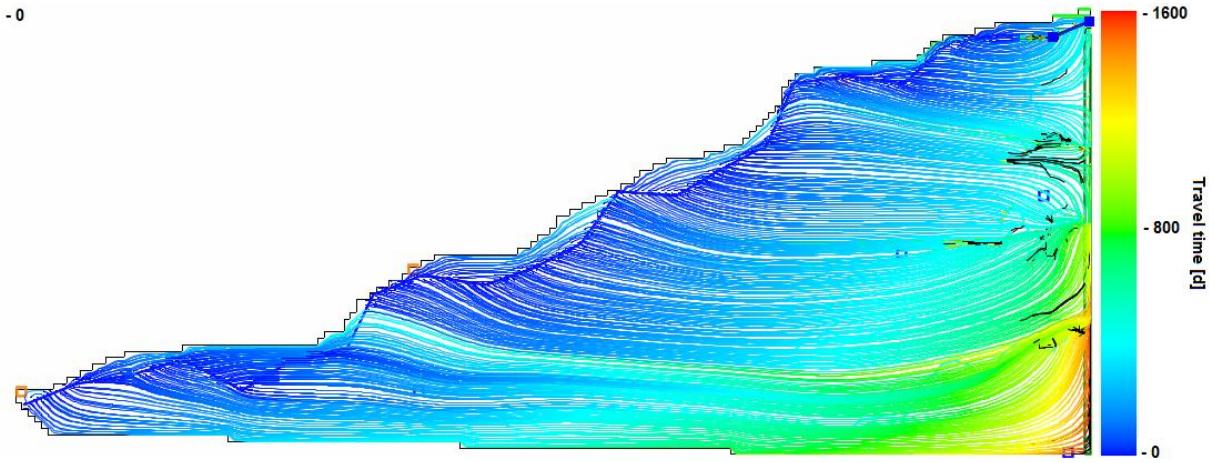

Figure 27: Modpath flow field for PEST Pilot Point calibration

Note that the travel times are not really relevant for the purpose of this analysis. The figure visualizes the flow field of water from the river through the aquifer into the low-land springs and the constant flux boundary. One can see the pretty horizontal flow from west to east through-out the model domain. Furthermore, groundwater flow is channelled into the low-land springs in the eastern part of the model domain. In the south-eastern corner, groundwater flow is changing direction from eastward to north-eastward to exfiltrate through the southern streams. This leads to the higher groundwater head at Murphys Road. Extruding more water through the constant flux boundary in the south-eastern corner would possibly counter this. Since this flux boundary is only a rough estimate and the spatial distribution of water flow out of the model area eastward is even less known, changing this boundary condition might be an option to overcome this problem of the model structure.

The river-groundwater flux between the gauging stations at Wratts Road and State Highway 1 is problematic due to the west-ward shift of the SH1 gauging station in the model: in reality, this measurement point lies further downstream of the river. Therefore, the negative net flux from the groundwater back into the river that is observed in reality can not be reproduced in the model. Modelling results imply that the river still losses water into the aquifer directly after the Wratts Road gauging station, and only closer to the SH1 gauging station starts gaining water again – therefore, the shift of the gauging station shortens the gaining part of the river dramatically (see later for more details on this).

For the prediction of the flow into the southern streams, the results from all three models are quite similar. In itself, this prediction is not very interesting, but in can be used as a target in the latter uncertainty analysis and might be of interest should future measurements of the real-world fluxes become available.

The fourth prediction is not a necessarily numerical but more of a graphical one: Modflow's output allows a cell-by-cell analysis of the river-groundwater exchange fluxes. While such information can not be obtained in reality, the overall distribution of the model output can be examined for anomalies or unrealistic results.

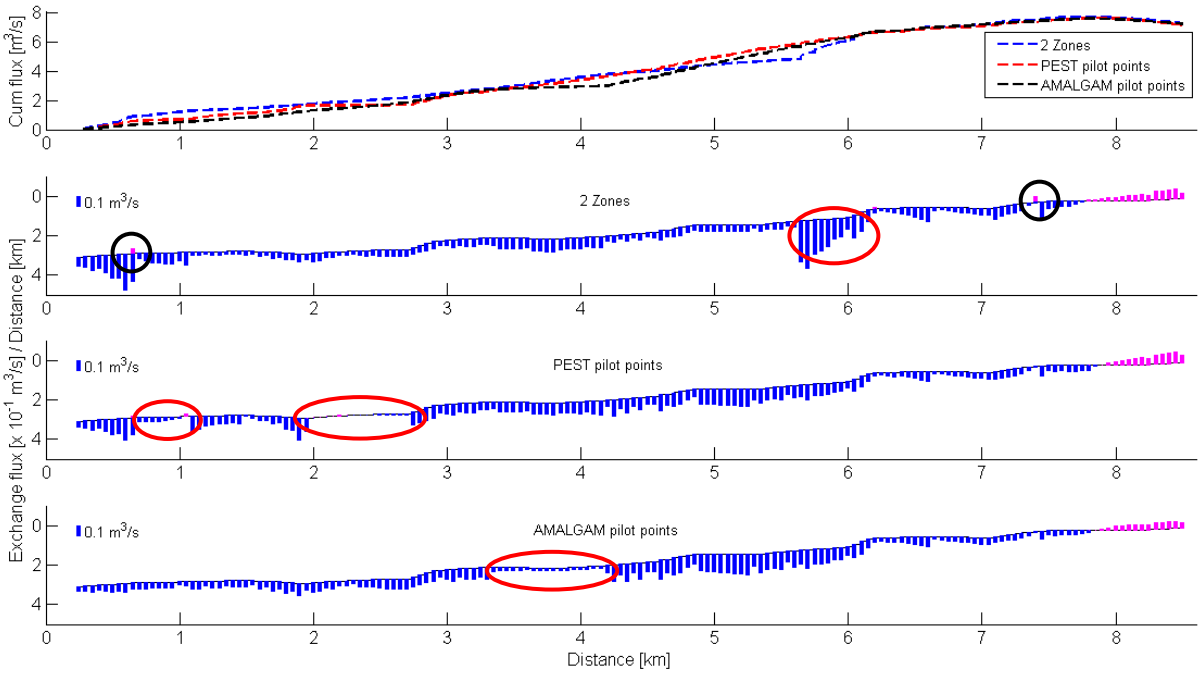

Figure 28: River-groundwater exchange fluxes

Figure 28 above shows the cell-by-cell information on the river-groundwater exchange fluxes. The top plot depicts the cumulative fluxes along the river for all three model calibrations. They are pretty similar, which makes sense since they are restricted for two thirds by the flux observations. The other plots show the cell-by-cell flux into (blue) or out (pink) of the aquifer from the river. Overall, all three plots show groundwater recharge for most of the river distance, and the starting return flux from groundwater into the river at the end. Prediction-wise, one would expect fairly smooth fluxes, since the river bed composition is probably rather uniform on the model's scale. Of the three model calibrations, the AMALGAM version has the smoothest exchange fluxes, with only one area of somewhat low fluxes (marked red). The Pilot Point calibration with PEST has two such low-flux zones (red again) and some irregularity in the first third of the river. While the plot for the two-zone approach shows no area of very low exchange fluxes, there is an area with an unlikely spike (circled red) and two small zones where negative fluxes occur in otherwise groundwater-to-river flux areas. In summary, the AMALGAM calibration has the smoothest distribution of exchange fluxes and the two-zone model the worst.

To sum up the information gained from looking at model predictions about the different calibrations used, only the distribution of river-groundwater exchange fluxes gives some insight into the different calibrations. For all predictions, though, further knowledge about them or relevant parameters (southern stream flux measurements, information on river bed composition, etc.) would make them more valuable to assess model performance (i.e. using them as observations). The value of the predictions as are will be clearer from the uncertainty analysis.

Concluding the synthesis of the different model calibrations, all three fit the measurements well enough. The two Pilot Point approaches include some data into the model (aquifer tests, assumed aquifer conductivity heterogeneity instead of river bed conductivity heterogeneity) that the two-zone calibration is not capable of employing.

4.2 PREDUNC uncertainty analysis

The uncertainty analysis with PREDUNC is split into two parts: first, parameter and boundary conditions are examined for their effect on the uncertainty of the different observations (and predictions). Second, the effect of a possible new head measurement in the model area on these uncertainties is analysed.

4.2.1 PREDUNC: PEST Pilot Point calibration

At the beginning of the PREDUNC analysis, the basic uncertainties for all observations and predictions were computed. All uncertainties are given as standard deviations in this analysis. PREDUNC allows the comparison of pre- and post-calibration uncertainties for this matter. They are summarized in the following Table 18.

Table 18: PEST Pilot Point calibration: base uncertainties

Name	Value	Unit	Uncertainty	
			Pre-calibration	Post-calibration
Condors Recharge	40.320	m	28.367	0.084
Condors No2	34.276	m	20.785	0.084
Wratts Road	12.262	m	3.641	0.084
Selmes Road	7.702	m	1.292	0.081
Murphys Road	9.686	m	2.542	1.393
Rock Ferry - SH6	2.644	m^3/s	3.971	0.067
SH6 - Wratts Rd	4.266	m^3/s	5.672	0.067
Wratts Rd - SH1	0.237	m^3/s	1.567	1.059
Spring Creek	-3.127	m^3/s	4.383	0.020
Southern Streams	-0.839	m^3/s	2.024	0.751

The pre-calibration uncertainties of the head measurements are all around 1/3 of the measurement value, while higher than one measured value for all the fluxes in the system. Post-calibration uncertainties are very small for all calibrated values. The remaining uncertainty is probably mostly a representation of measurement error which can not be reduced through model calibration. For the three non-calibrated predictions (Murphys Road head, Wratts Rd – SH1 flux and the southern Streams), the post-calibration uncertainties remain significantly higher.

Table 19: PEST Pilot Point calibration: parameter / boundary condition uncertainty reduction

Name	HKs	River bed Cond	Drain bed Cond	Wells	Flux BC	SFR functions
Condors Recharge	< 1%	< 1%	< 1%	< 1%	< 1%	< 1%
Condors No2						
Wratts Road	31.4%					
Selmes Road	71.7%					
Murphys Road	95.3%	6.5%				
Rock Ferry - SH6	7.0%	< 1%				
SH6 - Wratts Rd	8.1%					
Wratts Rd - SH1	69.1%	10.9%				
Spring Creek	1.5%	< 1%				
Southern Streams	92.5%	10.4%				

Table 19 lists the same observations and predictions. This time, the uncertainty reductions (in percent of the post-calibration uncertainties) due to knowledge of different parameters or boundary conditions are shown. As one can see, knowledge of following categories lead to almost no uncertainty reduction:

- Drain bed conductivities: while influencing the flux into the drains by a certain amount (and therefore definitely influencing the Spring Creek and Southern Streams observations), the underlying aquitard has probably a higher influence on the overall fluxes. Therefore, knowledge of the drain bed conductivities is irrelevant for the reduction of the remaining uncertainty of all observations and predictions.
- Wells: the amount of abstraction (and its distribution through the different zones) out of the aquifer should obviously have an influence on the water balance, and therefore the head (and to a degree, the flux) observations. Due to the very small amounts abstracted during the model period, changes during the uncertainty analysis are correspondingly small. Therefore, they result in no relevant uncertainty reduction.
- Flux BC: the constant flux boundary condition at the eastern model boundary is rather small compared to the overall fluxes in the system. Again, this results in congruently small changes during the uncertainty analysis. These small changes to the boundary condition do not have a strong effect on the model predictions.
- SFR functions: the knowledge of all factors and exponents of the river functions for SFR has no relevant influence on the predictions. This is surprising, but probably due to the rather small range (standard deviations of 0.01) with which they were allowed to fluctuate. This range could not be increased without leading to occasional malfunction of the `SFR package`, though.

On the other hand, knowledge of all river bed conductivities led to some measureable uncertainty reduction for the three predictions (7 – 11%). Knowing the hydraulic conductivities (for the uncertainty analysis, this includes the aquifer field, the aquitard conductivity and the K_{xy}/K_z factor) resulted in partially very high uncertainty reductions. Obviously, the three prediction's uncertainties could be reduced the furthest, due to their high base uncertainties – between 70 to 95%. But even the small uncertainties of the

observations could be reduced somewhat in most cases by the hydraulic conductivities (between 1 to 70%). This shows that the hydraulic conductivity is the most important parameter for groundwater modelling.

For the uncertainty analysis of potential new heads, a total of 118 dummy head measurements were laid out in a structured grid over the model domain. Then, successive knowledge of one head was assumed and the resulting uncertainty reduction computed. For the observations, knowledge of one additional head resulted in no real uncertainty reduction (< 5%). For the three predictions, uncertainty reductions through new head measurements are shown in the following figures:

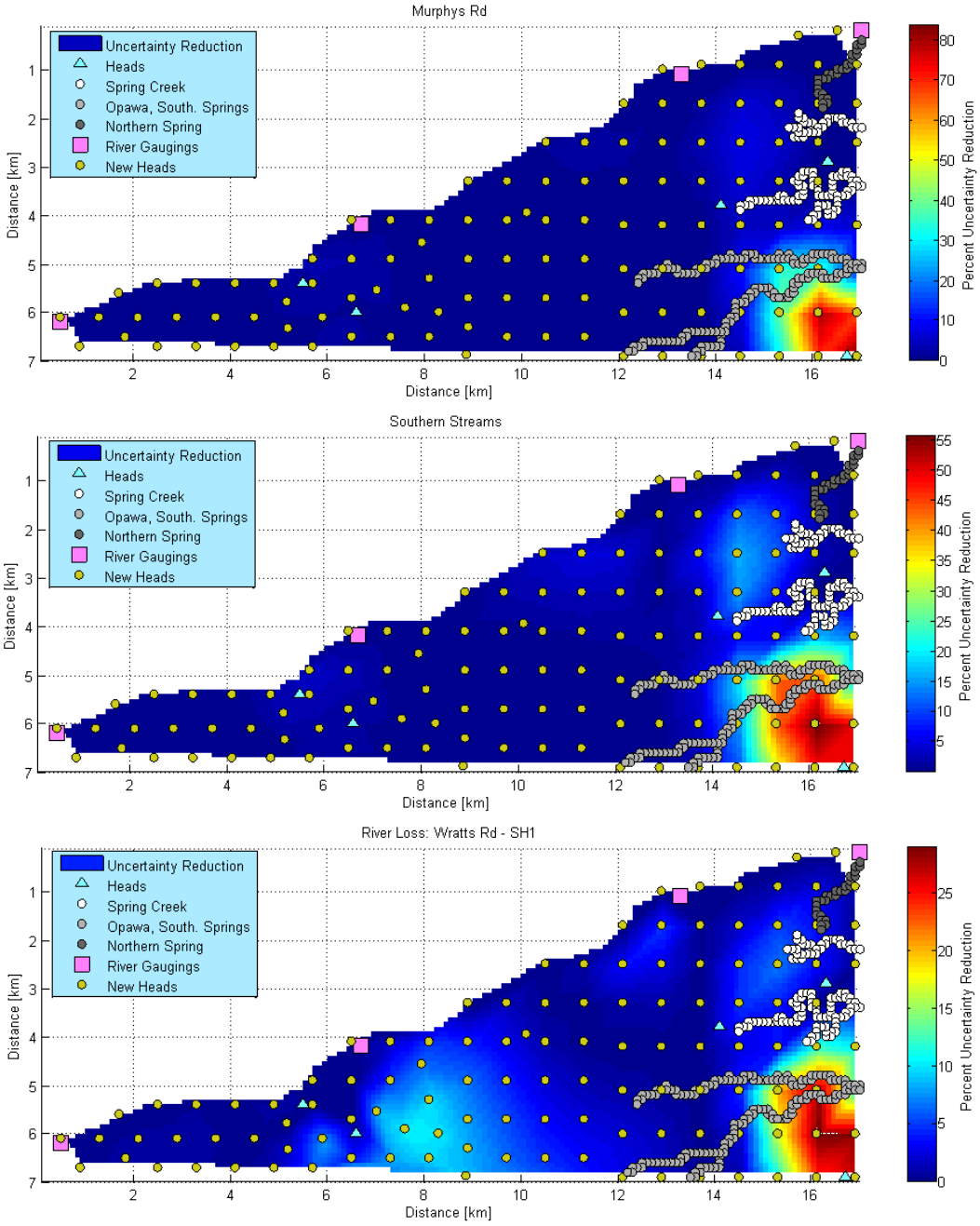

Figure 29: PEST Pilot Point calibration: new head uncertainty reduction

All three predictions depend highly on the knowledge of the groundwater head in the south-east corner. With a relatively parallel head field throughout the model domain, this corner is

the only part where this changes, with the groundwater flow bending from east to north-east. Therefore, this spot is generally important. This is highly problematical, since the first of the three “predictions”, Murphys Road, is a real head measurement that is situated in this very area but was excluded from the calibration process due to problems meeting it. This shows that these problems have an impact on the whole model and should be regarded further.

4.2.2 PREDUNC: AMALGAM Pilot Point calibration

Again, pre- and post-calibration uncertainties for all observations and predictions were computed first and are summed up in the following Table 20.

Table 20: AMALGAM Pilot Point calibration: base uncertainties

Name	Value	Unit	Uncertainty	
			Pre-calibration	Post-calibration
Condors Recharge	40.469	m	41.306	0.425
Condors No2	34.601	m	42.260	0.425
Wratts Road	12.228	m	4.693	0.417
Selmes Road	7.228	m	0.786	0.198
Murphys Road	9.240	m	2.751	1.276
Rock Ferry - SH6	2.779	m^3/s	5.198	0.169
SH6 - Wratts Rd	4.356	m^3/s	8.474	0.169
Wratts Rd - SH1	0.300	m^3/s	8.782	2.789
Spring Creek	-3.339	m^3/s	5.208	0.049
Southern Streams	-0.838	m^3/s	2.853	0.954

The patterns and absolute values in both pre- and post-calibration uncertainties are similar to the ones of the PEST calibration.

The same can be said mostly about the uncertainty reductions through the different parameter and boundary condition groups, as obvious from Table 21. There are a few exceptions here, though:

- Uncertainty reductions through knowledge of the hydraulic conductivities for the heads at Wratts Road and Selmes Road are significantly higher (72.0% instead of 31.4% and 95.5% instead of 71.7%, respectively). This might be due to higher post-calibration uncertainties for the heads in the AMALGAM case, allowing more reduction.
- The influence of the river bed conductivities is higher for all observations and predictions. This is most significant for the predictions (19.9% instead of 6.5%, 51.6% instead of 10.9% and 32.9% instead of 10.4%), probably due to their overall higher uncertainties. Why the river bed conductivities have higher influences for the AMALGAM calibration then for the PEST calibration is not so clear. It might be due to the higher hydraulic conductivities under the river in the AMALGAM case, amplifying changes to the river bed conductivities.

Table 21: AMALGAM Pilot Point calibration: parameter / boundary condition uncertainty reduction

Name	HKs	River bed Cond	Drain bed Cond	Wells	Flux BC	SFR functions
Condors Recharge	< 1%	1.7%	< 1%	< 1%	< 1%	< 1%
Condors No2		< 1%				
Wratts Road	72.0%	1.0%				
Selmes Road	95.5%	6.3%				
Murphys Road	92.9%	19.9%				
Rock Ferry - SH6	1.7%	6.6%				
SH6 - Wratts Rd	2.8%	< 1%				
Wratts Rd - SH1	69.8%	51.6%				
Spring Creek	< 1%	< 1%				
Southern Streams	91.2%	32.9%				

The uncertainty reduction through potential new heads is mostly similar to the PEST calibration's results as well. Therefore, the figures for the predictions are not shown here, but can be found under Figure 40 in the Appendix. Again, the most relevant new head measurements with potential uncertainty reductions of up to 60% are in the south-east corner, where the unused Murphys Road measurement is already available.

4.3 Null-Space Monte-Carlo (NSMC) uncertainty analysis

4.3.1 NSMC: PEST Pilot Point calibration

For the Null-Space Monte-Carlo analysis of the PEST Pilot Point calibration field, a total of 200 random fields with null-space departure of parameter values were created. From these, 135 could be recalibrated with acceptable precision (equal to or smaller than the objective function of the standard PEST Pilot Point calibration). In the following, these 135 calibrated fields will be analysed.

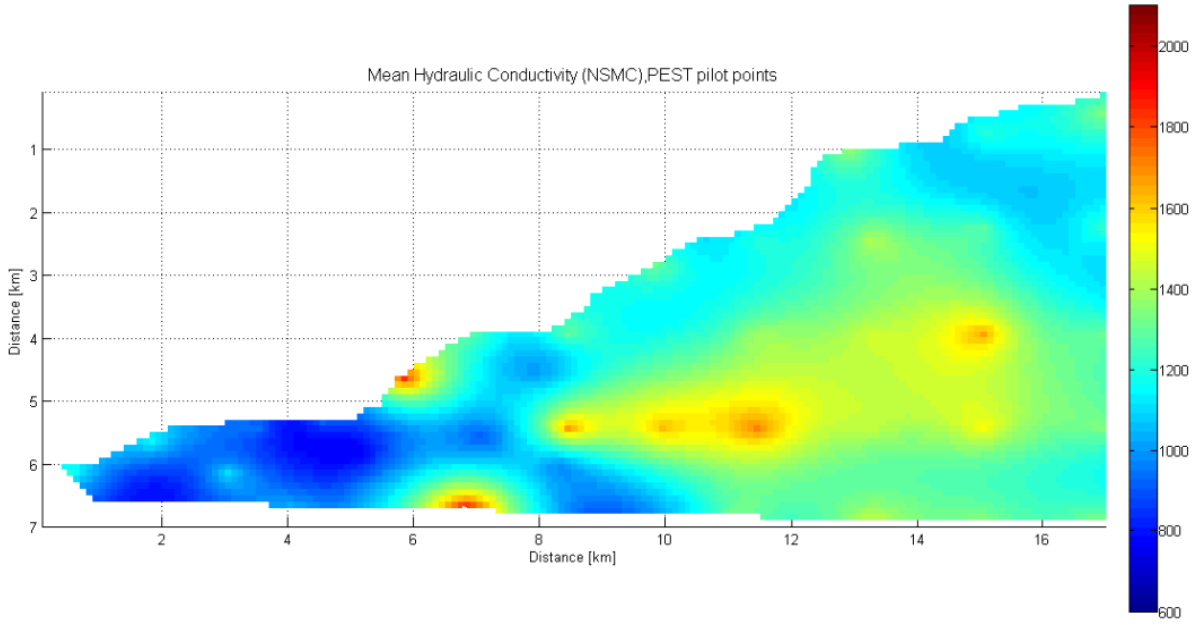

Figure 30: NSMC PEST Pilot Point calibration: mean hydraulic conductivity

The above Figure 30 shows the mean hydraulic conductivity of the 135 fields. It exhibits much of the same features the one realisation that came out of the calibration does: low-conductivity at the western part of the model, high conductivity zone in the middle of the wide eastern part, the two high-conductivity points situated 7km to the east. The mean values are around the same as the calibrated field, too.

Figure 31 depicts the coefficient of variation (in %) of the hydraulic conductivity of the 135 fields. There is a big area of low variation (~ 10%) from 6 to 10 km eastward from the western edge. This is where most of the aquifer tests lie, restricting parameter variability. Another low variation point (15 – 20%) is in the eastern part of the model. Here, high conductivities seem to be necessary to relay the groundwater quickly into Spring Creek, the main water outlet of the system. High coefficients of variation (up to 45%) can be found along the model edges, where no or few information is available.

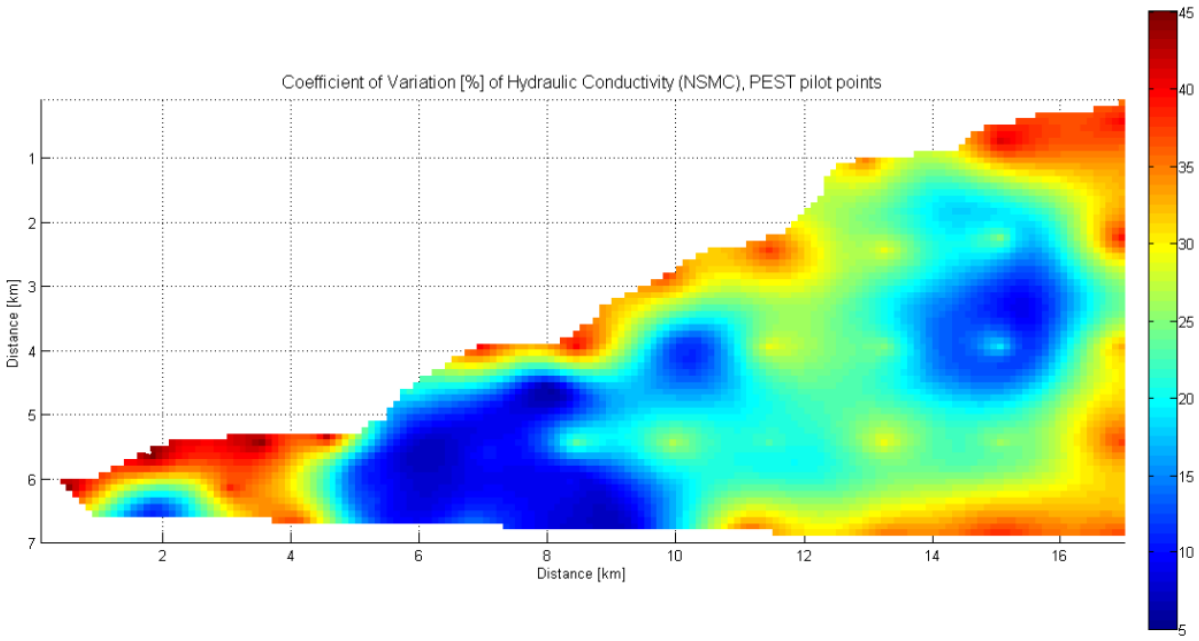

Figure 31: NSMC PEST Pilot Point calibration: Coefficient of Variation of Hydraulic Conductivity

Looking at the other parameters, their differing values for all 135 realizations are summed up as histograms in Figure 32. A green dot illustrates the value of the parameter in the original PEST Pilot Point calibration.

The ratio of horizontal to vertical hydraulic conductivity, K_{xy}/K_z, is fairly normally distributed between values of 5 to 10, with the original value of around 7 in this distribution. While a somewhat higher number of ratios around 5 occur, no realizations have values smaller than around 4, even though the parameter boundaries allowed K_{xy}/K_z to be as small as 1. A slower vertical groundwater flow seems necessary for the model.

The hydraulic conductivity of the aquitard is at the upper boundary of 20 m/d for most realizations, and some values between 10 m/d and 20 m/d (like the original calibration). Interestingly, more than 10 realizations have a value at the lower parameter boundary of around 1 m/d, with almost none in between. This value is the main regulator for the groundwater flow out of the system. Therefore, high values seem necessary to outlet all the water quickly enough to not raise groundwater heads in the eastern model part (where two head observations are).

Analysing the bed conductivities of the drains (K spring (S) to K drain (N)), the histograms mostly show somewhat normal distributions, with the original calibrated values around the means. The Southern Drain and Opawa River have some higher outliers (>1 m/d and >2m/d, respectively), indicating possible solutions where more water exfiltrates through these southern drains. The small Northern Drain has a high amount of very small conductivities at the lower range of its parameter boundary (<0.1 m/d). Lying between Spring Creek and the western part of the Wairau River, water in this area is redirected to those two outlets. This is partially forced due to the flux observation at Spring Creek.

The bed conductivities of the eleven river sections show mostly very low values below 0.1 m/d. Some of them seem to almost always have the low values (up to 100 realizations), while some others are higher more often. Since two thirds of the river is restricted by flux observations so are the river bed conductivities. This means that in a between two gauging

stations, a high river bed conductivity in one segment probably enforces lower river bed conductivities for the other segments. This is probably interchangeable in between the gauging stations, with always one higher and several lower river bed conductivities. That would explain why almost all sections have high parameter values sometimes, but not very often.

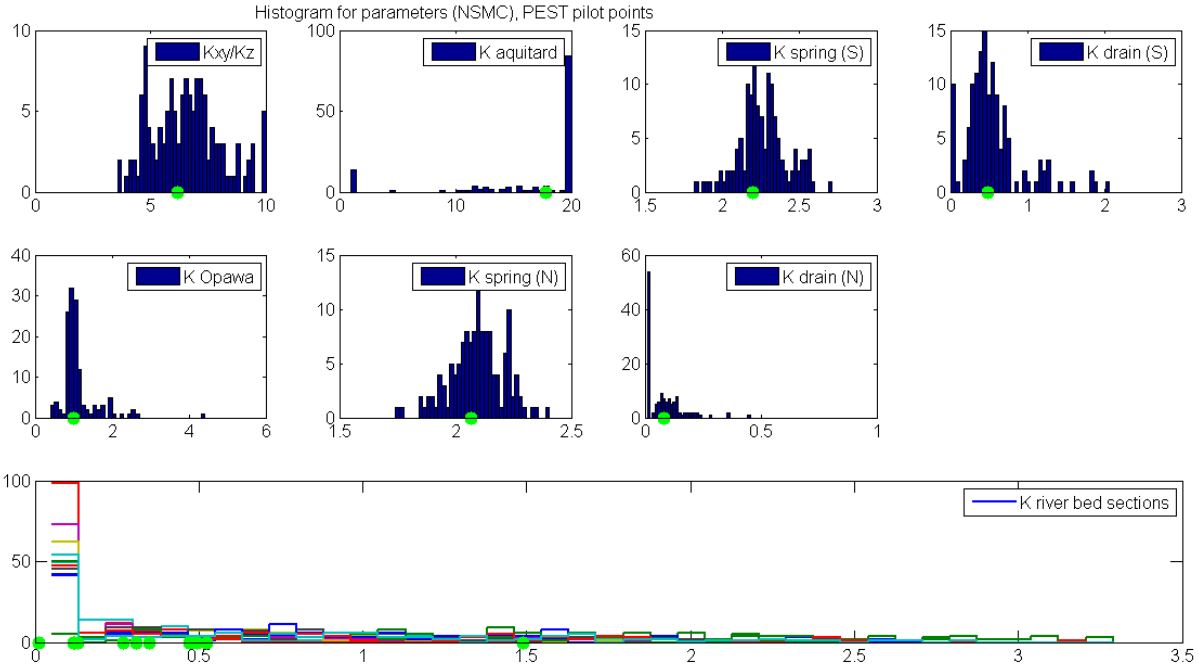

Figure 32: NSMC PEST Pilot Point calibration: Histogram of parameters (green dots indicate original values)

After the parameters, the distribution of the head field is of interest. Looking at the mean heads, the field shows all signs seen earlier in the originally calibrated one: west-to-east flow, bend head field around the river where water infiltrates the aquifer, and the mound at the south-eastern model edge. Looking at the standard deviations, the downstream part of the model is quite fixed regarding its groundwater head distribution. Here the standard deviations are smaller than 1 m. The same goes for the upper part where the two head measurements (Condors Recharge and Condors No2) lie. Upstream of that, head distribution is rather uncertain (up to 3 m standard deviation), since no real data is available for the western model edge. Furthermore, heads along the river are more uncertain. As they depend strongly on the hydraulic conductivity fields and river bed conductivities (both uncertain as well), they are influenced by more parameters than heads at other locations. Therefore, the uncertainties rise at these locations.

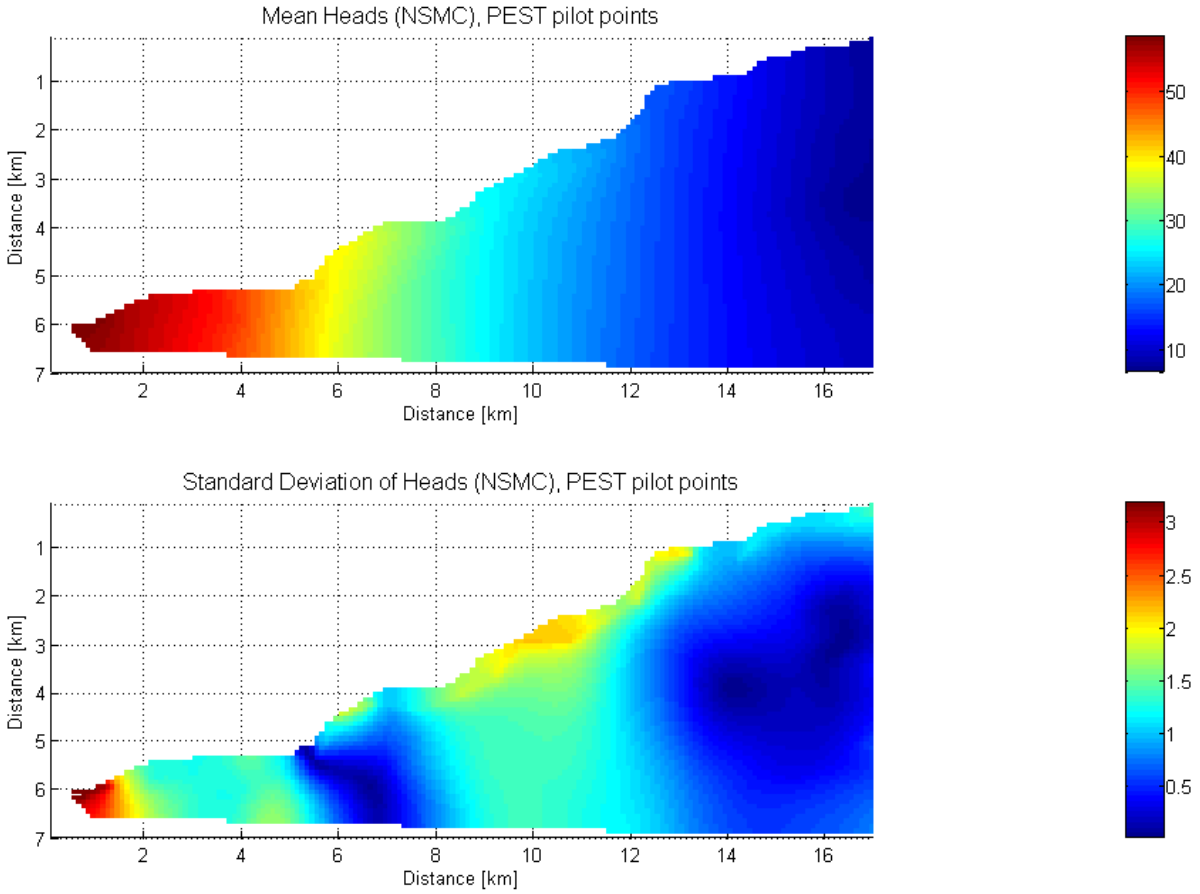

Figure 33: NSMC PEST Pilot Point calibration: Mean heads and standard deviation

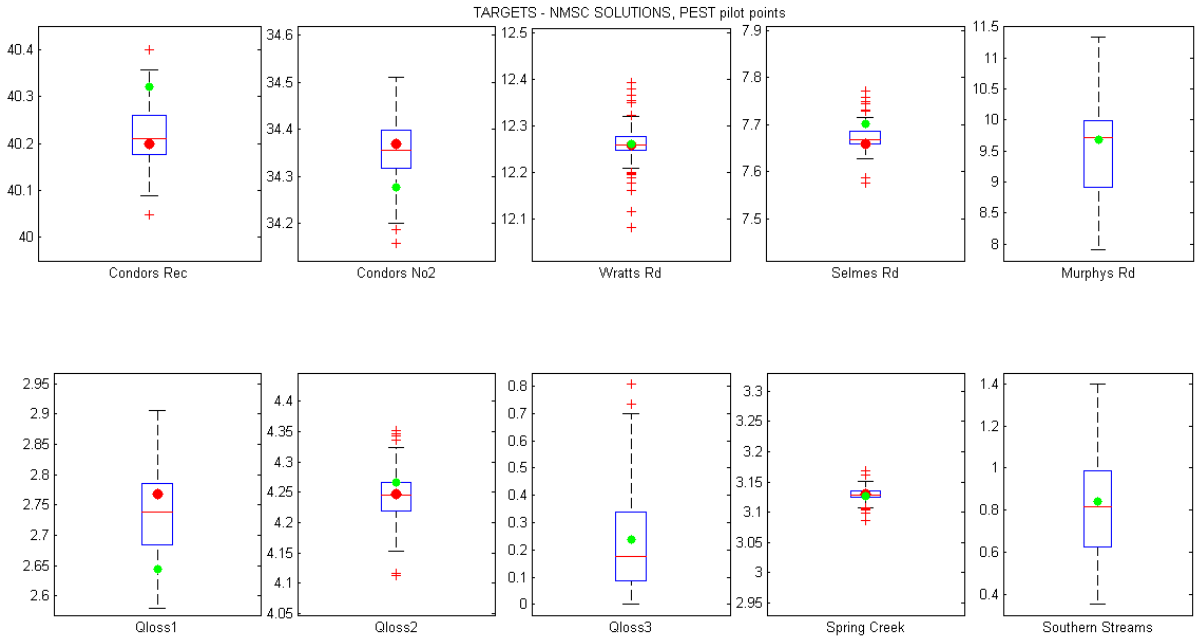

Figure 34: NSMC PEST Pilot Point calibration: Observations and predictions (green: original values, red: measured values)

Finally, the distribution of observation and prediction values for the 135 realizations is shown in the above Figure 34. In the plots, the edges of the boxes are marked by the 25th and 75th percentiles, respectively. The median is marked with the red line. The whiskers mark the

extent of the other data points, expect for data points marked with red crosses, which are considered outliers. The green dot marks the original calibrated value, and red dots are the measured values for the observations.

Looking at the four measured heads, measured values are always inside the boxes and close to the medians. Since the models were calibrated on these heads, the distribution of simulations around these values is unsurprising. For three of the four heads, the original calibration lies outside of the boxes, though. The width of the distribution range for all head measurements depends on the value: the higher the head, the wider the distribution. For the two measured river fluxes (Qloss1 and 2), the measured values are again within the boxes. For Qloss1 (between Rock Ferry and SH6), the original calibration not part of the 25th to 75th percentile, though. Since both river fluxes were calibrated on, the shown variability (0.1 m³/s or smaller for the 25 % to 75 %) is small. The last measurement for calibration, Spring Creek, shows a very small variability, with the original calibration and the measurement almost at the median. This shows that meeting the Spring Creek target seems to be unproblematic, since all model realizations fit this observations almost perfectly.

For the first prediction, groundwater head at Murphys Road, the 25 to 75 percentile box ranges over a meter, with more extreme values from 8 to 11.5 meters. The value from the original calibration lies right at the median here. Keeping in mind that for this head, an actual measurement of around 7 meters exists, this shows that a total of 135 realizations fail to meet this measurement for even the most extreme cases, again signifying the structural problems in the model. The second prediction, river flux from Wratts Road to SH1 gauging station (Qloss3), shows similar things. It exhibits a rather wide distribution, ranging up to 0.8 m³/s while going down to 0 m³/s. The aimed-for value here of -1.2 m³/s is not realistic for the model (due to the SH1 shifting), but small or even slightly negative values would probably be more realistic than some of the high values seen here. The third prediction, flux into the southern streams, ranges quite widely as well (a range of 1 m³/s for the extremes). Again, the original calibration value is right at the median of the distribution. Overall, the predictions are less certain than the calibration targets.

4.3.2 NSMC: AMALGAM Pilot Point calibration

Again, 200 random fields were created and recalibrated to ensure an acceptable fit. This time, only 71 of those fields remained for the ensuing analysis after recalibration.

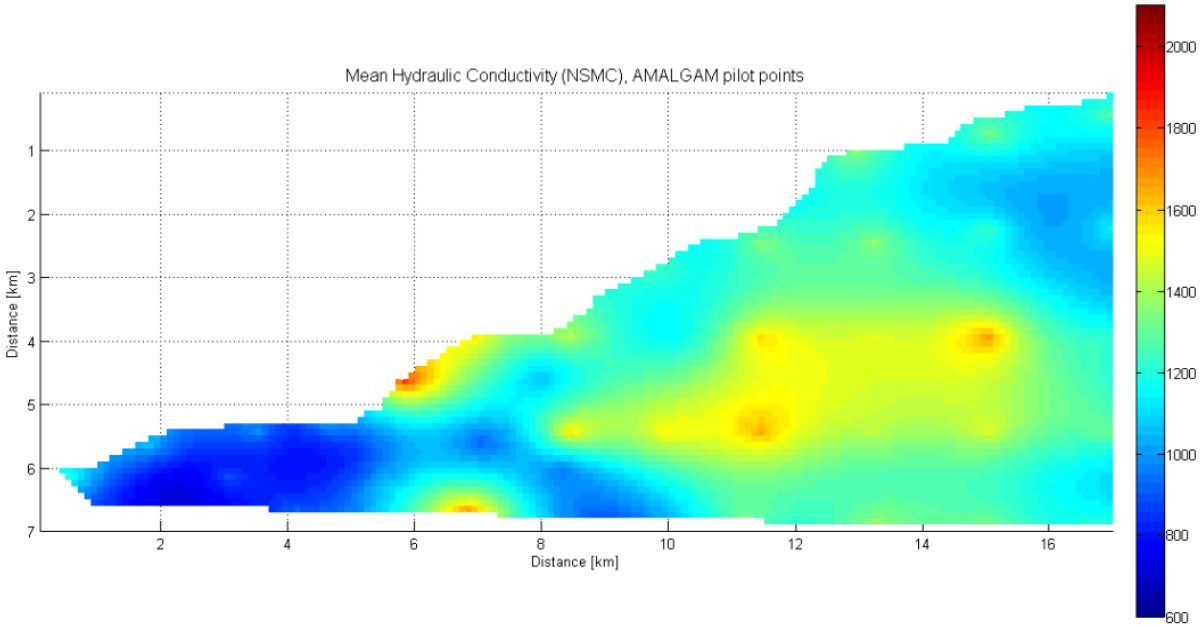

Figure 35: NSMC AMALGAM Pilot Point calibration: mean hydraulic conductivity

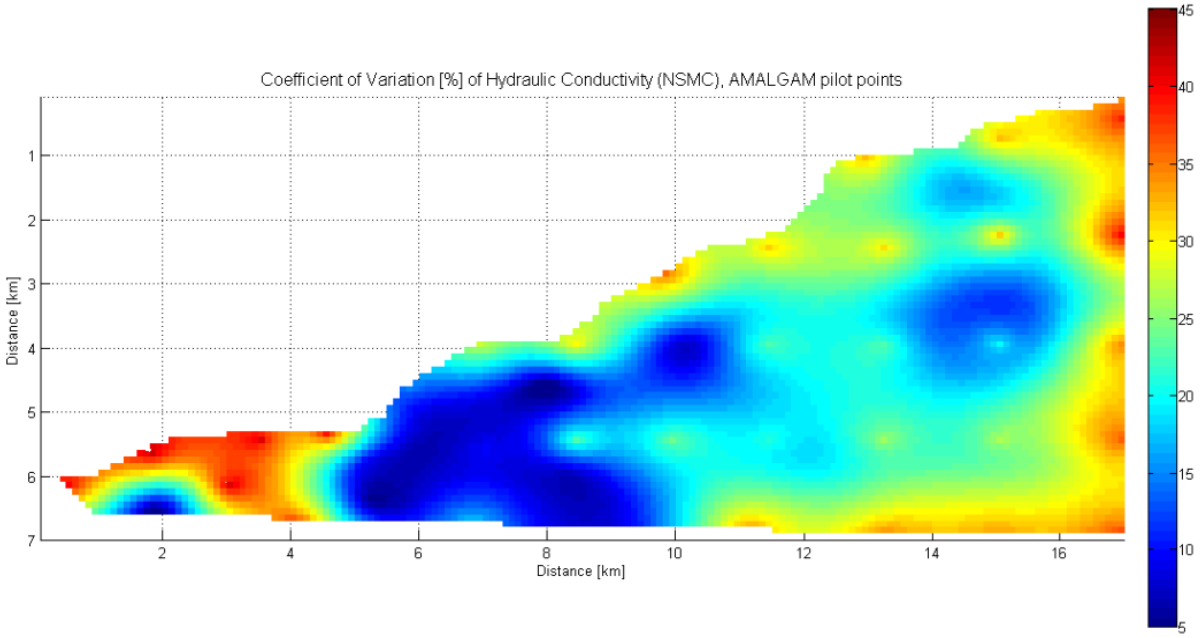

Figure 36: NSMC AMALGAM Pilot Point calibration: Coefficient of Variation of Hydraulic Conductivity

The mean hydraulic conductivity and its coefficient of variation show similar patterns and overall values to the results of the NSMC PEST analysis. Hydraulic conductivity is small at the western edge (< 800 m/d) and high in the middle of the eastern part (up to 1600 m/d). The variation is low where the aquifer tests are situated and where Spring Creek is (below 15 %), very high along the model boundaries in the west, east and south-east (up to 45 %) and somewhat high along the river (25 to 35 %). Again, the originally calibrated field is similar to the mean hydraulic conductivity field, so not an extreme field.

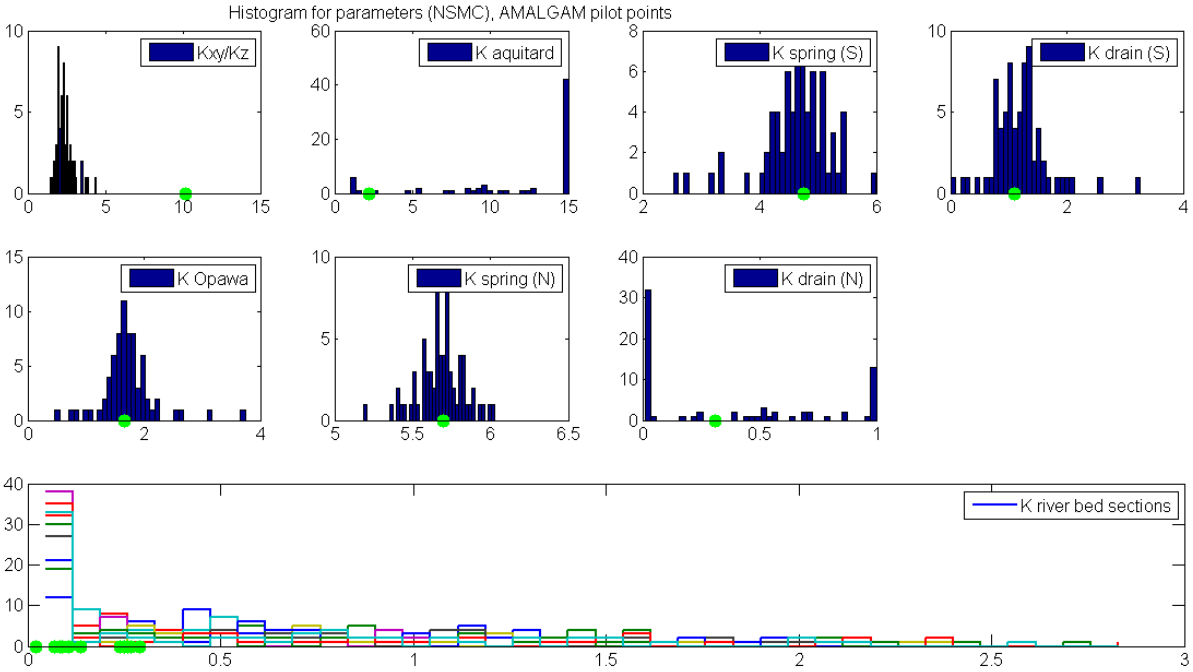

Figure 37: NSMC AMALGAM Pilot Point calibration: Histogram of parameters (green dots indicate original values)

For the histograms of the other parameter values, some differences can be seen to the PEST results, while for some parameters it looks similar.

In this case, the values for the ratio of horizontal to vertical hydraulic conductivity are all situated between 1 and 5, with the original value of around 10 a total outlier of the distribution. The high original value is similar to the PEST results, but there all values were distributed between 5 and 10. While it was assumed during the PEST NSMC results that this value has to be high, the AMALGAM NSMC results show the opposite. Maybe the observations are rather insensitive to the parameters and the fact that its values are distributed at the upper bound for PEST and the lower bound for AMALGAM is only random. Running more than the 71 or 135 different fields used for NSMC in this study could verify this.

For the hydraulic conductivity of the aquitard, the distribution of the NSMC results looks very similar to the results based on the PEST calibration.

The distribution of bed conductivities of the drains is comparable with the results based on the PEST calibration, with the positions of the original values also similar to the PEST NSMC analysis. The absolute values of the distribution for the two Spring Creek drains (spring (S) and (N)) are very different. They were considerably lower for the PEST NSMC analysis (ranging from 1.8 to 2.6 m/d and 1.8 to 2.4 m/d, respectively) than for the results based on the AMALGAM field (ranging from 4 to 5.5 m/d and 4.4 to 6 m/d, respectively). Why the AMALGAM NSMC analysis produces higher Spring Creek bed conductivities is inexplicable. The distribution of the other related parameters for the spring flux (hydraulic conductivity field, hydraulic conductivity of the aquitard) is very similar to the PEST NSMC results. Again, the Spring Creek observation might be more sensitive to other parameters than the bed conductivity, something the PREDUNC analysis also hinted at.

Finally, results of the river bed conductivity distribution are very similar to the ones found in the PEST analysis.

Since the figure of the mean heads and the standard deviation from those means for the Null-Space Monte-Carlo analysis of the AMALGAM calibration is almost identical to the one examined earlier in the PEST section, it is not stated here. The figure can be found under Figure 41 in the Appendix and its explanation would be the same as in the above chapter 4.3.1.

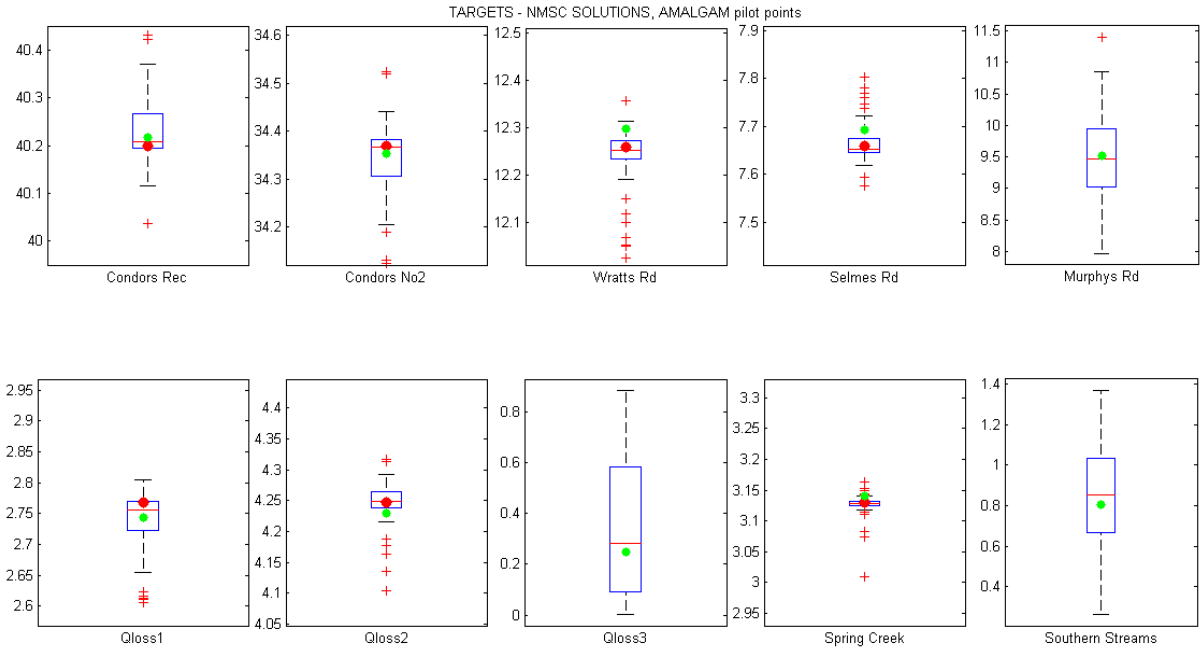

Figure 38: NSMC AMALGAM Pilot Point calibration: Observations and predictions (green: original values, red: measured values)

Figure 38 shows the box plots for the observations and predictions. Again, the red line is the median, the box marks the 25th to 75th percentile, red crosses are considered outliers, green dots are the original calibrated values and red dots the measurement values.

For the four calibrated heads, the pictures look somewhat similar to the PEST results. The values of the original calibration defer. They are closer to the measured value for the first two heads compared to the PEST results. The distribution of head simulations are very similar to the PEST NSMC analysis, though, in position of the median and range of the percentiles. The same can be said for the two river flux measurements (Qloss1 and 2) – while original values are different, the distributions are very similar to the PEST NSMC analysis. This is also true for the Spring Creek flux measurement. Again, the distribution of the NSMC simulations is very tight, and all simulated values are lower than the actual measurement by between 0.05 to 0.1 m³/s. These similarities for both NSMC analyses regarding the observations are not surprising. Good model-to-measurement fit (as well or better than the originally calibrated models) was guaranteed for all realizations used in the analysis. This greatly restricts variability for observations.

The three boxplots for the predictions are very similar to the PEST results. Again, all three predictions are much more variable than the observations. The two predictions with measured values (Murphys Road head ~ 7m, Qloss3 at -1.2 m³/s) are not meeting those

values with even the most extreme results of the NSMC analysis. The model structural reasons for this were explained earlier. The third river flux prediction between Wratts Road and SH1 gauging station (Qloss3) shows a tendency for higher values in both the PEST and AMALGAM NSMC analyses.

As is apparent from the above analysis, for some models out of the NSMC analysis the two measured predictions (head at Murphys Road, river flux between Wratts Road and SH1 gauging station) are closer to their measured values than the original calibrated models. While even the most extreme realizations are still off (and understandably so, as was explained), it might be of interest to analyse such an extreme realization. This is done in the following:

Table 22: NSMC AMALGAM Pilot Point realization #69: model fit

Observation / Prediction	Type [Unit]	Observed Value	Original Calibration	NSMC Field #69
Condors Recharge	Groundwater Heads [m]	40.199	40.216	40.213
Condors No.2		34.370	34.353	34.408
Wratts Road		12.258	12.298	12.222
Selmes Road		7.662	7.692	7.699
Rock Ferry – SH6	River Flow Losses [m³/s]	2.768	2.744	2.762
SH6 – Wratts Road		4.247	4.230	4.263
Spring Creek	Drain Flux [m³/s]	-3.128	-3.141	-3.128
Murphys Road	Groundwater Head [m]	7.022	9.525	7.959
Wratts Road – SH1	River Flow Losses [m³/s]	-1.282	0.248	-0.133
Southern Streams	Drain Flux [m³/s]	-	-0.807	-0.392

The table above shows the observation and prediction values for a single realization taken out of the AMALGAM NSMC results as well as the calibration results of the original AMALGAM Pilot Point realization used in this thesis. This model was specifically chosen for its extremely low head at Murphys (closer to the measured 7 m) and negative net flux between the Wratts Road and SH1 gauging stations. As obvious from the table, the model fit to the observations is similar to the original model. For the predictions (the three bottom entries), the simulated values are less off for the two measured predictions: Murphys Road is only about 1 meter off instead of 2.5 m, and the river flux between the last two gauging stations is overall negative instead of positive.

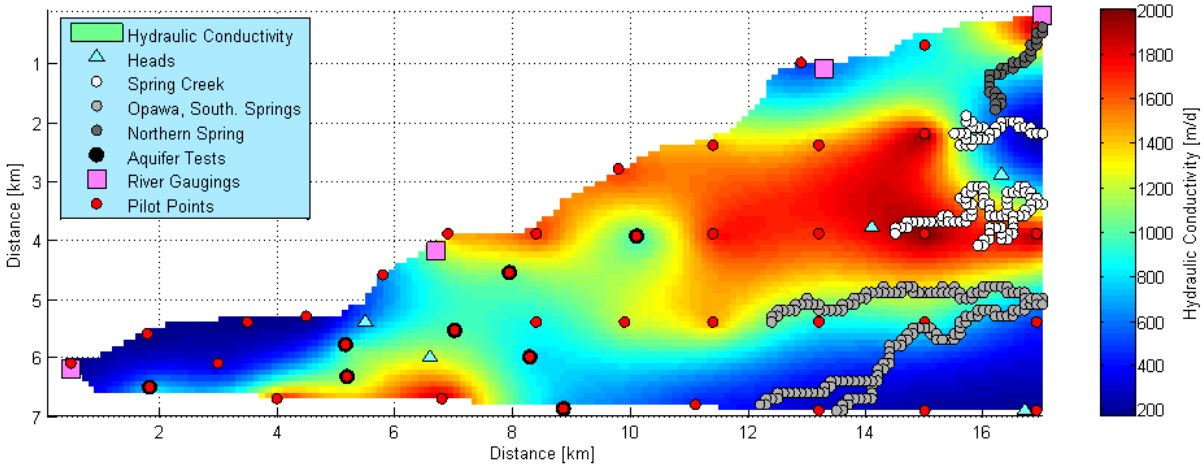

Figure 39: NSMC AMALGAM Pilot Point realization #69: hydraulic conductivity field

The Figure 39 above shows the hydraulic conductivity field of this realization. One can easily see that the span of values is much higher, from 200 m/d up to 2000 m/d (instead of 700 m/d to 1500 m/d). This obviously allows a lot more flexibility. While these values were allowed in the calibration method, regularisation constraints were supposed to penalize models with such fields, since they are deemed less realistic. During the NSMC recalibration, no regularisation was constraining the models, allowing the creation of such fields. Since the aim of the applied Null-Space Monte-Carlo method is the uncertainty analysis of parameters, observations and predictions, such fields are acceptable. This field would probably not be chosen as a preferred outcome of a calibration, tough.

In Table 23, parameter values of the original calibration and the random NSMC model are compared. As stated during the discussion of the model calibrations, no real information about these parameter values is given as a guideline for acceptance or dismissal of a model on behalf of these values. Only for the river bed conductivity, a certain uniformity was preferred on its probable realism. Regarding this benchmark, the random NSMC model does well. While rather small values (mostly < 0.1 m/d), the river bed conductivities are distributed uniformly. Therefore, these parameter values give no further insight for the assessment of the random NSMC model realization.

Table 23: NSMC AMALGAM Pilot Point realization #69: parameter values

Parameter Name / Group	Unit	Parameter Number	Original Calibration	Random NSMC Model
Aquitard Hydraulic Conductivity	m/d	-	2.24	2.01
K_{xy}/K_z	-	-	10.19	15
River bed conductivity	m/d	1	0.08	0.65
		2	0.11	0.01
		3	0.07	0.01
		4	0.09	0.16
		5	0.02	0.01
		6	0.14	0.02
		7	0.25	0.04
		8	0.26	0.04
		9	0.29	0.02
		10	0.27	0.01
		11	0.24	0.01
Drain bed conductivity	m/d	1 (Spring Creek)	4.75	4.44
		2 (Southern Drain)	1.11	1.26
		3 (Opawa River)	1.67	1.80
		4 (Spring Creek)	5.70	5.41
		5 (Northern Drain)	0.31	0.01

Similar analysis could be done for different NSMC realizations at extreme ends of the spectrum. This test case was chosen to show that a potential better fit to the two measured predictions could only be achieved with certain limitations on other aspects of the model. In this case, the hydraulic conductivity field was not restricted by the regularisation constraints and therefore varies widely compared to the original calibrated model. This was assessed as unrealistic.

4.4 Discussion of the steady-state models

In this chapter, all results of the analysis of the three model calibrations and the corresponding uncertainty analysis are summarized and discussed. This should provide condense insight on what information the models convey, how they compare to each other and what problems or shortcomings of the models are.

The main results of the model calibration are three calibrated models with an acceptable model fit to the observations. A hydraulic conductivity parameterization of only two zones for the model domain led to reasonable results. On the other hand, two Pilot Point parameterizations creating interpolated fields were generated using two different methods of model calibration: PEST's linear approach and choosing a single field out of the Pareto solutions of AMALGAM. These two models were calibrated to satisfactory simulation-measurement fit, too.

Further analysis of the two-zone model revealed certain deficiencies. Its calibrated parameters turned out to be less realistic with the model's inability to include aquifer test information into the hydraulic conductivity field and its necessity to compensate for missing hydraulic conductivity heterogeneity with fluctuating river bed conductivities. This seemed unwanted on this scale of abstraction and led to less homogenous river water infiltration into the aquifer. Therefore, the model is probably less suitable for further analysis of the Wairau

system, especially if new data on hydraulic conductivity or river bed composition should be available and included into the model. Furthermore, it is doubtful that the parameterization of the hydraulic conductivity is complex enough for further studies on transient water flow and river-groundwater interaction in the system.

The other two models do not have these deficiencies in those areas and therefore seem more suitable for further analysis. Unfortunately, highly parameterized modelling has the problem of non-uniqueness of its calibration result which had to be addressed. For both modelling cases, this problem was diminished by adding regularization constraints for certain members of the parameters, including spatial information of hydraulic conductivity distribution and adding numerical stability to the calibration process. Nonetheless, calibration with PEST only results in a single model, which is not a unique solution. The AMALGAM calibration yields a variety of models and allows the analysis of objective trade-off and parameter distribution, therefore addressing this factor. In this thesis, only one solution of those results was chosen for comparison with the PEST result.

While different in calibrated parameter values and calibration method, both Pilot Point models show a lot of similarities, suggesting that the data lead to certain structures and are therefore informative for model calibration. This was supported by the uncertainty analysis: parameter variability was mostly low, and the two calibrated cases were mostly close to the mean values. Furthermore, the influence of boundary conditions based on assumptions (well abstraction, eastern constant flux, SFR parameterization) on model observations and predictions was very small, verifying the appropriateness of these assumptions.

Unfortunately, these models still showed certain problems: the head measurement at the south-eastern corner, Murphys Road, and the third river-groundwater flux estimate (between the Wratts Road and SH1 gauging stations) could not be met and were therefore excluded from model calibration. Again, this impression was verified by the results of the uncertainty analysis, where both were highly uncertain. While certain single realizations of the Null-Space Monte-Carlo analysis fit both the observations acceptably and these two predictions somewhat better than the chosen calibrated models, it was shown that this is due to a trade-off between the predictions and other model assumptions like the regularization of the hydraulic conductivity field and the river bed conductivities. This led to the question of usability of these two predictions as observations in the model calibration.

The first prediction is the negative net flux between the gauging stations Wratts Road and SH1. The change from river-to-groundwater flow back to groundwater-to-river flow in the last segment of the modelled river could be verified by all models. The magnitude implied by the measured value could not be reproduced, though. This is due to the upstream shift of the SH1 gauging station relative to its real location in the model. This shift reduces the overall distance in between the two gauging stations. Furthermore, this distance reduction is taking place in the most eastern part of the model domain, where water flux from groundwater back into the river is starting. The model includes the part where river-to-groundwater flow is taking place, while excluding a main part of the domain where groundwater-to-river flow happens. This greatly changes the net flux between the two gauging stations. Therefore, the model can not reproduce this estimated flux and it is not usable as an observation due to the relocation of the SH1 gauging station. Potentially, the model could be extended eastward to include the real SH1 gauging station to change this situation.

For Murphys Road, the problem is a combination of certain details in the model. First of all, the location of the groundwater head measurement is at the south-eastern border of the model domain, right between two boundary conditions: the no-flow boundary in the south and the constant flux boundary in the east. Both are rough estimates based on expert knowledge, and therefore uncertain. Obviously, a shift of the no-flow boundary (or even a change to a flux boundary, either positive (northward) or negative (southward)) would obviously change the system locally, affecting this head measurement. The constant flux out of the model domain in the east and its spatial distribution are only rough estimates. Therefore, changing the overall flux or allowing more water to exfiltrate in the south-east is a real possibility to address this problem. Furthermore, the head at Murphys Road is very dependent on the interaction with the two southern streams, the Southern Drain and Opawa River. Both are included, but without any real knowledge besides their locations. The inclusion of potential flux measurements at these springs, as well as more information about their drain bed composition, would greatly help knowing this local part of the model. Not even the drain's surface elevation is known – with the coarseness of the DEM, the drain channels are not represented in the model surface. Therefore, the drains were lowered into the 4^{th} layer on the assumption of their position about 1.5 meters below the surrounding surface. Measurements of drain channel elevations could change the vertical location of these drains dramatically, thereby strongly influencing the flow field and heads in the area. As is, the groundwater head at Murphys Road can not be used as a measurement in the model.

5 Conclusion and outlook

The original purpose of the model is the analysis of river-groundwater interactions between the Wairau River and its aquifer in the system. Therefore, main focus of the analysis of the models and of future structural changes, potential data acquisition and problem trade-off should be this interaction. Under this aspect, the inability to fit the head at Murphys Road far away from the river is less important than the model's ability to fit other heads and the river flux estimates. Furthermore, the two Pilot Point models exhibited nice distributions over river-to-groundwater flow along the river, as well as the mentioned flow back into the river in the lowest part of the model. In conclusion, both calibrated models are adequate bases for further modelling analysis of the Wairau system.

The next step in this further modelling is arguably the transient modelling. While the summer period on which the steady-state model was based is relatively homogeneous in river flow, well abstraction and recharge, the ability of a transient model to project daily fluctuations is nonetheless a key factor. Increasing modelling time to the seasons of highly fluctuating surface hydrology of autumn, winter and/or spring would further increase difficulty to fit measurements, while resulting in a much more robust model for long-term prediction of aquifer response to those weather-controlled events. Therefore, this might be done step by step, with some recalibration of both already calibrated parameters (like the hydraulic conductivity field and the river bed conductivities) as well as the addition of new parameters relevant for transient modelling (storativity of the aquifer).

Furthermore, it might be of interest to implement transport schemes into the model to analyse water travel times in the system, for example. This might then be used as predictions for further uncertainty analysis, for example doing a data worth analysis for the expansion of the existing monitoring network. Transport could be equally interesting in the case of solute movements for analysis of agricultural influences, for example.

6 References

Cunliffe, J.J., (ed.) 1988. Water and soil resources of the Wairau Volume Two: Groundwater. Marlborough Catchment and Regional Water Board, 107p.

Doherty, J., 2010a. PEST: Model-Independent Parameter Estimation – User Manual: 5th Edition. Watermark Numerical Computing. Document can be found via the following web address: http://www.pesthomepage.org/Downloads.php.

Doherty, J., 2010b. Addendum to the PEST manual. Watermark Numerical Computing. Document can be found via the following web address: http://www.pesthomepage.org/Downloads.php.

Doherty, J., 2010c. Methodologies and Software for PEST-Based Model Predictive Uncertainty Analysis. Watermark Numerical Computing. Document can be found via the following web address: http://www.pesthomepage.org/Downloads.php.

Doherty, J., 2013. Groundwater Data Utilities – Part B: Program Descriptions. Watermark Numerical Computing. Document can be found via the following web address: http://www.pesthomepage.org/Downloads.php.

Harbaugh, A., 2005. MODFLOW-2005, the U.S. Geological Survey Modular Ground-Water Model – the Ground-Water Flow Process. U.S. Geological Survey Techniques and Methods, 6-A16, variously paginated.

Niswonger, R.G. and Prudic, D.E., 2005. Documentation of the Streamflow-Routing (SFR2) Package to include unsaturated flow beneath streams--A modification to SFR1. U.S. Geological Survey Techniques and Methods, Book 6, Chap. A13, 47 p.

Vrugt, J.A. and Robinson, B.A., 2007. Improved evolutionary optimization from genetically adaptive multimethod search. *Proceedings of the National Academy of Science*, 104 (3), p. 708-711.

Wilson, S. and Wöhling, T., 2015. Wairau River-Wairau Aquifer Interaction. Lincoln Agritech Ltd. Report 1003-5-R1, 49 p.

Winston, R.B., 2009. ModelMuse-A graphical user interface for MODFLOW-2005 and PHAST. U.S. Geological Survey Techniques and Methods, 6-A29, 52 p.

7 Appendix

Table 24: Well Zones and Abstractions

Zone Number	Well Abstraction [m/d]
1	$-5.6571*10^{-3}$
2	$-4.777143*10^{-3}$
3	$-3.771429*10^{-3}$
4	$-3.771429*10^{-3}$
5	$-4.651429*10^{-3}$
6	$-4.651429*10^{-3}$
8	$-4.651429*10^{-3}$
9	$-4.902857*10^{-3}$
10	$-4.651429*10^{-3}$
11	$-4.902857*10^{-3}$
13	$-6.914286*10^{-3}$
14	$-6.914286*10^{-3}$
15	$-5.531429*10^{-3}$
16	$-4.148571*10^{-3}$
17	$-4.777143*10^{-3}$
18	$-4.525714*10^{-3}$
19	$-4.525714*10^{-3}$

Table 25: Calibrated Pilot Point Values

Pilot Point Number	Pilot Point Location		PEST calibrated value [m/d]	AMALGAM calibrated value [m/d]
	x [cells]	y [cells]		
1	70.17	55.40	1000.00	985.82
2	88.63	68.68	943.75	936.13
3	79.47	45.61	1077.26	1067.59
4	82.92	59.91	1096.87	966.45
5	51.90	63.25	750.00	1032.43
6	51.63	57.80	575.625	942.40
7	101.07	39.41	1071.28	1128.82
8	18.36	65.05	971.875	713.06
9	150.00	7.00	790.04	1146.22
10	169.00	4.00	897.32	936.83
11	114.00	24.00	886.58	1225.21
12	132.00	24.00	966.02	1312.84
13	150.00	22.00	1200.54	1177.45
14	169.00	22.00	1293.97	1071.66
15	69.00	39.00	1021.37	1488.88
16	84.00	39.00	825.96	1117.96
17	114.00	39.00	1186.25	1274.94
18	132.00	39.00	1310.86	1510.54
19	150.00	39.00	1825.93	1349.55
20	169.00	39.00	1349.32	1179.66
21	18.00	56.00	801.35	708.25
22	35.00	54.00	802.15	738.93
23	58.00	46.00	1813.94	1455.75
24	84.00	54.00	1705.64	1169.47
25	99.00	54.00	1711.84	1141.48
26	114.00	54.00	1678.21	1260.59
27	132.00	54.00	1014.67	1269.83
28	150.00	54.00	1232.80	1222.78
29	169.00	54.00	1025.06	1101.07
30	132.00	69.00	1449.51	1083.79
31	150.00	69.00	1181.88	1321.78
32	169.00	69.00	1302.90	976.31
33	5.00	61.00	799.717	640.17
34	111.00	68.00	1240.61	1242.51
35	68.00	67.00	2164.60	1462.78
36	40.00	67.00	934.59	756.70
37	98.00	28.00	1133.56	1103.68
38	30.00	61.00	1426.91	794.98
39	45.00	53.00	762.18	828.03
40	129.00	10.00	846.70	1252.27

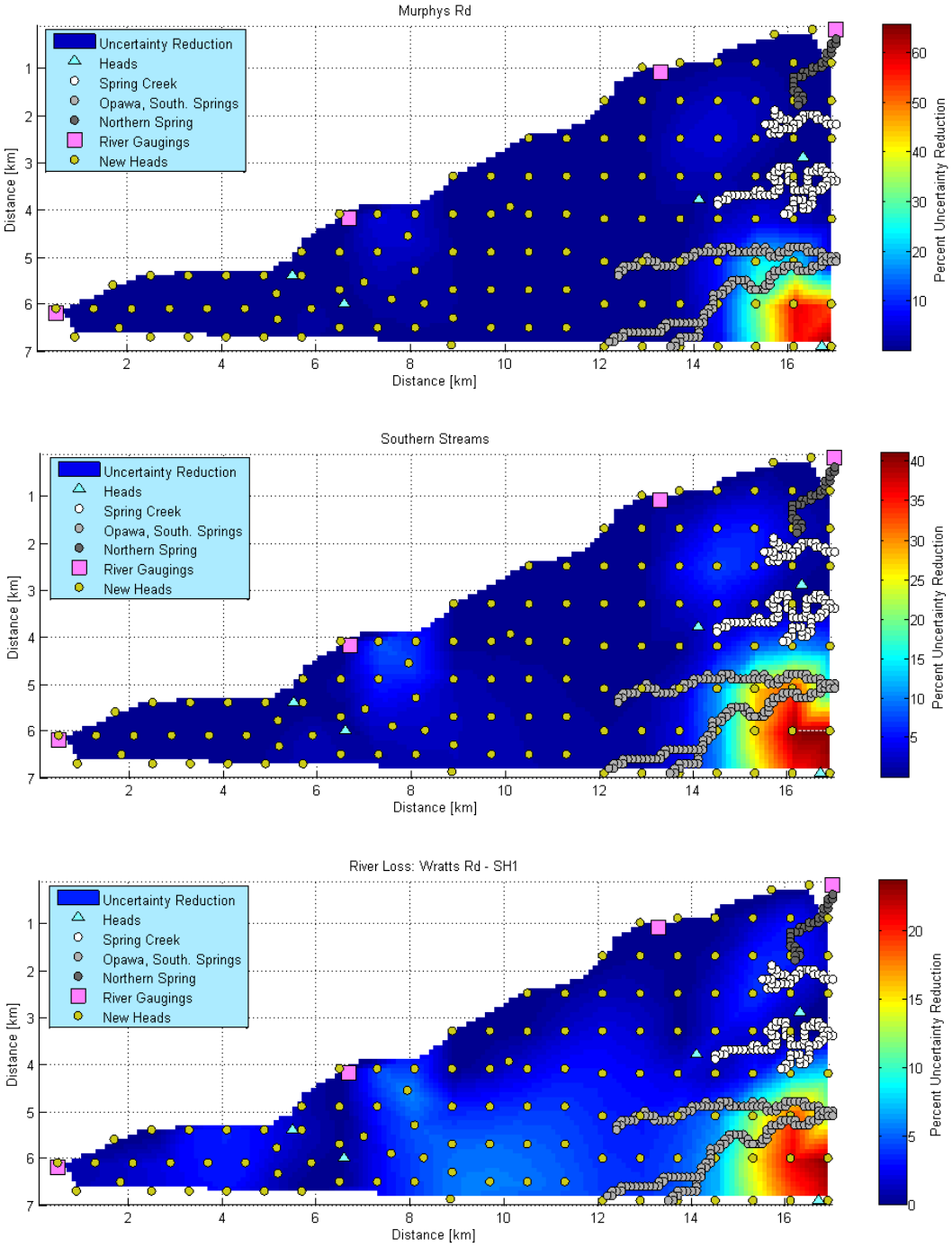

Figure 40: AMALGAM Pilot Point calibration: new head uncertainty reduction

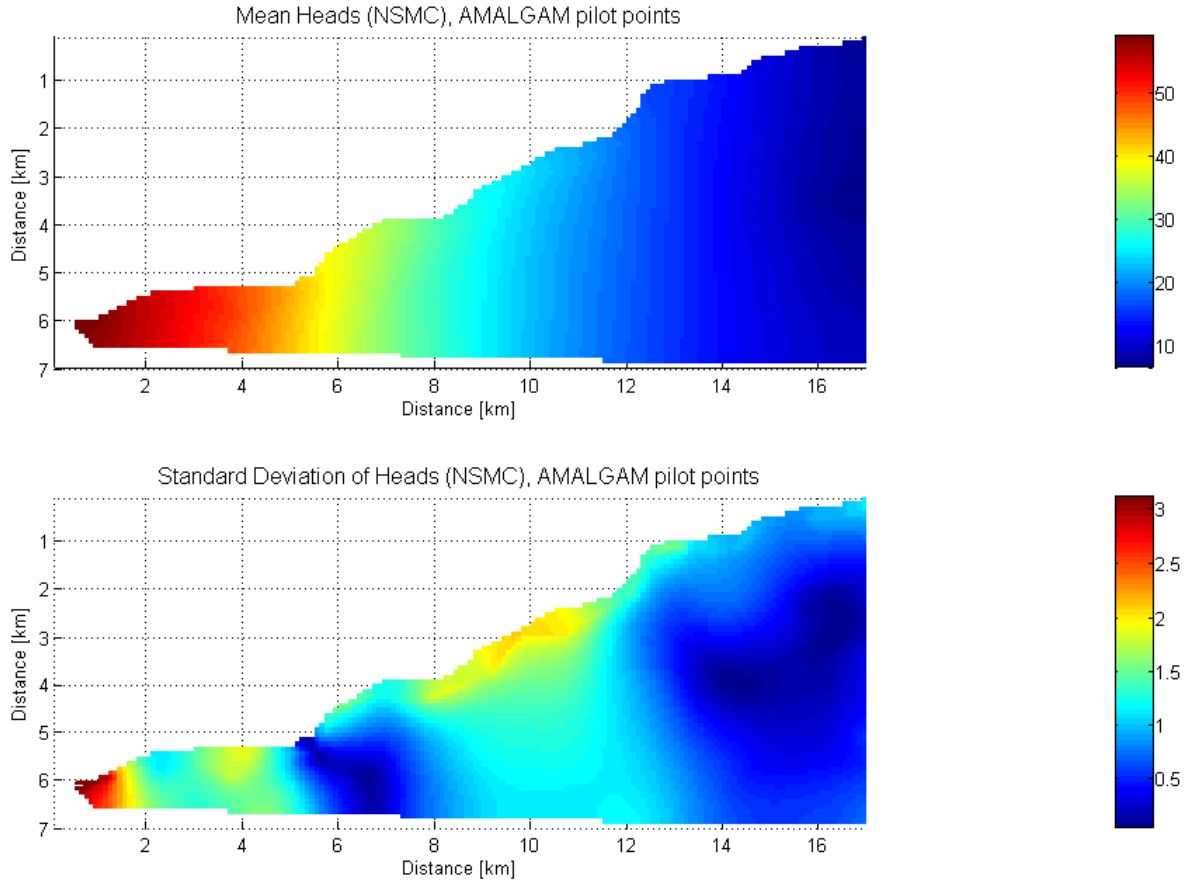

Figure 41: NSMC AMALGAM Pilot Point calibration: Mean heads and standard deviation