Quantifying river-groundwater interactions of New Zealand's gravel-bed rivers: The Wairau Plain.

Thomas Wöhling[1,2], Moritz Gosses[1], Scott Wilson[2], Peter Davidson[3]

[1]Technische Universität Dresden, Department of Hydrology, 01069 Dresden, Germany.
[2]Lincoln Agritech Ltd., Ruakura Research Centre, Hamilton 3240, New Zealand.
[3]Marlborough District Council, Blenheim, New Zealand

Abstract

New Zealand's gravel rivers have deposited coarse, highly conductive gravel sediments that are covered with relatively poorly developed thin soils. The shallow groundwater in these stratified gravel aquifers is predominantly fed by river water. Recharge mechanisms in these rivers are poorly understood and the management of the groundwater resources is challenging, particularly under a more variable future climate. To better understand the river-groundwater exchange processes in these rivers, we investigate the Wairau Plain aquifer which is pumped for drinking water and irrigation. A three-dimensional surface water - groundwater flow model (MODFLOW) has been set up using a revised geological model of the Wairau Plain. The model was calibrated using targeted field observations, "soft" information from experts of the local water authority, and the model-independent parameter estimation software PEST. We determined the trade-off between data fit and parameter homogeneity using regularization techniques and the PEST Pareto analysis tool. The calibrated model performs well for both the calibration data set and independent data. Flux-weighted transit-time distributions for the largest spring at the Wairau Plain were calculated using particle tracking. Mean transit times of less than 1 yr suggest very young water for Spring Creek. The uncertainty of the model simulations was evaluated using Null-space Monte-Carlo methods and is larger for mean travel times than the for river-groundwater exchange flows. Our analysis suggests that the river is hydraulically perched above the regional water table in its upper reaches and is gaining downstream where marine sediments overlay the unconfined gravels and where most of the springs of the Wairau Plain emerge. The net recharge to the Wairau aquifer is on average 7.3 m^3s^{-1}. Although the river discharge is highly dynamic and regularly exceeds 1000 m^3s^{-1}, the net exchange flow is capped and rarely exceeds 12 m^3s^{-1}. Changes in aquifer storage are mainly affected by the frequency and duration of low-flow periods in the river. This study significantly improved our understanding of Wairau river-groundwater exchange mechanisms. We

hypothesise that the methodology and general mechanisms are transferable to other New Zealand rivers with similar characteristics.

1 Introduction

Many New Zealand rivers flow from mountain valleys onto alluvial plains where they have deposited Quaternary gravel sediments of varying thickness (Rosen and White, 2001). These rivers lose water to shallow, unconfined aquifers formed by the alluvial fans and gain water near the coast as groundwater moves into confined aquifers and returns to the surface (e.g. Larned et al., 2008). Lowland aquifers are often an important water resource for municipal, agricultural, and industrial uses (e.g. Brown et al., 1999; Rosen and White, 2001). The management and protection of these water resources requires a good understanding of the interacting processes, particularly the quantification of river-groundwater exchange rates and their prediction under changing environmental conditions. River recharge can be the major source for groundwater in gravel-bed river systems and land-surface recharge is typically much lower. For the Heretaunga Plains aquifer, for example, the annual rainfall recharge is only 3% of the river recharge (Dravid and Brown, 1997). Less than 20% of estimated Avon River base flow is rainfall recharge (White, 2009).

Although the research on surface-subsurface exchange processes has increased dramatically towards the end of the last century (Stanley and Jones, 2000), the understanding and quantification of the interaction processes present still a major challenge (Sophocleous, 2002; Brunner et al., 2011; Lamontagne et al., 2014). Field techniques to quantify river-groundwater exchange rates encompass, amongst others, the measurement of the hydraulic gradient between the river and the adjacent groundwater, dilution tests with chemical or heat tracers, pumping or slug tests, and mass balance approaches (for a comprehensive review c.f., Kalbus et al., 2006; Rosenberry and LaBaugh, 2008; González-Pinzón et al., 2015). Some of these methods are rather elaborate and time-consuming and also difficult to transfer to larger scales. Differential stream flow gauging is a more readily applied mass balance method for the river-reach scale, where the net loss/gain over the length of a river section is determined by the difference between gauged flows at an upstream and a downstream cross-section. However, the flow channels of gravel-bed rivers in New Zealand are typically braided which could require simultaneous flow measurements in multiple braids (White et al., 2001). In addition, the mass balance approach requires the quantification of all other sources and sinks along the river reach such as tributaries, potential water takes, and underflow within the confines of the active river bed.

A controlling factor to determine the exchange rates between surface water and groundwater is the state of connection between the two compartments (Brunner et al., 2009, 2011). This is often poorly understood in the field. In addition, the state of connectivity might change due to the more dynamic nature of river flows and a delayed reaction of groundwater levels. Therefore, Brunner et al. (2011) called for more field studies dealing with the state of disconnection. In New Zealand, several studies have been dedicated to investigate river-groundwater connectivity and river-groundwater exchange flows (e.g., Brown et al., 1999; Larned et al., 2008; Rupp et al., 2008; White, 2009; White et al., 2012; Close et al., 2014). The Selwyn River is a prime example for highly complex spatio-temporal flow patterns and various states of connectivity and large river water losses in the alluvial plains (Larned et al., 2008; Rupp et al., 2008). Other New Zealand river systems showed consistent flow patterns over larger periods of times. Differential discharge measurements taken between 1957 and 1995 at a 3 km section of the Ngaruroro River show a consistent loss of 4.3 m^3s^{-1} for river flows below 35 m^3s^{-1} (Dravid and Brown, 1997). Similarly, consistent flow losses were reported for sections of the Rakaia River and the Waimea River by White et al. (2001).

Since field measurements are time-consuming, expensive and often not at the targeted time/space scale, the estimation of river-groundwater exchange rates are often complimented by hydrological modelling. Numerical models can be used to integrate field observations of various types and to investigate scenarios for (regional) water management (e.g., Panday and Huyakorn, 2004; Kollet and Maxwell, 2006; Jones et al., 2008; Spanoudaki et al., 2009; Maxwell et al., 2015). Several competing models and modelling schools have been discussed in the scientific literature (e.g., LaBolle et al., 2003; Furman, 2008) which is not repeated here. A comprehensive review of regional integrated models has recently been presented by Barthel and Banzhaf (2016). Integrated models, that simulate both saturated and unsaturated flow, as well as surface water, groundwater and the full coupling between them in a physical way (Brunner et al., 2010), can be highly accurate. Yet they require a large amount of data for their parametrization and their practical application is often restricted by large run-times (von Gunten et al., 2014). This is particularly challenging in braided river systems where the flow channel geometry is extremely complex and frequently changing over time. Some attempts were made to generate braided river terrain models in New Zealand for the Rees River and the Waimakariri River using airborne photography, LIDAR, and multi-point statistics (Pirot et al., 2014; Williams et al., 2014, 2016) but these methods are far from being routinely applied in surface water - groundwater modelling.

On the other hand, conceptual models that treat subsurface compartments as reservoirs are less data-hungry, have less parameters and are typically much faster. In the New Zealand context, Yang et al. (2017) introduced an additional conceptual groundwater store to the national hydrological model TopNet (Bandaragoda et al., 2004; Clark et al., 2008) to account for water transfer from rivers and also for cross/inter-catchment groundwater flow. However, river losses/gains are inputs to the TopNet model and are considered to be constant over time. This limits the potential application of TopNet to river basins where the river-groundwater exchange flows are known and time-invariant.

The numerical model MODFLOW is most frequently used to simulate surface water - groundwater interactions (Furman, 2008). MODFLOW (Harbaugh, 2005) in its calculations distinguishes between hydraulically connected and disconnected states and generally constitutes a good compromise between fully coupled models and conceptual models. Brunner et al. (2010) have revised the assumptions of MODFLOW in the context of simulating surface-water - groundwater interactions and provided some guidance about its application. In a later study it was concluded that the behaviour of disconnected river can often be approximated by neglecting the unsaturated zone (Brunner et al., 2011). MODFLOW has previously been applied in New Zealand (e.g. Fenemor, 1989; Baalousha, 2012; Gusyev et al., 2013).

In our study, we also use MODFLOW as a simulation tool to analyse and quantify surface water - groundwater interaction in gravel-bed rivers. The aims of this study are summarized as follows:

- to present modelling techniques for integration of hydrological data of various types with the specific focus on understanding river-groundwater exchange flows,
- to perform a detailed investigation of the spatial and temporal variation of the exchange flows, their dependence on river flows, and the state of connection between river and groundwater, and
- to assess parametric and predictive uncertainty on simulated groundwater heads, net exchange flows, as well as spring flows and transit times using rigorous yet pragmatic methods suitable for highly-parametrized models.

We will demonstrate our approach for a section of the Wairau River on the Wairau Plain, which interacts strongly with the shallow, stratified and highly-conductive gravel aquifer. The aquifer is managed by the Marlborough District Council (MDC) and is of regional importance because it supplies all of the municipal water requirements for Blenheim, Renwick, and Woodbourne, together with most of the vineyard irrigation supply.

The reminder of the paper is structured as follows. First we present the study area and the corresponding MODFLOW model. Then we describe the calibration strategy including the various calibration targets and our proposed uncertainty quantification methodology. In Section 3 we present the results of the model calibration, the analysis of the river-groundwater exchange mechanisms, as well as model predictions of transit-time distribution for the largest Spring on the Wairau Plain. The paper is concluded by a synthesis of our findings.

2 Materials and Methods

2.1 Wairau Plain

The study site is located in the lower reaches of the Wairau River catchment in the Marlborough District of the northern South Island, New Zealand. The Wairau River basin drains an area of 3430 km^2 which is covered by a mix of exotic pine and native beech forest in the northern and western ranges (elevation up to 2300 m) and pasture and shrub-lands in the southern hills. Just prior to discharging into the Pacific sea, the Wairau River enters the Wairau Plain which is New Zealand's largest wine growing area. Here, the braided gravel-bed river flows since modern times in a 100 - 200 m wide floodway at the northern edge of the Plain with constructed stop-banks as much as 1 km apart (Figure 1). The elevation of the Plain ranges from 72 m.a.s.l. in the West to sea level in the East over a distance of roughly 27 km. The river almost exclusively feeds the underlying Wairau aquifer which serves as the major resource for drinking water and irrigation in the region. It is the most extensive and important resource in the region by far and ranks amongst the most significant aquifers in New Zealand (Davidson and Wilson, 2011). A slow, but constantly declining trend in aquifer levels and spring flows have been observed over the past decades, which has triggered this investigation aimed at a better understanding of the recharge mechanism and the river-groundwater interactions.

In this study, we focus on a 22 km long section of the Wairau River that encompasses the entire recharge area and the majority of the Wairau Plain. It covers the river reach from downstream of the Waihopai River confluence to the SH1 bridge upstream of the Tuamarina River confluence (Figure 1).

Geology The earliest investigations of the Wairau Plain geology were carried out by Brown (1981). More recently, a detailed 3D geological model of the coastal Wairau Plain geology and its deeper aquifer structure was presented by Raiber et al. (2012). The basement geology of the Wairau basin consist of schist in the North and greywacke in the South. These rocks are overlain by a sequence of Pliocene to Pleistocene glacial outwash gravels interspersed with interglacial marine horizons at the coast. The youngest of these gravels is the Speargrass Formation, which is considered to form the base of the Wairau Aquifer. The Wairau Aquifer is hosted by high permeability Holocene sediments of the Rapaura Formation. These gravels have been formed by alluvial reworking of the Speargrass Formation and are orders of magnitude more transmissive. Towards the coast, the Rapaura Formation is overlain by marine silts of the Dillons Point Formation, which form a confining horizon to the Wairau Aquifer (Brown, 1981). More recently, Wilson (2016) reviewed the geological records of the Rapaura Formation for a more detailed analysis of its internal structure. Structure contours of the Speargrass Formation surface indicated that the Rapaura Formation has a maximum thickness of 30 to 35 m, and is typically 20 m over most of the aquifer. Three lithological members were distinguished, with some lateral variability evident in the uppermost member:

- Upper Member: 8 ± 3m of mostly stratified gravels of moderate permeability incised locally by facies of high permeability associated with recent alluvial channels.
- Low Permeability Member: clay-rich gravels 3-9 m thick, deposited as over-bank flow deposits when sea levels began to stabilise about 6.5 ka.
- Lower Member: high permeability alluvial gravels 9.5 ± 5m thick deposited 9.5 to 7 ka during a period of warming global temperatures.

Based on the identified stratigraphy, a new conceptual model for the internal structure of the part of the Rapaura Formation that underlies the study area was developed (Figure 2). The soils of the Wairau Plains are typically shallow, stony and well-draining. They can be classified in nine groups as depicted in Figure 1.

Hydrological data The flow record of the Wairau River close to the SH1 bridge dates back several decades. In June 2014, MDC staff installed three additional temporary recorder sites upstream of SH1, where the river flows in a single braid. These sites are subsequently referred to as Rock Ferry, SH6, and Wratts Rd (Figure 1). River stage is measured at these sites and then converted into discharge using elaborately established rating curves. The flow ratings at these sites have been renewed after each larger flow event because of changes in the braided river bed geometry. Spot gaugings of Wairau River flow were conducted at these and other sites as far back as in the 1970s. The gaugings were conducted usually under low-flow conditions and at the same day. The differential gauging allows to examine river losses and gains. Both historic and recent gaugings show consistently the river losing water between Rock Ferry and Wratts Rd and then gaining between Wratts Rd and SH1 (Figure 3).

At the intersection of the Rapaura and Dillons Point Formations, groundwater is forced to the surface and emerges as springs across the Wairau Plain. The major spring on the

Plain is Spring Creek which has a mean flow of about 4.0 m^3/s at the Motorcamp recorder site (Figure 3). The record of manual gaugings dates back to 1990 and was complemented in 2013 by an automatic recorder. However, only manual gaugings are used in this study since the automatic recorder had frequent malfunctions and was at least in some cases influenced by channel blocking and weeds.

Groundwater levels are observed at four permanent (3009, 3821, 3954, 4577) and six temporary MDC wells (903, 7007, 1685, 1690, 1696, 10426) distributed over the length of the Wairau Plains (Figure 3). The wells are screened at different depths and across all three of the main facies of the Rapaura Formation. The temporary wells were equipped in January 2016 specifically for this project. Data gaps occurred in wells 7007, 1685, and 10426 in December 2016 and January 2017. Some additional spot measurements (manual dipping) were taken in wells 7007 and 10426 prior to January 2016. All permanent wells are part of the MDC core monitoring program and have continuous long-term records.

Meteorological data The Wairau Plain receives on average 650 mm of annual precipitation. The mean annual temperature is 12.8 °C and the sun shines on average 6.7 hours per day. This unique climate makes the area so attractive for winegrowers. The meteorological data required for our calculations was sourced from the Blenheim Research Station.

All the continuous data was aggregated / averaged to daily values for use in our model simulations. To integrate the new information from the temporary logger sites, we have chosen the time period between 1/7/2013 and 20/02/2017 in our investigation.

2.2 Wairau Aquifer model

A transient surface water - groundwater model for the study area was set up in MODFLOW-NWT which is designed to solve problems involving drying and re-wetting non-linearities of unconfined groundwater-flow (Niswonger et al., 2011). The graphical user interface ModelMuse (Winston, 2009) was used to set up the model domain and boundary conditions. A plan view of the model boundaries is shown in Figure 1. The total model area is 84.8 km^2. To the North, the domain is bounded by the northern bank of the Wairau River. In the West, the domain starts at Rock Ferry, a natural rock constriction of the Wairau valley. The southern boundary is normal to the regional groundwater level contours shown in Davidson and Wilson (2011). The eastern boundary is drawn at the SH1 bridge, approximately 5 km off the coast, because groundwater in the Rapaura Formation is forced up through the confining Dillons Point Formation which forms a natural boundary. Deep groundwater flow over the eastern face of the Rapaura Formation is considered constant throughout time at a rate of - 0.7 m^{3}s^{-1} as estimated from spring flows in Grovetown Lagoon to the East. The FHB package (Leake and Lilly, 1997) is used to implement the corresponding boundary condition in the model.

The top elevation of the model domain is derived from a high resolution LIDAR image which was interpolated at the grid nodes of the MODFLOW computational grid. The bottom of the model domain is defined by the elevation of the Speargrass Formation (Figure 2) and is considered an impermeable zero-flux boundary. The northern, western, and southern boundaries of the model domain are considered no-flux boundaries too.

Figure 4 depicts the computational MODFLOW grid. As a result of a preliminary sensitivity analysis with different grid sizes, we selected a regular cell size of 200 $\times$ 200 m in our model. The geology was implemented by three computational layers matching the formation boundaries. Thus the model domain consists of 3$\times$ 2120 = 6360 active cells. The first layer of the grid is considered the Upper member of the Rapaura Fm in the West and the Dillons Point Fm in the East. On the surface, the intercept of the two formations is marked as aquitard boundary in Figures 1 and 4. A minimum grid-cell thickness of 1.0 m was assumed for the cells of the intercept and also for all other grid cells for numerical efficiency. According to the geological record, the Lower and Low permeability members of the Rapaura Formation outcrop underneath the Wairau River for a relatively short section in the West.

Wairau River The Wairau River is implemented by the streamflow routing package SFR (Niswonger and Prudic, 2005) which can be used to simulate connected and disconnected streams. Because of the highly permeable sediments of the gravel-bed river, we consider head-dependent stream leakage when the river is connected with groundwater and unit-gradient flow when the river is disconnected. Further, we defined 12 different sections along the river, with locations corresponding to locations of a detailed survey of the river geometry conducted at 25 cross-sections between Rock Ferry and SH1. The SFR package requires as input the time series of river discharge at Rock Ferry which was calculated from the discharge record at SH1 plus a a constant 7.64 m^3s^{-1} that was determined from a correlation analysis using the concurrent record of stream flow ($R^2 = 0.98$). MODFLOW calculates the actual river length for each river cell. The SFR package further requires the parametrization of the stream-bed hydraulic conductivity, thickness of the river bed, and two geometric functions describing the functional relationship between river stage $h\,[\mathrm{m}]$, wetted perimeter of the stream channel $L_{wp}\,[\mathrm{m}^2]$, and discharge in the river $Q_{riv}\,[\mathrm{m}^3/\mathrm{s}]$:

$$h(Q_{riv}) = aQ_{riv}^b\,; \quad L_{wp}(Q_{riv}) = cQ_{riv}^d\ , \tag{1}$$

where, a, b, c, d are empirical constants. These constants were initially derived for each of the 25 river cross-sections, simultaneously put into the model, and simulation results compared to a single set of parameters derived for an average cross-section and applied to all cross-sections in the model. The sensitivity to the more detailed representation of the river geometry was low, because of the parameter interactions to stream-bed conductivity and the underlying hydraulic conductivity field. This allowed us to reduce the parametrization effort and we applied a single set of river geometry parameters at 12 river cross-sections in all our simulations: $a = 0.192$, $b = 0.16$, $c = 4.83$, $d = 0.239$. The thickness of the river-bed was assumed to be 1 m in all river sections. A detailed (and transient) parametrization of the geometry functions in the model is possible but in fact not desirable for practical reasons, because acquiring the channel-bed geometry information involves a significant experimental effort at regular intervals in gravel-bed rivers (after each major flood). Modelling approaches to simulate the transient evolution of river-bed morphology (Pirot et al., 2014; Williams et al., 2014, 2016) are still far from being routinely used.

Springs and Streams The stream network and the springs emerging at the eastern Plain are depicted in Figure 4. They are simulated using the DRN package (Harbaugh, 2005) which describes head-dependent flux boundaries. If the head in a drain cell falls below a certain threshold, the flux from the drain to the model cell drops to zero. The DRN package requires the specification of drain bed conductivity K_D and drain elevation. The latter is derived from the LIDAR image and was offset by -1 m for the channel depth. Five different sections of the springs and streams are distinguished (Figure 4): ND describes the northern drain, a spring that discharges into the Wairau River just East of the aquitard boundary. Spring Creek is divided at the flow gauging station in a western and an eastern part (SC1 and SC2, respectively). Further, the Omaka River and the eastern Opawa River at the southern model boundary (OR1) are distinguished from the western Opawa River (OR2). Each drain section is parametrized separately.

Recharge and Irrigation Groundwater recharge from the land surface is considered as a specified flux boundary at the top of the model domain using the RCH package (Harbaugh, 2005). The land use of the Wairau Plain is predominantly vineyards which are irrigated using groundwater that is pumped locally from the aquifer. Irrigation abstraction is simulated using the WEL package which applies a specific flux boundary to internal cells, here specifically cells in layer three. Groundwater recharge and irrigation demand are computed for each of the nine soil types of the Wairau Plain using a soil water balance model which is described in detail in Section 2.3.

2.3 Land surface recharge model

The landuse at the Wairau Plain is almost exclusively vinyards. Land surface recharge and irrigation demand are simulated using a daily soil moisture balance model, which has been modified from the Rushton model (Rushton et al., 2006). The Rushton model is a simple two-layer soil model which uses a near-surface soil store to enable evapotranspiration to occur during soil moisture deficit conditions on days following rainfall events. Without this near surface soil storage, evapotranspiration values following rainfall events would be underestimated (de Silva and Rushton, 2007). The proportion of rainfall infiltration that is partitioned to the near surface soil store is determined by an empirical coefficient, f_s. Values of f_s are related to soil texture and drainage, and are zero for coarse sandy soils, 0.4 for sandy loams, and 0.75 for clay loams (Rushton et al., 2006). For the soils of the Wairau Plain we estimated values of f_s, ranging from 0.1 (gravelly sand) to 0.75 (deep clay loam). Soil texture and Total Available Water (TAW) values were sourced from the New Zealand Fundamental Soil Layer database (Landcare Research, 2000). Readily Available Water (RAW) for vineyard grapes was assumed a value of 45% of TAW, following Allen et al. (1998). Soil moisture in the deeper soil layer is calculated after near-surface evapotranspiration has been accounted for in the near surface soil store, S. The soil moisture deficit, SMD for each day with index i is calculated for each soil type as follows:

$$\mathrm{SMD}_i = \mathrm{SMD}_{i-1} - \Delta_i + S_i + AE_i \tag{2}$$

where, Δ_i is the balance of daily inputs to the soil:

$$\Delta_i = P_i - R_i + S_{i-1}\,, \tag{3}$$

and P_i and R_i are precipitation and surface runoff, respectively. Daily rainfall and potential evapotranspiration (PET) values were taken from the record for Blenheim Research Station. PET is derived by the Penman-Monteith equation (Allen et al., 1998) as grass reference evapotranspiration ET_0. A seasonally-varying crop factor, K_c, was applied for vineyard grapes based on sap flow measurements in a Marlborough vineyard (Green et al., 2014). Actual evapotranspiration (AE) is assumed to equal PET when soil water is readily available. For RAW $<$ SMD $<$ TAW, the vineyard becomes water-stressed and transpires at a reduced rate unless inputs to the soil exceed PET. This situation is represented in the model by applying a water stress factor:

$$AE_i = K_{S,i} K_{c,i} ET_{0,i}\,, \tag{4}$$

with

$$K_{s,i} = \frac{\mathrm{TAW} - \mathrm{SMD}_{i-1}}{\mathrm{TAW} - \mathrm{RAW}}. \tag{5}$$

If the soil moisture content reaches the value of TAW, the roots are unable to extract water, and AET $= \Delta$. Drainage to groundwater occurs only in the model when SMD is negative, i.e. when there is surplus water in the soil moisture reservoir. Soil moisture calculations are started during winter conditions so that an initial soil moisture deficit of zero can be assumed. This enables a lead-in time for the model to establish a suitable initial condition for the beginning of the first calendar year. Surface runoff R is calculated by the SCS method (Soil Conservation Service, 1972). It is assumed that 2.2 mm vineyard irrigation occurs on days when soil moisture is less than 70% of RAW during the irrigation season

(October to April). This irrigation threshold was determined by comparing modelled irrigation demand with water meter data from vineyards on the Wairau Plain.

2.4 Parametrization

The model domain and the boundary conditions described in Section 2.2 require various parameters to be specified. The corresponding parametrization scheme is described in this section.

The three lithological members of the Rapaura Formation have different hydraulic properties which are considered in the model by an independent parametrization of the hydraulic characteristics for the three layers. There is also considerable horizontal heterogeneity of aquifer properties (Davidson and Wilson, 2011; Wilson and Wöhling, 2015) which is implemented in each layer using a pilot point parametrization technique (e.g., Doherty 2003; Doherty et al. 2010) for the hydraulic conductivity and specific yield fields. Pilot points are discrete, user-defined locations throughout the model domain that are used here for cell-by-cell parametrization of the saturated hydraulic conductivity K_H, and of the specific yield S_y through interpolation from the pilot points to the model grid. Corresponding to only regional changes in horizontal heterogeneity we used a exponential variogram with range 5 km and 26, 31, and 33 pilot points at a regular spacing for the Upper, Middle, and Lower member of the Rapaura Formation, respectively. Given the much lower hydraulic conductivity of the confining Dillons Point Formation compared to the Rapaura fm, uniform properties are assumed for the confining layer. Further, a uniform anisotropy factor for the hydraulic conductivity, f_a, and uniform specific storage, S_s, was assumed for each of the four geological units.

Other parameters to be considered in the model are the vertical hydraulic conductivity for each of the 12 defined river sections, K_R, and the drain bed conductivity, K_D, of each of the five drain sections as defined in Section 2.2.

2.5 Model calibration and uncertainty analysis

In total, there are 207 parameters for the Wairau Plain model (Table 1). These parameters can't be measured directly at the required spatial and temporal scales and thus effective parameter values need to be estimated trough model calibration. For highly parametrized models like the one presented, automatic model calibration is the only feasible option. In this study we used the model independent parameter estimation software PEST (Doherty, 2016b,c) which is ideally suited for highly parametrized inversion problems (Doherty et al., 2010).

Calibration data The objective of the model calibration in general is to minimize the discrepancy between model simulations and measured data. In our study we used observations of groundwater head, Spring Creek flows, a spot measurement of differential river flow gauging, and three "soft targets" which contain expert knowledge from MDC groundwater scientists. The data set is separated in 123-day lead-in period, a 925-day calibration period and a 284-day evaluation period (Table 2). Approximately 70% of the head observations from the four permanent observation bores are used for model calibration and the reminder (30%) for model evaluation. In contrast, the majority of the head observations from the six temporary bores (between 59 and 72%) are used for model evaluation. Another calibration target was formed on the basis of the historic differential river gaugings (Figure 3). It follows the rationale that the average river losses and gains between Rock Ferry and SH1 have been observed to be almost constant during low flow periods and consecutive dates. A low flow period is present in the calibration data set between 31/01/2014 and 15/03/2014 ($\bar{Q}_{riv} = 14.4\,\mathrm{m^3 s^{-1}}$ at SH1). The mean river exchange flux for that period and the river section between Rock Ferry and Wratts

Rd (losing section), $Q_{ex,13}$, is assumed to correspond to the mean loss from the historic measurements which is 5.73 m^3/s. This constitutes the first soft target in our model calibration. Secondly, the mean flows in the river reach between Wratts Rd and SH1, $Q_{ex,3}$, is targeted at a net gain of -0.5 m^3/s. Please note that flows out of the model domain are negative numbers and fluxes into the model domain are positive numbers. Other calibration targets specified from expert knowledge are a mean gain of Spring Creek flows downstream of the gauging station at Motorcamp of $Q_{SC2} = -0.5\,\mathrm{m^3s^{-1}}$ and a mean gain of all the southern streams of $Q_{SS} = -1.5\,\mathrm{m^3s^{-1}}$.

Objective function PEST uses a sum-squared error (SSE) objective function, that can be weighted by the measurement error (Doherty, 2016b). Using different physical quantities (data types) with different numerical ranges and different observation numbers in a SSE objective function leads typically to unequal weighting of the different data types. The weighting of the individual observations is therefore of great importance for the outcome of the calibration. Weighting of data expresses the degree of belief the modeller has in the individual pieces of information and is therefore to some extent subjective. The weighting of the different data types was determined by trial-and-error to obtain a balance between "hard" and "soft" calibration targets. The weights of individual data points are reported in Table 2 and used in all our model calibration runs.

Parameter regularization Regularization techniques are used in order to constrain potential solutions of the model calibration and to avoid unrealistic artefacts in spatially correlated data (e.g., Doherty 2003; Moore 2005). By regularization, parameter fields are penalized when deviating from the spatial correlation defined, for example, by a variogram. We applied Tikhonov regularization to the K_H and S_Y fields as well as to K_R using the variogram described in Section 2.4. The smoothness of a parameter field can be expressed by a weighted sum of parameter differences at neighbouring pilot points with weighting factors determined by the variogram. Deviations from "smoothness" is measured by a penalty objective function that has an optimum value of zero for a homogeneous field. The PEST groundwater utilities PPK2FAC, FAC2REAL (Doherty, 2016a), and the PEST utility ADDREG1 (Doherty, 2016c) are used to calculate the weighting factors for the pilot point locations and to implement a corresponding regularization objective function into the parameter estimation process with PEST.

Pareto optimization For reasons not further discussed here, there is typically a trade-off between the model's ability to correctly reproduce the data of local measurements and the smoothness of parameter fields. The aim of the regularized parameter inversion (i.e. the model calibration technique used here) is to find a compromise between data and regularization objective functions and thus avoid overfitting. For example, it is not desirable to place too much confidence in data that might not be represented in the model (e.g. by subscale effects). On the other hand, we want to include as much spatial heterogeneity in the calibrated model, as is legitimately supported by the data. The ideal weighting between data and regularization objective functions is difficult to define. It could well be argued that ideal weighting does not exist because the choice always involves some degree of subjectivity by the modeller. To guide the choice, the trade-off between the objective functions can be determined using multiobjective calibration methods (e.g., Wöhling et al., 2008; Reed et al., 2013) which result in a set of Pareto efficient solutions. These solutions have the property that moving from one to another along the tradeoff surface results in the improvement of one objective while causing deterioration in at least one other objective (see Gupta et al., 1998 and others for further information on the Pareto optimality). The Pareto optimization concept was adapted for highly parametrized inversion (Moore et al., 2010) and implemented in PEST. The method is used in this study to simultaneously calibrate the model parameters and calculate the trade-off between data and regularization objective functions described above. Subspace projection techniques to increase the computational efficiency of the highly-parametrized model inversion were also trialled here (SVD-Assist, Doherty et al., 2010; Doherty, 2016c), but the combination of techniques lead to parameters being frozen at their boundaries, which was an undesirable effect. All our calibration runs were conducted using the parallel computing tool BEOPEST. The parameter ranges were derived from geological information and expert knowledge (Table 1).

The result of the calibration is a Pareto efficient set of solutions which was filtered to meaningful trade-offs by the concept of ϵ-dominance (Kollat et al., 2012). Finally, a compromise solution that exhibits both a good data fit and realistic parameter fields was subjectively selected from the Pareto set which then constitutes the calibrated model.

Uncertainty analysis After model calibration, an uncertainty analysis was performed to assess the robustness of the model calibration and the reliability of model simulations and predictions. Highly parametrized model calibration rarely leads to unique parameter estimates, because of the insensitivity of model outputs corresponding to historical observations of system state to some parameters, excessive correlation with other parameters, or both (Doherty and Hunt, 2009). Conceptually, the parameter space can be divided into two subspaces, the solution space and the null space. The solution space comprises parameter combinations that are informed by the available data set. The null space comprises parameter combinations that have little effect on model outputs when superimposed on the calibration parameter set (Moore and Doherty, 2005; Doherty and Hunt, 2009). Note, however, that these parameter combinations may have an effect on model outputs that are not contained in the calibration data set. Sampling the parameter null space and analysing the resulting model simulations is therefore an effective means to determine non-linear predictive uncertainty and is superior to linear first-order second moment (FOSM) methods. Null space Monte-Carlo (NSMC) sampling utilities and FOSM predictive uncertainty estimation utilities are readily implemented in PEST (Tonkin and Doherty, 2009; Doherty et al., 2010; Doherty, 2016c). NSMC sampling is applied in this study to estimate post-calibration predictive uncertainty.

Model predictions Two types of model predictions are distinguished in this study. The first type consists of data types that are already contained in the data set (here: groundwater heads and Spring Creek flows) where predictions are made for different times and different model forcings. These predictions are subsequently referred to as type I predictions. The root mean squared error (RMSE) and the coefficient of determination ($\mathrm{R}^{\mathbf{2}}$) were used as metrics to summarize the model performance for these prediction types.

The second prediction type comprises model predictions / data types that are not fully contained in the calibration data set. These predictions are subsequently referred to as type II predictions. In this study, we predict the transient net-exchange flows between Rock Ferry and SH1 (only low-flow means were used as soft-target in the calibration) as well as the transit time distribution and mean transit time for Spring Creek water upstream of the flow gauge.

Transit time distributions were calculated using reverse particle tracking methods with MODPATH (Pollock, 2012). The resulting particle tracks and residence times were post-processed to calculate cumulative flux-weighted transit time distributions:

$$cdf_{TT} = \frac{1}{Q_T} \int_{i=1}^{N_p} \tau_i \cdot q_i \ , \tag{6}$$

where, $i = 1 \ldots N_p$ denotes the particle index, N_p is the total number of particles, τ_i is the particle travel time, q_i is the flux in the cell where the particle originates, and Q_T is the total flux, i.e. the sum of all q_i. The flux-weighted mean transit time (MTT) is then calculated as the 50% quantile of the cdf_{TT}.

3 Results and Discussions

First up in this section, the performance of the calibrated model and the uncertainty of type I predictions is analysed. In the second sub-section, we discuss parameter uncertainty and the plausibility of calibrated parameter values. Then, the river-groundwater exchange mechanisms for the considered section of the Wairau River (type II prediction) are analysed. Finally, we present the results of the other type II prediction, namely the transit time distribution and mean travel time of Spring Creek flows.

3.1 Model calibration and evaluation

3.1.1 Trade-off between data and regularization objective functions

The model was calibrated using a data objective function (OF_{dat}) and a regularization objective function (OF_{reg}). Figure 5 shows the trade-off between the two objective functions. Open circles depict all Pareto solutions obtained by the model calibration, while the orange solutions depict the ϵ-dominant solutions which were used for the analysis. One purpose of ϵ-dominance is to truncate meaningless solutions at the ends of a Pareto front, where a small change in one objective function leads to a large change in at least one other objective function. This is the case in here along the x-axis in Figure 5, where a small improvement in the data fit leads to a rather large distortion of the parameter fields as penalized by the regularization objective function.

Overall, we observed a large trade-off between data and regularization objective functions which is demonstrated by the rather curved shape of the Pareto front. A more angular shape of the Pareto front would indicate less trade-off between the two. The visual inspection of the parameter fields of the ϵ-dominant solutions revealed strongly distorted fields for $OF_{reg} > 600$ that are a strong indication of over-fitting (results not shown). On the other hand, parameter fields became unrealistically smooth for $OF_{reg} < 200$ while the data fit deteriorated quickly. Therefore, we subjectively selected a compromise between the two objective functions ($OF_{dat} < 60.6$, $OF_{reg} = 378.3$, indicated in blue in Figure 5). Note, that the compromise solution can be selected in an objective manner by determining the Pareto solution with the least Euclidean distance to the origin (for more details see e.g., Wöhling et al., 2013). The parameter set of the compromise solution is subsequently referred to as the calibrated model and used subsequently for analysing model performance. NSMC simulations were conducted with that solution as described in Section 2.5 to access predictive uncertainty. The performance of the calibrated model is reported in the next section while the parameter set and the corresponding uncertainty is discussed in Section 3.1.3.

3.1.2 Model performance

Groundwater heads Simulated groundwater heads obtained with the calibrated model are compared to observations at the permanent and temporal MDC wells. Results are summarized in Figures 6 and 7 and in Table 3 and are subsequently described. The groundwater wells are presented by location from West to East in the figures, following the gradients of the land surface and the groundwater table. Model simulations and observed groundwater heads are indicated by the blue and orange lines (dots), respectively. The 95% uncertainty bounds determined from the NSMC simulations are shaded grey. A vertical dashed line indicates the divider between the calibration period (left) and the evaluation period (right). To facilitate an better comparison of the temporal dynamics between wells, a constant y-axis spacing of 6 m was used in all figure panels.

Overall, the calibrated model represents the regional groundwater surface well. There is a gradient between approximately 57 m.a.s.l. in the West at well 903 (Figure 6) and 7 m.a.s.l. in well 3954 in the East (Figure 7) which is well reproduced by the model simulations. The temporal variability of the groundwater heads as well as the depth to the water table generally decreases from West to East. The variability is largest in wells located close to the river (wells 903, 1690, 1696, 7007) and lowest in wells that are located underneath the confining layer (wells 4577 & 3954). The detail of the observed groundwater head variability is reproduced satisfactorily for most wells. However, some discrepancies remain in wells 903 and 1690 (Figure 6) which can be explained by a relatively short calibration data record for these wells and by model structural uncertainty at the western boundary. The structural uncertainty includes a (too) narrow model domain with surrounding no-flow boundaries in the East, potential groundwater inflow from the Waihopai River, and/or an influence from Gibsons Creek (Figure 1). Correspondingly, the model-to-measurement misfit is larger for these wells compared to the other wells (Table 3).

Taking the perspective of a regional analysis, the performance of the model is considered satisfactory for these other wells, which is confirmed by low RMSE values (ranging between 0.05 and 0.31 m) and large R^2 values (ranging between 0.68 and 0.91) for the calibration period (Table 3). The model performance during the evaluation period is similar, but shows a slightly larger variability with RMSE values ranging between 0.05 and 0.49 m and R^2 values ranging between 0.66 and 0.91. The 95% uncertainty bounds generally cover the observations except for the wells at the western boundary where simulated heads are generally overestimated (biased) and exhibit the largest model-to-measurement misfit. The uncertainty tends to increase with the temporal variability of the groundwater heads and is lower for the wells under the confining Dillons Point Formation.

Spring Creek The largest spring on the Wairau Plain is Spring Creek with a mean flow of about 4 m^3s^{-1} at the Motorcamp gauging station. Spring Creek is fed by upwelling groundwater and originates at the interface between the highly conductive Upper member of the Rapaura Formation and the confining Dillons Point Formation (Figure 4). The relatively large variability of the flow record and the correspondence to the Wairau River flows shown in Figure 8 suggest the existence of rapid subsurface flow paths. These are not uncommon for New Zealand's gravel-bed rivers which form highly transmissive networks called open-framework gravels (Dann et al., 2009). Recent field work in the Wairau floodway supports the existence of open-framework gravels in the Upper member of the Rapaura Formation.

The model simulations match the observed Spring Creek flows well, although the variability of the flows seems to be overestimated when evaluated by the manual spot gaugings (Figure 8). Data taken by a continuous stage recorder installed in January 2013, however, showed very fast responses of Spring Creek flows to Wairau River floods and that the variability of simulated spring flows could be realistic. The recorder data were not used in the model calibration, though, because of continuing experimental challenges (e.g., weeds) that lead to drift and bias in the flow record.

The RMSE values of simulated Spring Creek flows are 0.22 and 0.32 m^3s^{-1} for the calibration and evaluation period, respectively. It should be noted that only seven data points were available in the evaluation period (Table 3). The 95% uncertainty bounds are too narrow to capture all the observations which is potentially a result of the chosen uncertainty quantification method. Following the NSMC procedure presented by Tonkin and Doherty (2009), we applied a re-calibration step for the underlying parameters which might in this case lead to an overly optimistic contraction towards the calibrated model parameters. In addition, the sample of 100 NSMC simulations might be simply too small. On the other hand, alternative uncertainty quantification methods based on stochastic parameter sampling techniques are too time-consuming for application to highly-parametrized models and therefore not further investigated here.

Soft Targets The fitness of the calibrated model to the soft targets is summarized in Figure 9. The box plot shows the 50/95% uncertainty bounds by the boxes and whiskers, respectively. Also shown are the target values in blue and the simulation of the calibrated values in orange. A very good agreement between targeted expert knowledge and the model and narrow uncertainty ranges are obtained for the flow in the downstream branch of Spring Creek, Q_{SC2}, as well as for the average river-groundwater exchange flows under low-flow conditions, $Q_{ex,1}$ and $Q_{ex,13}$. The results suggests that the model reproduces both the upwelling of groundwater through the confining Dillons Point Formation and the behaviour observed in the historic differential flow gaugings (Figure 3).

The flow target for the southern streams, Q_{SS}, is overestimated by the calibrated model and has larger uncertainty bounds. The target was based on historic stream gaugings in the ephemeral Opawa River prior to a diversion scheme into Gibsons Creek became operational. Smaller springs and drains that exist South of Spring Creek are not considered in the model, which would in part explain the discrepancies together with structural uncertainties of the southern no-flow boundary. However, the focus of the study is on the river-groundwater exchange fluxes and the soft targets are weighted less compared to other types of data in accordance to the subjective belief (or its counterpart uncertainty) of the information (Table 2). We have found that the inclusion of expert knowledge in our model calibration is highly valuable for both constraining the parameter space, and for establishing a degree of trust in the calibrated model.

3.1.3 Parameter uncertainty

The uncertainty of type I predictions was presented in the previous sections along with the performance of the model for the calibration and evaluation data set. The underlying parameter uncertainty of the calibrated model is presented in this section and is summarized in Figures 10, 11 and 12.

Hydraulic conductivity fields The left column of Figure 10 depicts the hydraulic conductivity fields of the three members of the Rapaura Formation. It is reiterated here that hydraulic conductivity (and specific yield) is only estimated at pilot point locations which are then used for interpolation onto the MODFLOW grid which is presented in the corresponding figures. Also shown are the Wairau River and the considered stream network for orientation and the groundwater observation wells in the respective facies. Consistent with the geological expertise, the Upper member of the Rapaura Formation exhibits the largest K_H values while the low permeability member in the middle has a somewhat lower permeability. Hydraulic conductivity seems to increase underneath the Wairau River from West to East with a high-conductive zone downstream of Giffords Rd connecting the river, Wratts Rd well and the Spring Creek area. The prediction of high-conductive zones in the Lower member of the Rapaura Formation is not easily understood, but the overall pattern is consistent with earlier investigations of transmissivities derived from well specific capacity by Davidson and Wilson (2011). The hydraulic conductivity of the confining Dillons Point Formation is about three orders of magnitude smaller than the maximum values in the Rapaura Formation, which is consistent with geological knowledge and exploration results.

The uncertainty of the hydraulic conductivity fields is presented as one standard deviation of K_H of the NSMC runs in the right panels of Figure 10. The uncertainty is relatively large for some areas of the Upper and Lower members of the Rapaura Formation. It is interesting to note that this does not result in an equally large uncertainty for type I model predictions as was shown in the previous subsections. In some areas, the large uncertainty is likely to be caused by insensitivity to model outputs (e.g., in the eastern part of the Lower member). In other areas it may be caused by trade-offs in the fit to different pieces of information in the calibration data set (in the Upper member). Since the absolute value of K_H for the Dillons Point Formation is orders of magnitude smaller, the uncertainty appears to be zero in Figure 10. This is not the case as shown below by normalized parameter ranges.

Specific yield fields Figure 11 shows the specific yield fields of the calibrated model (left panels) and their respective uncertainty (right panels). The S_y values are within expected ranges for the coarse gravel materials of the Rapaura aquifer. Only little variability can be seen in the parameter fields with two distinctive exceptions in the Northeast of the Upper member and the West of the low permeability member. However, the uncertainty of the S_y-fields is relatively large and uniform in all three members of the Rapaura Formation with one standard deviation exceeding $^1/_3$ of the entire parameter range. This is also the case for the Lower member, where the parameter values remained close to their starting values. This suggests that the sensitivity of specific yield to the model outputs set is relatively low in light of the calibration data set and that there is potential for parameter simplification.

Other parameters The uncertainty of parameters that are not spatially correlated over the entire model domain are depicted by box plots in Figure 12. Note that the parameters are normalized by their respective ranges which are listed for convenience at the top of the graph. The boxes and whiskers show again the 50% and 95% uncertainty bounds, respectively. Also shown are the median (red lines) and the parameter values of the calibrated model (blue dots). In some cases, the parameter values of the calibrated model fall on their upper or lower boundary (Figure 12). These bounds represent meaningful physical limits even though parameters are effective parameter values for the grid-cell scale of 200×200 m. Although a better data fit would be possible, we didn't want to increase the parameter ranges or introduce more fine-scale detail to the model, mainly because we wanted to avoid overfitting. Some peculiarities of the calibration parameter set are subsequently discussed.

The effective (uniform but unisotropic) hydraulic conductivity of the Dillons Point Formation is with $K_{aq} = 11.7\,\mathrm{m^3s^{-1}}$ about two orders of magnitude smaller than the average in the Rapaura Formation. The value seems to be relatively high for the fine-textured marine sediments. However, it should be noted that K_{aq} is an effective value that accounts for both flow through the pore matrix and flow along faster vertical passageways for upwelling groundwater through the sediments. The existence of these pathways causes the springs on the Wairau Plain to still gain water along their course to the East. Correspondingly, the effective value for the specific yield of the marine sediments is relatively large ($S_{y,aq} = 1E^{-3}$) but the 95% uncertainty bounds for both K_{aq} and $S_{y,aq}$ cover almost the entire range of expected values.

The specific storage for the three members of the Rapaura Formation S_{S1-3} is small and insensitive because the unit hosts unconfined groundwater. These parameters can be omitted from the model calibration - unlike the corresponding parameter for the confining Dillons Point Formation (S_{S4}).

The Upper and Lower members of the Rapaura Formation exhibit no significant difference in vertical vs. horizontal hydraulic conductivity ($F_{xz1} = 1.1$, $F_{xz3} = 1.4$). This is somewhat contradictory to data from bore logs and aquifer tests and suggests that groundwater head data perhaps isn't well suited to constrain anisotropy in unconfined sediments. In contrast, the Low Permeability member has a factor $F_{xz2} = 4.4$ lower vertical hydraulic conductivity (Figure 12) which corresponds well to layers of finer material interbedded in this unit as described in Section 2.1.

A regularization constraint was placed onto the spatial variability of river-bed hydraulic conductivity of the 12 river sections ($K_{b,R1} \ldots K_{b,R12}$) to avoid overfitting and to make the model more robust to predictive bias. The calibration resulted in a deviation from the optimal regularization constraint, i.e. from all river-bed conductivities having the same value. In other words, the data has forced the pattern of the river bed conductivities which has a direct impact on river-groundwater exchange rates in the different river sections. In general, the river-bed conductivities increase from West to East (Figure 12) and are largest in the Wratts Rd area that coincides with the high-conductive zone in the Upper member of the Rapaura Formation described above. Together, these features form a highly transmissive passage of Wairau River water to Spring Creek.

3.2 River-groundwater exchange mechanisms

The results of the previous section demonstrated that the calibrated model performs well to historic data (both for calibration and independent data sets) and that the obtained parameter set is in agreement with expectations, previous data and expert knowledge. This is a prerequisite for a trustworthy model in general and specifically if type II predictions are to be made by the model. The river-groundwater exchange flows are such a prediction. Results are summarized in Figures 13, 14, and 15 and are subsequently discussed.

3.2.1 Net exchange flows

Daily values of net river-groundwater exchange flows for the Wairau River section between Rock Ferry and SH1, Q_{ex}, are presented in Figure 13b). The top panel depicts the corresponding Wairau River flows during the considered simulation period. The net exchange flow is always positive and most of the time $Q_{ex} > 5\,\mathrm{m}^3\mathrm{s}^{-1}$ which means that overall, the Wairau River is always losing water to the aquifer. Figure 13b) also shows that the exchange flow is highly dynamic and correlated with the river flow. Large flood events in the Wairau River typically also result in peaks for the exchange flow. However, smaller flood events of less than $250\,\mathrm{m}^3\mathrm{s}^{-1}$ at the end of prolonged low-flow periods in summer also result in strong recharge peaks. One example is the relatively small flood event ($Q_{riv} = 261\,\mathrm{m}^3\mathrm{s}^{-1}$) on 17/03/2014 which occurred after a 7-week recession period without any floods and caused a relatively large recharge peak of $16.2\,\mathrm{m^3s^{-1}}$. The much larger river flood peak one month later (18/04/2014, $Q_{riv} = 967\,\mathrm{m^3s^{-1}}$) resulted in a recharge peak that was similar in size ($Q_{ex} = 18.0\,\mathrm{m^3s^{-1}}$) compared to the previous event. Similar examples can also be found in summer 2015 (08/03/2015, $Q_{riv} = 274\,\mathrm{m^3s^{-1}}$, $Q_{ex} = 16.2\,\mathrm{m^3s^{-1}}$) and autumn 2016 (13/05/2016, $Q_{riv} = 330\,\mathrm{m^3s^{-1}}$, $Q_{ex} = 14.5\,\mathrm{m^3s^{-1}}$). It is interesting to note that the relatively large parametric uncertainty (see previous Section) has only little effect on the predictive uncertainty of the net exchange flow. The 95% uncertainty bounds are very narrow and hardly discernible in Figure 13b).

Recharge flows greater than the $5\,\mathrm{m^3s^{-1}}$ base line seem to be triggered already by even smaller flood events and since they occur more frequent in winter and less frequent in summer, the aquifer is mainly recharged in the winter months. To analyse this further, we have depicted the exchange flow anomaly in Figure 13c). The anomaly is calculated as the deviation of the cumulative net exchange flow from its mean during the simulation period. Negative/positive gradients in the anomaly curve indicate exchange fluxes below/above the mean, respectively. The seasonality is clearly visible in this representation of model results. During summer, the gradient is negative indicating lower than average recharge. During April - September (autumn/winter in the southern hemisphere) the gradient is reversed indicating higher recharge and that the aquifer storage is re-filled during that time. Groundwater heads are responding accordingly and show the same seasonality (Figures 6 and 7).

If the seasonal pattern of groundwater recharge from the river would be equal for each consecutive year, the anomaly curve would exhibit the same maximum and minimum value in each year. This is apparently not the case as seen in Figure 13c). There is inter-annual variability of rainfall in the Wairau catchment and thus also of aquifer recharge. 2014 and 2015 were particularly dry years on record which causes the seasonal maximum of the recharge flow anomaly to decrease for these years. However, the summer 2016 brought two major flood events in an usually dry period and was followed by a particularly wet winter and spring, which caused above-average aquifer recharge. It can be concluded from the analysis that time periods with frequent, consecutive river floods with return periods in the order of only weeks lead to enhanced aquifer recharge while prolonged dry periods cause lower aquifer recharge.

3.2.2 Spatial variability of river-groundwater exchange flows

To study the spatial variability of river-groundwater exchange flows along the Wairau River, a snapshot of the model simulations was taken on 17/02/2014 which relates to a low-flow period and a date where differential gaugings were conducted in the river. Figure 14a) shows the groundwater head contours for that particular day. They are mainly oriented from West to East following the gradient of the land surface. Some groundwater mounding can be seen under the river between Rock Ferry, SH6, and half-way through to Wratts Rd, which corresponds to a less transmissive area in the Upper Member of the Rapaura Formation (Figure 10) and lower river bed conductivities in the upstream region (Figure 12).

The simulated exchange flows for all river and drain cells in the model domain are depicted in Figure 14b). Yellow and green colours indicate losses while blue colours indicate gains. The analysis revealed that the largest river losses are to be found in an area half-way between SH6 and Wratts Rd where the Upper member of the Rapaura Formation is relatively thick. East of the line of confinement formed by the Dillons Point Formation, all rivers and streams are gaining. Particularly high fluxes are visible at the origin of Spring Creek and the lower reach of the Opawa River. To analyse the spatial pattern further, we plotted in Figure 14c) the river-groundwater exchange flows along the path of the Wairau River. The length and the up/downward direction of the bars indicates the flow rate and losing/gaining conditions, respectively. Also shown are the river stage and the groundwater table underneath the river. The picture confirms that the river is losing in the first 18 km of the modelled section and that it is gaining downstream of the location where the Dillons Point Formation is outcropping at the surface. The analysis revealed further, that the river appears to be hydraulically disconnected (perched) over long distances in the losing section. Head observations are mainly not located close to the river (Figure 14a), but the projection onto the groundwater table underneath the river shows a good agreement also for the confined area in the East, where the river is connected to groundwater (Figure 14c). The largest river losses were predicted for the section between SH6 and Wratts Rd ($Q_{ex2} = 4.77\,\mathrm{m}^3\mathrm{s}^{-1}$) while the losses in the upstream section are lower ($Q_{ex1} = 1.20\,\mathrm{m}^3\mathrm{s}^{-1}$). Downstream of Wratts Rd, the river is at first still losing and then gaining all the way to SH1, which results in a net gain of $Q_{ex3} = -0.54\,\mathrm{m}^3\mathrm{s}^{-1}$. These values are consistent with the differential gauging data taken on that day (Figure 3).

3.2.3 Correlation with river flow

Important information for resource management purposes is the functional relationship between Wairau River flows and the net river-groundwater exchange, $Q_{ex,13}$. In order to cover a larger range of hydrological situations, we have extended the model forcings of the calibrated model and performed a forward simulation for the time period 1/1/2000 to 20/2/2017. For each day in this simulation, $Q_{ex,13}$ is plotted over the corresponding river discharge at SH1 in Figure 15. For clarity, the graph is truncated at Wairau river flows of $120\,\mathrm{m}^3\mathrm{s}^{-1}$. There seems to be a cap on the net exchange flows which don't exceed $12\,\mathrm{m}^3\mathrm{s}^{-1}$ for the data depicted here and rarely exceed $15\,\mathrm{m}^3\mathrm{s}^{-1}$ even for the larger flows in the simulation period (not shown). On the other hand, the net exchange flows are relatively stable above a $5\,\mathrm{m}^3\mathrm{s}^{-1}$ threshold for river flows greater than $20\,\mathrm{m}^3\mathrm{s}^{-1}$. However, when the river discharge at SH1 falls below $20\,\mathrm{m}^3\mathrm{s}^{-1}$, a steep decrease of the net exchange flows can be observed. This is an interesting result because the exchange flows seem to vary throughout the year within a relatively narrow range of $5-8\,\mathrm{m^3s^{-1}}$ but are markedly decreasing during low flow periods (Figure 15).

River recharge is the major source of water for the Wairau Plain aquifer. In the considered 3.5-year simulation period, the land surface recharge was only 1% of the total water balance and exhibited a strong seasonality (no recharge in summer). In wetter years, this value might be slightly larger. However, the impact on aquifer storage seems to depend almost exclusively on the river exchange flows, in particular on the frequency and duration of low-flow periods. A single large flood event doesn't counterbalance the net storage decrease of extended dry periods which is also supported by the analysis presented in the previous section (Figure 13). This means that extended dry periods could lead to a net aquifer storage decrease which would require an above-average wet period with frequent, but not necessarily large river floods to refill.

3.3 MTT predictions to Spring Creek

More than half of the estimated mean river exchange flow of $7.3\,\mathrm{m^3s^{-1}}$ re-emerges in Spring Creek. The spring is of great value for the community of the city of Blenheim for recreational activities and its discharge and water quality is an important indicator for the state of the shallow Rapaura aquifer. The age of the water is a supplementary measure for estimating the risks and negative impacts associated with hydrological extremes (such as droughts), catastrophic events (e.g., contaminant spills), and changes in land-use and climate. The simulated flux-weighted transit-time distribution of the calibrated model is depicted in Figure 16. Most of the water in Spring Creek appears to be older than 190 days. The mean transit time evaluated at the 50%-quantile of the cumulative density function (*cdf*) is less than a year (MTT = 344 d). A distinctive tailing of the *cdf* suggests a small contribution of water being older than two years.

The comparatively young age of the Spring Creek water is caused by the highly transmissive subsurface zone between the Wairau River and the Spring Creek area as described in Section 2.4 and depicted in Figure 10. Qualitatively, the MTT is in good agreement with the analysis of previous and current water chemistry and isotope data (e.g., Davidson and Wilson, 2011). The uncertainty of this type II model prediction is relatively large as shown by the spread of the transit-time *cdf*s from the NSMC runs (grey lines in Figure 16). The MTTs range between 222 and 421 days (histogram) and the mean of the MTT *cdf* is at 315 days lower that the MTT of the calibrated model.

The young age of Spring Creek makes it vulnerable to hydrological extremes. Scenario simulations with the calibrated model show that without the recharge from the Wairau River, Spring Creek would run dry within approximately 300 days (dash-dot line in Figure 8). These results demonstrate that Spring Creek is very closely related to the Wairau River and any changes in the flow statistics of the river will result in a matching change at Spring Creek with little time delay.

4 Summary and Conclusions

In this study, we presented a model-based approach to analyse the surface water - ground water exchange mechanisms in one of New Zealand's gravel-bed rivers. A highly parametrized numerical model was set up for a 23 km long section of the Wairau River and calibrated using different data types, regularization techniques, and the parameter estimation software PEST. The trade-off between data fit and parameter homogeneity was investigated with Pareto analysis methods. Null-space Monte-Carlo sampling techniques were applied to estimate predictive uncertainty for data types that were used in the calibration (type I predictions) and data types that were not included in the calibration data set (type II predictions).

Based on the results of this analysis, the following main conclusions can be drawn:

- The gravel aquifer underneath the river is almost exclusively recharged by river water. Land surface recharge accounts for only 1% of the long-term water balance.

- The river is disconnected from groundwater and constantly losing over 80% of the considered river section. This causes the net exchange flows to be always positive.

- Since the stage variation in braided rivers is relatively low compared to channelized rivers, the seepage rates from the disconnected sections of gravel-bed rivers also varies only within a narrow range. Net exchange-flow rates have a lower threshold which is typically exceeded and are capped at larger flows ($\sim$5 and 15 $m^3\,s^{-1}$, respectively, for the Wairau River).

- During low-flow periods, the active channel area is reduced and river-exchange flows decrease exponentially. Thus, prolonged dry periods lead to strongly reduced aquifer recharge. Single flood events typically do not refill the aquifer storage due to the cap on exchange flows.

- Shallow gravel aquifers under New Zealand's gravel-bed rivers can be extremely transmissive. The mean transit time of the major spring at the Wairau Plain is estimated to be less than 1 year. This makes the spring vulnerable to hydrological extremes.

- Groundwater resources in the shallow gravel aquifer are vulnerable too. Climate variation and particularly an increase in the frequency and duration of droughts will cause a drastic depletion of aquifer storage. However, the system is resilient to some degree, i.e. a sequence of wet years would increase aquifer storage again.

Previous reports on other New Zealand gravel-bed rivers suggest that the river-groundwater exchange mechanisms found here for the Wairau River are typical for rivers with similar settings. Local differences remain due to the specific geological settings and hydraulic characteristics of the aquifer materials. However, the topographical setting of the of the South Island of New Zealand has caused the formation of relatively similar low-land river systems, particularly along the East coast. The rigour and transferability of our findings to other gravel-bed river systems should be investigated in future studies.

Acknowledgements

This work is a joint effort between the University of Dresden, Germany, Lincoln Agritech Ltd. and the Marlborough District Council, New Zealand. It has been supported by core funding of these organizations, by the International Bureau of the Federal Ministry of Education and Research, Germany (grant 01DR13023), the German Research Foundation (grant WO 1781/1-1), the Royal Society of New Zealand on behalf of the New Zealand Ministry of Business, Innovation and Employment, and the MBIE Transfer Pathways programme (contract LVLX1502). Special thanks are extended to Mike Ede and the entire field team at MDC for collecting the data for this study and to people from the MDC's hydrology and rivers group for fruitful insights and discussions.

References

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1998). Crop evapotranspiration. FAO Irrigation and Drainage Paper No. 56, Rome, Italy.

Baalousha, H. M. (2012). Modelling surface-groundwater interaction in the Ruataniwha basin, Hawke's Bay, New Zealand. *Environmental Earth Sciences*, 66(1):285–294.

Bandaragoda, C., Tarboton, D. G., and Woods, R. (2004). Application of TOPNET in the distributed model intercomparison project. *Journal of Hydrology*, 298(1-4):178–201.

Barthel, R. and Banzhaf, S. (2016). Groundwater and surface water interaction at the regional-scale: A review with focus on regional integrated models. *Water Resources Management*, 30(1):1–32.

Brown, L. (1981). Late quaternary geology of the wairau plain, marlborough, new zealand. *New Zealand J. Geol. Geophys.*, 24:477–490.

Brown, L. J., Dravid, P. N., Hudson, N. A., and Taylor, C. B. (1999). Sustainable groundwater resources, heretaunga plains, hawke's bay, new zealand. *Hydrogeology Journal*, 7(5):440–453.

Brunner, P., Cook, P. G., and Simmons, C. T. (2009). Hydrogeologic controls on disconnection between surface water and groundwater. *Water Resour. Res.*, 45(1).

Brunner, P., Cook, P. G., and Simmons, C. T. (2011). Disconnected surface water and groundwater: From theory to practice. *Ground Water*, 49(4):460–467.

Brunner, P., Simmons, C. T., Cook, P. G., and Therrien, R. (2010). Modeling surface water-groundwater interaction with MODFLOW: Some considerations. *Ground Water*, 48(2):174–180.

Clark, M. P., Rupp, D. E., Woods, R. A., Zheng, X., Ibbitt, R. P., Slater, A. G., Schmidt, J., and Uddstrom, M. J. (2008). Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. *Advances in Water Resources*, 31(10):1309–1324.

Close, M., Matthews, M., Burbery, L., Abraham, P., and Scott, D. (2014). Use of radon to characterise surface water recharge to groundwater. *Journal of Hydrology (NZ)*, 53(2):113–127.

Dann, R., Close, M., Flintoft, M., Hector, R., Barlow, H., Thomas, S., and Francis, G. (2009). Characterization and estimation of hydraulic properties in an alluvial gravel vadose zone. *Vadose Zone Journal*, 8(3):651–663.

Davidson, P. and Wilson, S. R. (2011). *Groundwaters of Marlborough*. Number ISBN 978-1-927159-03-3. Marlborough District Council, The Caxton Press, Christchurch, New Zealand.

de Silva, C. and Rushton, K. (2007). Groundwater recharge estimation using improved soil moisture balance methodology for a tropical climate with distinct dry seasons. *Hydrological Sciences Journal*, 52(5):1051–1067.

Doherty, J. (2003). Ground water model calibration using pilot points and regularization. *Ground Water*, 41(2):170–177.

Doherty, J. (2016a). *Groundwater data utilities: Part B: Program descriptions*. Watermark Numerical Computing.

Doherty, J. (2016b). *PEST, Model-independent parameter estimation, User manual Part I*. Watermark Numerical Computing, Brisbane, Australia, 6th edition.

Doherty, J. (2016c). *PEST, Model-independent parameter estimation user manual Part II: PEST utility support software*. Watermark Numerical Computing, 6th edition.

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

Doherty, J. E., Fienen, M. N., and Hunt, R. J. (2010). Approaches to highly parameterized inversion: Pilot-point theory, guidelines, and research directions. Scientific Investigations Report 2010-5168, U.S. Geological Survey. 36 p.

Dravid, P. N. and Brown, L. J. (1997). Heretaunga Plains groundwater study. Vol 1: Findings. Technical report, Hawke's Bay Regional Council and Inst Geol Nucl Sci. 254 pp.

Fenemor, A. D. (1989). Groundwater modelling as a tool for water management: Waimea Plains, Nelson. *Journal of Hydrology (NZ)*, 28(1):17–31.

Furman, A. (2008). Modeling coupled surface-subsurface flow processes: A review. *Vadose Zone J*, 7(2):741–756.

González-Pinzón, R., Ward, A. S., Hatch, C. E., Wlostowski, A. N., Singha, K., Gooseff, M. N., Haggerty, R., Harvey, J. W., Cirpka, O. A., and Brock, J. T. (2015). A field comparison of multiple techniques to quantify groundwater-surface-water interactions. *Freshwater Science*, 34(1).

Green, S., Agnew, R., and Grevin, M. (2014). Monitoring nitrate loss under vineyard soils on the Wairau Plains, Marlborough. Research report spts 10328, Plant & Food.

Gupta, H. V., Sorooshian, S., and Yapo, P. O. (1998). Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research*, 34(4):751–764. DOI:10.1029/97WR03495.

Gusyev, M. A., Toews, M., Morgenstern, U., Stewart, M., White, P., Daughney, C., and Hadfield, J. (2013). Calibration of a transient transport model to tritium data in streams and simulation of groundwater ages in the western Lake Taupo catchment, New Zealand. *Hydrology and Earth System Sciences*, 17(3):1217–1227.

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., USGS.

Jones, J. P., Sudicky, E. A., and McLaren, R. G. (2008). Application of a fully-integrated surface-subsurface flow model at the watershed-scale: A case study. *Water Resour. Res.*, 44:–.

Kalbus, E., Reinstorf, F., and Schirmer, M. (2006). Measuring methods for groundwater - surface water interactions: a review. *Hydrology and Earth System Sciences*, 10(6):873–887.

Kollat, J. B., Reed, P. M., and Wagener, T. (2012). When are multiobjective calibration trade-offs in hydrologic models meaningful? *Water Resour. Res.*, 48(3):W03520.

Kollet, S. J. and Maxwell, R. M. (2006). Integrated surface-groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. *Advances in Water Resources*, 29:945 – 958.

LaBolle, E. M., Ahmed, A. A., and Fogg, G. E. (2003). Review of the integrated groundwater and surface-water model (IGSM). *Ground Water*, 41(2):238–246.

Lamontagne, S., Taylor, A. R., Cook, P. G., Crosbie, R. S., Brownbill, R., Williams, R. M., and Brunner, P. (2014). Field assessment of surface water-groundwater connectivity in a semi-arid river basin (Murray-Darling, Australia). *Hydrol. Process.*, 28(4):1561–1572.

Landcare Research (2000). *New Zealand Fundamental Soil database*. Landcare Research, Lincoln, New Zealand. http://lris.scinfo.org.nz/.

Larned, S. T., Hicks, D. M., Schmidt, J., Davey, A. J. H., Dey, K., Scarsbrook, M., Arscott, D. B., and Woods, R. A. (2008). The Selwyn River of New Zealand: a benchmark system for alluvial plain rivers. *River Res. Applic.*, 24(1):1–21.

Leake, S. and Lilly, M. (1997). Documentation of a computer program (FHB1) for assignment of transient specified-flow and specified-head boundaries in applications of the modular finite-difference ground-water flow model (MODFLOW). Open-File Report 97-571, U.S. Geological Survey. 50 p.

Maxwell, R. M., Condon, L. E., and Kollet, S. J. (2015). A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. *Geoscientific Model Development*, 8(3):923–937.

Moore, C. (2005). *The use of regularized inversion in groundwater model calibration and prediction uncertainty analysis*. PhD thesis, The University of Queensland.

Moore, C. and Doherty, J. (2005). Role of the calibration process in reducing model predictive error. *Water Resour. Res.*, 41(5):W05020.

Moore, C., Wöhling, Th., and Doherty, J. (2010). Efficient regularization and uncertainty-analysis using a global optimization methodology. *Water Resources Research*, 46:W08527.

Niswonger, R., Panday, S., and Ibaraki, M. (2011). MODFLOW-NWT, A Newton formulation for MODFLOW-2005. Techniques and Methods 6-A37, U.S. Geological Survey. 44 p.

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. Techniques and methods, U.S. Geological Survey. Book 6, Chap. A13, 47 p.

Panday, S. and Huyakorn, P. S. (2004). A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. *Advances in Water Resources*, 27:361 – 382.

Pirot, G., Straubhaar, J., and Renard, P. (2014). Simulation of braided river elevation model time series with multiple-point statistics. *Geomorphology*, 214:148–156.

Pollock, D. W. (2012). User guide for MODPATH version 6 - A particle-tracking model for MODFLOW. USGS Numbered Series 6-A41, U.S. Geological Survey, Reston, VA.

Raiber, M., White, P. A., Daughney, C. J., Tschritter, C., Davidson, P., and Bainbridge, S. E. (2012). Three-dimensional geological modelling and multivariate statistical analysis of water chemistry data to analyse and visualise aquifer structure and groundwater composition in the Wairau Plain, Marlborough District, New Zealand. *Journal of Hydrology*, 436-437:13–34.

Reed, P., Hadka, D., Herman, J., Kasprzyk, J., and Kollat, J. (2013). Evolutionary multiobjective optimization in water resources: The past, present, and future. *Advances in Water Resources*, 51(0):438–456.

Rosen, M. and White, P., editors (2001). *Groundwaters of New Zealand*. New Zealand Hydrological Society Inc., Wellington. 489 pp.

Rosenberry, D. and LaBaugh, J. W. (2008). Field techniques for estimating water fluxes between surface water and ground water. Technical Report 4-D2, USGS.

Rupp, D. E., Larned, S. T., Arscott, D. B., and Schmidt, J. (2008). Reconstruction of a daily flow record along a hydrologically complex alluvial river. *Journal of Hydrology*, 359(1-2):88–104.

Rushton, K., Eilers, V., and Carter, R. (2006). Improved soil moisture balance methodology for recharge estimation. *Journal of Hydrology*, 318(1-4):379–399.

Soil Conservation Service (1972). *National Engineering Handbook: Section 4, Hydrology*. SCS, US Department of Agriculture.

Sophocleous, M. (2002). Interactions between groundwater and surface water: The state of the science. *Hydrogeology Journal*, 10:52–67.

Spanoudaki, K., Stamou, A. I., and Nanou-Giannarou, A. (2009). Development and verification of a 3-d integrated surface water-groundwater model. *Journal of Hydrology*, 375(3-4):410–427.

Stanley, E. H. and Jones, J. B. (2000). *Streams and ground waters.*, chapter Surface-subsurface interactions: past, present, and future, pages 405–417. Academic Press, San Diego.

Tonkin, M. and Doherty, J. (2009). Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques. *Water Resources Research*, 45:W00B10.

von Gunten, D., Wöhling, Th., Haslauer, C., Merchan, D., Causape, J., and Cirpka, O. (2014). Efficient calibration of a distributed pde-based hydrological model using grid coarsening. *Journal of Hydrology*, 519, Part D(0):3290–3304.

White, P., Kovacova, E., Zemamsky, G., Jebbour, N., and Moreau-Fournier, M. (2012). Groundwater-surface water interaction in the Waimakariri River, New Zealand, and groundwater outflow from the river bed. *Journal of Hydrology (NZ)*, 51(1):1–24.

White, P. A. (2009). Avon river springs catchment, christchurch city, new zealand. *Australian Journal of Earth Sciences*, 56(1):61–70.

White, P. A., Clausen, B., Hunt, B., Cameron, S., and Weir, J. J. (2001). *Groundwaters of New Zealand*, chapter Groundwater-surface water interaction, pages 133–160. New Zealand Hydrological Society. ISBN 0-473-07816-3.

Wöhling, Th., Gayler, S., Priesack, E., Ingwersen, J., Wizemann, H.-D., Högy, P., , Cuntz, M., Attinger, S., Wulfmeyer, V., and Streck, T. (2013). Multiresponse, multiobjective calibration as a diagnostic tool to compare accuracy and structural limitations of five coupled soil-plant models and CLM3.5. *Water Resources Research*, 49(12):8200–8221. (submitted).

Wöhling, Th., Vrugt, J. A., and Barkle, G. F. (2008). Comparison of three multiobjective algorithms for inverse modeling of vadose zone hydraulic properties. *Soil Science Society of America Journal*, 72(2):305–319.

Williams, R. D., Brasington, J., and Hicks, D. M. (2016). Numerical modelling of braided river morphodynamics: Review and future challenges. *Geography Compass*, 10(3):102–127.

Williams, R. D., Brasington, J., Vericat, D., and Hicks, D. M. (2014). Hyperscale terrain modelling of braided rivers: fusing mobile terrestrial laser scanning and optical bathymetric mapping. *Earth Surf. Process. Landforms*, 39(2):167–183.

Wilson, S. and Wöhling, Th. (2015). Wairau River - Wairau aquifer interaction. Envirolink Report 1514-MLDC96, Lincoln Agritech Ltd.

Wilson, S. R. (2016). Wairau aquifer stratigraphy review. Technical Report 1053-1-R1, Lincoln Agritech Ltd., Chistchurch, New Zealand.

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

Yang, J., McMillan, H., and Zammit, C. (2017). Modeling surface water-groundwater interaction in New Zealand: Model development and application. *Hydrol. Process.*, 31(4):925–934.

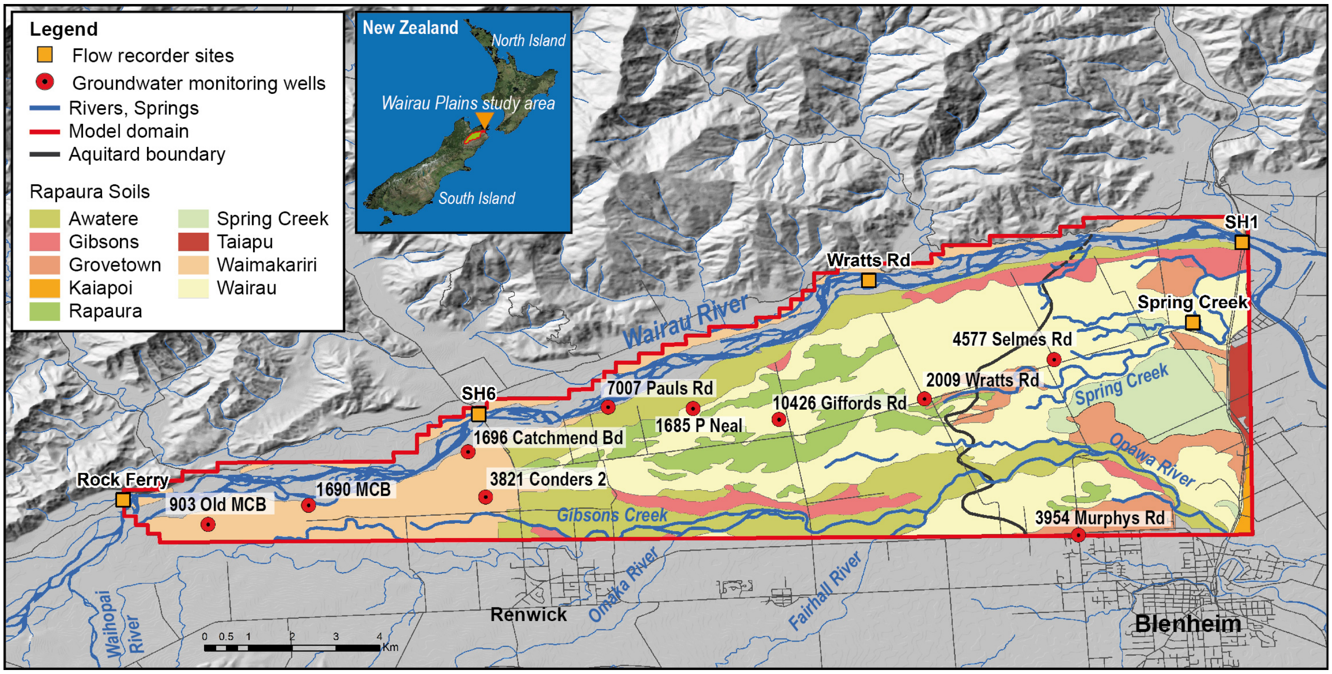

Figure 1: The Wairau Plain study site and model domain.

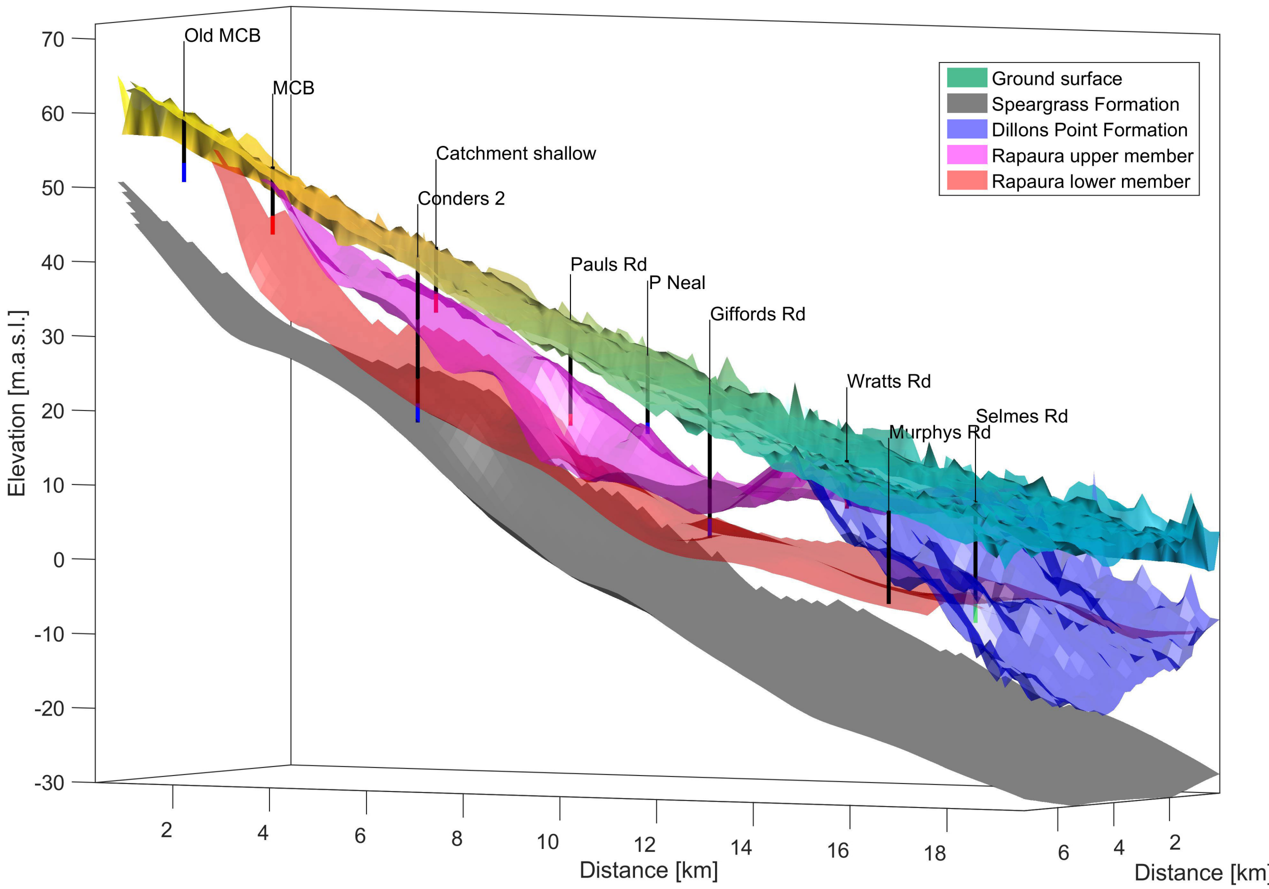

Figure 2: Conceptualization of the geology in the Wairau Plain model domain.

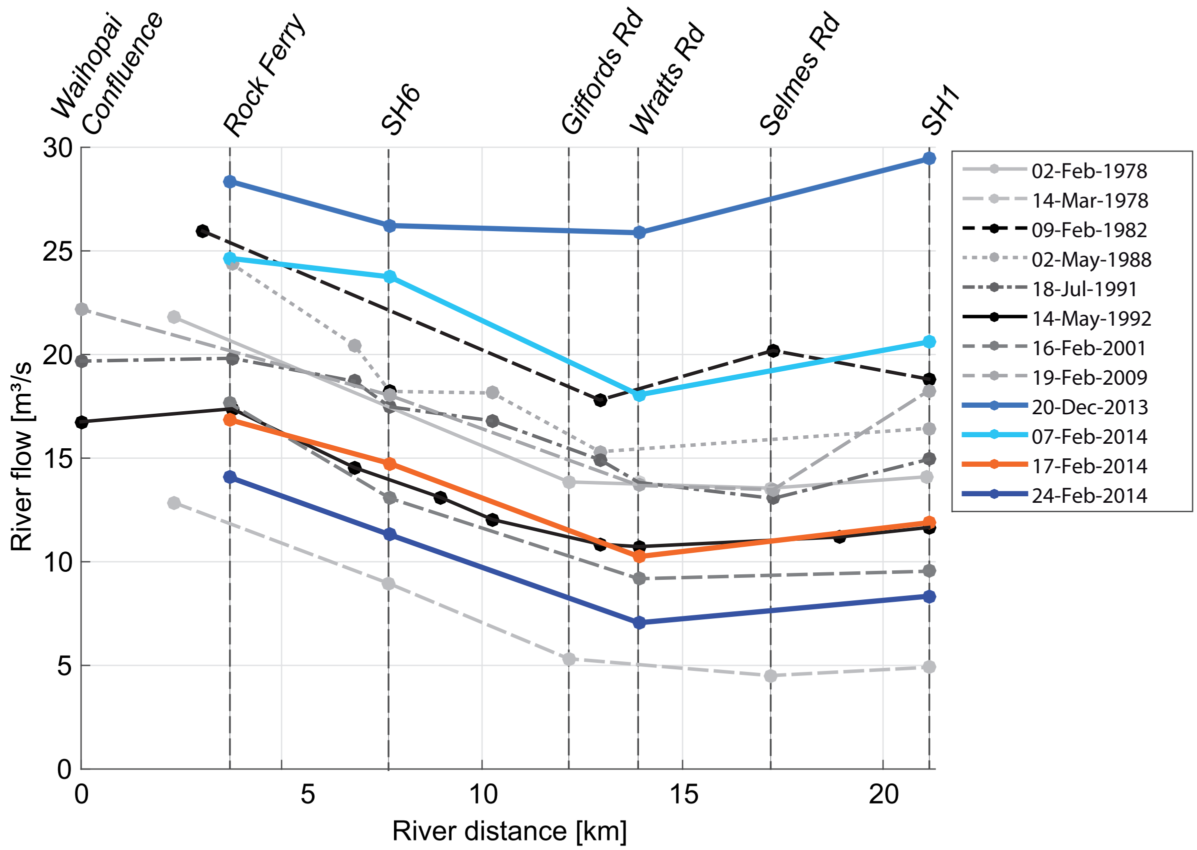

Figure 3: Historic and contemporary differential flow gaugings of the Wairau River.

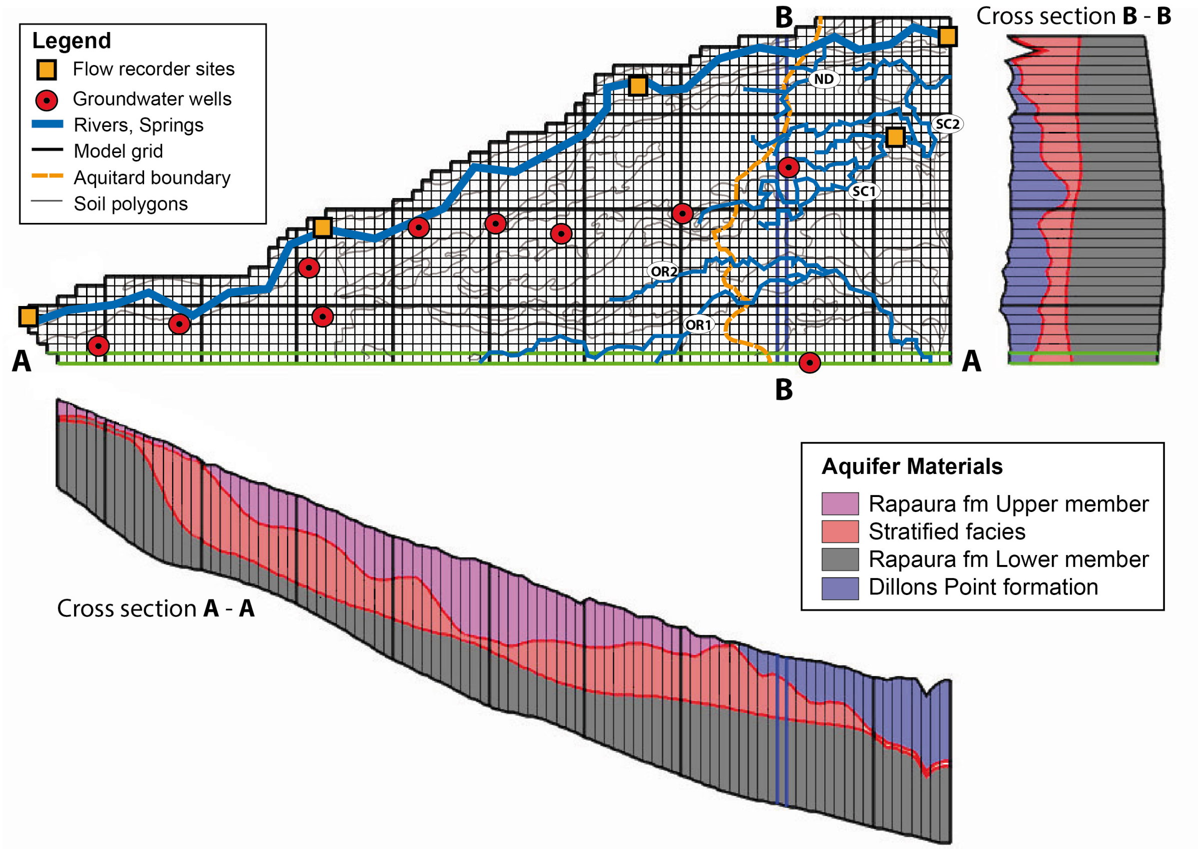

Figure 4: Wairau Plain surface water - groundwater flow model.

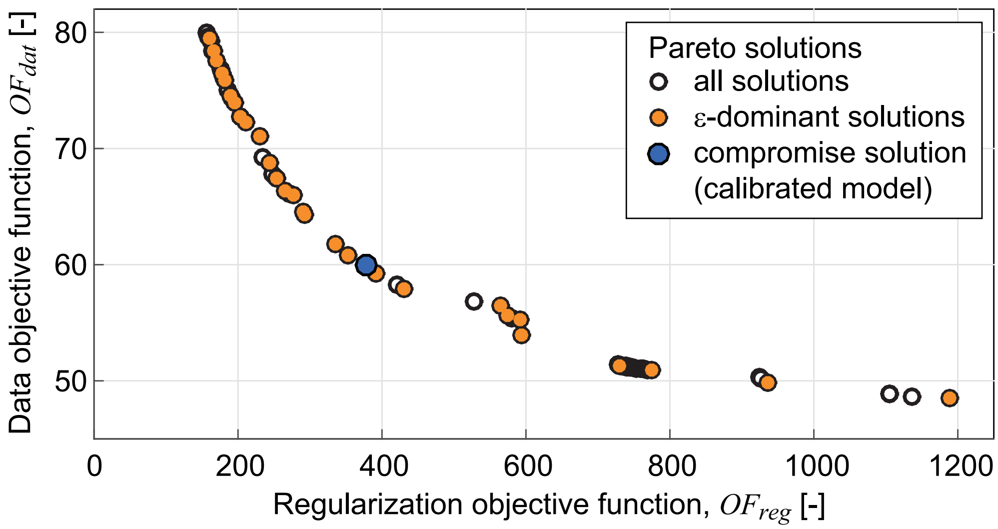

Figure 5: Model calibration: Trade-off between measurement and regularization objective function.

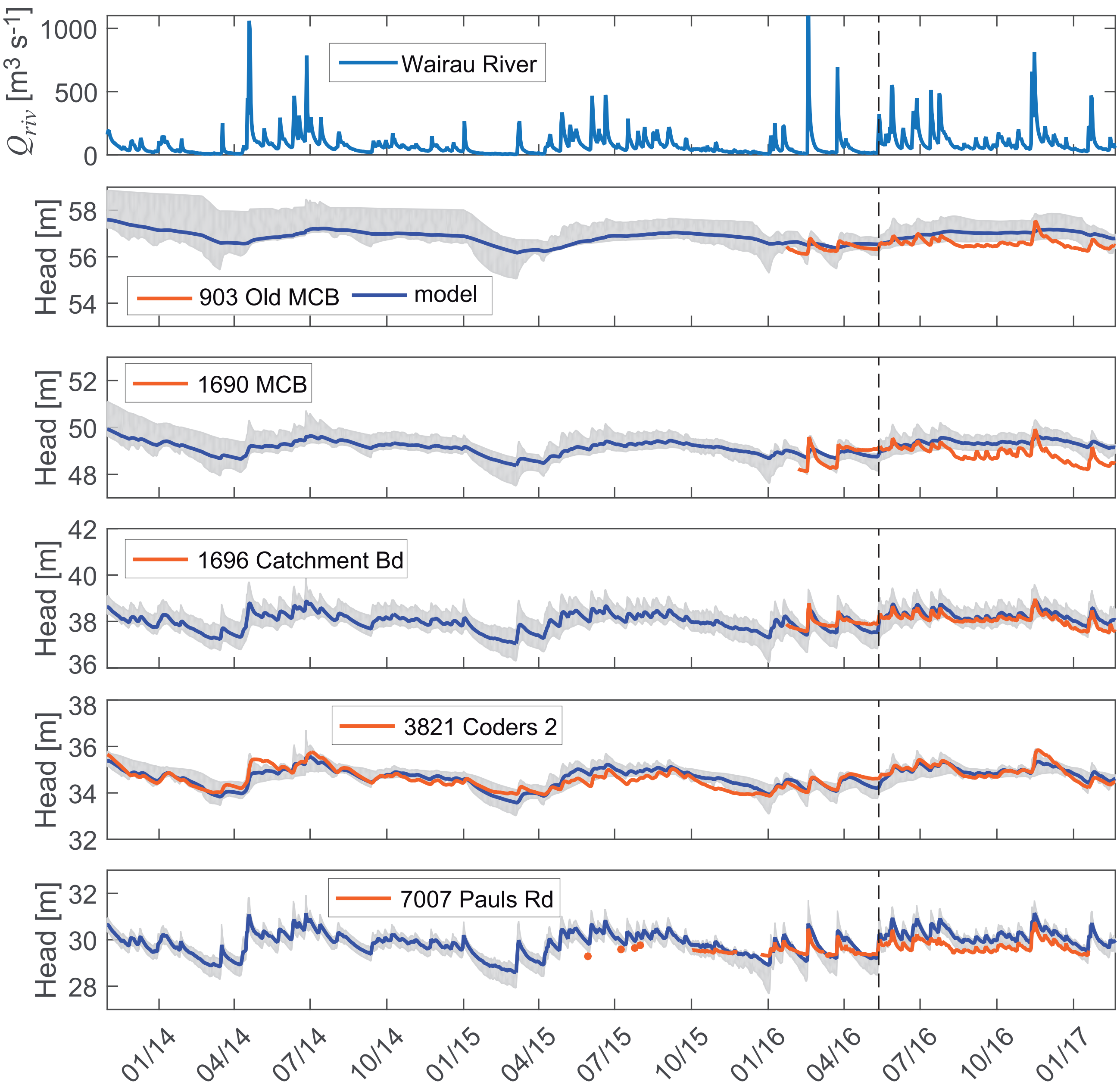

Figure 6: Measured and predicted groundwater heads for the Old MCB, MCB, Catchment Board, Conders, and Pauls Rd wells. The corresponding Wairau River discharge at SH1 is shown in the top panel and 95% uncertainty bounds are shaded gray.

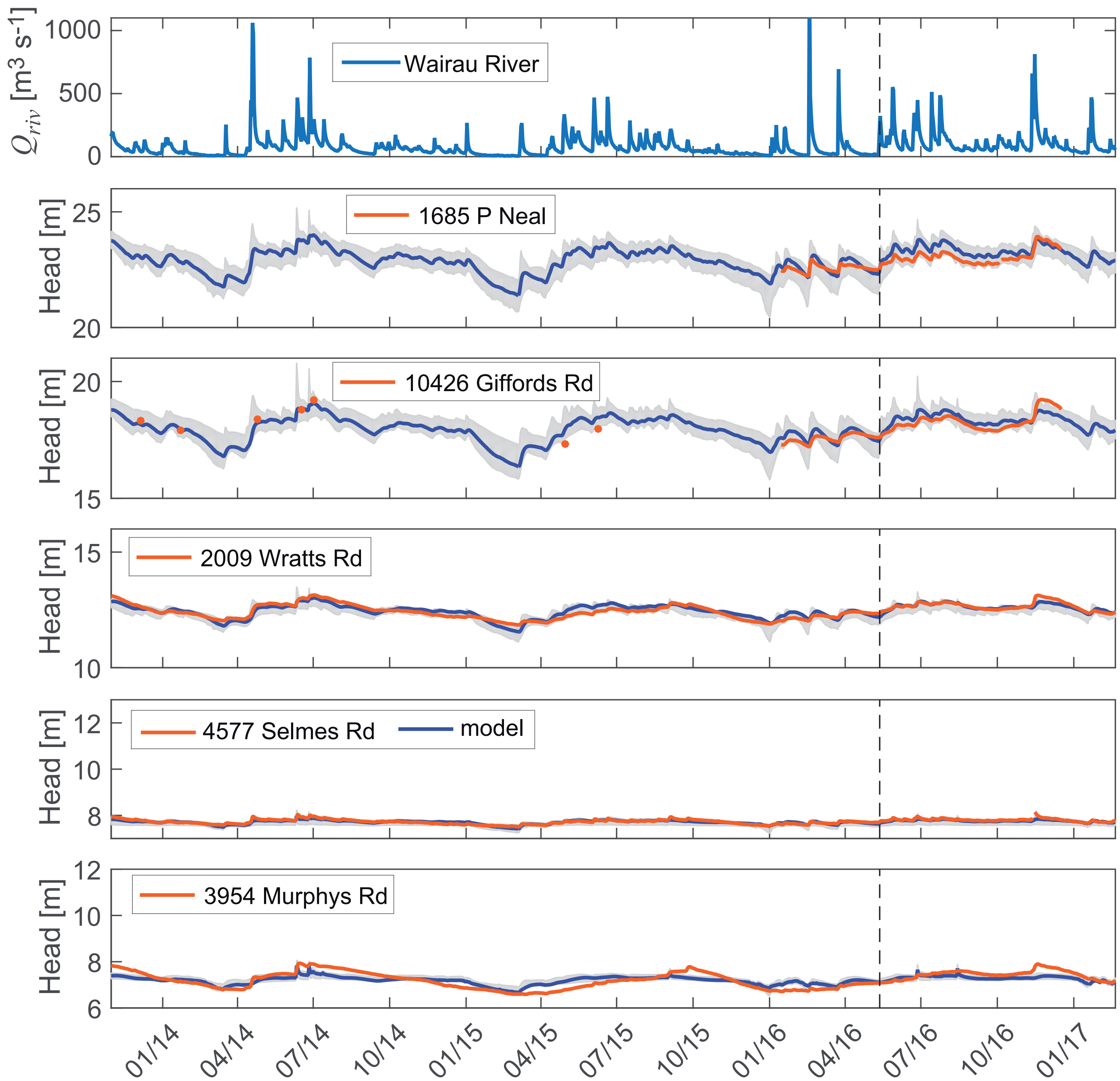

Figure 7: Measured and predicted groundwater heads for the P Neal, Giffords Rd, Wratts Rd, Selmes Rd, and Murphys Rd wells. The corresponding Wairau River discharge at SH1 is shown in the top panel and 95% uncertainty bounds are shaded gray.

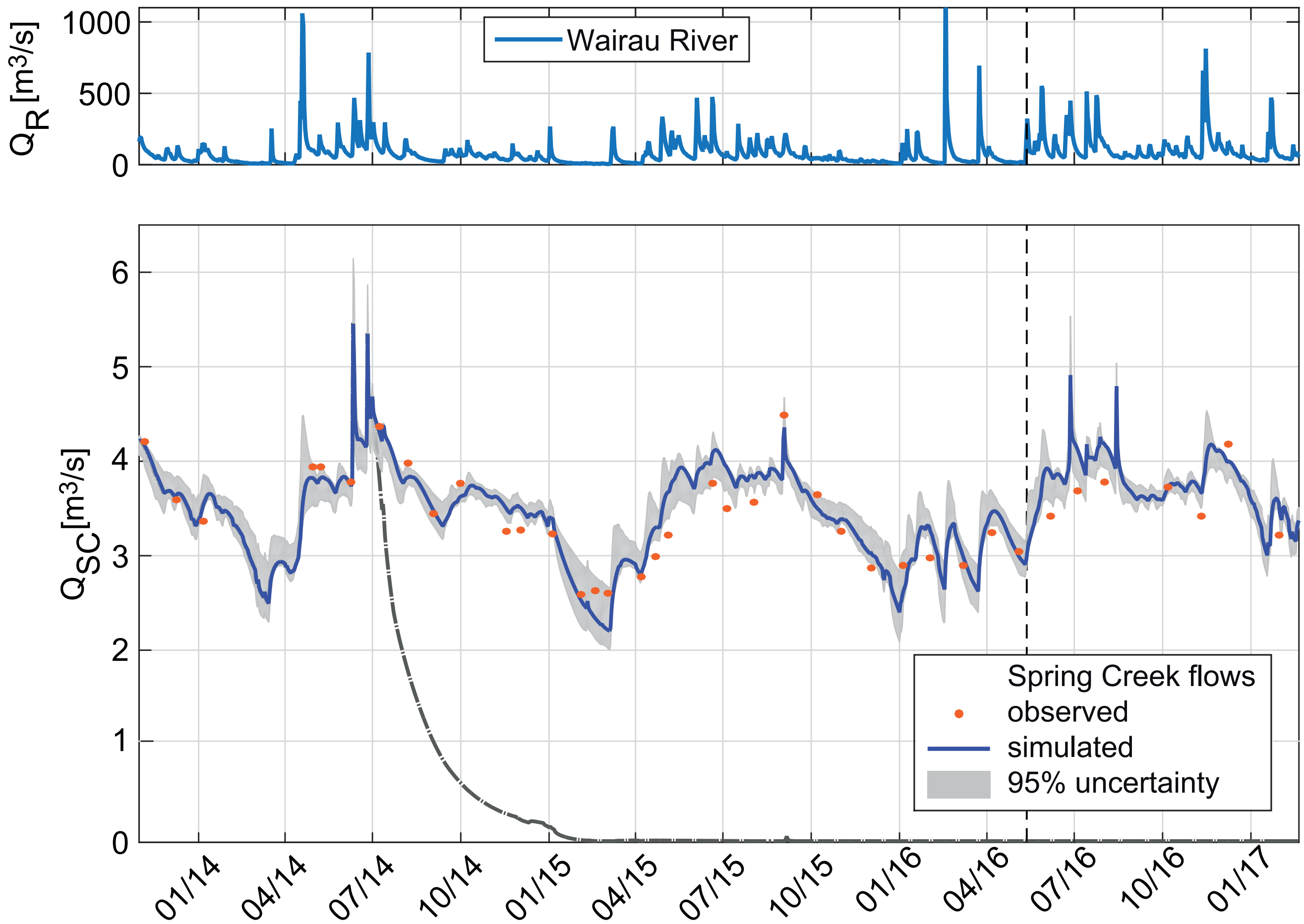

Figure 8: Measured and predicted Spring Creek flows including the 95% predictive uncertainty bounds.

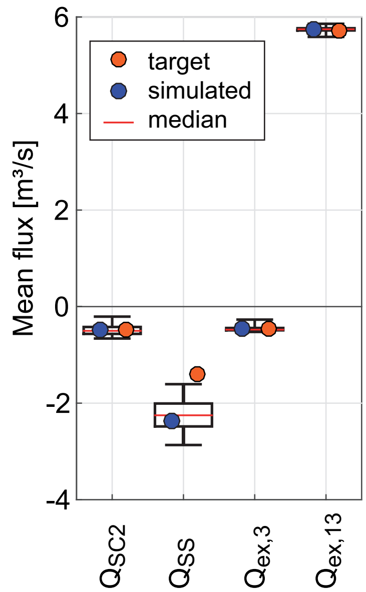

Figure 9: Boxplot of the calibration "soft targets", the corresponding simulations with the calibrated model and the 50%/95% uncertainty bounds (box/whiskers).

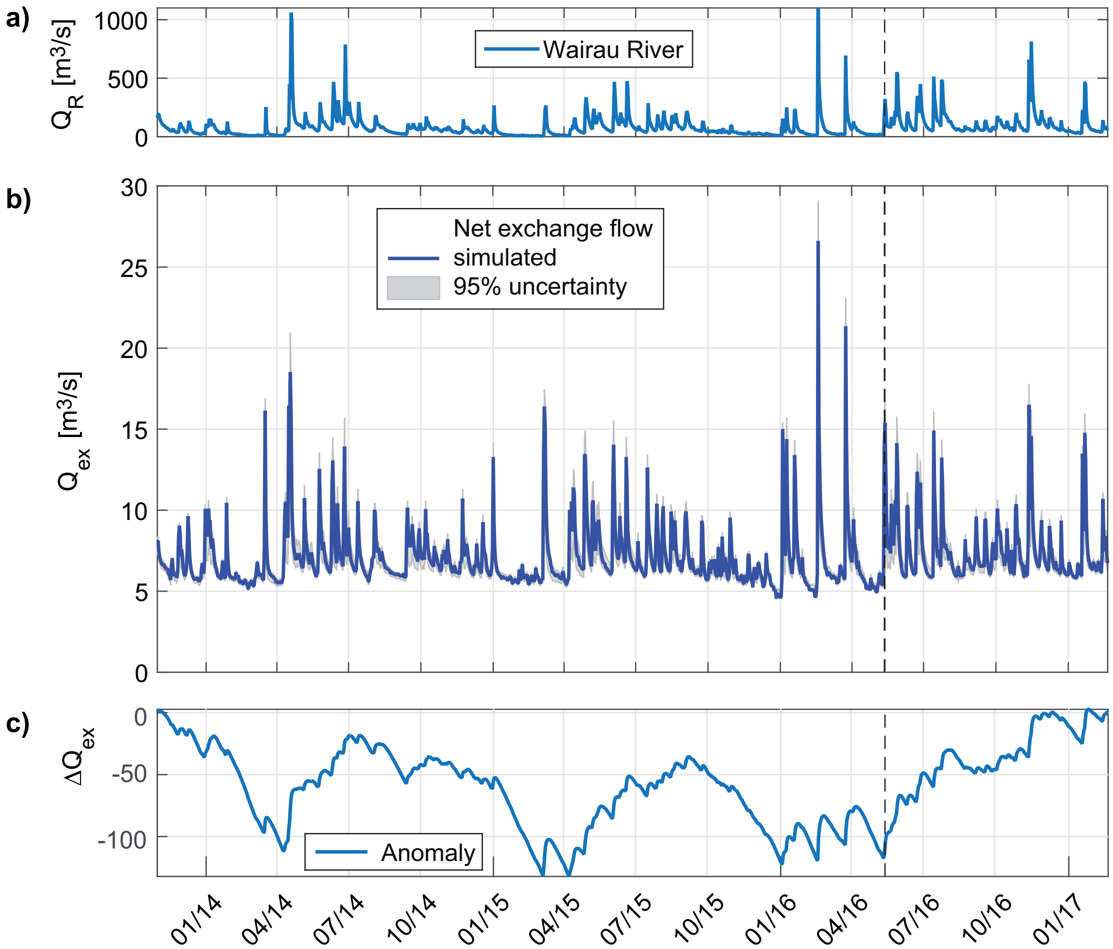

Figure 13: Simulated net river - groundwater exchange flow: a) the Wairau River flow, b) the net exchange flow, and c) the anomaly of net exchange flows.

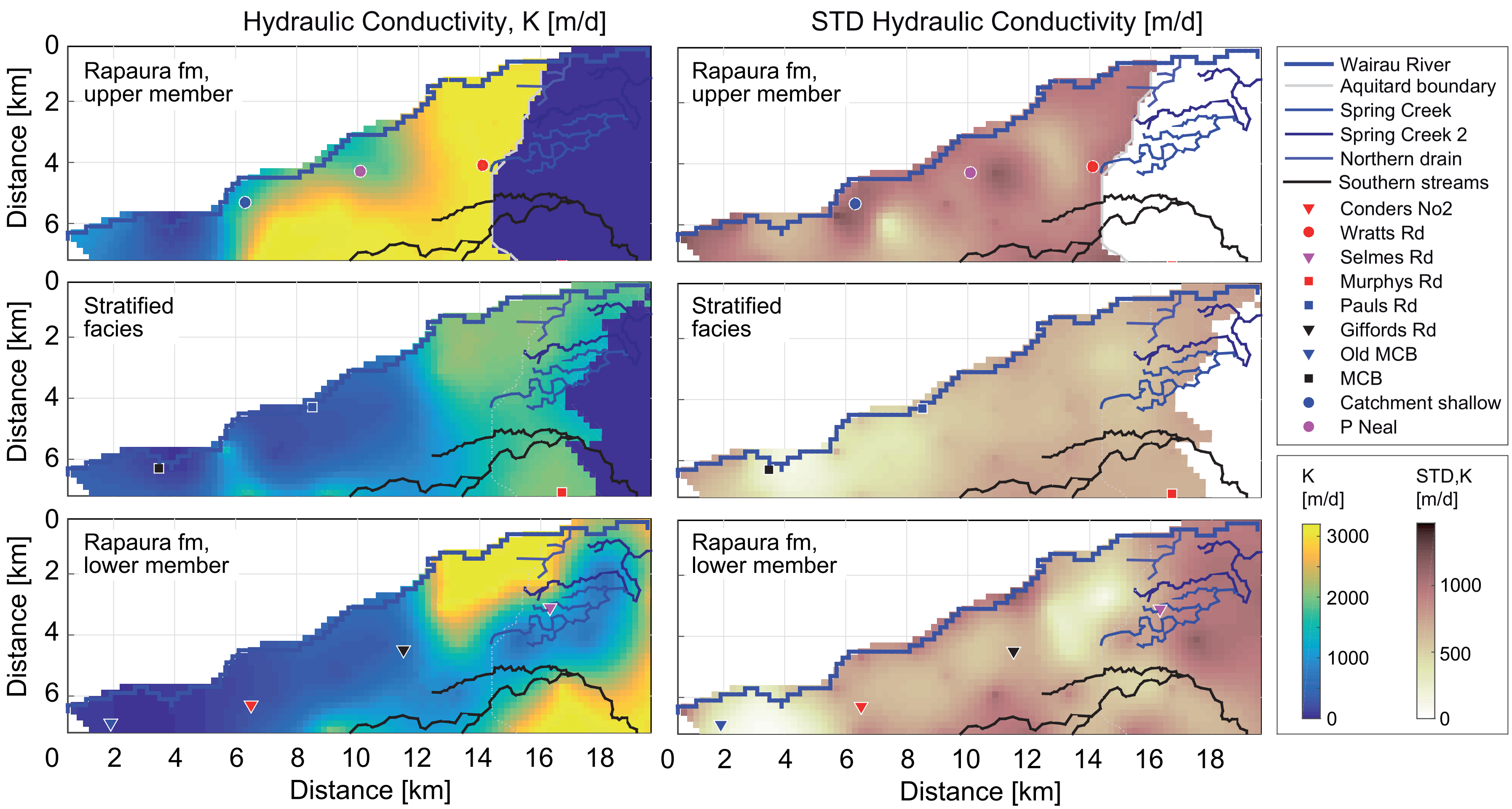

Figure 10: Hydraulic conductivity fields of the calibrated model and the three different layers (left panels) and their corresponding uncertainty (right panels).

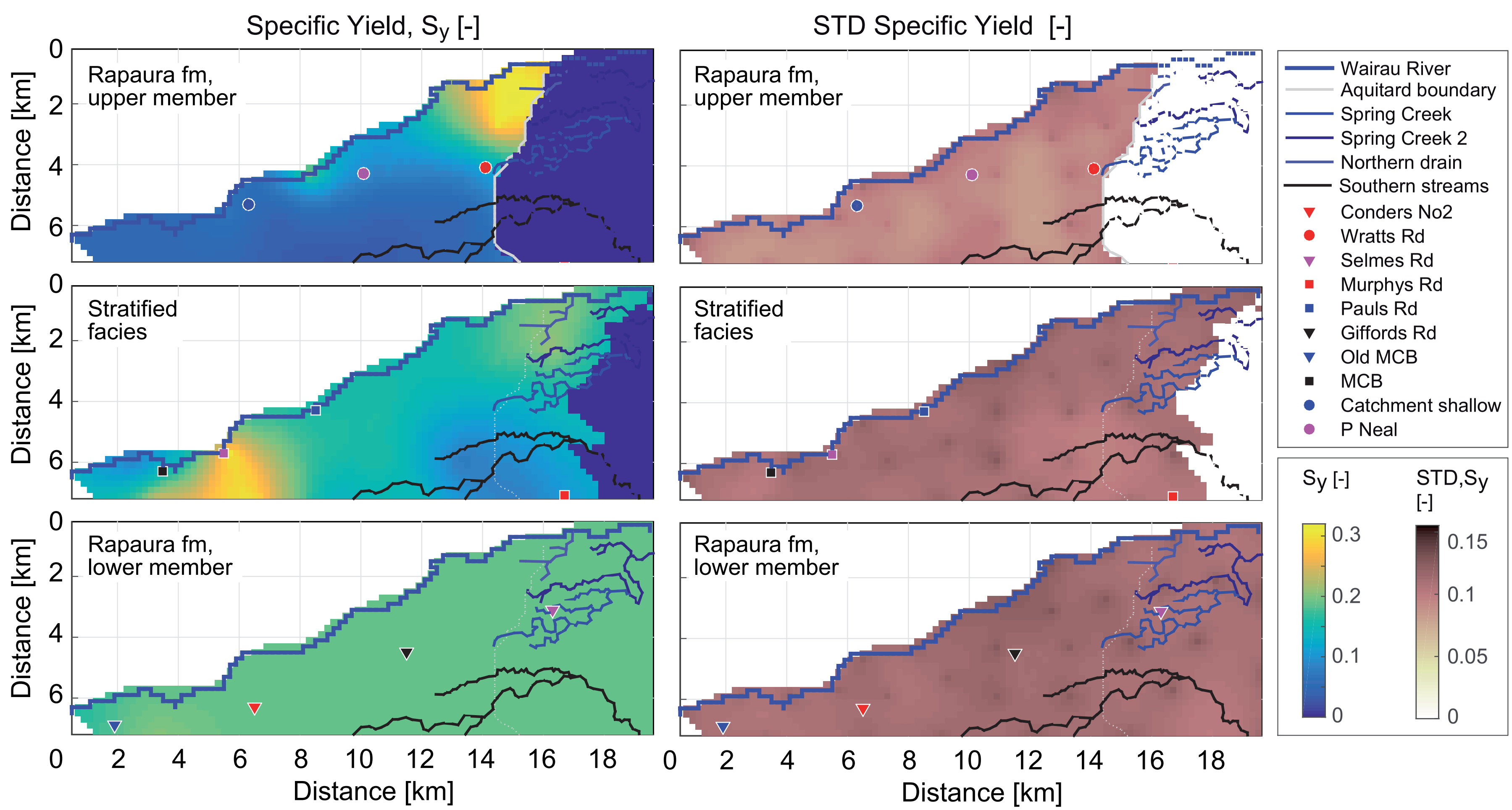

Figure 11: Specific yield fields of the calibrated model and the three different layers (left panels) and their corresponding uncertainty (right panels).

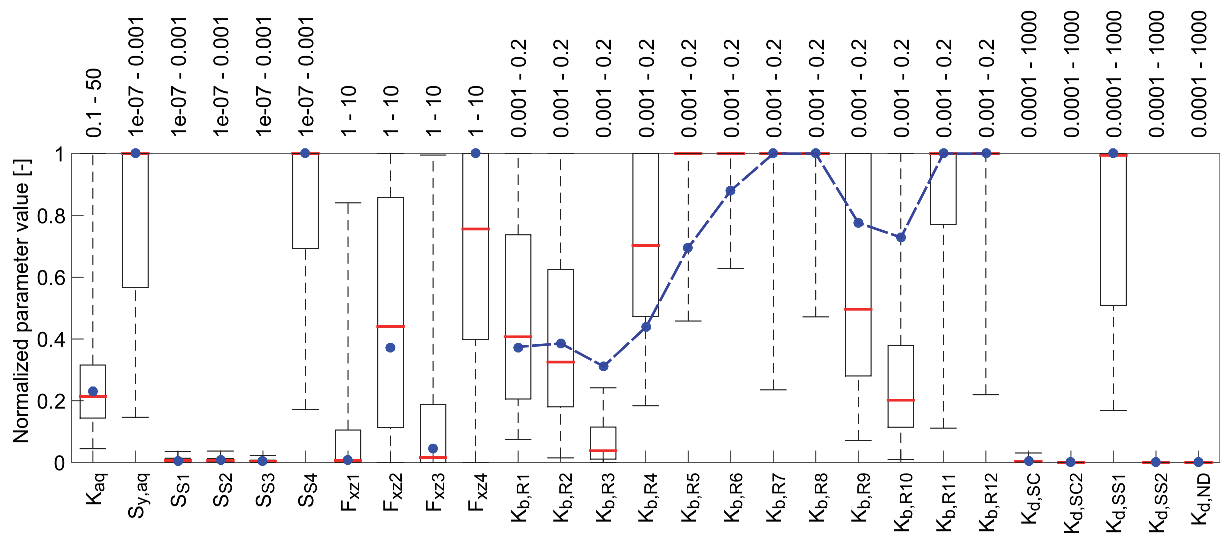

Figure 12: Boxplot of normalized parameter uncertainty with the 50%/95% uncertainty bounds shown by the boxes and whiskers, respectively. The parameter values of the calibrated models are indicated by the blue dots. The uncertainty of the K and S_y parameter fields are shown in Figures 12 and 13.

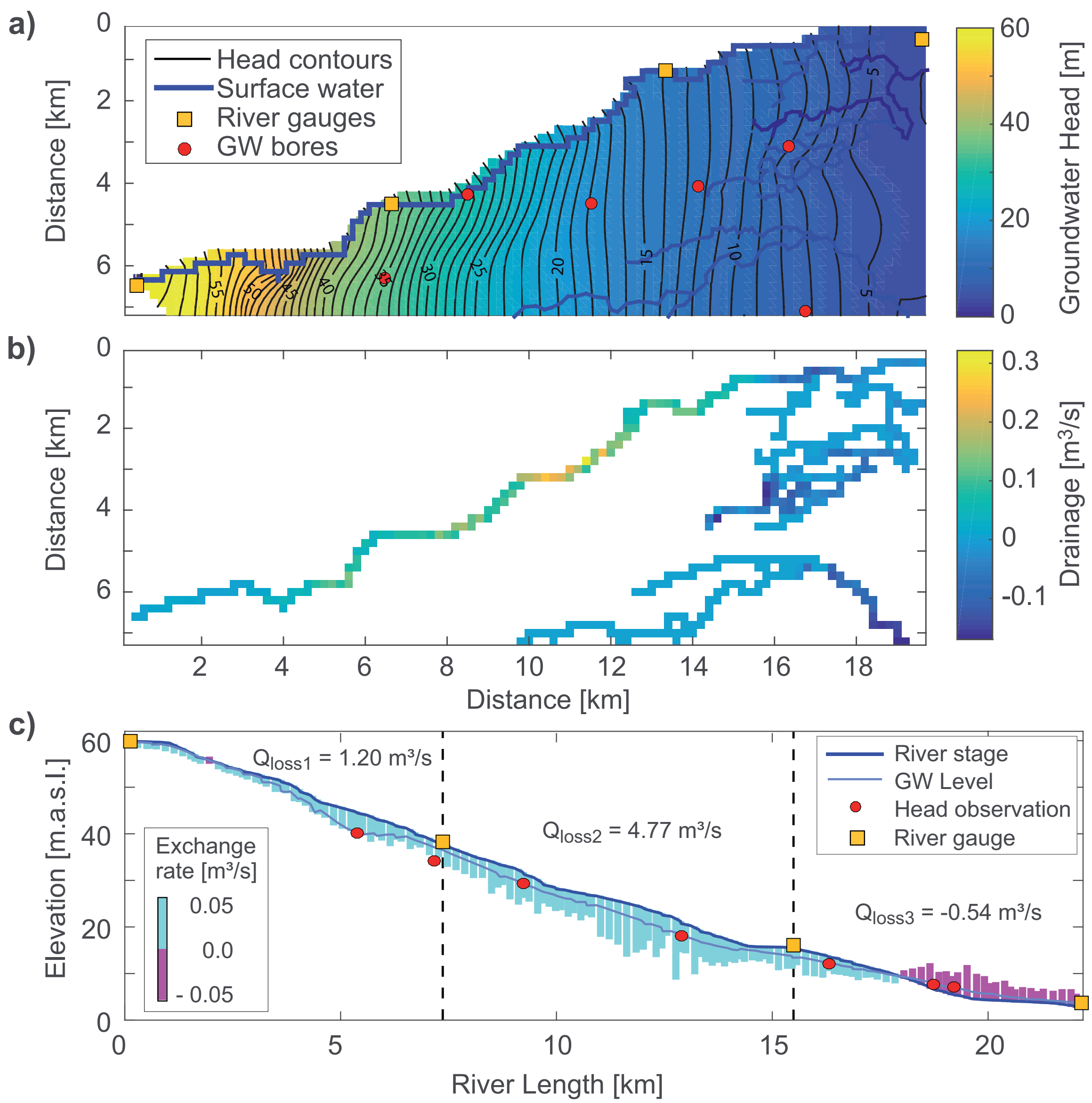

Figure 14: Snapshot of the transient model simulations for the low flow period (17/02/2014): a) the hydraulic head field, b) the spatial distribution of the drainage rates of the Wairau River and the spring/stream network of the Wairau Plain, and c) the connectivity between Wairau River and groundwater and the cell-by-cell drainage rates projected along the river length.

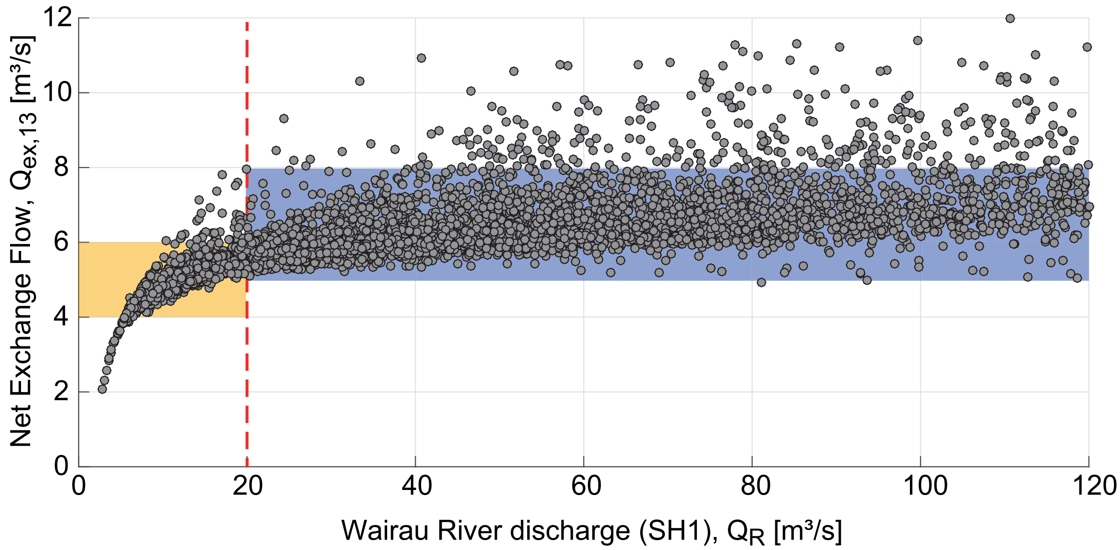

Figure 15: Relationship between Wairau River discharge and net river-groundwater exchange flow.

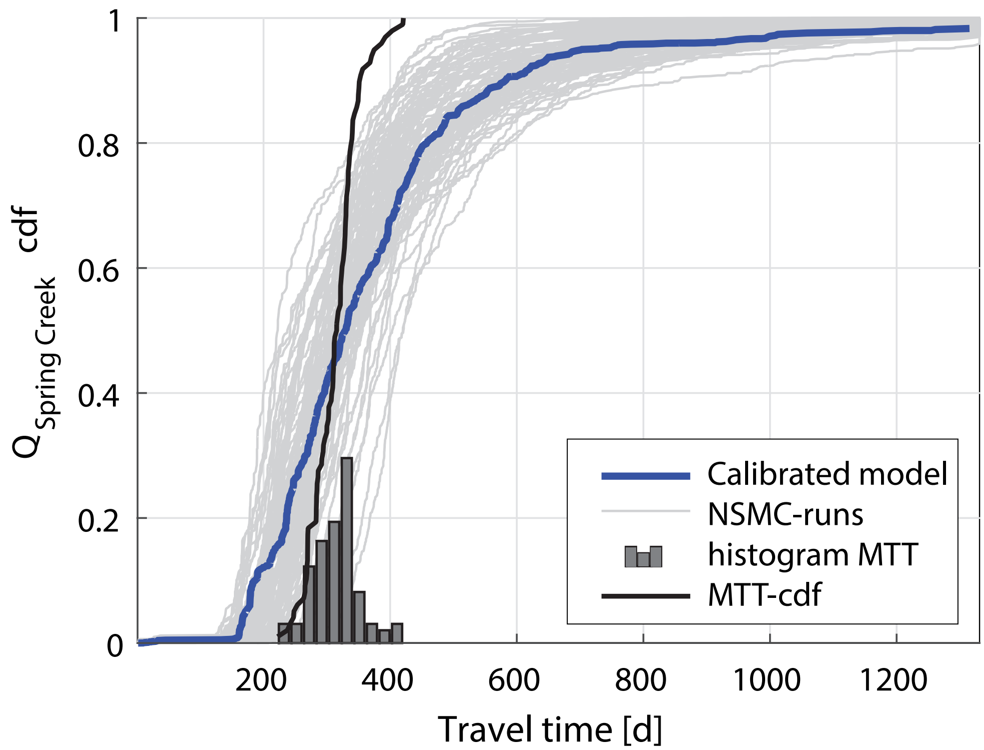

Figure 16: Cumulative travel time distribution of Spring Creek flows of the calibrated model and the corresponding uncertainty from the NSMC runs. Also shown is the cumulative distribution function of the mean travel time of these runs.

Table 1: Parameters of the Wairau Plain model and corresponding ranges used in the model calibration. Symbols are described in the text.

Name	# of parameters	Range
K_H (Rapaura Fm) [$\mathrm{m\,d^{-1}}$]	90	$1E^{0}$ - $1E^{3}$
K_H (Dillons Pt. Fm) [$\mathrm{m\,d^{-1}}$]	1	$1E^{-1}$ - $5E^{1}$
S_y (Rapaura Fm) [$\mathrm{m^3\,m^{-3}}$]	90	$1E^{-4}$ - $3E^{-1}$
S_y (Dillons Pt. Fm) [$\mathrm{m^3\,m^{-3}}$]	1	$1E^{-7}$ - $1E^{-3}$
S_S [$\mathrm{m^{-1}}$]	4	$1E^{-7}$ - $1E^{-3}$
f_a [-]	4	$1E^{0}$ - $1E^{1}$
K_R [$\mathrm{m\,d^{-1}}$]	12	$1E^{-3}$ - $2E^{-1}$
K_D [$\mathrm{m\,d^{-1}}$]	5	$1E^{-4}$ - $1E^{3}$

Table 3: Performance of the calibrated model for groundwater head and Spring Creek flow data.

	Calibration Period			Evaluation Period		
Target	# of obs	RMSE	R^2	# of obs	RMSE	R^2
Old MCB	111	0.19	-0.22	284	0.45	0.12
MCB	97	0.35	0.53	284	0.55	0.36
Catchment Bd	111	0.31	0.68	284	0.21	0.89
Conders 2	925	0.24	0.84	284	0.17	0.91
Pauls Rd	199	0.30	0.84	252	0.49	0.90
P Neal	119	0.23	0.83	217	0.34	0.79
Giffords Rd	126	0.20	0.79	219	0.28	0.78
Wratts Rd	925	0.15	0.87	284	0.10	0.88
Selmes Rd	925	0.05	0.87	282	0.05	0.84
Murphys Rd	925	0.23	0.82	284	0.21	0.66
Spring Creek Flows	31	0.22	0.91	7	0.32	0.67

Table 2: Number of observations used for model calibration and evaluation and their weight in the objective function.

Data type	Number of observations: Calibration period (925 days)	Number of observations: Evaluation period (284 days)	Observation weight (single data point)	Calibrated model objective function [-]
Groundwater head (permanent wells)	3700	1134	0.465	39.4
Groundwater head (temporary wells)	763	1540	0.465	
Differential river gauging	2	0	20.0	0.4
Spring Creek flows	31	7	$1.056\,E^{-4}$	19.3
Mean Spring Creek flow (section 2)	1	0	$6.056\,E^{-5}$	$6.2\,E^{-3}$
Mean flow southern streams	1	0	$1.056\,E^{-5}$	1.6
Total	4498	2681	-	60.7