Applications of digital baseflow separation techniques for model validation, Wairarapa valley, New Zealand

Lucas Charles Everitt

A thesis submitted in fulfilment of the requirements for the degree of Masters of Science in Physical Geography, at the Victoria University of Wellington, New Zealand

September 2018

Abstract

The representation of groundwater processes in hydrological models is crucial, as the connectivity between groundwater and surface water is significant. It is particularly important for regions such as the Wairarapa that experience high water stresses. Intensified agriculture has increased demand for irrigation, which can lead to depletion and degradation of reservoirs. This study compared observed streamflow records to TopNet-0 and TopNet-GW model outputs at points along the Mangatarere stream, a sub-catchment in the Wairarapa valley, New Zealand. Model performance was assessed using a suite of quantitative and qualitative comparisons. This analysis aimed to assess the similarities and differences between observed flow and the model outputs with respect to their model structures. Baseflow estimates from recursive digital filters were also compared at these sites to assess the groundwater representation of the models. The investigation can be considered representative of the wider Ruamahanga catchment, as the geology and hydrology in the region is relatively analogous.

Flow infilling and baseflow separation was undertaken at 13 Wairarapa flow gauges to provide considerations to the model outputs. Options investigated for flow infilling included a straight infill or calculation of the flow difference at each point. Potential multipliers included a long-term or a monthly option. The difference infill, coupled with the long-term multiplier, was found to be the optimum method. Independent baseflow estimates included a Q90/Q50 flow duration curve index and indices generated from the Eckhardt and Bump & Rise recursive digital filters. The two digital filters produced similar statistics but were found to employ uncertain parameters that significantly affect outputs.

TopNet-GW benefitted from up-to-date calibrations and as such produced generally excellent simulations in comparison to observed streamflow. With the addition of the deep groundwater conceptual reservoir in the structure of the model, simulated flow baseflow index estimates and graphical assessment of flow recession curves indicate TopNet-GW reproduces groundwater processes well despite potential over-representation of baseflow at the expense of high flow periods during peak flows. These findings highlight the importance of combining subsurface and surface flow dynamics to resolve water management issues and improve model performance at the catchment scale.

Acknowledgements

There are a number of people and organisations I would like to acknowledge and thank for their support and assistance in the completion of this research thesis:

Table of Contents

Table of Figures

Table of tables

Chapter 1
Introduction

Groundwater represents a significant portion of the global hydrological budget making up approximately 89% of the world's unfrozen freshwater (Younger, 2009). It is no surprise, then, that groundwater has been an important resource for human use and consumption for centuries. However, elevated anthropogenic abstraction can alter natural subsurface processes, rapidly changing the quantity and quality of groundwater bodies along with the surface water systems they interact with (Sophocleous, 2002). Estimating this level of aquifer-river interaction to validate predictive hydrological models is the main focus of this research.

In New Zealand, intensification of agricultural land use in areas with relatively low annual rainfall has resulted in an increased demand for freshwater supplied by groundwater aquifers. Consequently, issues relating to water quality, depletion and equitable access have arisen (Fenwick *et al.*, 2004). For shallow aquifers that connect with nearby surface water, the increasing abstractions can lead to the same unit of water being allocated more than once (double allocation). This is particularly important because over-allocation can result in significant stream flow reductions (depletion). Stream depletion leads to warmer water temperatures, increased nuisance plant and algae growth and decreased dissolved oxygen levels (Sophocleous *et al.*, 1995; Hunt, 1999). Perceiving both ground and surface waters as a single, interconnected flow system is now becoming the focus of many investigations (Winter, 1999; Jones and Gyopari, 2006; Yang *et al.*, 2017).

Regional councils are seeking better tools to manage resources with respect to the recharge and availability of groundwater and surface water reservoirs across entire catchments (Winter, 1998; Bunny *et al.*, 2014; Moreau and Bekele, 2015). Investigations at these broad scales need to be low-cost, efficient, and utilize readily available data (GNS, 2018). Catchment scale hydrological models have been identified as a means to generate accurate and relevant information under these conditions. Model investigations provide insights into factors such as flow variability, residence time, quality and quantity (Jones and Gyopari, 2006; Yang *et al.*, 2017; Zammit and Yang, 2017). In particular, conceptually based hydrological models generate outputs across large areas relatively quickly without the need for intensive field investigations

1

(Blöschl and Sivapalan, 1995). However, efforts must be made to ensure the model outputs are of desired accuracy and precision, whilst retaining computational efficiency. The interaction between ground and surface water is influenced by a range of meteorological conditions, geological formations and physiographic processes, making it difficult to quantify (Winter, 1999). Therefore, the representation of groundwater flow is inherently difficult to incorporate into catchment-scale modelling investigations (Brodie and Hostetler, 2005; Yang *et al.*, 2017). Digital baseflow separation techniques estimate the contribution of groundwater to a stream using frequency analysis on continuous streamflow data. They allow for an independent assessment of groundwater in regional hydrological systems and as such have potential as part of a validation process to ensure reliability in catchment scale hydrological model outputs (Biondi *et al.*, 2012; Collischonn and Fan, 2013). While model validation is relatively common, validation in the realm of low-flow scenarios appears to be scarce within the literature.

The Smart Models for Aquifer Management (SAM) research project is a GNS Science led programme with a primary focus of identifying optimal groundwater-surface water models that address large scale, specific environmental problems, using data at hand (GNS, 2018). In the context of the wider SAM project, this research looks to evaluate two catchment scale hydrological models, TopNet-0 and TopNet-GW, with specific regard to their representation of groundwater by using readily available data and simple comparative indices. In terms of performance comparisons for subsurface connectivity, the baseflow index represents the contribution of groundwater to a stream. This can be calculated from readily available streamflow data. Common baseflow investigation techniques attempt to separate baseflow from total flow through flow duration curves (e.g. Al-Faraj & Scholz, 2014; Collischon & Fan, 2013), graphical analysis (e.g. Brodie & Hostetler, 2005; Hewlitt & Hibbert, 1967), recursive digital filters (e.g. Eckhardt, 2005; Stewart, 2015), and/or chemical tracer analysis (e.g. Michel et al., 2015). There are often gaps in streamflow datasets however, and flow infilling from regional rainfall data or from nearby flow gauges can be required to complete the flow record (Harvey *et al.*, 2012). The main aim of this research was to use existing streamflow records as a resource at the model validation stage and assess the model's ability to match the model-generated discharge sequence with the

2

observed discharge time series in the Wairarapa valley, New Zealand. In order to achieve this aim, a number of specific research objectives were defined:

1. Establish a robust process for infilling flow gauges using highly correlated donor gauges because continuous datasets are required for the successful implementation of Objective 2

2. Identify an optimal method of quantifying groundwater contribution to streams by applying recursive digital filters to readily available continuous flow datasets.

3. Compare results from the second objective with two catchment-scale hydrologic models (TopNet-0 and TopNet-GW) at three sites along the Mangatarere River to assess their application for model validation purposes.

The investigations presented here focus on a selection of long-term flow datasets sourced from the Wairarapa valley of New Zealand. Many river systems throughout the valley are considered to display significant links to underlying groundwater aquifers (Jones and Baker, 2005). Land use in the Wairarapa valley is dominated by pastoral agriculture, which is placing significant pressure on groundwater and surface water reserves (Jones and Baker, 2005; Bunny *et al.*, 2014). Over-allocation of surface water and groundwater reservoirs high connectivity has the potential for significant deleterious environmental outcomes (Bunny *et al.*, 2014; Hughes and Gyopari, 2014). The extent of surface and groundwater interaction in the Wairarapa valley and the influence of abstractions from these systems is not well understood (Hughes and Gyopari, 2014). Adding to the growing body of hydrological understanding and supporting the development of robust model predictions will offer insights into future possibilities for the region. These contributions will eventually allow for comprehensive, catchment-scale water management strategies that could be applied nation-wide.

This research is divided into eight chapters. Chapter two introduces scientific concepts of hydrological flow systems drawn from relevant literature. Included is a presentation of the drivers that influence the contribution of groundwater to streams, along with an overview of baseflow analysis techniques. Chapter three provides a general overview

of hydrological models with respect to TopNet-0 and TopNet-GW, followed by a description of qualitative and quantitative dataset comparison techniques in the context of streamflow analysis and model validation. Chapter four presents an overview of the geological history, hydrogeology, climate, and land use of the Wairarapa valley. Chapter five describes the observed flow datasets used in this investigation before focusing on the various options for flow infilling. Chapter six delivers an account of the base-flow separation techniques investigated, beginning with a detailed description of the methodologies employed followed by an analysis and interpretation of the results. Chapter seven provides a description of the model validation exercise employed at three sites along the Mangatarere River. An initial presentation of the relevant model parameters and outputs is given alongside a description of the methodology, followed by an examination and discussion of the results. Finally, Chapter eight presents an overview of the main findings of this research along with a range of recommendations and avenues for future research.

Freshwater makes up just 2.5% of total global water resources. The major portion of that freshwater is held in glaciers and ice caps, while 30.1% is stored as groundwater (Figure 2.1) (Shiklomanov, 1993). This means groundwater is arguably the largest readily available freshwater resource for human development and consumption. Groundwater reservoirs interact with and are influenced by atmospheric and surface water processes (Freeze and Cherry, 1979; Sophocleous, 2002). These interactions occur at several temporal scales, from daily weather variations to seasonal and decadal climate fluctuations (Schwartz and Zhang, 2003).
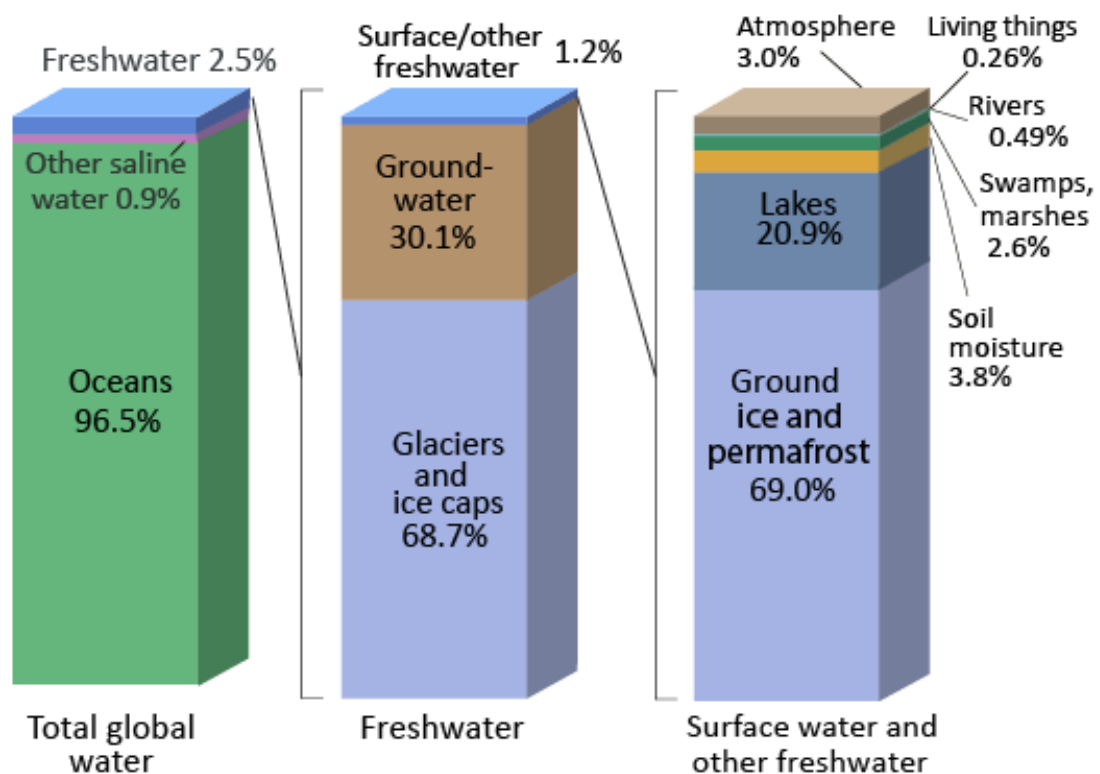
Figure 2.1. The world's water resources. Presented without change from Shiklomanov (1993).

This chapter presents a general overview of groundwater flow dynamics, followed by a description of the drivers and characteristics of groundwater flow systems with respect to atmospheric and surface water interactions. The final section of this chapter focuses on the flow from groundwater reservoirs to streams, commonly referred to as baseflow, and describes a selection of methods utilised to quantify this interaction.

2.1 Physical hydrogeology

Water is present throughout the subsurface in several principal zones (Winter, 1998). The shallowest of these is the soil-water zone, which has direct links to the surface (Figure 2.2). Here, water is held by surface tension of soil particles and is readily available to plants (Freeze and Cherry, 1979). Water accumulates in the soil-water zone up to a point known as field capacity, beyond which gravitational forces overcome surface tension and water propagates downward (Bear, 1979). Directly below the soil-water zone is the intermediate vadose zone. This zone contains both water and air within the pores between soil particles (Todd and Mays, 2005). The soil-water zone, intermediate vadose zone and the lower capillary fringe make up the total vadose zone, the area of unsaturated sub-surface material above the water table.
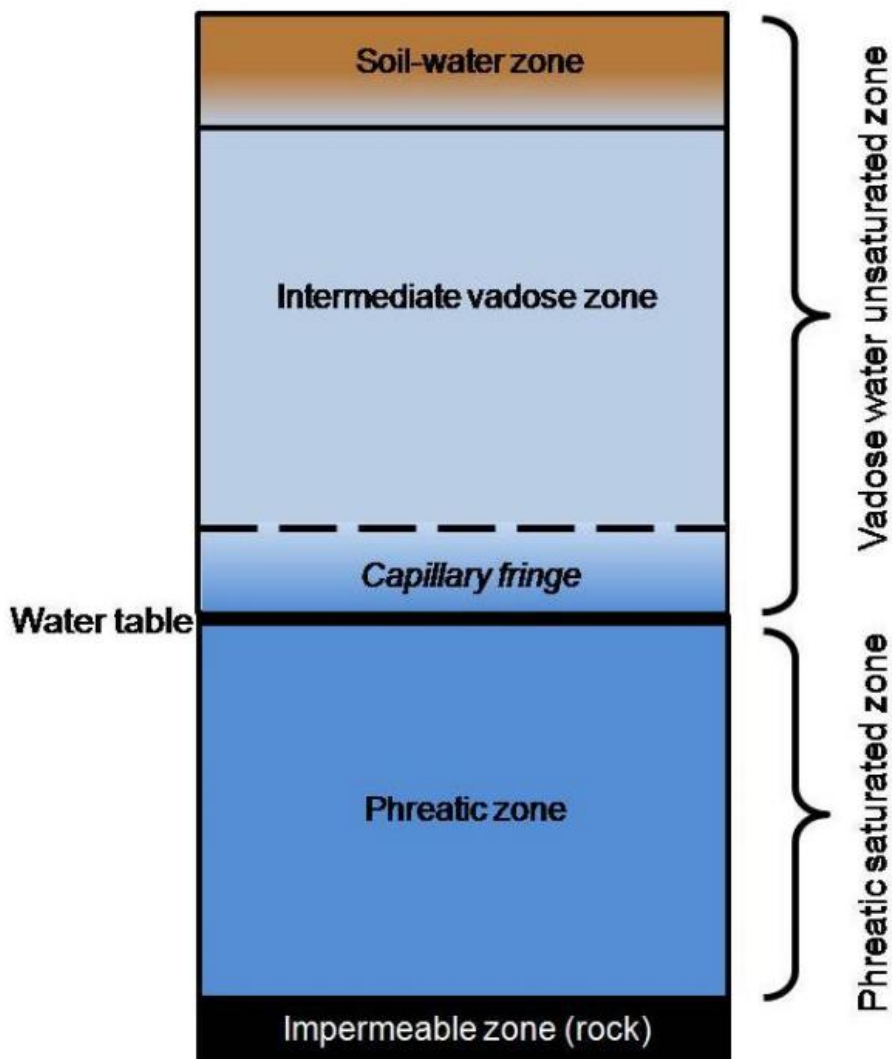
Figure 2.2. Schematic representation of subsurface water zones (Guggenmos, 2010).

The water table is generally defined as the plane at which the fluid pressure (p) is exactly atmospheric (1 bar) (Zhang *et al.*, 2013). This line of equal pressure marks the boundary between the vadose zone (unsaturated) and the phreatic (saturated) zone (Schwartz and Zhang, 2003). In the phreatic zone, pores are (theoretically) entirely filled with water. The connectivity and subsequent exchange of groundwater is driven by pressure gradients associated with recharge and discharge zones, as well as topographic variability (Sophocleous, 2002). The magnitude and velocity of this exchange is limited by the porosity of the sub-surface medium (rock, soil, etc.). The phreatic zone is the region below the water table within which groundwater flows (Freeze and Cherry, 1979).

2.2 Groundwater flow systems

Groundwater moves through porous media as a function of differences in hydraulic head (pressure) across a region. Groundwater movement can be expressed in the following equation, called Darcy's flow law (Todd and Mays, 2005).

$$q_G = -KA\frac{dh}{dl}$$ Eqn. 2.1

where q_G is groundwater flow in cumecs (m³/s), K is the hydraulic conductivity of the respective porous medium in metres per second (m/s), A is the cross sectional area of flow in square metres (m²), dh is the change in hydraulic head (h_1 - h_2) in metres and dl is the relative distance between the two measurements of hydraulic head in metres (m). The hydraulic conductivity (K) is the ease with which groundwater moves through a porous medium. Factors that influence K include particle size and arrangement of the medium (Todd and Mays, 2005). Under full saturation, permeable units like sand and gravel are associated with high values for K, while low K values are associated with low permeability materials such as clay (Schwartz and Zhang, 2003). General hydraulic conductivity values for some common physiographical mediums are provided by Schwartz and Zhang (2003) in Table 2.1.

Materials	Hydraulic Conductivity (m/s)
Gravel	$3 \times 10^{-4} – 3 \times 10^{-2}$
Coarse sand	$9 \times 10^{-7} – 3 \times 10^{-3}$
Fine sand	$2 \times 10^{-7} – 2 \times 10^{-5}$
Clay	$1 \times 10^{-11} – 4.7 \times 10^{-9}$
Sandstone	$1 \times 10^{-10} – 6 \times 10^{-6}$
Permeable basalt	$4 \times 10^{-7} – 2 \times 10^{-2}$
Fractured metamorphic rock	$9 \times 10^{-9} – 3 \times 10^{-4}$
Unfractured metamorphic rock	$3 \times 10^{-14} – 2 \times 10^{-10}$

The hydraulic gradient represents the energy available for groundwater to flow. It is a function of the difference in hydraulic head (d*h*) across a specific distance (d*l*) (Freeze and Cherry, 1979), where hydraulic head can be measured as the height of a water table relative to a common datum (e.g. sea level). Groundwater flow is directed along hydraulic gradients from areas of high hydraulic head to low hydraulic head (Figure 2.3) (Winter, 1999; Schwartz and Zhang, 2003).
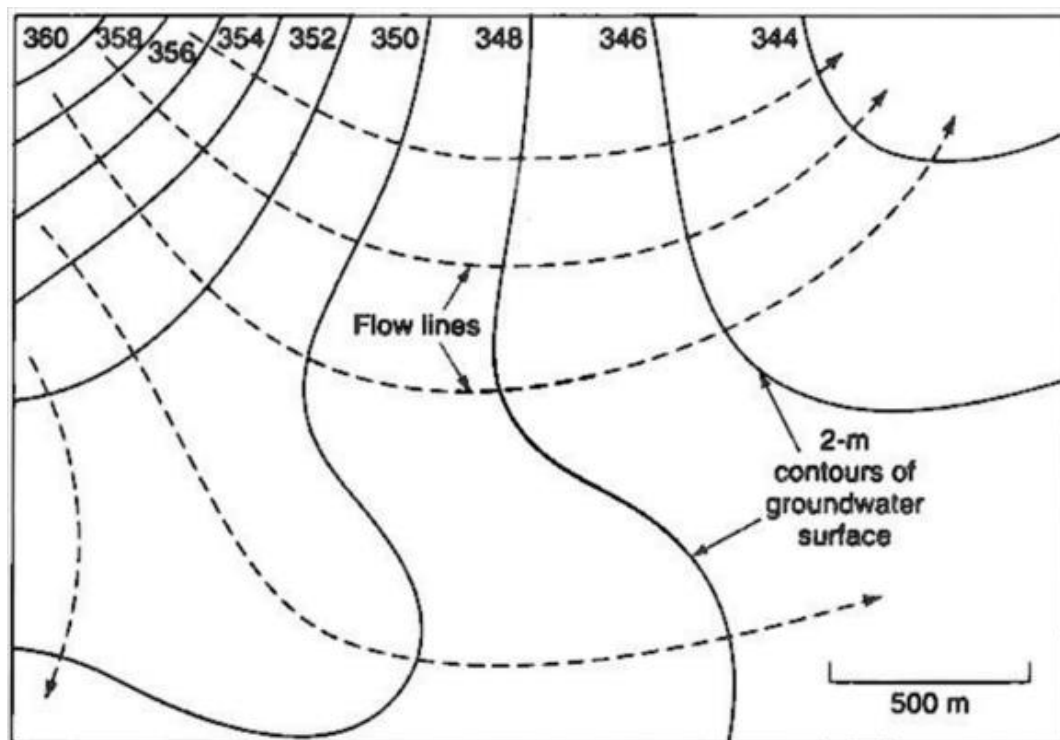
Figure 2.3. Groundwater flow lines are perpendicular to isopycnal lines, indicating direction of groundwater movement. Modified from Todd and May (2005).

2.2.1 Zones of recharge and discharge

Groundwater is replenished from the surface at recharge areas, while upwelling of groundwater to the surface occurs at discharge areas (Figure 2.4) (Sophocleous, 2002). Recharge areas are zones of high hydraulic head induced by precipitation, elevation or irrigation. At discharge areas groundwater flows towards low pressure zones and is lost from the groundwater system in the form of seeps, springs and streams (Schwartz and Zhang, 2003). The push out of recharge areas and the pull into discharge areas, due to the differences in hydraulic head/pressure, means groundwater flow is in perpetual motion (Freeze and Cherry, 1979; Schwartz and Zhang, 2003).
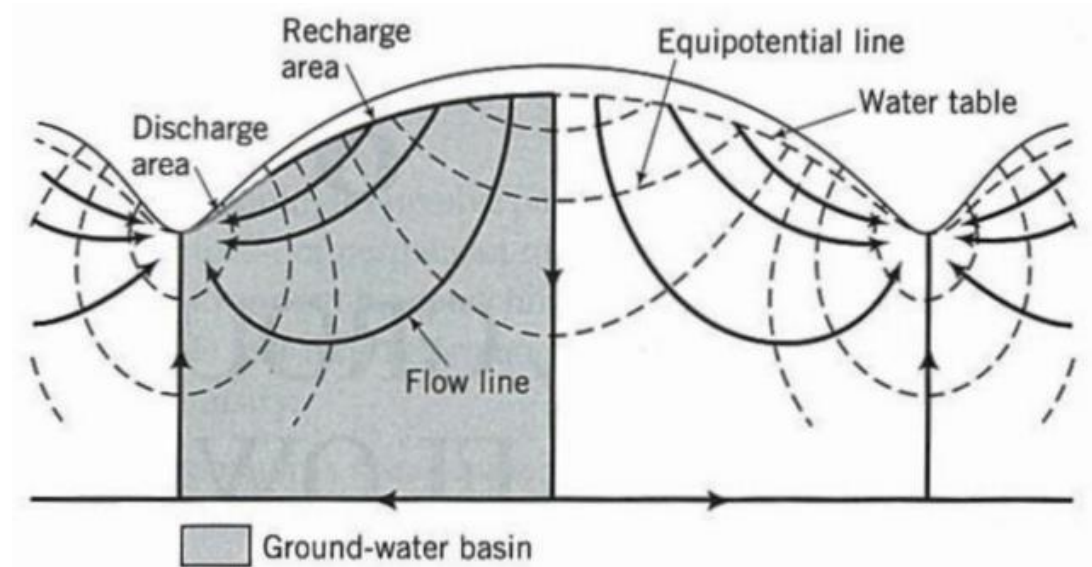
Figure 2.4. Simplified schematic representation of groundwater recharge and discharge areas. Presented without change from Schwartz and Zhang (2003).

2.2.2 Groundwater flow extent

The extent of groundwater flow systems can be differentiated at regional, intermediate, or local scales (Sophocleous, 2002). This categorisation is made with respect to the spatial distribution of recharge and discharge areas (Figure 2.5). At the largest scale, regional flow systems occur where recharge areas are present along major topographic highs and groundwater divides, with corresponding discharge areas being situated along major draining divides or in the lower reaches of a basin (Sophocleous, 2002). Regional groundwater systems cover large areas, sometimes entire surface catchments and discharge into major rivers, lakes or oceans (Sophocleous, 2002). Intermediate flow systems are present where one or more small topographic highs or lows separate

recharge and discharge zones. Lastly, local flow systems occur in areas where recharge and discharge points are immediately adjacent to each other with no significant topographic dividers. Processes occurring at the local scale are generally the most active of the three flow systems due to a high connectivity with the surface influences (Todd and Mays, 2005).
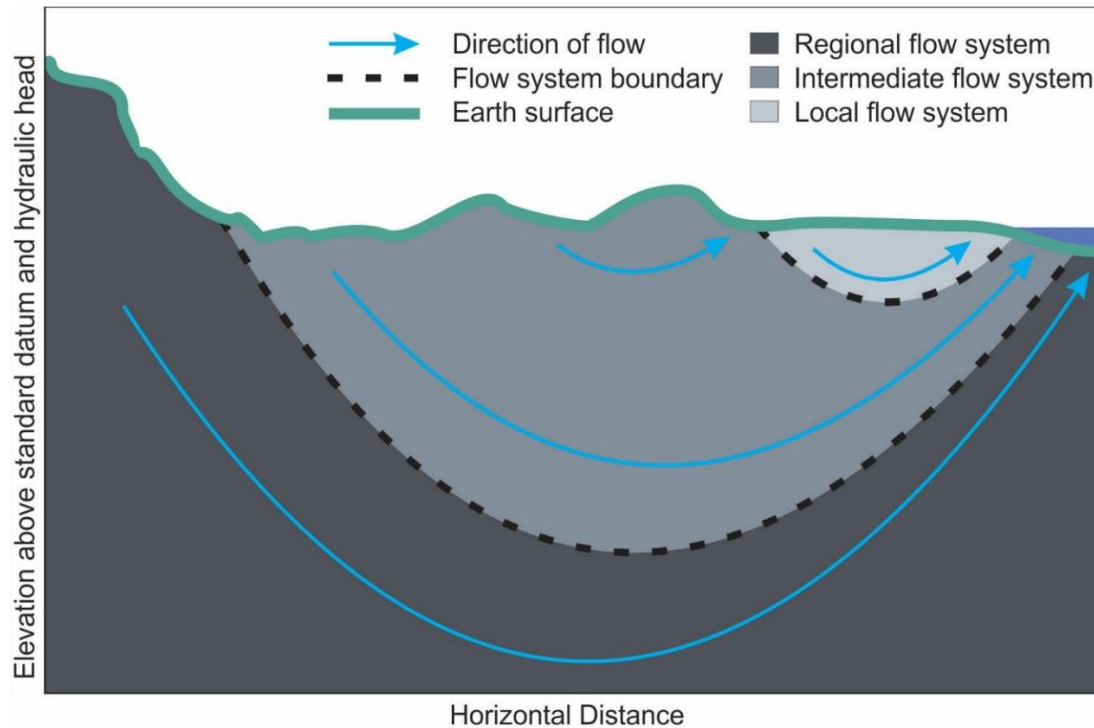
Figure 2.5 Schematic representation of regional, intermediate and local groundwater flow systems. Modified from Schwartz and Zhang (2003).

2.3 Factors influencing groundwater flow systems

2.3.1 Topographic and geological influences

The direction of groundwater flow generally mimics surface topography, from steep terrain toward topographic depressions that feed streams and rivers (Haitjema and Mitchell-Bruker, 2005). Topography also influences groundwater recharge zones due to the orographic effect on spatial rainfall patterns. Generally, catchments that begin at high elevations and terminate on low lying plains will have higher rainfall at the top of the catchment (recharge) and less rainfall in the lower reaches (discharge) (Yang *et al.*, 2017).

Groundwater flow in an aquifer is further controlled by the hydraulic conductivity of the soil and/or geology (Schwartz and Zhang, 2003). Rock or soil with a high hydraulic conductivity readily accepts water and forms the basis of groundwater aquifers (Figure 2.6). Conversely, layers of low hydraulic conductivity form impermeable layers, known as aquitards (Schwartz and Zhang, 2003). The vertical distribution of aquitards such as clay or metamorphic rock layers form the upper and lower boundaries of an aquifer. Confined aquifers are constrained above and below by impermeable layers of rock, and are recharged by up-catchment regional flow, whereas an unconfined aquifer occurs in areas where permeable material has no impermeable upper boundary. These unconfined aquifers interact directly with surface hydrology (Bear, 1979).
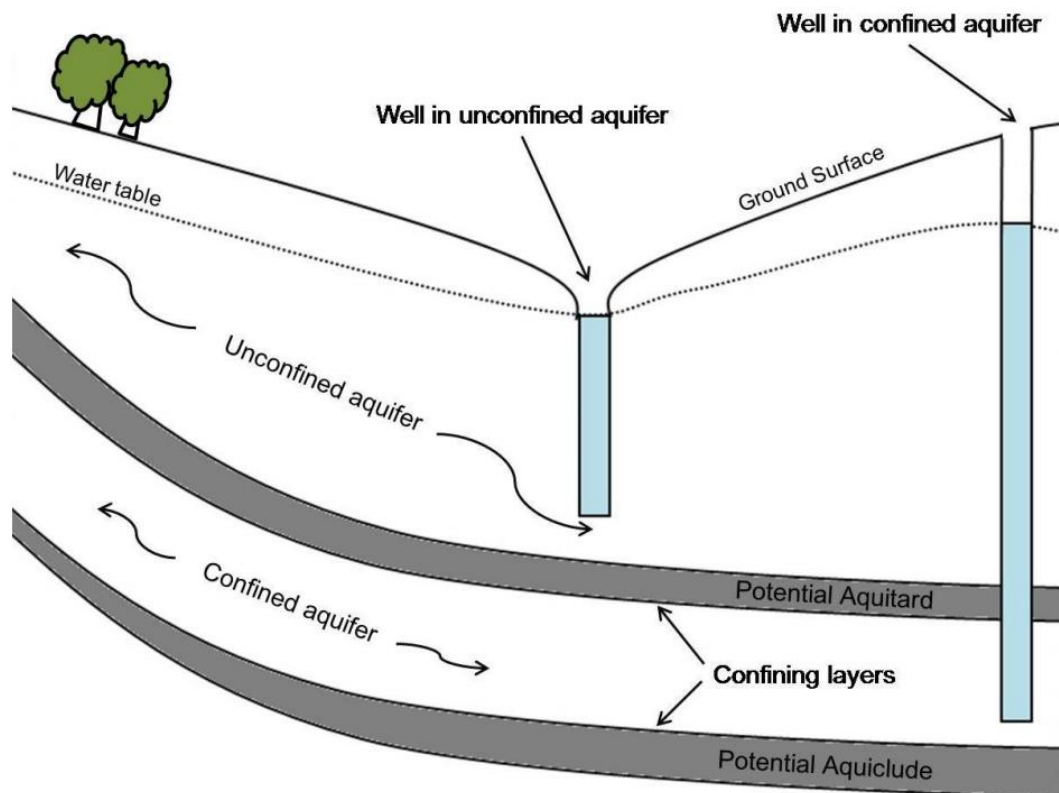
Figure 2.6. Schematic representation of a confined and unconfined aquifer (Guggenmos, 2010).

Unconfined groundwater aquifers interact with surface water across a variety of physiographic and climatic landscapes. In these areas, development, abstraction, or contamination of one reservoir can affect the other (Sophocleous, 2002).

2.3.2   Recharge mechanisms and variability

Recharge into groundwater systems is from infiltration to the phreatic zone across the surface and from river systems, both of which are fundamentally driven by precipitation (Schwartz and Zhang, 2003). High precipitation at the surface leads to elevated hydraulic pressure, an exceedance of field capacity and permeation of water downward to the phreatic zone (Freeze and Cherry, 1979). At discharge areas, where the water table intersects with the surface, groundwater flow can be accelerated by evaporation (Dahm *et al.*, 1998). This indicates a certain level of atmospheric control on groundwater flow systems. High rainfall generally results in more recharge, and significant evaporation – a function of sun, temperature and wind – will promote discharge (Schwartz and Zhang, 2003). If groundwater flow is so heavily influenced by climate, then it stands to reason that groundwater flow variability occurs at similar temporal scales. In fact, single rainfall events and seasonal patterns can alter regional, intermediate and local flow systems (Sophocleous, 2002). However, the response of groundwater flow to these drivers can be heavily attenuated by the hydraulic conductivity of the medium and topographic influences on hydraulic gradient profiles (Woessner, 2000).

2.3.3   Anthropogenic factors

Anthropogenic influences also have significant effects on groundwater flow conditions (Winter, 1998). Abstraction from an aquifer at a rate faster than recharge will generate a point of low pressure in the associated aquifer and flow will be redirected from its natural path towards the abstraction point (Sophocleous, 2002). This can lower the aquifer's water table and reverse the hydraulic gradient in an area. In reverse, irrigation to paddocks can result in excess water percolating into the groundwater, raising the water table above natural levels (Sophocleous, 2002).

2.4   Stream and aquifer interaction

Flow between an aquifer and surface waters can be categorised as either influent, effluent, or neutral (Sophocleous, 2002). As stated above, groundwater flow is dictated by hydraulic pressure differences (Schwartz and Zhang, 2003). The direction of transfer between the two reservoirs is ultimately dependent on the height of the stream stage (water level) relative to the adjacent groundwater table (Winter, 1998). Under influent conditions, a stream loses water to underlying groundwater systems (Figure

2.7a). Influent conditions occur when the stream stage is higher than the water table. Influent conditions can also occur when stream and groundwater systems are not directly connected hydraulically in what is known as a disconnected influent stream (Figure 2.7b). Groundwater systems that discharge water to surface water reservoirs are known as effluent systems (Figure 2.7c). Effluent conditions occur when the adjacent water table is at or higher than stream stage and a stream can switch from effluent to influent conditions throughout the course of one rainfall event (Winter, 1999).
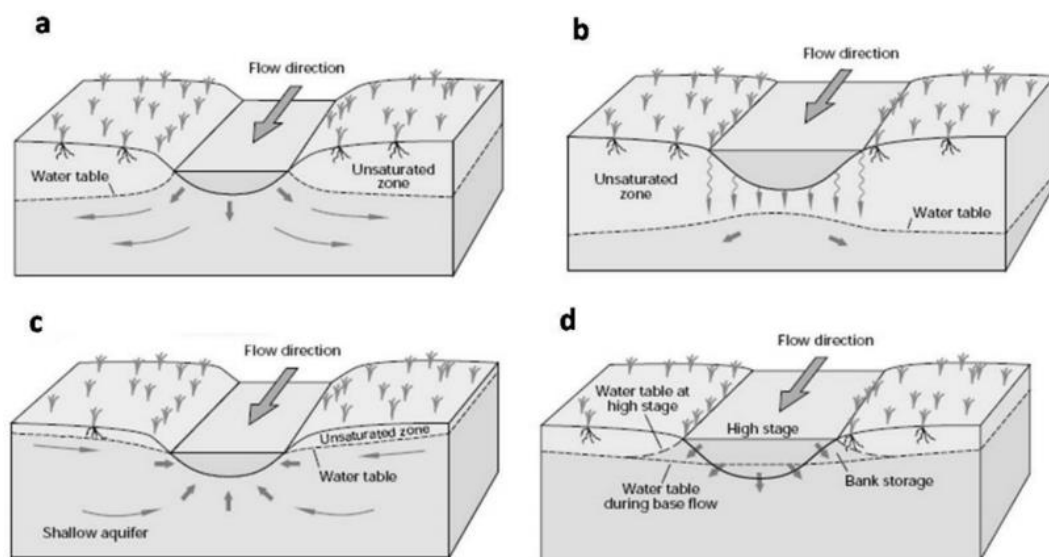
Figure 2.7. Schematic representation of (a) Influent (losing) stream, hydraulically connected to the groundwater system, (b) disconnected losing stream; (c) hydraulically connected effluent (gaining) stream system and (d) storage of excess water in neighbouring river banks. Presented without change from Winter et al. (1998).

High precipitation events lead to an overall excess of water in a catchment, raising both stream stage and groundwater table. This surplus water infiltrates stream banks as a form of temporary storage (Figure 2.7d) (Winter, 1999). Following the event peak, water stored in the banks will return to the river system or percolate into the aquifer below. This is an important buffering mechanism for flood events (Sophocleous, 2002). The volume of water stored in the bank depends on the duration and intensity of the precipitation event as well as the transmissivity and storage capacity of the banks and aquifer (Brunke and Gonser, 1997).

2.4.1 Baseflow

The contribution of water from an aquifer to a stream is commonly referred to as baseflow (Brodie and Hostetler, 2005). Generally, baseflow is understood to maintain the flow of a stream in between the high intensity, short duration inputs that occur due to rainfall events (Tallaksen, 1995). Actual measurement of baseflow is fundamentally impossible and the definition remains relatively arbitrary. Despite this, it is commonly estimated and used as an indicator for surface water-groundwater interaction (Stewart, 2015). It is only recently, with the help of chemical tracer analysis, that quantitative estimates of baseflow contribution to streams are being produced (Stewart, 2015). While chemical tracer separation provides many new insights, these investigations are expensive and uncommon in comparison to widely used streamflow observation based methods (Brodie and Hostetler, 2005). The concept of baseflow is best understood by examining a single-event stream hydrograph. A hydrograph is a time-series record of discharge at a gauging site. Total flow is a considered to be made up of different flow components. These components are defined by Brodie and Hostetler (2005) as:

(1) Quickflow – the direct response to a rainfall event including overland flow (runoff), lateral movement in the soil profile (interflow) and direct rainfall onto the stream surface (direct precipitation), and;

(2) Baseflow – the longer-term discharge derived from natural groundwater storages.

The relative contributions of quickflow and baseflow change through time on a hydrograph record, particularly during rainfall events (Brodie and Hostetler, 2005). The single event hydrograph shown below represents the theoretically understood response of both components to a single rainfall event (Figure 2.8). It can be divided into three main stages along the time axis:

(1) pre-event low-flow dominated by the baseflow component;

(2) increasing quickflow as a response to the rainfall event (rising limb), with a reduction of baseflow due to the rapid rise of the stream level relative to surrounding groundwater levels, generating influent flow, and

14

(3) recession of the quickflow component (falling limb), along with the delayed response of a rising baseflow component due to infiltrating rainfall (Brodie and Hostetler, 2005).

At some point along the falling limb, the quickflow component ceases and baseflow becomes dominant. Baseflow also recedes as natural storages are gradually drained following the rainfall event. While this is the current understanding, it was originally thought that baseflow was relatively constant and unresponsive to short term rainfall events (Brodie and Hostetler, 2005). It was only after modern tracer separation analysis was applied to streamflow by Wittenberg and Sivapalan (1999) that, even in flood periods (peak flows), outflow from groundwater was found to be a significant contributor to stream flow. Additionally, the analysis showed a significant response to immediate rainfall events that mimicked total flow peaks (Brodie and Hostetler, 2005).
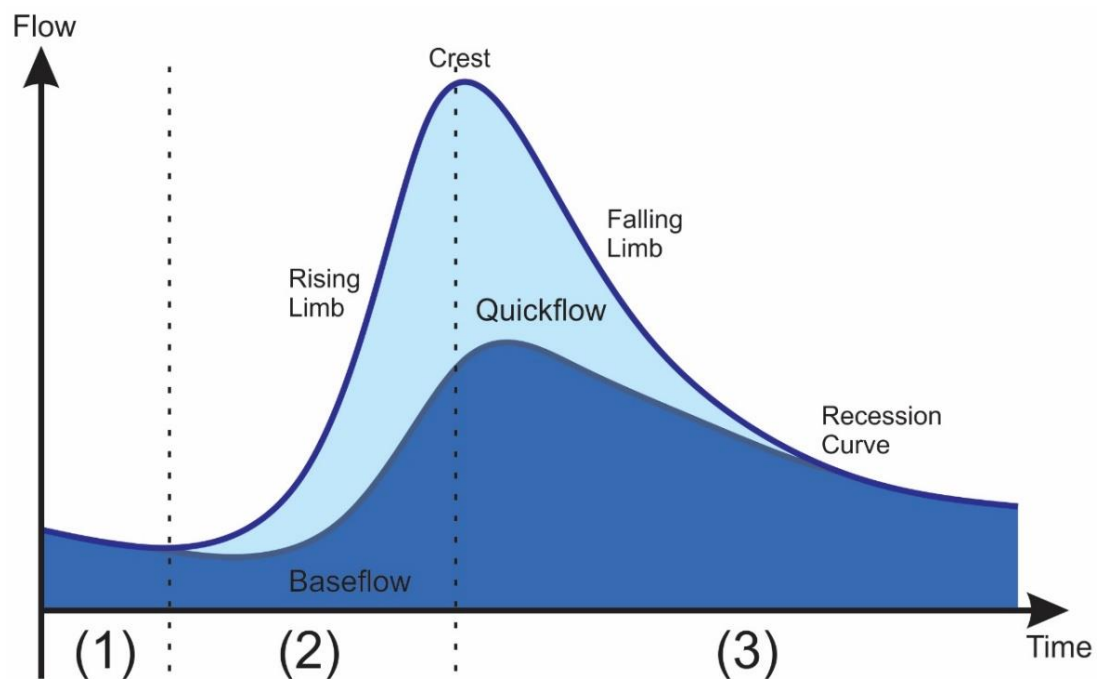
Figure 2.8 Baseflow and quickflow components of a storm hydrograph. Adapted from Brodie and Hostetler (2005)

Inferences of baseflow contribution to a stream have been used to inform on the natural storage properties of a catchment, residence times and more (Brodie and Hostetler, 2005). Groundwater discharge from shallow unconfined aquifers is commonly assumed to be the main portion of the baseflow component (Sophocleous, 2002). For this to occur, the aquifer must be sufficiently replenished (usually on a seasonal basis),

have a shallow water table that is, for the most part, higher than the stream stage, and have adequate water storage and transmission properties to maintain flow to the stream (Smakhtin, 2001a). For a stream where the associated aquifer satisfies these criteria, analysis of the baseflow component can indicate the magnitude and timing of groundwater contribution (Brodie and Hostetler, 2005).

It is important to note the assumption that baseflow equates to groundwater discharge is not always valid. Water can be released over different timeframes from alternative storages such as connected lakes or wetlands, snow-pack, or stream-banks (Schwartz and Zhang, 2003). The hydrographic record represents a net water balance, such that the numerical value for baseflow is influenced by any water losses from the stream such as direct evaporation, transpiration from riparian vegetation, or seepage into underlying aquifers (Kalbus *et al.*, 2006). Further, water use or management activities such as stream regulation, direct water extraction and nearby groundwater pumping can significantly alter the baseflow component (Sophocleous, 2002).

2.4.2 Baseflow analysis

The proportion of base flow in relation to the total flow of a stream at any point is known as the base flow index (BFI) (Brodie and Hostetler, 2005). The BFI is a useful index because it is a value that can be compared across streams and between catchments (Collischonn and Fan, 2013). Calculating the BFI is the focus of many groundwater flow investigations that aim to discern and quantify the interaction between streams and unconfined aquifers (Spongberg, 2000; Eckhardt, 2005; Collischonn and Fan, 2013; Zhang *et al.*, 2013; Al-Faraj and Scholz, 2014).

Base flow analysis methods generally fall within the following categories: tracer separation, frequency and recessional analysis methods, or empirical hydrograph separation (Hall, 1968; Tallaksen, 1995; Gonzales *et al.*, 2009; Stewart, 2015). This section will describe each of these categories in turn before briefly discussing the methods selected for this investigation. In depth descriptions of the selected methods are provided in the baseflow investigation chapter (Chapter six).

Tracer separation

Tracer separation methods use measurements of dissolved chemicals and/or stable isotopes to identify unique components relative to the mass balance of water (Stewart

et al., 2007). Waters from different sources are assumed to have unique chemical compositions (Pinder and Jones, 1969; Sklash and Farvolden, 1979; McDonnell *et al.*, 1990). By measuring the relative concentrations of chemicals that are known to be derived by groundwater, surface water or soil water, tracer methods allow for objective separation of flow components (Stewart, 2015). Evidence produced by tracer analysis has further been utilised for optimisation of empirical hydrograph separation techniques (Chapman and Maxwell, 1996; Chapman, 1999; Stewart, 2015). For example, Chapman and Maxwell (1996) used evidence from isotope tracer separation methods to infer a base flow component that was more responsive during significant rainfall events than previously established.

Analytical separation

Analytical methods, based on the fundamental theories of water flows and natural state equations, include frequency and recession analysis (Stewart, 2015). Frequency analysis derives relationships between the magnitude and the frequency of streamflow discharges, this includes using streamflow frequency data to generate flow duration curves (FDC) (Welderufael and Woyess, 2010).

Flow duration curves assess the frequency of flow volumes over the length of a dataset and provide catchment specific estimations for low-flow and baseflow metrics (Al-Faraj and Scholz, 2014). There are many indicator metrics that can be drawn from the flow duration curve such as the ratio of Q90 (flow value which is equalled or exceeded 90% of the time) to Q50 (median flow) is often used as a comparative estimate of baseflow contribution to a stream (Tallaksen, 1995; Collischonn and Fan, 2013).

Recession analysis focusses on the receding limb of the hydrograph following the stream peak. The timing of which presents insights into relative reservoirs and physical catchment characteristics (Stoelzle *et al.*, 2013). Recession analysis includes correlating or matching up multiple recession curves in a dataset in order to generate a master recession curve (Chapman, 1999). Equations can be applied to fit the recession segments to predict flow recession (Collischonn and Fan, 2013). Classic exponential decay functions, used to represent heat flow, diffusion, or radioactivity, are often applied. This assumes a linear relationship between storage and outflow which is contemporarily less applicable, so alternative functions are under investigation (Chapman, 1999; Stewart, 2015).

Empirical hydrograph separation

Brodie and Hostetler (2005) provide a useful review of empirical hydrograph separation methods. These are the most commonly used techniques due to the wide availability of streamflow data (Zhang *et al.*, 2013). Empirical hydrograph separation methods include manual graphing methods (Sloto and Crouse, 1996) and low frequency digital filtering of flow (Eckhardt, 2005; Zhang *et al.*, 2013). Both the frequency analysis and the digital filter methods work best when continuous flow datasets with no gaps is utilised (Harvey *et al.*, 2012).

Empirical separation methods are often thought to be rather subjective, however contemporary methods are becoming more objective with the aid of chemical and tracer separation investigations (Chapman and Maxwell, 1996; Stewart, 2015). The Eckhardt and Bump & Rise digital filters estimate baseflow at every point in the dataset before calculating a mean baseflow index (Eckhardt, 2005; Stewart, 2015). Both algorithms have been informed by quantitative chemical tracer analysis and can be considered more objective than traditional filter methods. The digital filter methods used in this investigation can be applied to readily available historical flow data, making them extremely cost-effective for hydrologic investigations.

This investigation will be utilizing empirical hydrograph separation methods, as well as some basic analytical methods to generate useful indices for intra-catchment comparisons. Techniques include calculating Q90/Q50 from flow duration curves (Collischonn and Fan, 2013), as well as baseflow estimation by the Eckhardt recursive digital filter (EK) (Eckhardt, 2005) and the Bump & Rise recursive digital filter (BR) (Stewart, 2015). These methods are common in the field of hydrological investigation and straightforward to apply as they only require a continuous record of flow.

The indices generated from base flow separation can be used as tools for the validation and evaluation of catchment scale hydrologic models. They are especially useful in this sense as the metrics are representative over large areas. Hydrograph separation methods can be considered useful for comparison against catchment scale models as they are efficient and use readily available data. Alternative field investigations and chemical/tracer analysis offer quantitative measurements for groundwater-surface water interactions, however they are often employed at small scales and expensive (Stewart, 2015), thus less suited to model validation at broad scales.

2.5 Summary

Groundwater flow in the saturated phreatic zone is a function of topography, geologic and soil permeability and climatic conditions. The numerous influencing factors and their inherent variability means the groundwater flow system in any region can be extremely complex across both time and space (Sophocleous, 2002). Generally, rock or soil with high permeability (hydraulic conductivity) are potential aquifers, whereas low permeability rock will yield low volumes of water and can even act as a flow barrier or aquitard (Freeze and Cherry, 1979). Unconfined aquifers that possess a high hydraulic conductivity are likely to exhibit strong connectivity with surface hydrologic processes, including streams (Schwartz and Zhang, 2003). Understanding this interaction, and the degree of connectivity is important for water resource management as abstraction from one such aquifer will likely have significant detrimental effects on the associated stream and *vice versa* (Sophocleous, 2002).

Baseflow is considered to be the contribution of groundwater to a stream or river (Brodie and Hostetler, 2005). While difficult to directly measure, it can be estimated as a ratio of total flow, known as the baseflow index (BFI). Values for BFI can be used to inform on sub-surface flow characteristics and subsequently as a comparison between different catchments (Collischonn and Fan, 2013). Although many methods for calculating BFI exist, empirical baseflow separation techniques are the most common because of the ease and accessibility of streamflow data (Stewart, 2015). However, these methods are rather arbitrary and often lack any physical basis (Szilagyi, 2004). Recent empirical methods for base flow separation have been informed by quantitative tracer and chemical separation studies and therefore can be considered more objective (Chapman and Maxwell, 1996; Stewart, 2015).

Chapter 3
Streamflow analysis and model validation

The basis of this research is a comparison of datasets to establish whether recorded discharge time series are reliably reproduced by the hydrological models. Catchment scale hydrologic models (also known as distributed models) are used as tools to assess the flow dynamics of entire catchments, however they possess their own inherent challenges (McMillan *et al.*, 2016). Hydrological models that are implemented at the catchment scale are often computationally intensive and require huge amounts of data in order to generate accurate outputs (Yang *et al.*, 2017). There is a need in New Zealand to generate model predictions that are both computationally efficient, as well as accurate enough to be useful for management decisions (GNS, 2018). In order for model outputs to be considered reliable, they must first go through a model validation stage (McMillan *et al.*, 2016). Traditional model validation processes employ a host of quantitative statistics that compare observed flow to model outputs over a validation period. Certain graphical comparisons are also drawn to qualitatively assess the modelled outputs similarity to observed flow (Moriasi *et al.*, 2007).

Evaluation of model performance typically involves a comparison between simulated outputs and a set of observations that were not used for model calibration (Biondi *et al.*, 2012). Thus, testing the ability of the model to objectively emulate real-world processes. The validation process helps to identify shortcomings in parameter estimation or model structure (McMillan *et al.*, 2016). Robust evaluation allows for model inter-comparisons and development, effectively optimising decision making at management levels (Beven and Young, 2013).

Section 3.1 introduces the models relevant to this investigation, TopNet-0 and TopNet-GW (Bandaragoda *et al.*, 2004; Yang *et al.*, 2017). This is followed by a description of standard model validation procedures and dataset comparison tools in section 3.2 and 3.3 (Moriasi *et al.*, 2007; McMillan *et al.*, 2016). This includes a consideration of the quality of datasets used for validation, along with descriptions of the qualitative and quantitative performance measures employed for comparing flow (Biondi *et al.*, 2012). This research uses these tools not only for comparisons between observed and modelled flow (Chapter 7), but also to compare observed flow datasets in the flow infilling investigation (Chapter 5).

3.1 Hydrological models

Hydrologic models integrate a series of inputs; climate data, digital elevation models (DEM), soil data, field observations and geological characteristics to produce a range of outputs that predict flow at varying spatial scales (Beven *et al.*, 1995). Catchment modelling of flow is useful for flood forecasting, contaminant tracing, investigating the long-term effects of management decisions on catchment hydrology and further defining the relative dominance of environmental drivers (Beven, 1989; Beven and Freer, 2001; Bandaragoda *et al.*, 2004; McMillan *et al.*, 2013; Yang *et al.*, 2017; Zammit and Yang, 2017).

Broad, New Zealand-centric hydrologic models must be able to capture the significant variability in climate, topography and geology present in many of the larger catchments (Yang *et al.*, 2017). The need for an all-encompassing model must be compromised to ensure computational efficiency and informative outputs (Robinson, 2004). As such, conceptual models are considered applicable to this type of modelling endeavour (Yang *et al.*, 2017). Conceptual models are less computationally intensive compared to physically based models. They generally consist of a series of reservoirs that are linked by different flow pathways. Transfer between the reservoirs is based on a conceptual understanding of the physical processes taking place (Yang *et al.*, 2017). When a driver such as rainfall is introduced the water is distributed into the various reservoirs depending on a set of boundary conditions and parameters (Bandaragoda *et al.*, 2004).

Lumped-parameter models, hereon described as conceptual models, evaluate the transfer between the different hydrological reservoirs (Yang *et al.*, 2017). This section describes the development of the TOPMODEL, TopNet-0 and TopNet-GW hydrological models, which are commonly used throughout New Zealand and around the world for water resource management and theoretical investigations (McMillan *et al.*, 2013; Yang *et al.*, 2017; Zammit and Yang, 2017).

3.1.1 *Conceptualisation*

TopNet-0 – commonly TopNet – is a conceptually based, semi-distributed, catchment-scale hydrological model (Bandaragoda *et al.*, 2004). It is a development of the original TOPMODEL, which reproduces the hydrological behaviour of sub-catchments using a set of conceptual tools (Beven and Kirkby, 1979). Catchment delineation is carried

out using a digital elevation model (DEM) and strahler order river network (Clark *et al.*, 2008). Sub-catchments are generally defined with respect to strahler order one streams. TopNet-0 then amalgamates the sub-catchment outputs from TOPMODEL using a kinematic wave channel routing algorithm (Goring, 1994). This effectively allows for larger catchments to be modelled using the combined outputs of smaller sub-catchments.

TopNet-0 incorporates soil moisture deficit, potential evapotranspiration, interception and soil zone components (Bandaragoda *et al.*, 2004). The physical processes considered at each sub-catchment are presented in Figure 3.1. For a detailed description of the components that make up TopNet, refer to Bandaragoda *et al.* (2004). Complete model equations are given in Clark *et al.* (2008).
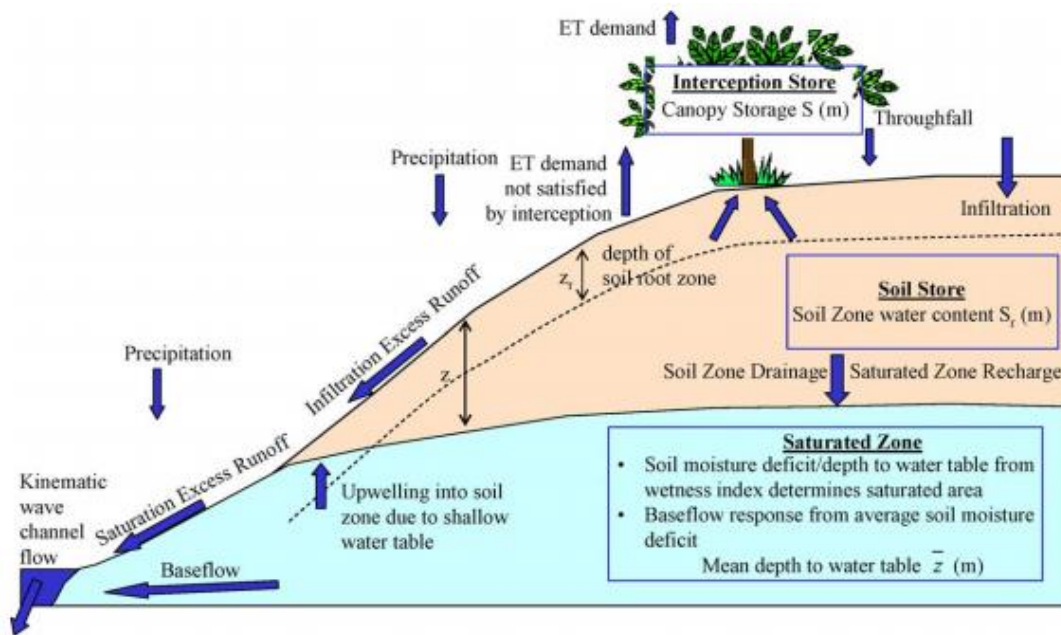
Figure 3.1. Schematic representation of the physical processes represented by the TopNet modeling system (Bandaragoda et al., 2004).

The water balance model relies heavily on catchment topography information to simulate surface flows, shallow groundwater flow and the interaction between the two reservoirs (Clark *et al.*, 2008). This reliance on topography means that in large flat areas like the Wairarapa valley the model begins to have difficulties in simulating the integrated surface and groundwater system (Yang *et al.*, 2017). In these flat areas water management decisions are crucial for sustainable agriculture and significant exchanges

22

between surface water and groundwater are likely to take place. TopNet-GW is an adaptation of the TopNet-0 model which includes an additional conceptual groundwater layer underlying the original TopNet-0 conceptual stores (Yang *et al.*, 2017). This is represented in Figure 3.2 by the red box. The additional groundwater component (S_D) allows for exchange between surface and subsurface reservoirs, as well as subsurface (aquifer) flow between adjacent catchments (Yang *et al.*, 2017). Calculation of the flux between shallow aquifer (S_S), river (S_R) and the added groundwater layer (S_D) is described in Yang *et al.* (2017). The various flow pathways between the three reservoirs in TopNet-GW are presented in Figure 3.2 and described in Table 3.1.
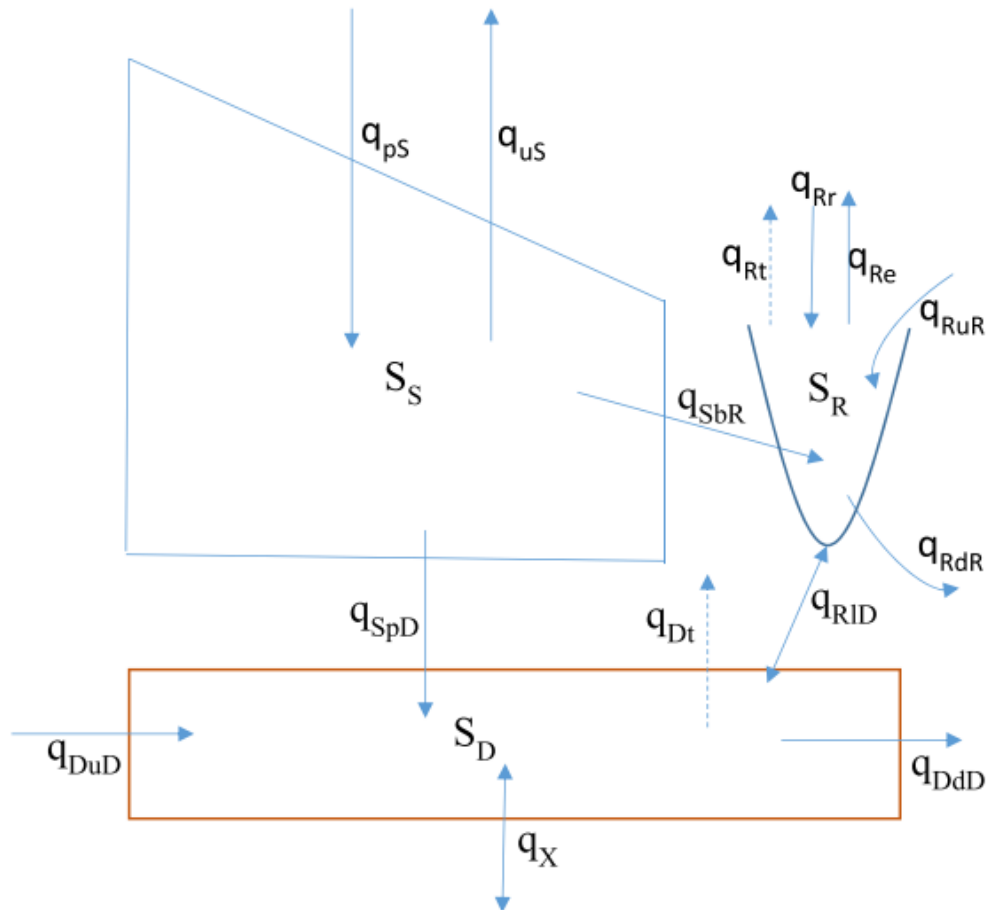
Figure 3.2. Schematic representation of surface water and groundwater interaction of one sub-catchment in the TopNet-GW model. Boxes S_S, S_R and S_D are shallow aquifer store, river store and groundwater store respectively. Solid arrows denote fluxes between the stores and dashed arrows are water takes. See Table 3.1 for full description of parameters. The original TopNet-0 only had S_S and S_R stores (Yang et al., 2017).

Table 3.1. State variable and fluxes among shallow aquifer, river, and groundwater in TopNet-GW (Yang et al., 2017).

Name	Meaning	Name	Meaning
S_S	Shallow aquifer storage (L)	q_{RID}	Flux between river and groundwater (L/T)
S_R	River storage (L)	q_{Rr}	Precipitation falling in the river (L/T)
S_D	Groundwater storage (L)	q_{Re}	Evaporation from the river (L/T)
q_{Su}	Upward movement from groundwater to soil zone (L/T)	q_{Rt}	Water take from the river (L/T)
q_{pS}	Percolation from soil zone into shallow aquifer (L/T)	q_{Dt}	Groundwater take (L/T)
q_{SbR}	Shallow aquifer outflow q_S to the river (L/T)	q_{DuD}	Groundwater flow from upper groundwater store (L/T)
q_{SpD}	Shallow aquifer outflow q_S to the groundwater (L/T)	q_{DdD}	Groundwater flow to the next groundwater store (L/T)
q_{RuR}	Flux from upstream to river (L/T)	q_x	Intercatchment groundwater flow (L/T)
q_{RdR}	Flux from river to downstream (L/T)		

3.1.2 *Parametrisation*

Parameters describe the unchanging properties of the sub-catchments or model elements (Bandaragoda *et al.*, 2004). There are 31 parameters used in the original TopNet-0 model that represent the physical properties of a sub-catchment (Clark *et al.*, 2008). This includes factors such as soil properties, topography, land cover and other characteristics that are considered time invariant (Bandaragoda *et al.*, 2004). Information for parameter estimates is provided by national datasets on catchment topography (i.e., LINZ digital elevation model), physical (Land Cover Database version 3-LCDB3, Land Resource Inventory, Newsome et al., 2012), soil (Fundamental Soil Layer – FSL, Wilson & Giltrap) and hydrological properties (River Environment Classification, Snelder & Biggs, 2002). The detailed derivation of such parameters from these nationally available datasets can be found in Clark *et al.* (2008). The new TopNet-GW model adds four additional parameters (f_s, q_{RID}, k_D and q_x) to account for the additional groundwater store (S_D) (Yang *et al.*, 2017).

TopNet-0 and TopNet-GW are driven by time series of precipitation and temperature data (NIWA, 2018). Outputs include a time series of modelled river flow throughout the modelled river network, as well as evaporation. TopNet also produces a series of diagnostic outputs for each sub-catchment, this includes estimates of baseflow.

24

3.2 Validation of hydrologic models

A review by Biondi *et al.* (2012), asserts the need for a universally applicable set of validation protocols due to the large number of existing hydrological models with varied structures. There is a growing discrepancy between the complexity of modelling tools and the capacity of modellers to assess the reliability of that model (Hug *et al.*, 2009). This issue is compounded by a general lack of sufficiently informative data (Biondi *et al.*, 2012). For example, measured variables are often point values, whereas modelled variables are often averaged in time and/or space (Moriasi *et al.*, 2007). Furthermore, measured variables like observed streamflow possess their own uncertainties with respect to monitoring technology, assumptions and approaches (McMillan *et al.*, 2016). A conventional model validation process is described and discussed with respect to comparative performance metrics, considerations of uncertainty in comparison datasets and clarification of model validation terminology.

3.2.1 Performance Metrics

Model performance can be addressed by a suite of qualitative and quantitative tools (Biondi *et al.*, 2012). Qualitative model validation uses graphical comparisons between observed and simulated data. This includes standard hydrographs, scatter plots, residual plots, flow duration curves and cumulative flow graphs to infer similarities and differences at various scales and flow volumes (Yang *et al.*, 2017; Zammit and Yang, 2017). Quantitative model validation is based on numerical performance metrics. There are many dataset comparison metrics used in the field of hydrology such as the Nash-Sutcliffe Efficiency Index (NSE), Percent Bias Measure (PBIAS), and RMSE-Observations Standard Deviation Ratio (RSR) (Moriasi *et al.*, 2007; Biondi *et al.*, 2012). Each index is a unique representation of how well the modelled data fits the observed data. Qualitative and quantitative comparison approaches should be considered as fundamental tools for model validation and used in a complementary fashion since they are each able to capture unique aspects of model performance (Biondi *et al.*, 2012).

3.2.2 Measured data uncertainty

An important consideration of assessing modelled flow against observed flow is the uncertainty of the measured dataset. This can arise from differences in measurement conditions and techniques across watersheds (Moriasi *et al.*, 2007). Léonard *et al.*

(2000), Schmidt (2002) and Herschy (2002) demonstrated errors in streamflow measurements of approximately 6% of the flow value measured by the current meter. Further, upon a comprehensive review of the literature, Pelletier (1988) found the uncertainty of discharge measurements might be as high as 20% of the observed value.

At the validation stage, the quality of the independent dataset naturally comes under scrutiny. It is important to ensure that the model outputs are being validated against unbiased and precise data (McMillan *et al.*, 2010). Therefore, a robust understanding of how flow data is produced is crucial. The most common method for recording continuous discharge (flow) is by measurement of the stream stage (height), which is then converted to discharge using a flow rating curve (Rantz, 1982; Schmidt, 2002). The rating curve is a model of the stage-discharge relationship at that particular gauging site. The rating curve is developed by fitting a curve through a set of concurrent stage and flow measurements on a graph (Tomkins, 2014). Robust streamflow datasets are maintained by taking regular measurements of the stage-discharge relationship, especially after significant changes to river profiles (McMillan and Westerberg, 2015). In terms of model validation, Moriasi *et al.* (2007) states that modelled streamflow can be characterised as "good" if it is within 10% to 15% of measured streamflow data of typical quality.

3.2.3   Terminology

It is important to note; the correct use of terminology is under discussion within the literature. Disagreement is drawn from the argument that models cannot be wholly verified or validated due to the fact that they are and always will be approximations of the real system (Beven and Young, 2013). In other words, open-system models such as hydrological models cannot be confirmed with absolute certainty, only falsified (Refsgaard and Henriksen, 2004). While it is outside the scope of this literature review, discussions on this topic include Refsgaard and Henriksen (2004) and Beven and Young (2013). Rykiel (1996) asserts that *validation* means to test whether a model is acceptable for its intended use. According to Beven and Young (2013), this process is *conditional validation*, provided the conditions of the model evaluation are set out and explained *a priori* to the analysis taking place. For the purposes of this research, 'validation' shall be defined as the ability of a model to match an independent data set.

3.3 Dataset comparison tools

For the purposes of dataset comparisons and to produce robust and reproducible results, a suite of graphical and statistical tools have been identified for use (Moriasi *et al.*, 2007; Biondi *et al.*, 2012). The underlying principles of dataset comparison techniques mean the following methods can be applied to assess the similarities of any two datasets (Harvey *et al.*, 2012). This section presents the performance metrics, quantitative and qualitative, used as dataset comparison tools for the flow infilling and model validation investigations in Chapters 5 and 7.

3.3.1 Qualitative analysis

Graphing techniques allow for subjective and qualitative evaluation of two datasets. According to Biondi *et al.* (2012), visual inspection represents a fundamental step in dataset comparisons, despite many goodness of fit metrics existing. Graphical inspection allows for the study of temporal differences and facilitates the identification of patterns in error occurrence (Biondi *et al.*, 2012).

Standard visual inspection of hydrographs can identify bias and differences in timing and magnitude of peak flows and the shape of recession curves. However they become less applicable when datasets are long (Moriasi *et al.*, 2007). Other common graphical comparisons include scatterplots, flow duration curves and flow frequency distributions. Scatter plots provide an objective reference given in the form of a line of best fit (Biondi *et al.*, 2012). Flow duration curves (FDC), along with flow histograms, can illustrate how well the datasets reproduce the frequency of measured flows (Zammit and Yang, 2017). General visual agreement between flow frequency distributions indicates adequate simulation over the range of the conditions examined (Biondi *et al.*, 2012).

3.3.2 Quantitative analysis

Numerical performance metrics provide quantitative and comparable estimates of dataset relationships. The three quantitative statistics described here are considered among a suite of similar statistical metrics in reviews by Moriasi *et al.* (2007) and Biondi *et al.* (2012) to be the most applicable and most widely used metrics for dataset comparison and model validation. Quantitative performance criteria and their acceptable bounds must be established and outlined before the validation process can be carried out to ensure conditional validation as defined by Refsgaard (1997).

The Nash-Sutcliffe efficiency index is a normalised statistic that examines the relative magnitude of the residual variance (noise) compared to the measured data variance (information) (Nash and Sutcliffe, 1970). NSE indicates how well the two datasets fit the 1:1 line on a scatter plot (Moriasi *et al.*, 2007). NSE is computed as:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Y_i^{obs} - Y_i^{sim})^2}{\sum_{i=1}^{n}(Y_i^{obs} - Y_i^{mean})^2} \qquad (3.1)$$

where, for a given constituent being evaluated, Y_i^{obs} is the *i*th observation, Y_i^{sim} is the *i*th simulated value, Y_i^{mean} is the mean of the observed data and *n* is the total number of observations.

Regarding performance criteria, values for NSE are between -∞ and 1.0, the optimum value being NSE = 1. Values ≤ 0 indicate that the mean observed value is a better predictor than the simulated value, which points to model failure.

The use of NSE as a quantitative statistic for streamflow comparisons is justified as it is recommended by numerous review papers (Moriasi *et al.*, 2007). Historically, it is a commonly used hydrological metric for reflecting the overall fit of a hydrograph (Sevat & Dezetter, 1991). However, according to McCuen *et al.* (2006), the NSE index has a tendency to be influenced by the extremes or outliers of a dataset. This is a significant caveat for investigations that are focussed on capturing low-flow dynamics in a dataset with large variations in flow. Corrections have been made in an investigation by Zammit and Yang (2017), in which NSE calculations were carried out on flow data that had been transformed into the log domain to better capture the low-flow dynamics of the model outputs.

RMSE-observations standard deviation ratio (RSR):

Singh *et al.* (2004) have suggested a comprehensive statistic to qualify what is considered a low RMSE by looking at standard deviation of observations while also incorporating an error index. The statistic, RMSE-observations standard deviation ratio (RSR), is calculated as the ratio of the RMSE and the standard deviation of measured data:

$$RSR = \frac{RMSE}{STDEV_{obs}} = \frac{\sqrt{\sum_{i=1}^{n}(Y_i^{obs} - Y_i^{sim})^2}}{\sqrt{\sum_{i=1}^{n}(Y_i^{obs} - Y_i^{mean})^2}} \qquad (3.3)$$

RSR incorporates the benefits of error index statistics and includes a scaling/normalisation factor, so that the resulting statistic and reported values can be compared between sites of varying flow magnitudes (Moriasi *et al.*, 2007). RSR varies from the optimal value of 0 to a large positive value. The lower the RSR, the lower the RMSE, and the better the fit of the two datasets. According to Moriasi *et al.* (2007), an RSR value ≤ 0.70 can be considered a threshold for satisfactory performance. It is important to note that the RSR is structurally identical to the NSE calculation, and as such the NSE can be considered the foremost statistic and will be predominantly referred to.

Percent Bias (PBIAS):

Percent Bias measures the average tendency of a dataset to be higher or lower than the other (Gupta et al., 1999). Computation of PBIAS is shown in the following equation:

$$PBIAS = \frac{\sum_{i=1}^{n}(Y_i^{obs} - Y_i^{sim}) * (100)}{\sum_{i=1}^{n}(Y_i^{obs})} \qquad (3.2)$$

where PBIAS represents the positive or negative deviance of the two datasets being evaluated as a percentage. The optimal value of PBIAS is zero and can propagate in a positive or negative direction (Moriasi *et al.*, 2007). Positive values indicate underestimation, whereas negative values indicate overestimation (Gupta et al., 1999). Low-magnitude values indicate similar results across the datasets. Moriasi *et al.* (2007) suggests satisfactory performance when the value for PBIAS is ±25% for daily streamflow data.

PBIAS is favourable because it has the ability to clearly indicate poor performance (Gupta et al., 1999). However, PBIAS values for streamflow tend to vary more during dry years than during wet years. A fact that should be considered when attempting to do a split-sample (summer vs winter) evaluation (Moriasi *et al.*, 2007).

3.4 Gaps in the literature and research justification

Model validation is an important aspect of the model development process. When low-flow dynamics are specifically under investigation, it is common practice to compare modelled and observed flow after they have been transformed into the log domain (Zammit and Yang, 2017). This enables more weight to be placed on the low-flow dynamics as compared to the high flow extremes. However, besides this small adjustment to the traditional validation process, low-flow model validation is rare in the current literature. As such, this research investigates the value of independent baseflow separation estimates to traditional model validation techniques that may have specific interest in low-flow or groundwater-surface water interaction dynamics.

Robust analysis of streamflow and baseflow allows for comparisons across different catchments and helps to inform catchment scale water management decisions. Observed total streamflow data is commonly used to calibrate and validate hydrological models and continuous flow records are readily available across New Zealand (Moriasi *et al.*, 2007; McMillan *et al.*, 2016; Zammit and Yang, 2017). However estimates of baseflow are more complex to produce and rarely used as validation tools. As management decisions are increasingly considering the interaction between surface water and groundwater, comparative estimates of baseflow have potential to inform hydrological models in conjunction with standard model validation procedures using total streamflow datasets.

This research coincides with current efforts to increase groundwater representation in catchment-scale hydrological models. TopNet-GW is one such model where a focussed groundwater component has been built into the original TopNet-0 model (Yang *et al.*, 2017). It aims to produce model outputs while considering groundwater and surface water reservoirs as a fully integrated hydrological system. TopNet-GW has significant implications for holistic water management applications in New Zealand.

In this investigation, total stream flow as well as multiple baseflow estimates are used to validate two catchment scale hydrologic models; TopNet-0 and TopNet-GW. The overall aim is to assess the efficacy of the outputs produced by the two models as well as using the baseflow analysis to inform on general groundwater flow characteristics in the Wairarapa region to further improve groundwater modelling techniques.

A clear understanding of the geological history, climate processes and makeup of the Wairarapa valley is necessary in order to understand the flow dynamics and evaluate the efficacy of hydrological models in the region. Geologic and geomorphic processes define the boundary conditions for groundwater flow, as well as points of interaction with surface flow. Eroded material has a significant influence on the distribution of permeable and impermeable layers. Variations in climate and rainfall impart a control on groundwater flow. Human induced land-use changes alter the natural flow conditions within a catchment, usually reducing the connectivity between surface and groundwater flow systems. Chapter 4 describes the geomorphic history and climatic conditions of the Wairarapa valley. A description of more recent anthropogenic factors, including water use and restrictions, is also presented. The region has been intensively studied and much of the following description is based off investigations by Begg *et al.* (2005), Jones and Gyopari (2006), Guggenmos (2010), Jones and Baker (2005) and Zammit and Yang (2017), along with publications and reports sourced from GWRC and NIWA websites, with a focus on groundwater – surface water dynamics.

4.1 Geological History

The Wairarapa valley is a structural depression that rests overtop a subduction interface between the Australian and Pacific plates (McConchie, 2000). Active plate tectonism, initiated approximately 280-150 million years before present (MYBP), has transferred stresses to the Earth's surface through a range of active faults, folds and uplift blocks (Morgan and Hughes, 2001). The valley is confined by a sequence of axial formations; on the western side by the Tararua and Remutaka Ranges, and to the south-east by the Aorangi ranges (Figure 4.1). The ranges are generally heavily vegetated and composed largely of Triassic-Jurassic greywackes. The highest relief and elevation in the region is in the west, up to 1500m (Kamp, 1992). The eastern ranges consist of Early-Pleistocene/Late-Tertiary marine deposited sediments, rising to 800m in some areas. They are the result of compression, faulting and uplift offshore approximately 13-6 MYBP (McConchie, 2000).
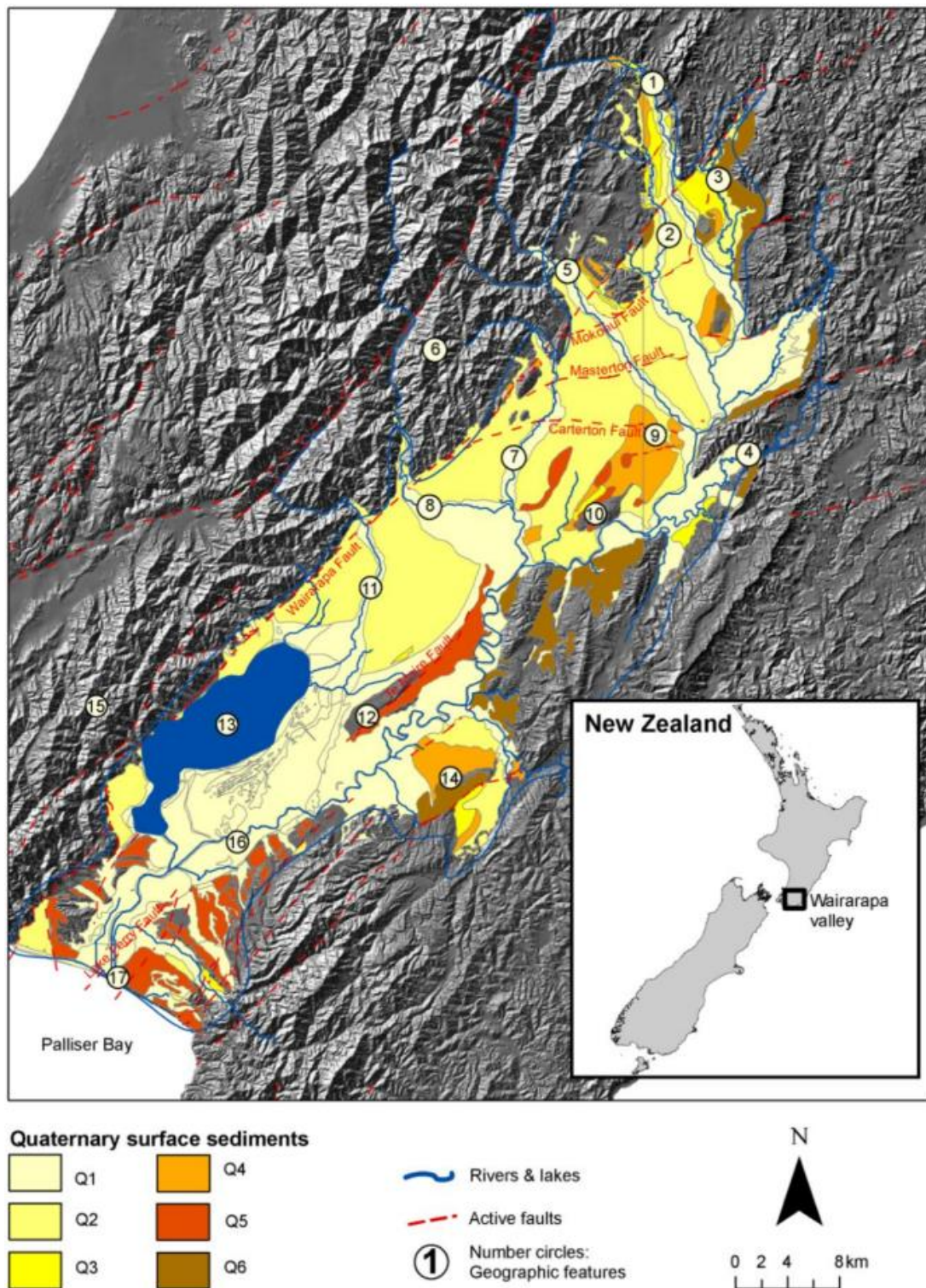
Figure 4.1. Location and geological map of the Wairarapa valley, New Zealand showing Quaternary surface sediments, active fault systems, major river and water bodies and a number of geographic features. Refer to Table 3.1 for Quaternary surface sediment ages. Circled numbers indicate major geographic features: 1) Ruamahanga River, 2) Waipoua River, 3) Whangaehu River, 4) Taueru River, 5) Waingawa River, 6) Tararua Ranges, 7) Mangatarere River, 8) Waiohine River, 9) Fernhill, 10) Tiffen Hill, 11) Tauherenikau River, 12) Te Maire Ridge, 13) Lake Wairarapa, 14) Martinborough Terrace, 15) Remutaka Ranges, 16) Lower Ruamahanga River, 17) Lake Onoke (McConchie, 2000; Guggenmos, 2010).

The floor of the valley itself is predominantly infilled with Quaternary (< 2 MYBP) aged sediments from the Tararua and Remutaka Ranges (Kamp, 1992). There is a varied distribution of sediment that has a significant impact on the dispersal and permeability of aquifer bodies within the region (Guggenmos, 2010). Relatively sparse vegetation and steep slopes throughout the Quaternary induced high levels of physical weathering and subsequent high sediment loads within rivers (Kamp, 1992). The sediment was transported by fluvial systems out of the ranges and deposited in the valley in a sequence of fluvial fans (Q2 in Figure 4.1 and Table 4.1). A succession of cold and warm climate erosional processes took place, resulting in fluvial fans that are a mix of poorly sorted gravel, sand and silt deposits. The larger river systems (Tauherenikau, Waipoua and Ruamahanga) produced fans that stretched all the way across to the eastern edge of the valley (Kamp, 1992; Jones and Gyopari, 2006). Simultaneously, well-sorted, fine sediments from the eastern marine derived ranges were predominantly deposited on the eastern margins of the valley.

Table 4.1. Timescale (Stage, epoch, period and age) of common Quaternary surface sediments (oxygen isotope stages) from the Wairarapa Valley, New Zealand. Adapted from McConchie (2000) and Guggenmos (2010).

Geological Units	Stage	Epoch	Period	Age
Q1				
		Holocene		
Q2				10ka
Q3	- Last Glacial Max -			
Q4			Quaternary	
Q5				80ka
		Pleistocene		
Q5	- Last Interglacial -			100ka
Q5				
Q6				130ka

Warmer inter-glacial temperatures during the Holocene led to an intensified hydrological regime and subsequent reworking of the Quaternary fan deposits (Guggenmos, 2010). Revegetation of the ranges, along with higher rainfall acted to reduce the relative sediment supply to fluvial systems thus increasing their erosive and

entrainment ability. Subsequent incision, reworking and redistribution of the Quaternary layers occurred (Q1 in Figure 4.1 and Table 4.1) (Kamp, 1992). Meanwhile, significant marine, estuarine and lacustrine layers were deposited in the southern parts of the valley near Palliser Bay during the mid-Holocene. This was a function of cyclical climate related sea level fluctuations and a general trend of subsidence in the lower portion of the Wairarapa valley (Begg *et al.*, 2005).

4.2 Hydrogeology

The confining base of the groundwater system in the Wairarapa valley is believed to be a middle Quaternary (mQa) layer of clay and silt sediments at depths varying from 50-100m (Begg *et al.*, 2005). In the upper (northern) area, the regional groundwater system is crosscut by the Mokonui, Masterton and Carterton faults (Jones and Gyopari, 2006). Here, the less permeable mQa sediment layers have been pushed up towards the surface along these fault lines to create barriers that restrict the movement of groundwater and compartmentalise the regional system into sub-regional flow systems (Guggenmos, 2010). Elsewhere, localised zones of uplift and subsidence have resulted in elevated basement and mQa units – Tiffen Hill and Fernhill – as well as localised depressions – Te Ore Ore Basin. The aerial extent of these landforms and the variable distribution of sedimentary units has created numerous flow systems within the Wairarapa Valley (Jones and Gyopari, 2006). The GWRC has divided the Wairarapa groundwater systems into five unique units, which have been identified based on lithology, aquifer yields and aquifer properties (Table 4.2).

Table 4.2. Hydrostratigraphic units, their general hydraulic nature and distribution. 'K' denotes hydraulic conductivity (Jones & Gyopari, 2006).

Unit	General hydraulic nature	Spatial distribution
Sub-basin gravels	Medium to high K, generally thin localised zones.	Distal environment – lower valley, eastern side of valley, sub-basins (Te Ore Ore, Parkvale).
Q1 Holocene gravels	High K, reworked, strong connection with rivers	Main river channels, Waiohine floodplain, Ruamahanga floodplain, lower valley.
Q2 Alluvial fans/outwash gravels	Low K, poor yields	Major fan systems on western valley side of Tauherenikau, Waiohine, Waingawa, Waipoua rivers.
Uplifted blocks	Very low or low K. low bore yields. Form flow barriers	Lansdowne, Tiffen Hill, Fernhill, Te Maire ridge, Martinborough terraces.
Lower valley sub-basin estuarine and lacustrine deposits	Very low K; occasional thin high K gravel layers.	Lower valley, Lake Wairarapa.

4.2.1 Groundwater flow direction

Groundwater flow within the Wairarapa valley is largely regional, moving from high elevation in the north, towards Lake Wairarapa in a south-easterly direction (Figure 4.3) (Begg *et al.*, 2005). While this is the general flow direction, smaller scale variations in groundwater flow direction occur due to the interference of several impermeable barriers (Jones and Gyopari, 2006). Waters in the upper valley propagate south-east through alluvial fans toward the Ruamahanga River and Te Ore Ore plains. Further down the valley, south of the Carterton fault, groundwater is redirected by the impermeable Tiffen Hill-Fernhill formation to flow in union with the Parkvale and Carterton basins. Rounding the southern edge of the basins, groundwater flow energy ebbs as a result of the subdued relief and is redirected toward the Ruamahanga River (Begg *et al.*, 2005). Significant groundwater (piezometric) gradients force flow out from the Tauherenikau and Waiohine gorges south toward Lake Wairarapa and also south-east to join the regional groundwater sourced from further up the valley (Guggenmos, 2010).

The next significant control on groundwater flow is the occurrence of Te Maire Ridge, which directs regional flow in a south-east direction as it is channelled between the ridge and neighbouring Martinborough Terraces. Te Maire Ridge also acts as a flow barrier for groundwater that flows south-east in association with the Tauherenikau outwash fans, directing this flow south (Begg *et al.*, 2005). The impermeable sediments that make up Te Maire ridge forces groundwater systems upwards. Subsequently, a number of springs discharge there (Jones and Gyopari, 2006).

Approaching Lake Wairarapa, this section of the valley is characterised by low-relief and the piezometric gradient is further reduced (Begg *et al.*, 2005). The large lower valley sub-basin is a dominant regional feature that continues to subside and thus groundwater is directed here. Finally, an uplifted impermeable layer along the coast prevents much of the groundwater from reaching the sea, forcing it to the surface (Jones and Baker, 2005).
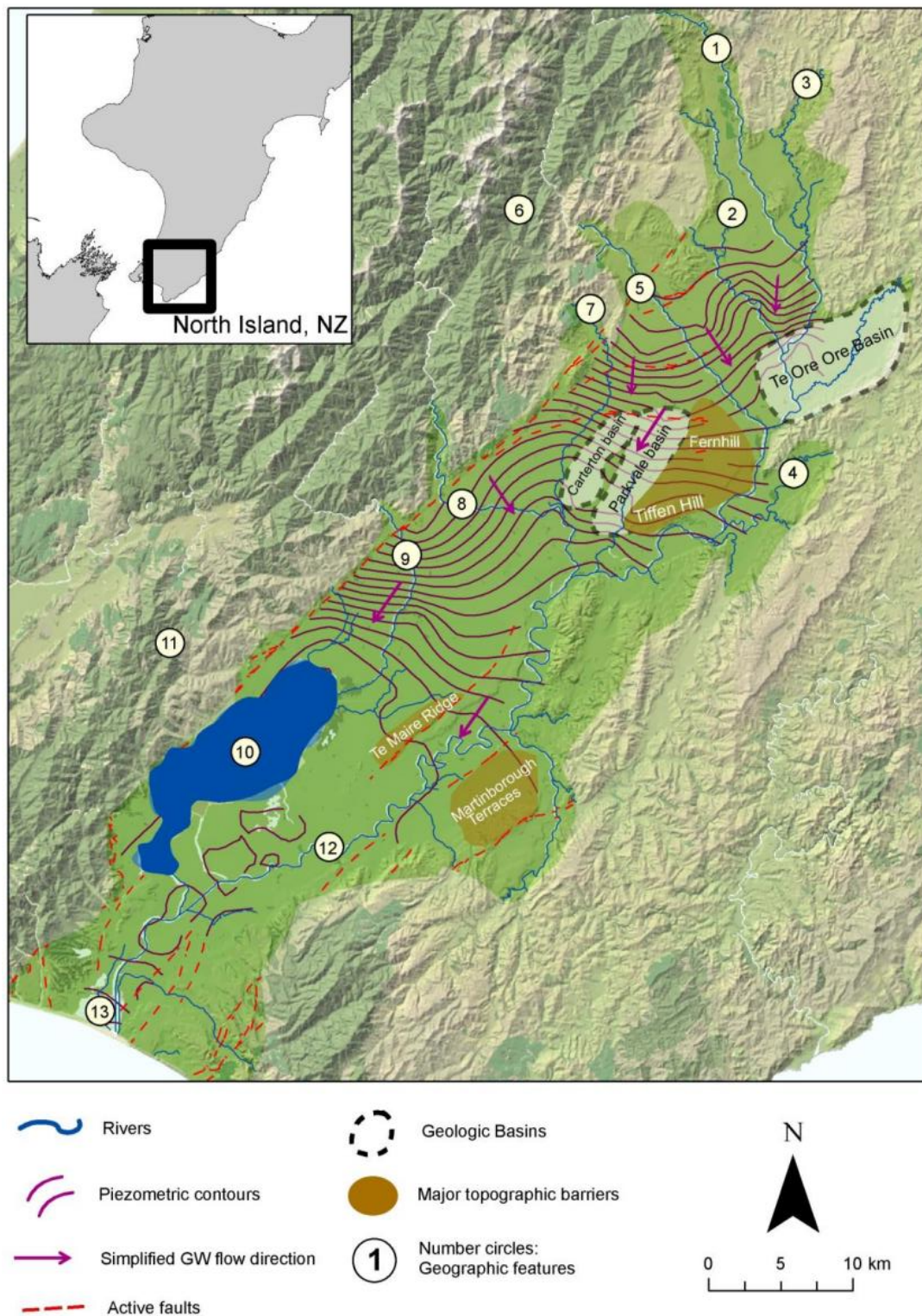
Figure 4.2. Piezometric contour map of the Wairarapa valley, new Zealand showing regional groundwater flow direction, topographic flow barriers, geologic basins, active faults and a number of geographic features and rivers: 1) Ruamahanga River, 2) Waipoua River, 3) Whangaehu River, 4) Taueru River, 5) Waingawa River, 6) Tararua Ranges, 7) Mangatarere River, 8) Waiohine River, 9) Tauherenikau River, 10) Lake Wairarapa, 11) Remutaka Ranges, 12) Lower Ruamahanga River, 13) Lake Onoke (Begg et al., 2005).

4.2.2   Wairarapa groundwater recharge mechanisms

Work undertaken by Jones and Gyopari (2006) found that the underlying groundwater systems displayed significant connectivity with rivers throughout the Wairarapa valley. This investigation classified the main recharge mechanisms and characterised the major rivers in the region into three categories with respect to their interaction with groundwater reservoirs; influent, effluent or neutral. These results are presented in Figure 4.4.

Groundwater systems in the Wairarapa valley are predominantly recharged by rivers and rainfall (Figure 4.4), while other minor recharge mechanisms include lakes and snowmelt (Jones and Gyopari, 2006). The spatial distribution of these mechanisms is largely controlled by the specific Quaternary surface sediments and the network of faults that span the Wairarapa valley (Guggenmos, 2010). Groundwater underlying permeable Q1 river gravels are largely recharged by their overlying river systems, this includes the upper Mangatarere, Tauherenikau and parts of the Ruamahanga. Influent river recharge provides the largest quantity of water to groundwater units in the valley (Jones and Gyopari, 2006).

Recharge by rainfall is also significant, with approximately 35% of total rainfall in the region ending up in groundwater systems (Jones and Gyopari, 2006). Rainfall recharge is also associated with deep aquifers in the lower portions of the valley and any other area not directly linked to a river system. In some areas, groundwater recharge can be provided by a mix of these two mechanisms, with rain-river recharge zones located near the Waiohine, lower Mangatarere, Waipoua and upper Ruamahanga Rivers (Jones and Gyopari, 2006). Zones of mixed recharge are predominantly associated with Q1, Q2, and Q4 sediments.

Discharge from aquifer to river also occurs. Groundwaters characterised as effluent are predominantly in the lower areas of the Wairarapa valley (Jones and Gyopari, 2006). It is these flat topographic zones where the groundwater table and the stream stage are likely to be at similar heights, promoting consistent baseflow conditions (Winter, 1999). Additionally, chemical analysis by Guggenmos (2010), suggests 30-60% baseflow provision at the Mangatarere stream from adjacent groundwater systems.
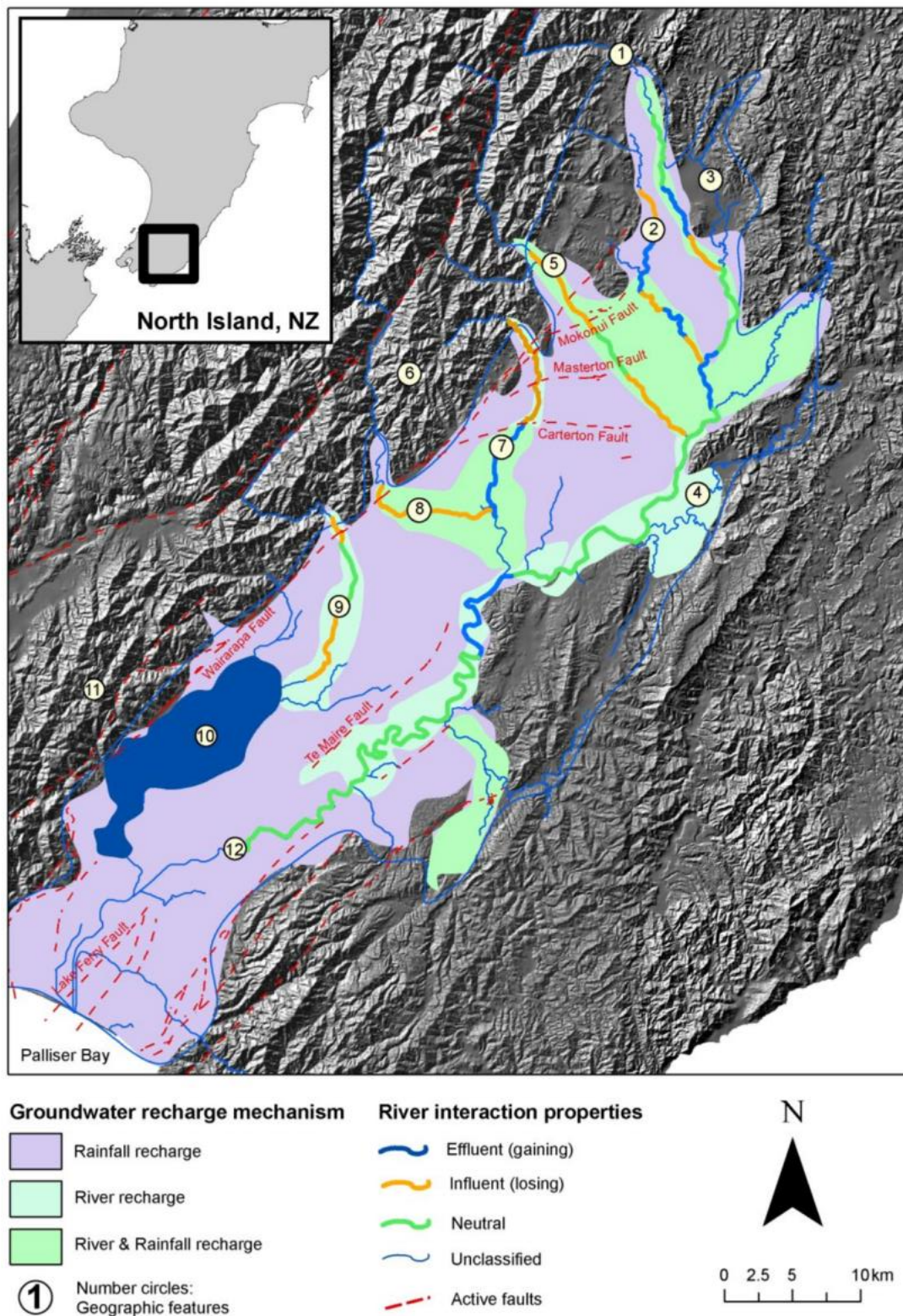
Figure 4.3. Groundwater - surface water interaction and recharge properties in the Wairarapa Valley, New Zealand as classified by the Greater Wellington Regional Council. Significant, unclassified rivers are also included. Circled numbers indicate major geographic features:1) Ruamahanga River, 2) Waipoua River, 3) Whangaehu River, 4) Taueru River, 5) Waingawa River, 6) Tararua Ranges, 7) Mangatarere River, 8) Waiohine River, 9) Tauherenikau River, 10) Lake Wairarapa, 11) Remutaka Ranges, 12) Lower Ruamahanga River (Jones & Gyopari, 2006).

4.3 Surface Hydrology

Numerous river systems flow across the Wairarapa Valley, with the majority of their headwaters sourced from the Remutaka and Tararua ranges. Water is primarily provided to these systems by rainfall, and secondarily by snowmelt. There is also a significant proportion supplied by underlying groundwater systems (see above). Flow values are known to display significant seasonal variability due to the numerous input sources. The highest river flows occur during the winter months (May to August) due to the high rainfall totals over this period and the late-winter snowmelt (MfE, 2018). Significantly lower flow rates during summer are due to lower rainfall totals coupled with increased anthropogenic abstraction (Table 4.3) (Hughes and Gyopari, 2014). With respect to groundwater interaction, these rivers generally lose water to underlying groundwater systems as they exit the surrounding hills and cross historic permeable outwash plains. Further downstream rivers tend to switch to effluent systems, whereby a significant proportion of total flow is provided by groundwater systems (Jones and Gyopari, 2006).

Table 4.3. Mean monthly stream flow (m^3/s) for a selection of major Rivers in the Wairarapa valley. Data range varies per site and is presented in the first column (GWRC, 2017).

Site location	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Kopuaranga at Stuarts	2.4	1.5	2.0	2.3	3.0	5.3	7.1	5.0	5.4	4.8	2.6	1.4
Mangatarere at SH2	2.7	1.1	2.4	4.6	4.7	6.5	7.5	6.9	8.0	6.1	3.2	1.6
Ruamahanga at Wardells	12.2	12.8	13.4	15.6	21.3	30.9	39.1	35.4	29.8	32.3	21.0	17.4
Waingawa at Kaituna	6.8	6.5	7.1	7.7	9.9	11.8	13.1	12.8	12.5	13.1	10.2	9.8
Waiohine at Gorge	17.7	16.3	17.0	18.4	22.6	26.8	30.1	28.8	28.5	33.6	26.1	25.6
Waipoua at Mikimiki	2.0	1.2	1.8	2.4	3.4	4.4	6.9	6.5	5.3	6.2	2.6	1.5

The Ruamahanga River is the largest in the valley in terms of flow, ranging from 11-20 m^3/s in summer up to 20-40 m^3/s in winter at the Wardells bridge flow gauge (GWRC, 2017). Headwaters of the Ruamahanga River initiate high in the northern Tararua Ranges, and flow in a south-easterly direction across poorly sorted gravels in the Wairarapa valley. The river turns south along the eastern boundary of the valley before meandering across the lower plains at the southern end of the catchment (Jones and Gyopari, 2006). The Ruamahanga River was once a tributary of Lake Wairarapa, however it was anthropogenically diverted from its original course to flow directly into

Lake Onoke. This was carried out as part of the Lower Wairarapa Valley Development Scheme that aimed to reduce flooding hazards and create land for agriculture (Perrie, 2005). The Ruamahanga shares confluences with most of the other significant rivers at various locations throughout the valley. It displays a high degree of interaction with underlying groundwater systems and has been identified as having both influent and effluent reaches according to concurrent flow measurements (Jones and Gyopari, 2006). There is also a change from effluent to influent over the reaches where the river flows across the Mokonui and Masterton faults. Further, groundwater contribution is significant as the Ruamahanga passes over the Te Ore Ore sub-basin and the nearby Greytown springs. The lower reaches appear to show little interaction with groundwater systems (Jones and Gyopari, 2006).

The Waiohine is another major river that drains the Tararua ranges, with an average flow of 11-30 m^3/s. The Waiohine flows in an easterly direction before being joined by the smaller Mangatarere stream just south of Carterton. The river system then flows in a south-east direction to join the Ruamahanga River near the Papawai Township (Guggenmos, 2010). The Waiohine has been characterised as influent from the point where it exits the Tararua Ranges to its confluence with the Mangatarere stream. Approximately 1800 m^3/s is lost to groundwater across Q1 gravels over this entire reach (Jones and Gyopari, 2006). Similarly, the Mangatarere River has been identified as influent as it flows onto the plains from the Tararua Ranges before switching to effluent as it flows over the Carterton fault (Jones and Gyopari, 2006). The Mangatarere stream has an average flow of 0.9-3 m^3/s and is a focus point for hydrological modelling and research in the region (Guggenmos, 2010; Zammit and Yang, 2017).

The Waingawa, Waipoua and Tauherenikau Rivers also drain the Tararua Ranges. Flowing south-east, these rivers dissect poorly sorted alluvial gravels and are known to have a variable interaction with groundwater systems. As with the Ruamahanga and Mangatarere, their degree of interaction is highly influenced by the Mokonui and Masterton fault lines (Jones and Gyopari, 2006). All three river systems are characterised by influent reaches that change as they cross the respective fault lines. The Waingawa is characterised as neutral along the reach between the Mokonui and Masterton fault traces. After flowing across the Masterton fault it returns to an influent

system. The Waipoua switches from influent to effluent multiple times as it flows south-east across the valley. The Tauherenikau maintains a neutral reach for ca. 4.5km, then switches back to an influent system in the lower catchment. The Waingawa and the Waipoua flow into the larger Ruamahanga River, whereas the Tauherenikau flows directly into Lake Wairarapa (Jones and Gyopari, 2006).

Lake Wairarapa is a significant hydrological feature in the region. It is a 76 km^2 shallow lake located in the lower subsiding section of the Wairarapa Valley. It receives the majority of its inflow from the Tauherenikau River as well as shallow groundwater discharge zones associated with neighbouring Q1 gravels (Guggenmos, 2010). Additional input is provided by shoreline springs that are sourced by deep confined aquifers located within the underlying lacustrine and estuarine sediments (Jones and Gyopari, 2006).

A number of smaller streams drain the north-eastern Wairarapa hills. The Whangaehu and Kopuaranga Rivers, and the Waingongoro and Huangarua Streams flow atop historic alluvial fans and join the Ruamahanga River. Likewise, the Taueru River drains into the Ruamahanga from the east after flowing out of the marine deposited ranges (Guggenmos, 2010). Little information is known regarding their interaction with groundwater systems. It is likely, using the rest of the catchment as an example, that groundwater interaction occurs where these systems flow over sections of Q1 river gravels. Excepting the Kopuaranga and Taueru Rivers, continuous discharge and flow data are not readily available as these systems are mostly monitored for flood control only (GWRC, 2017).

4.4 Climate

The Wairarapa region is characterised by a dry, warm climate. It is heavily influenced by the Remutaka and Tararua ranges on the western boundary (GWRC, 2014). These ranges act as a significant barrier, sheltering the valley and its plains from the predominant westerly winds (MfE, 2018). The prevailing wind and the ranges produce a strong orographic effect in the region. Consequently, there is relatively low annual rainfall (800-1000mm) across the Wairarapa valley (Figure 4.5). In contrast, the ranges themselves experience rainfall up to 6000mm per annum (NIWA, 2018). There is also significant seasonal variation in rainfall across the catchment, with highest rainfall during late-winter early spring, and low rainfall from January to March (Figure 4.6)

(GWRC, 2014). The seasonal rainfall variability results in increased river flows during the winter months compared to summer (Hughes and Gyopari, 2014). Mean annual temperature in the Wairarapa valley is approximately 13°C, with the highest temperatures occurring during the months between November and March (Figure 4.6) (NIWA, 2018).
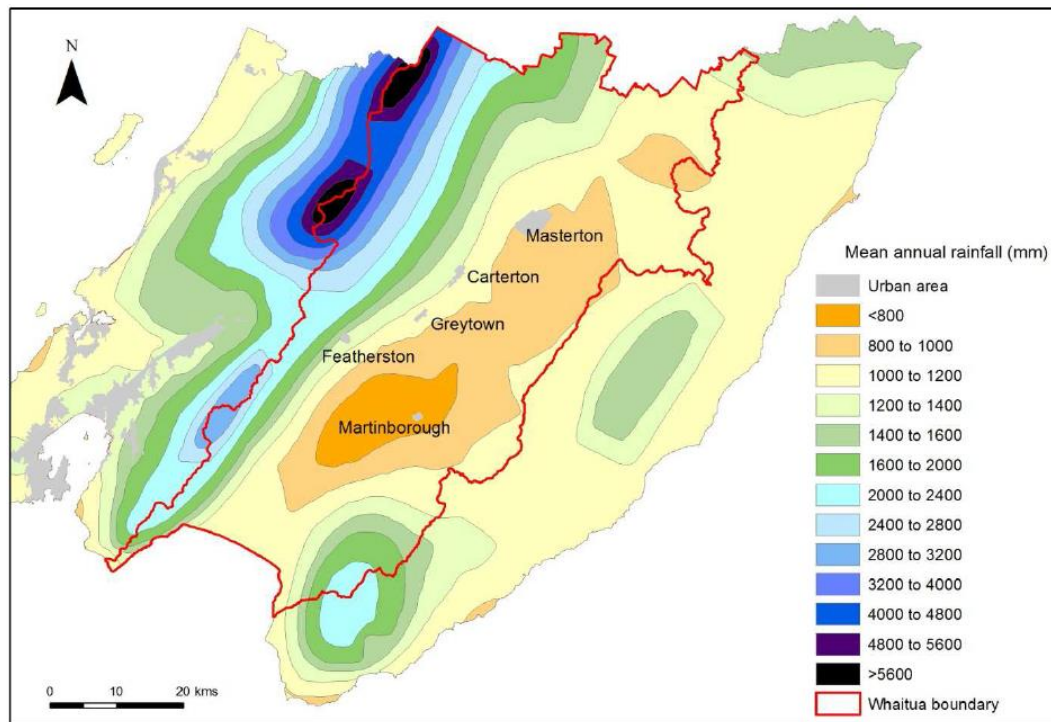
Figure 4.4. Mean annual rainfall for the Wellington region, New Zealand with Ruamahanga catchment delineated (GWRC, 2014).
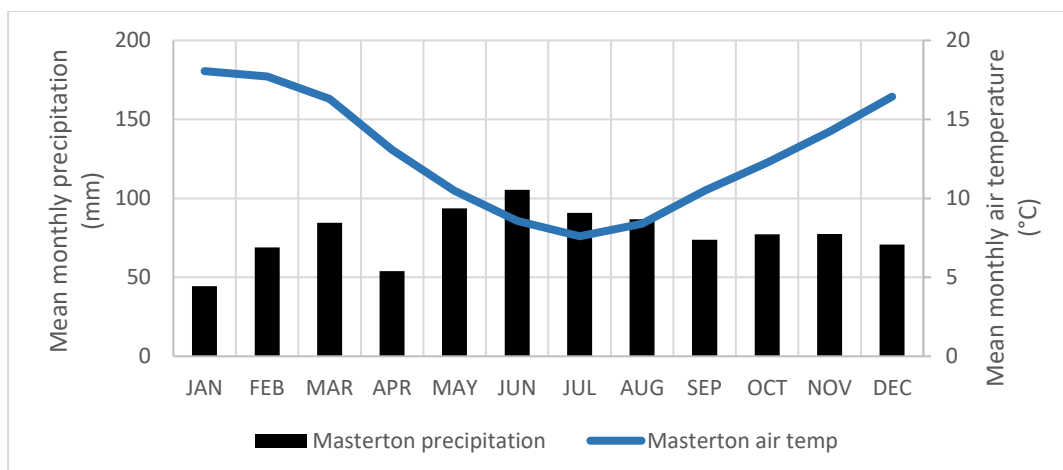
Figure 4.5. Mean monthly precipitation (mm) and air temperature (C) for Masterton (-40.957 S, 175.707 E). Based on 1981-2010 data (GWRC, 2014).

Climate modes across the South Pacific play a role in driving precipitation trends. Local precipitation trends in the Wairarapa Valley are likely to respond to these large scale mechanisms and display significant interannual and decadal variation depending on the climate mode and its phase (Jiang *et al.*, 2013). Climate modes that could affect precipitation variability in the Wairarapa region include the El Niño-Southern Oscillation (ENSO), the Southern Annular Mode (SAM) and the Interdecadal Pacific Oscillation (IPO) (Salinger and Mullan, 1999; Jiang *et al.*, 2013). Because of the combinations and interactions of these various climatic drivers, it is difficult to assess the regional impacts of individual climate modes. It is a combination of features, well as topographic influences, that produce the observed conditions (Jiang *et al.*, 2013).

El Nino Southern Oscillation (ENSO)

The ENSO is a coupled ocean-atmosphere phenomenon characterised by its El Nino and La Niña phases (Philander, 1983). In New Zealand the El Niño phase generally produces anomalously wet conditions in the south and west, and anomalously dry conditions in the north and east (Gordon, 1986). The La Niña phase generally produces conditions of the opposing nature (Gordon, 1986; Mullan, 1995). Significant El Niño events from the recent decades include the years of 1982/83 and 1997/98. For the Wairarapa region, an El Niño phase would result in increased precipitation due to an increase in southerly and westerly moisture laden winds whereas the La Niña phase would act in the reverse manner (Jiang *et al.*, 2013).

Southern Annular Mode (SAM)

The SAM indicates the latitude of the prevailing westerly wind belt that is generated in between the dominant high and low pressure systems that preside over the Southern Ocean and Antarctica respectively (Kidston *et al.*, 2009). When the SAM is in a negative phase, higher than normal pressure at high latitudes over Antarctica and lower than normal pressure at low latitudes over the Southern Ocean force the westerly wind belt north. The westerlies produce unsettled weather over New Zealand along with the Wairarapa region, while wind and storm activity ease over the southern oceans (Thompson and Solomon, 2002). The SAM fluctuates on a sub-annual basis and a record of variability is the mean monthly SAM Index (SAMI) from 1958-2008 (Marshall, 2003).

Interdecadal Pacific Oscillation (IPO)

The IPO is a long period oscillation that exhibits lower than normal precipitation southeast of the South Pacific Convergence Zone during its positive phase due to higher than normal mean sea level pressures (Salinger *et al.*, 2001). The IPO, which operates at a multi-decadal scale, has been classified positive for the period 1977-1998 and negative for the two periods 1958-1976 and 1999-2014 (Jiang *et al.*, 2013). It is understood that the Wairarapa region would experience more rainfall during a negative phase of the IPO (Jiang *et al.*, 2013).

Climate change impacts

The effects of climate change on the larger scale climate modes is a significant area of research. It is generally understood that extreme El Niño events will become increasingly common (Yeh *et al.*, 2009; Cai *et al.*, 2014). There is also an evident trend toward the positive phase of the SAM index over the recent decades (Marshall, 2003). Both of which would significantly influence precipitation and climate in the Wairarapa region. Factors such as these and many more were considered in an investigation carried out by the Ministry for the Environment (2012), which suggests temperatures in the Wairarapa are expected to be 0.7 to 1.1℃ warmer in 2040 compared to 1995 values. Further, the region is likely to experience up to 7% less rainfall, despite extreme rainfall events expected to become more frequent by a factor of 2.

These projections indicate that the Wairarapa region is likely to have more water over less time in its rivers and could experience more time in drought conditions (Tait *et al.*, 2002). This will likely result in greater pressure on groundwater resources and highlights the importance of sustainable water management in order to avoid significant water shortages (Keenan *et al.*, 2012; MfE, 2018).

4.5   Human history and land-use

As detailed in Chapter 2, groundwater-surface water interaction is affected by a range of physiographic and environmental processes. It is also important to document and understand the extent of human modification to vegetation, soil, and hydrological systems, as well as changes in land-use in the Wairarapa valley.

Historically, land cover in the Wairarapa valley changed significantly with the respective arrivals of Maori and European settlers. Dense native bush, predominantly

podocarp forest, was cleared or burned by Maori settlers to be colonised by native grasses, ferns, swamps and shrubs (Beadel *et al.*, 2000). Upon the arrival of European Pakeha, the extensive grassland on the plains made the Wairarapa an appealing prospect for farming (Beadel *et al.*, 2000). Much of the remaining native forest was cleared to promote fresh growth of grasses across the region. Many natural wetlands were drained to provide more land for farming practices and exotic plant species were introduced which quickly re-colonised the Wairarapa valley plains (Watts and Perrie, 2007).

Contemporary land use in the Wairarapa is predominantly pastoral crop and grasslands (Figure 4.7). Dairy, beef, horticulture and viticulture are major contributors to regional GDP (Statistics New Zealand, 2016). In particular, dairy farming has experienced intensification over recent decades and the Wairarapa region contains almost 50% of the dairy cattle in the Greater Wellington region (MfE, 2018). Most native forest is now confined to the axial ranges, with some small pockets scattered throughout the valley and along the shores of Lake Wairarapa. Small viticulture and market gardening projects are also clustered around Martinborough and other urban centres (Jones and Baker, 2005).
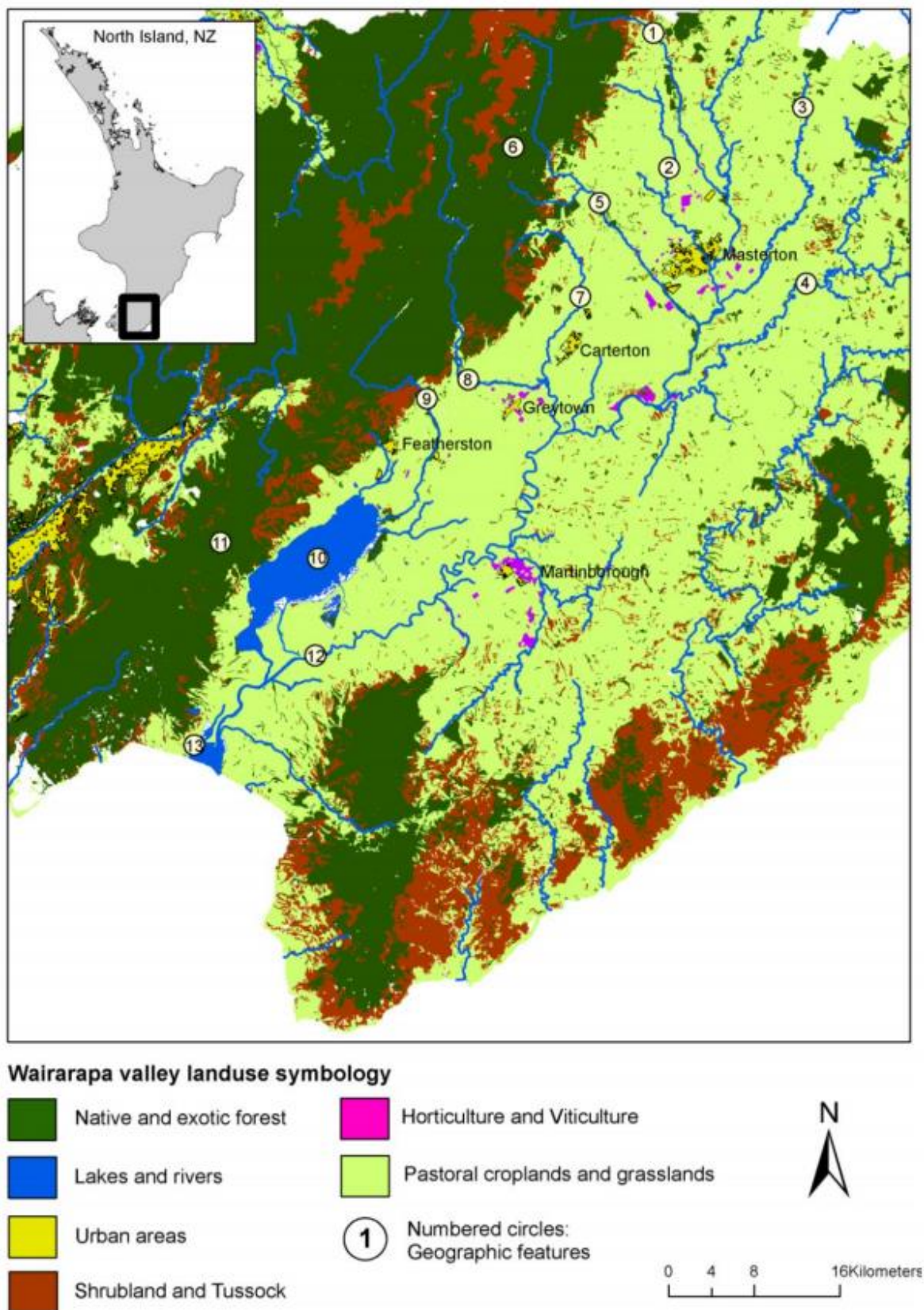
Figure 4.6. Land use map of the Wairarapa valley and surrounding areas identifying main land use types and major geographic features: 1) Ruamahanga River, 2) Waipoua River, 3) Whangaehu River, 4) Taueru River, 5) Waingawa River, 6) Tararua Ranges, 7) Mangatarere River, 8) Waiohine River, 9) Tauherenikau River, 10) Lake Wairarapa, 11) Remutaka Ranges, 12) Lower Ruamahanga River, 13) Lake Onoke (Jones & Baker 2005, Guggenmos, 2010).

4.5.1 Current hydrological monitoring in the Wairarapa valley

The Greater Wellington Regional Council (GWRC) undertake extensive hydrological monitoring throughout the year to monitor environmental trends in water quality and quantity, guide resource consent decision making and convey information about natural resources to the public (Watts & Gordon, 2008). Water stage is continuously monitored, usually at 15 minute intervals, within the majority of river systems (Guggenmos, 2010). Water stage is the height of the river and has a direct linear relationship with flow volumes (Le Coz, 2012). Stage-discharge rating curves, used to continuously estimate discharge, are assessed periodically throughout the year and after significant storm events to maintain robust estimates of flow (Zammit and Yang, 2017). Stage is also monitored at a number of selected groundwater bores throughout the Wairarapa valley. This raw stage data from both ground and surface water sites is uploaded, via telemetry, to the council's database every 2 to 3 hours (Guggenmos, 2010). This store of information can be used to document environmental change over time compared with State of the Environment (SoE) monitoring data from other regions of the country. Results from all hydrological monitoring programmes are summarised monthly and on an annual basis in summary reports compiled by GWRC (2017).

4.5.2 Surface and groundwater abstraction

Groundwater abstraction in the Wairarapa was initially used for small-scale stock and rural domestic supply (Morgan and Hughes, 2001). However, intensification of agriculture in the 1960's placed significant pressures on surface water abstractions. Comprehensive groundwater investigations carried out by the Wairarapa Catchment Board between 1981 and 1986 suggested that further comprehensive research was required to fully understand the dynamic nature of the Wairarapa groundwater system (Morgan and Hughes, 2001). At present, the GWRC administers surface water and groundwater allocations in the Wairarapa valley (GWRC, 2018).

According to Bunny *et al.* (2014) water allocation in the Wairarapa increased by 80% between 1990 and 2010, the majority of which is used for irrigation of dairy pastures. The growing demand for water is putting substantial pressures on surface and groundwater resources in the region (Bunny *et al.*, 2014; Hughes and Gyopari, 2014).

Currently, there are over 470 water consents in the Wairarapa region (approximately 150 surface water and 320 groundwater) (GWRC, 2018). Although the majority of groundwater takes are less than 500 m^3/day, several larger takes are over 4000 m^3/day. Additionally, significant seasonal variations in abstraction are present, with the majority of water used for irrigation during the summer months between October and March (Hughes and Gyopari, 2014). Thus the period of maximum abstraction coincides with the time of year most susceptible to drought or general low streamflow values (de Joux and Scarf, 2010). Accurate quantification of these abstraction rates is made more difficult in the Wairarapa region as most takes are not metered and consent descriptions only state maximum daily allowances (Jones and Baker, 2005; GWRC, 2018). The extent at which this significant trend of abstraction over the last half century has affected the interaction between surface water and groundwater is not well understood (Hughes and Gyopari, 2014)

Generally, the most highly allocated groundwater units are located adjacent to major river systems (Hughes and Gyopari, 2014). It is these abstractions from shallow Q1 river gravels that could lead to depletion of associated surface water flows. Hughes and Gyopari (2014) recommend that allocation of groundwater from an aquifer with high connectivity to surface water bodies, should be classed as 'surface water allocation' from the relevant river or stream. This would effectively combine unconfined aquifer and surface water allocation consents. According to one report, about half of the groundwater zones in the Wairarapa Valley are allocated at levels greater than 60% of their recommended yields (Bunny *et al.*, 2014). This statement is supported by a 2015 Proposed Natural Resources Plan (PNRP) that shows the majority of surface and groundwater allocations are at or near recommended limits (Tables 4.4 and 4.5, Figures 4.10 and 4.11) (GWRC, 2015).

Surface water allocation

Table 4.4. Proposed Natural Resources Plan (NRP) and actual surface water allocations for catchment management units in the Wairarapa region, New Zealand (GWRC, 2015).

Surface water unit	Proposed NRP limit (litres/sec)	Allocation (litres/sec)
Ruamahanga (total)	7535	7515
Kopuaranga	180	150
Waipoua	145	122
Waingawa	920	1184
Ruamahanga (upper)	1200	902
Parkvale	40	151
Booths	25	109
Mangatarere	110	465
Waiohine	1590	950
Papawai	65	337
Ruamahanga (middle)	1240	998
Huangarua	110	92
Ruamahanga (lower)	1475	2055
Lake Wairarapa	1800	1778
Otakura	30	140
Tauherenikau	410	233
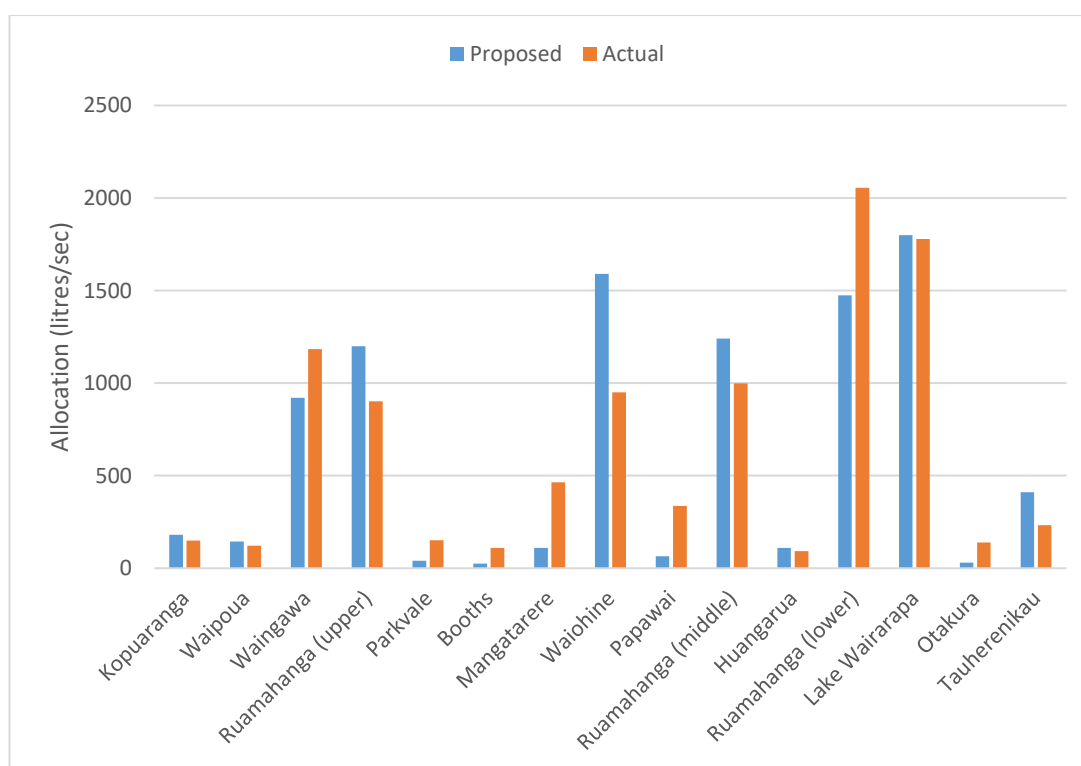
Figure 4.7. Actual allocation of surface water units and advised allocation under the Proposed Natural Resources Plan (2015). Data sourced from GWRC (2015).

Groundwater allocation

Table 4.5. Proposed Natural Resources Plan (NRP) and actual groundwater allocations for catchment management units in the Wairarapa region, New Zealand (GWRC, 2015).

Groundwater unit	Proposed NRP limit (m3/year)	Allocation (m3/year)
Te Ore Ore	480000	755485
Waingawa	1900000	994907
Upper Ruamahanga	3550000	582227
Fernhill-Tiffen	1200000	972228
Taratahi	1400000	491913
Parkvale (confined)	1550000	2162700
Parkvale (unconfined)	350000	340200
Mangatarere	2300000	2548691
Tauherenikau	6600000	6233757
Lake Wairarapa	6750000	8554757
Huangarua	650000	650002
Martinborough	800000	760422
Dry River	650000	570066
Onoke	2100000	2095200
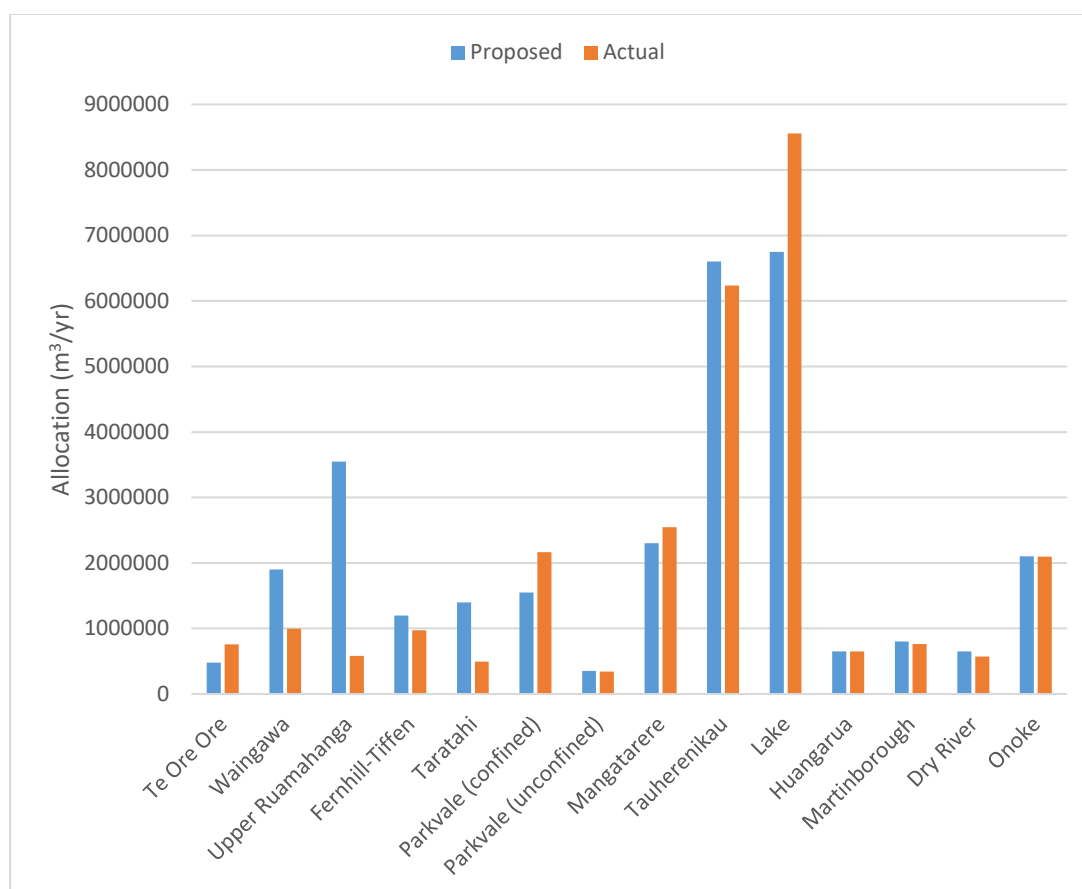
Figure 4.8. Actual allocation of groundwater units and advised allocation under the Proposed Natural Resources Plan (2015). Data sourced from GWRC (2015).

Restrictions

Allocations from surface water and groundwater reservoirs are restricted when flow goes below certain thresholds, at which point abstraction must be reduced or halted entirely (Tables 4.6 and 4.7) (GWRC, 2018). Defining these threshold values is managed on a case-by-case basis, and is dependent on the magnitude of the take, with priority given to municipal abstractions. Further annual restrictions are applied to both surface water and groundwater abstractions, capping continuous daily abstraction to approximately 5 months of the year (i.e. summer months) (GWRC, 2018).

Table 4.6. General restrictions that apply to surface water take consents in significant catchments in the Wairarapa valley, New Zealand. Partial restrictions vary, but tend to be 50% of consented take. Full restriction means cease take (GWRC, 2018).

Catchment	Monitoring Site	Partial restriction	Full restriction
Booths	Parkvale at Renalls weir	< 150 l/s	< 100 l/s
Kopuaranga	Kopuaranga at Palmers Bridge	-	< 270 l/s
Lower Mangatarere	Mangatarere at Gorge	< 270 l/s	< 200 l/s
Upper Mangatarere	Mangatarere at Gorge	< 330 l/s	< 240 l/s
Otakura	Otakura at Weir	-	< 95 l/s
Parkvale	Parkvale at Renalls Weir	< 150 l/s	< 100 l/s
Lower Ruamahanga	Ruamahanga at Waihenga Bridge	< 9200 l/s	< 8500 l/s
Upper Ruamahanga	Ruamahanga at Wardells	< 2700 l/s	< 2400 l/s
Tauherenikau	Tauherenikau at Gorge	< 1350 l/s	< 1100 l/s
Waingawa	Waingawa at Kaituna	< 1900 l/s	< 1700 l/s
Waiohine	Waiohine at Gorge	-	< 3040 l/s
Waipoua	Waipoua at Mikimiki Bridge	< 300 l/s	< 250 l/s

Table 4.7. General restrictions that apply to groundwater take consents in significant catchments in the Wairarapa valley, New Zealand. Partial restrictions vary but tend to be 50% of total consented take.

Catchment	Monitoring Site	Partial restriction	Full restriction
Lower Mangatarere	Mangatarere at Gorge	< 270 l/s	-
Otakura	Otakura at Weir	< 95 l/s	-
Lower Ruamahanga	Ruamahanga at Waihenga Bridge	< 8500 l/s	-
Upper Ruamahanga	Ruamahanga at Wardells	< 2400 l/s	-
Tauherenikau	Tauherenikau at Gorge	< 1350 l/s	-
Waingawa	Waingawa at Kaituna	< 1900 l/s	-
Waiohine	Waiohine at Gorge	< 3040	-
Waipoua	Waipoua at Mikimiki Bridge	< 250 l/s	-

4.6 Summary

The Wairarapa valley structural depression sits above the Pacific and Australian subduction zone. The valley is confined by Triassic greywacke ranges on its western margin and Pleistocene and late-Tertiary sedimentary ranges on the eastern margin. Successive glacio-fluvial layers from the Quaternary period have been deposited in the upper and middle section of the valley and deep layers of estuarine and marine sediment layers are present in the subsiding lower valley as a result of sea-level fluctuations. This variable assortment of sediments has created a complex regional groundwater system that is further compartmentalised by the Masterton, Mokonui and Carterton faults that strike north east through the valley, pushing impermeable flow barriers to the surface.

The valley plays host to a number of significant river systems that predominantly flow atop permeable Q1 alluvial gravels. These systems receive water from precipitation, snowmelt, and groundwater. Most experience their highest flows throughout winter months when precipitation is greatest. Concurrent flow gaugings carried out by the GWRC along with the presence of permeable Q1 alluvial gravels suggest these fluvial systems possess strong interaction properties with underlying, unconfined groundwater systems. The extent of this interaction is not well documented. It is generally understood that groundwater bodies directly below or adjacent to rivers are predominantly recharged by influent river systems, whereas groundwater bodies located farther afield are mostly recharged by rainfall. River recharge represents the largest mechanism in terms of magnitude, whilst rainfall recharge covers the largest areal extent.

There is a significant history of agricultural use in the Wairarapa valley and today over 70% of the region is occupied by pastoral agriculture. Accordingly, ground and surface water abstraction in the valley has increased over the last half century to accommodate for a surge in demand for freshwater. The impact of this extraction on ground and surface water interaction is an area of concern and not well quantified (Bunny *et al.*, 2014; Hughes and Gyopari, 2014).

Chapter 5
Flow infilling

The main aim of this research was to determine the extent to which observed flow records are applicable to groundwater focused model validation investigations in the Wairarapa valley, New Zealand. In order to achieve this, observed flow data provided by Greater Wellington Regional Council (GWRC) needed to go through a conditioning process to infill gaps in the streamflow record. This stage is especially important for baseflow investigations because continuous datasets are a requirement for the application of digital hydrograph separation, a common form of baseflow estimation (Chapman and Maxwell, 1996).

Streamflow records can provide information on significant hydrological drivers and reservoirs in a catchment (Harvey *et al.*, 2012), they are also readily available datasets for calibration and validation of hydrological models (McMillan *et al.*, 2016). It is relatively common for hydrological time series to have gaps or be incomplete and a process of flow infilling to complete a dataset is often required (Harvey *et al.*, 2012). Flow infilling is possible by interpolation from highly correlated, analogue gauging stations. Alternative methods such as manual inference, serial interpolation and hydrological modelling are usually applied at other temporal resolutions of a dataset (Harvey *et al.*, 2012).

This chapter presents the results from the flow infilling investigation and is divided into six main sections. The first section presents the original observed flow datasets that were employed throughout the various investigations. This is followed by Section two which describes the flow infilling methodologies applied, including the various options for flow infilling and gauge multipliers investigated. Section three presents the findings of the comparisons between infilled and observed flow at 13 gauges in the Wairarapa, based on quantitative statistics and graphical assessment. This is followed by a section discussing the flow infilling process carried out with respect to the three main stages: gauge correlations, gauge multipliers and infilling. Section five states the limitations surrounding this part of the investigation, while Section six presents a final summary of analogue gauge flow infilling and its application to gauges in the Wairarapa valley, New Zealand.

5.1 Observed flow data

Observed flow datasets were provided by the Greater Wellington Regional Council (GWRC) at 20 locations throughout the Wairarapa region (Table 5.1 and Figure 5.1). Hourly flow is recorded in cumecs (m^3/s), to three decimal places, by ARL flow gauges and uploaded via telemetry to a database. The observed records for each flow gauge vary significantly in duration. Five of the gauges record data from before 1980 and span the last four decades. Six gauges have datasets in the 15-39 year range. Nine of the datasets are less than 15 years in length, three of which possess less than three years of data. Most of the gauges continue to be operational at the time of this research.

Table 5.1. List of Ruamahanga flow gauge sites in alphabetical order.

No.	Site	Name	Lat. E	Long. S	Start	End
1	Booths Creek at Andersons Line	Booths	175.5361	-41.015	Jan-11	Dec-14
2	Dock Creek at U/S Otakura Junction	Dock	175.3467	-41.189	Jan-10	May-10
3	Kopuaranga at Palmers Bridge	Kopuaranga 1	175.6719	-40.8189	Mar-85	Dec-17
4	Kopuaranga River at Stuarts	Kopuaranga 2	175.6918	-40.8936	Aug-09	Dec-17
5	Mangatarere at Belvedere Bridge	Mangatarere 1	175.5098	-41.0127	Jan-04	Jan-18
6	Mangatarere River at Gorge	Mangatarere 2	175.5119	-40.935	Feb-99	Jan-18
7	Mangatarere River at State Highway 2	Mangatarere 3	175.4962	-41.0548	Sep-09	Jan-18
8	Otakura Stream at Weir	Otakura	175.3674	-41.1873	Dec-97	Jan-18
9	Papawai Stream at U/S Oxi Pond Confl	Papawai	175.4906	-41.1031	Dec-05	Aug-15
10	Parkvale Stream at Renalls Weir	Parkvale 1	175.5414	-41.0778	Jan-02	Dec-17
11	Parkvale tributary at Lowes Reserve	Parkvale 2	175.593	-40.9969	Mar-11	Nov-17
12	Ruamahanga River at Mt Bruce	Ruamahanga 1	175.598	-40.7546	Jan-75	Nov-17
13	Ruamahanga River at Waihenga Bridge	Ruamahanga 2	175.44	-41.1969	Jan-76	Jan-18
14	Ruamahanga River at Wardells	Ruamahanga 3	175.6724	-41.0048	Sep-77	Dec-17
15	Taueru River at Te Whiti Rd Bridge	Taueru	175.6681	-41.0634	Mar-02	Jan-18
16	Tauherenikau at Gorge	Tauherenikau	175.3574	-41.0681	Mar-76	Jan-18
17	Waingawa River at Kaituna	Waingawa	175.5229	-40.8867	May-76	Jan-18
18	Waiohine River at Gorge	Waiohine	175.4001	-41.0164	May-79	Jan-18
19	Waipoua River at Colombo Rd Bridge	Waipoua 1	175.6736	-40.9542	Dec-09	Mar-11
20	Waipoua River at Mikimiki Bridge	Waipoua 2	175.6174	-40.8452	Feb-07	Dec-17
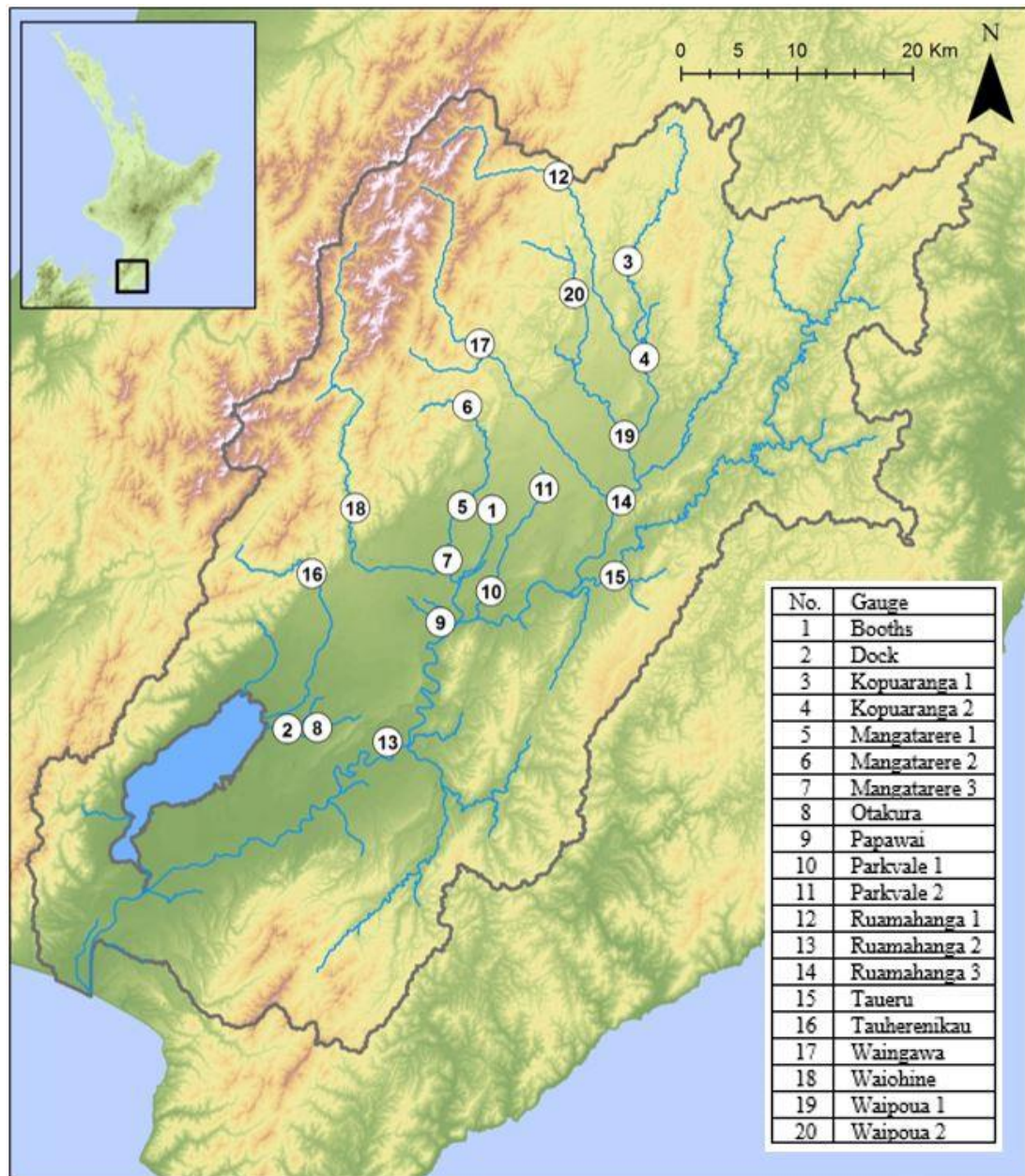
Figure 5.1. Locations of Wairarapa flow gauges assessed for use in this investigation (GWRC 2018).

Along with the variable record lengths, the datasets also contain numerous gaps, likely due to random gauge failure or large storm events etc (Zammit and Yang, 2017). The period 01/07/2010 to 30/06/2014 (four consecutive hydrological years) shows the best overlap of the most gauges and is forthwith referred to as the 'investigation period'. The dataset durations, their gaps and the investigation period are displayed in Figure 5.
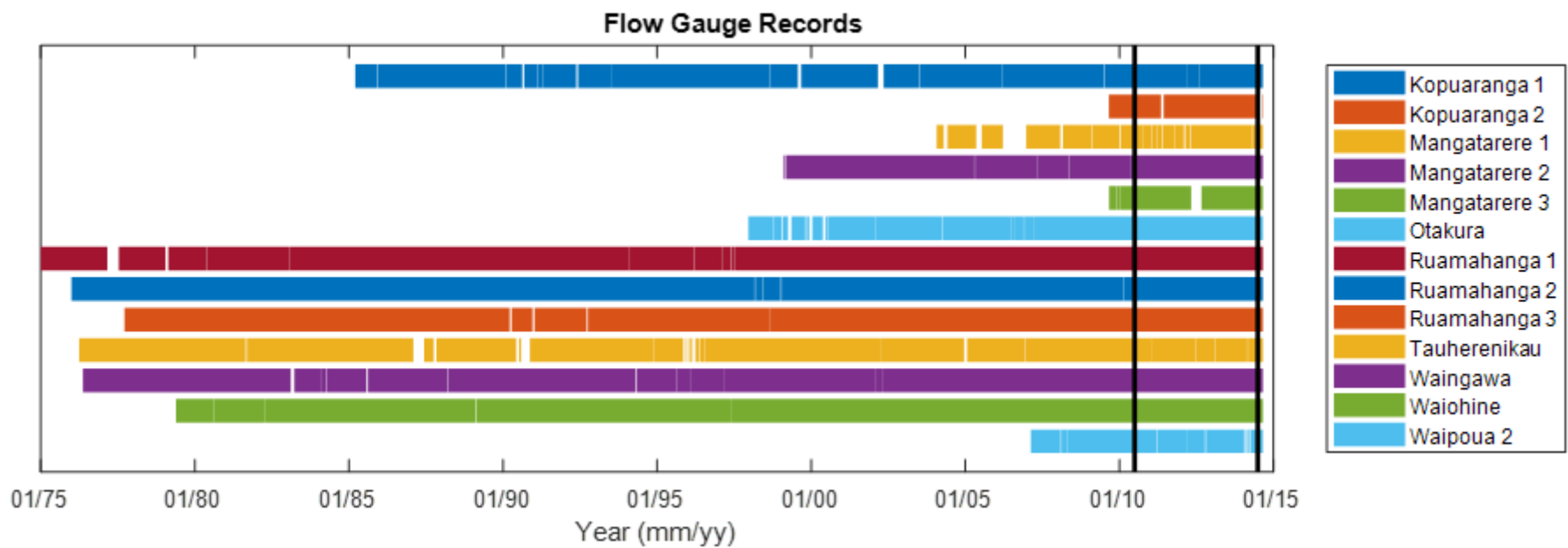
Figure 5.2. Duration of 20 Wairarapa flow gauges showing gaps in record and 2011-2014 hydrological years "Investigation Period" (black vertical lines)

5.1.1 *Measured streamflow quality*

For applications of model comparisons against measured streamflow, a careful review of data quality is an important consideration (McMillan *et al.*, 2010). Quality assurances provided by GWRC state that good quality discharge measurements are available at 14 of the 20 locations (GWRC, 2017). Many of the flow sites are fully rated (for high flows and low flows) from at least the mid-1970's onwards and have reliably maintained rating curves. Other sites have (for long periods in their history) been maintained primarily as flood warning sites only and low-flow records are less certain during these periods (GWRC, 2017). This has been taken into account during the validation process, with some gauges being removed from the comparison where measured streamflow has been considered too unreliable.

Of the 20 flow records provided by GWRC, six gauges with data quality characterised as moderate or low by GWRC were omitted from this investigation. Assessment of poor flow propagation of flow hydrographs supported this decision. A seventh gauge removed from the investigation was deemed to have a record that did not coincide well with the investigation period despite being characterised as "good" by GWRC. Leaving out the seven gauges was under the provision that their omission would not compromise the key aims of this investigation. Comments on record quality and factors that might affect natural flow conditions are based on notes that accompanied the original flow data and assessment of the hydrographs. With respect to duration of record and time considerations, it was decided that gauges that did not cover the majority of the investigation period would not provide representative or comparable results. Gauges left out of the analyses are listed below in Table 5.2 along with the reason for their omission.

Table 5.2. List of flow gauges removed from investigation. Quality assessed from GWRC comments.

Name	No.	Reason	Quality
Booths	1	Large gaps (>1 month) in time series	Moderate
Dock	2	Record less than one year	Low/moderate
Papawai	9	High flow affected by backwater flow	Moderate
Parkvale 1	10	Erratic record with large negative anomalies	Moderate/Good
Parkvale 2	11	Poor rating curve estimation, specifically at low flow	Low/moderate
Taueru	15	Modified by Kourarau Dam operation and large gaps	Moderate/Good
Waipoua 1	19	Short record with poor overlap into investigation period	Good

The 13 gauges used in this investigation are presented in Table 5.3 along with estimated catchment characteristics. Mean flow was calculated over the entire length of each dataset. Drainage area was estimated using a geographic information system (GIS) flow accumulation layer which was based off an 8m digital elevation model (DEM) (LINZ, 2016). An 8 point pour method was employed to evaluate flow upstream of any point, in this case the 13 gauges in the Wairarapa region (Figure 5.3). Specific discharge is calculated as the ratio of mean flow to catchment area. Aquifer classification was characterised by laying QMAP data (Begg and Johnston, 2000; Lee and Begg, 2002) over the 13 catchment delineations in the GIS. Predominant geology was used to infer between 'Hard Rock' or 'Porous' aquifers in a similar fashion to investigations by Collischonn and Fan (2013) and in accordance with classifications described by Eckhardt (2008). The Wairarapa valley is dominated by porous Quaternary age fluvial deposits (see Chapter four). Therefore, all of the catchments have been designated as having porous rather than hard rock underlying geology.

Table 5.3. List of flow gauge datasets used in this research, along with calculated long-term mean flow, estimated catchment size, specific discharge and predominant geology.

No.	Gauge	Mean Flow (m3/s)	Drainage Area (km²)	Spec. Dis. (m³/s/km²)	Aquifer
3	Kopuaranga 1	2.60	101.83	0.026	Porous
4	Kopuaranga 2	3.50	168.56	0.021	Porous
5	Mangatarere 1	1.97	56.38	0.035	Porous
6	Mangatarere 2	1.82	33.43	0.054	Porous
7	Mangatarere 3	4.51	118.25	0.038	Porous
8	Otakura	0.55	45.43	0.012	Porous
12	Ruamahanga 1	9.97	76.35	0.131	Porous
13	Ruamahanga 2	83.61	1743.55	0.048	Porous
14	Ruamahanga 3	23.52	397.97	0.059	Porous
16	Tauherenikau	9.16	111.27	0.082	Porous
17	Waingawa	10.17	76.19	0.134	Porous
18	Waiohine	24.37	182.36	0.134	Porous
20	Waipoua 2	3.75	79.68	0.047	Porous

Figure 5.3. Watersheds of 13 flow gauges in the Wairarapa valley, NZ. Grey represents rest of Ruamahanga catchment where no gauge is present.
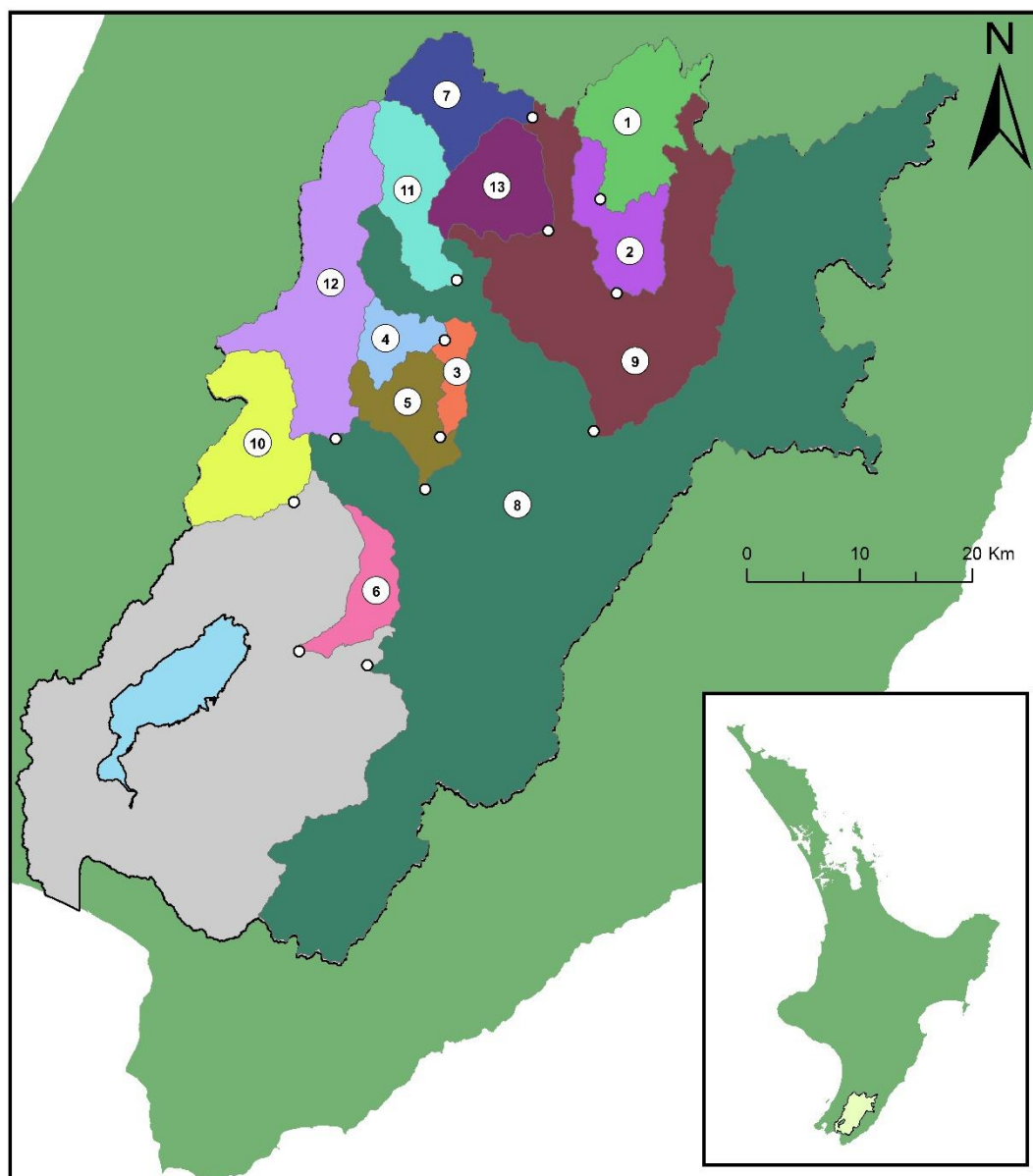
Legend

1. Kopuaranga 1	5. Mangatarere 3	9. Ruamahanga 3	
2. Kopuaranga 2	6. Otakura	10. Tauherenikau	
3. Mangatarere 1	7. Ruamahanga 1	11. Waingawa	
4. Mangatarere 2	8. Ruamahanga 2	12. Waiohine	
		13. Waipoua 2	

5.2 Flow infilling methodology

Flow infilling is a common practice in hydrology. There are many different methods depending on the desired temporal and spatial scale necessary for infilling (Harvey *et al.*, 2012). Infilling methods include interpolation from highly correlated analogue gauging stations, flow percentile matching using a flow duration curve, backward rainfall interpolation or hydrological modelling procedures (Kottegoda and Elgy, 1977; Hughes and Smakhtin, 1996; Rees, 2008; Harvey *et al.*, 2012). The following section presents a methodology for flow infilling with a focus on interpolation from highly correlated, analogue gauging stations. The reason this method was chosen is due to the selected gauges exhibiting high correlations to other gauges in the catchment, also the method is suggested for flow infilling at sub-daily frequencies over long period datasets (Rees, 2008; Harvey *et al.*, 2012).

Interpolation by alternative gauging stations is useful because it can be applied to longer sequences of missing data (Rees, 2008). The first step involves obtaining a relationship between the gauge needing infilling (candidate) and one or more analogous gauging stations (donor) (Harvey *et al.*, 2012). Donor gauging stations are ideally upstream or downstream of the candidate gauging station, but can also be drawn from nearby catchments or hydrologically similar catchments (Rees, 2008). Next, the simplest approach is to multiply the relevant data values from the donor by some scaling factor, which is usually the ratio of the two catchment areas or the ratio of the respective mean flows, before transposing that data into the gaps in the candidate record (Harvey *et al.*, 2012).

The infilling process had three significant stages: finding correlations, calculating a multiplier and finally infilling gaps. At each of these stages multiple options were investigated in order to ensure an optimal infilling method was employed. The stages are described in each of the subsequent paragraphs.

5.2.1 Correlations

Individual datasets were examined for cross-correlations to identify the flow gauges that were most similar to each other. Finding the highest correlation for each gauge was done by correlating all 20 gauges against all 20 gauges. The datasets were also shifted along the time axis from -12 hours to +12 hours, with a correlation carried out at hourly time steps. Separate gauges on the same river are known to exhibit significant

60

lags in hydrograph peaks in response to rainfall events. Visual comparisons of hydrographs from the Wairarapa gauges showed consistent lags in the same event peaks, some up to 5 hours difference. The lag time correlations were carried out to allow for gauges to demonstrate more significant correlations despite differences in time.

The donor selection criteria for analogue infilling method was based on a high correlation between the candidate and donor time series over identical periods. Where correlations were similar, donors on the same river were given preference. Further, the geographical proximity, physiographical similarity and relative hydrological response of the two catchments were qualitatively assessed to further inform the donor selection process.

5.2.2   *Scaling factor*

Applying a scaling multiplier corrects instances where two highly correlated flow datasets possess significantly different magnitudes of flow. This could be due to differences in catchment size, catchment shape, rainfall distribution, etc. In these instances infilling from the donor flow gauge would exhibit good propagation of flow (high correlation), but match up poorly with actual flow volumes. Scaling factors were calculated from the ratio of candidate to donor mean flow values at long-term (LT) and monthly time scales. The monthly average scalar method (MO) calculates a ratio of flow between the two gauges for each month in the year and applies the relevant multiplier to data at the relevant dates. This MO scale option was investigated as a potential solution to correct for sub-annual variability between the candidate and donor datasets. The two scaling factors were compared to the donor infilling where no scale factor (NO) was applied.

5.2.3   *Infilling*

The two options considered for infilling were to transpose the scaled data straight from the donor gauge (Straight Infill, STR), or calculate the difference of flow between each data point of the donor gauge, apply the multiplier and add to the flow from the previous time-step of the candidate data (Difference Infill, DIF). The second option aimed to smooth out steps in flow volume where the dataset switched from candidate to donor flow records. Where the receding limb produced large negative values in the calculated difference record, any negative value was reset to the flow of the previous

time step. The infilled datasets were finally rounded to 3 decimal places to ensure continuity with the limits of the observed flow gauge records.

Summary

The various combinations for infilling that were tested in this study are presented in Table 5.4. In order to assess the usability of this data conditioning process the methodology was compared over sections of measured data (i.e. data was purposefully deleted, infilled and compared to the original data). Three sections were removed from all 13 gauges, this included a nine day winter period with event peaks, a 20 day summer period with event peaks and a 20 day spring low-flow period. These periods were compared to the original candidate data using quantitative dataset comparison metrics (NSE) and graphical comparisons. The process described above was used to interpolate gaps in the data, not to extrapolate datasets outside their original temporal boundaries. The results of the flow infilling investigation are presented in Section 5.3.

Table 5.4. Infilling techniques tested by this study. Techniques were applied to datasets comprising time when observed flows existed for both the candidate and donor gauges.

Acronym	Name	Details
STR-NO	Straight infill, No mean scaling	Direct infill from donor to candidate using scale factor of 1.
STR-LT	Straight infill, Long term mean scaling	Direct infill from donor to candidate using scale factor based on long-term mean flows.
STR-MO	Straight infill, Monthly mean scaling	Direct infill from donor to candidate using scale factor based on monthly mean flows.
DIF-NO	Difference infill, No mean scaling	Infill adding donor differences to candidate using scale factor of 1.
DIF-LT	Difference infill, Long term mean scaling	Infill adding donor differences to candidate with scale factor applied based on long-term mean flows.
DIF-MO	Difference infill, Monthly mean scaling	Infill adding donor differences to candidate with scale factor applied based on monthly mean flows.

5.3 Flow infilling results

The donor gauge selected for each flow gauge is presented below in Table 5.5. As expected, flow gauges positioned on the same river generally showed the best correlations (Kopuaranga, Mangatarere and Ruamahanga gauges) and correlations improved when adjusted for downstream flow lag time. Flow gauges that were in relatively close proximity to each other, but on different rivers, also showed good correlations. The various gauges stationed at the foot of the Tararua ranges (Ruamahanga 1, Waipoua 2, Waingawa, Waiohine and Tauherenikau) correlated highly with each other.

Table 5.5. For each gauge, the highest correlated gauge along with the relevant lag-times, correlation coefficients and distance apart (Seven gauges left out – see section 4.1.1).

No.	Candidate	Donor	Lag (hrs)	R	Distance (km)
3	Kopuaranga 1	Kopuaranga 2	2	0.975	8.5
4	Kopuaranga 2	Kopuaranga 1	-2	0.975	8.5
5	Mangatarere 1	Mangatarere 3	1	0.972	4.8
6	Mangatarere 2	Mangatarere 1	1	0.967	8.6
7	Mangatarere 3	Mangatarere 1	-1	0.972	4.8
8	Otakura	Ruamahanga 2	1	0.643	6.2
12	Ruamahanga 1	Waingawa	0	0.941	15.9
13	Ruamahanga 2	Ruamahanga 3	-4	0.925	28.9
14	Ruamahanga 3	Ruamahanga 2	4	0.925	28.9
16	Tauherenikau	Waiohine	0	0.869	6.7
17	Waingawa	Ruamahanga 1	0	0.941	17.7
18	Waiohine	Waingawa	0	0.930	33.5
20	Waipoua 2	Ruamahanga 3	3	0.909	9.2

Correlations

Eleven of the 13 gauges exhibited correlation coefficients greater than 0.90 in relation to their prospective donor datasets. The Tauherenikau gauge showed a correlation of 0.869 to the Waiohine gauge. The lowest candidate-donor correlation was the Otakura gauge, which is relatively poorly correlated to the Ruamahanga 2 gauge (0.643).

The highest correlations came from gauges that were situated on the same river. Two gauges on the Kopuaranga River, three gauges on the Mangatarere River and the two downstream Ruamahanga River gauges all correlated highly (> 0.950) to gauges on their associated rivers. A further trend is apparent in the clustering of correlations of "hill-country" gauges with similar catchment sizes. This cluster includes a network of

correlations between the Ruamahanga 1, Waingawa, Tauherenikau and Waiohine gauges. These gauges have an average catchment size of 112.4 km^2, with a standard deviation (σ) of 42.62. Compared to the average catchment size of all 13 catchments ($\mu = 243$ km^2, $\sigma = 442.62$), this subset of smaller catchments all drain out of the eastern slopes of the Tararua ranges and have similar geological characteristics (Begg and Johnston, 2000).

Multiplier and infilling

After identifying suitable donor gauges, scale factors were calculated. The multipliers considered for the flow infilling process, were (1) a ratio of the two flow gauges calculated across the entire length of their datasets (LT) and (2) a ratio of the two flow gauges calculated for each month across the entire length of their datasets (MO). These multipliers were compared to the donor flow infilled using no multiplier, presented as NO in the following sections. The long-term multiplier for each gauge is presented below in Table 5.6 along with the mean flow of the 13 gauges and their relative donors. Monthly multipliers were relatively similar to LT values and therefore are presented in Appendix A.

Table 5.6. Mean flow of Candidate and Donor datasets. Long-term ratio used for LT flow infilling, calculated from the mean of flows from the Candidate gauge divided by the mean of flows from the Donor gauge.

No.	Candidate	Mean Flow	Donor	Mean Flow	Long-term Ratio
3	Kopuaranga 1	2.605	Kopuaranga 2	3.498	0.766
4	Kopuaranga 2	3.498	Kopuaranga 1	2.605	1.305
5	Mangatarere 1	1.970	Mangatarere 3	4.506	0.434
6	Mangatarere 2	1.816	Mangatarere 1	1.970	0.918
7	Mangatarere 3	4.506	Mangatarere 1	1.970	2.304
8	Otakura	0.555	Ruamahanga 2	83.610	0.007
12	Ruamahanga 1	9.974	Waingawa	10.173	0.983
13	Ruamahanga 2	83.610	Ruamahanga 3	23.522	3.503
14	Ruamahanga 3	23.522	Ruamahanga 2	83.610	0.285
16	Tauherenikau	9.162	Waiohine	24.374	0.370
17	Waingawa	10.173	Ruamahanga 1	24.374	1.017
18	Waiohine	24.374	Waingawa	9.974	2.415
20	Waipoua 2	3.745	Ruamahanga 3	10.173	0.166

The multipliers were applied to donor datasets before the infilling methods were employed to fill the artificial gaps. Table 5.7 presents the NSE scores for the two infilling methods with each of the two multipliers applied as well as just the donor

dataset (NO) in comparison to the observed flow. Excellent values for NSE, as prescribed in section 3.3.2, are in bold.

Table 5.7. Nash Sutcliffe Efficiency scores for Straight (STR) and Difference (DIF) infilling methods using one of three multipliers; Monthly (MO), Long-term (LT) or None (NO). The 13 infilled datasets were compared to the original datasets across three intentionally deleted periods in the original record.

NSE		STR			DIF		
Candidate	Donor	MO	LT	NO	MO	LT	NO
Kopuaranga 1	Kopuaranga 2	**0.920**	**0.926**	**0.944**	**0.926**	**0.932**	**0.950**
Kopuaranga 2	Kopuaranga 1	**0.857**	**0.878**	**0.962**	**0.852**	**0.876**	**0.957**
Mangatarere 1	Mangatarere 3	0.707	0.786	-4.887	0.660	0.739	-4.559
Mangatarere 2	Mangatarere 1	**0.875**	**0.932**	**0.897**	**0.887**	**0.940**	**0.910**
Mangatarere 3	Mangatarere 1	**0.869**	**0.892**	0.433	**0.836**	**0.857**	0.460
Otakura	Ruamahanga 2	-2.201	-2.154	-181683	-2.115	-2.190	-136113
Ruamahanga 1	Waingawa	**0.865**	**0.862**	**0.865**	0.757	0.746	0.753
Ruamahanga 2	Ruamahanga 3	0.509	0.549	0.361	0.522	0.558	0.531
Ruamahanga 3	Ruamahanga 2	0.787	0.797	-2.595	0.436	0.409	-1.560
Tauherenikau	Waiohine	0.711	0.726	-4.999	0.650	0.697	-4.027
Waingawa	Ruamahanga 1	**0.814**	**0.809**	**0.819**	0.601	0.600	0.612
Waiohine	Waingawa	**0.950**	**0.944**	0.484	0.818	0.811	0.396
Waipoua 2	Ruamahanga 3	0.794	**0.816**	-13.894	0.705	0.756	-9.810

Multiplier

From the above tables, NSE values are generally within the excellent or good categories for the LT and MO scale options. The infilling carried out with no multiplier (i.e. just the donor dataset) does not perform as well, especially at gauges that have lower correlations to their donor gauges (Otakura, Tauherenikau, Waipoua 2). While the two multiplier options are effective at improving NSE compared to the original donor dataset, the more complex MO multiplier does not produce perceptibly better results than the simple LT option. In fact, despite scores being relatively similar, the LT option produces superior results at more gauges. The similarity of the LT and MO datasets is further apparent on the hydrographs in Figures 5.4 and 5.5. The infill datasets calculated using the LT (yellow) and MO (purple) multipliers are almost indistinguishable on the graph, except at peak flows, and their respective scatter plots are also very similar.
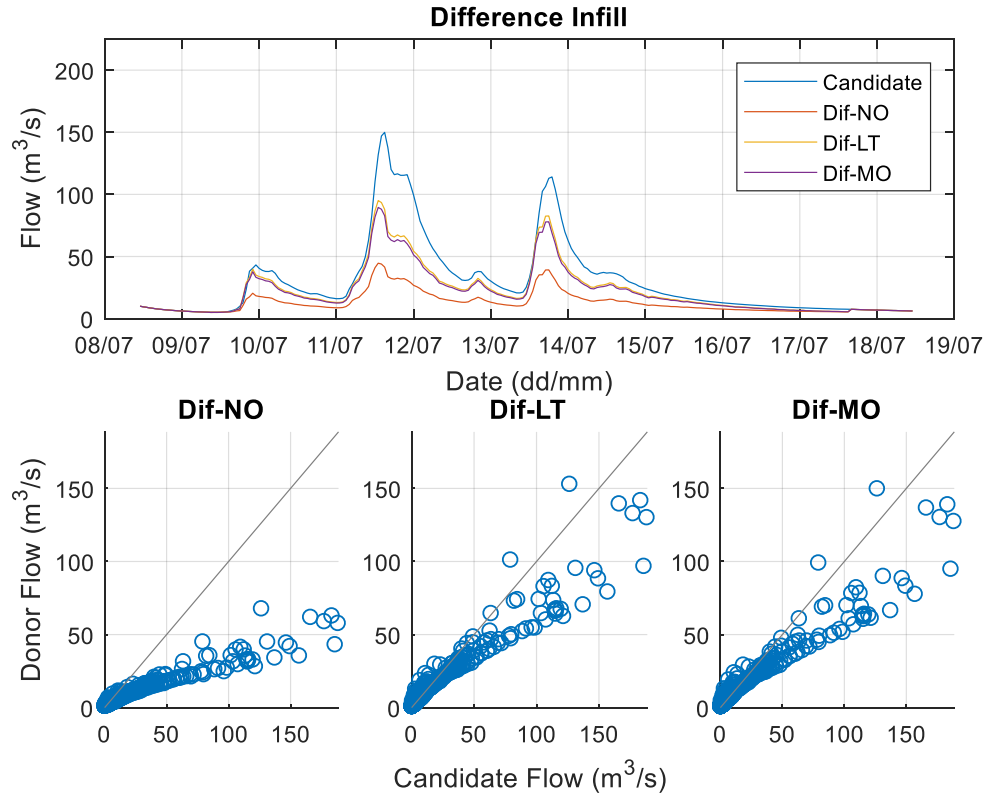
Figure 5.4. Difference infilling method at the Mangatarere 3 gauge using Mangatarere 1 donor data, compared to candidate data (blue line). Applied with long-term (LT), monthly (MO) and no (NO) multiplier with accompanying scatter plots for each multiplier option.

In the instance of flow infilling at the Kopuaranga 1 and Kopuaranga 2 gauges, NSE values indicate a best fit of donor data to the candidate data when the donor dataset is applied without a multiplier. The hydrograph and scatter plots in Figure 5.5 support this statement, showing the MO and LT infill options under-predicting hydrograph peaks compared with the candidate data and the NO dataset (i.e. just donor flow data). The two gauges have a high correlation at 0.975.

The scatter plots from Figures 5.4 and 5.5 show good correlation of donor and candidate flow at low-flow values (approx. < 25 m^3/s), however there is significant deviation from the goodness of fit line at higher flows. The underestimation at high flows is observable on the hydrograph in Figure 5.4. The peaks of the candidate dataset (blue) are much larger than the multiplier corrected donor datasets (yellow and purple). The deviation from a good correlation occurs at flows approximately above 50 m^3/s at the Mangatarere 3 gauge. This indicates the multiplier works best at low and medium flow but not at high flows. Across the entire dataset, low and medium flows are more

frequent than the high volume peak flows. As such, the multiplier calculated from the two datasets is likely to be largely influenced by the lower flows, which results in poor estimation of the extreme high flow volumes.
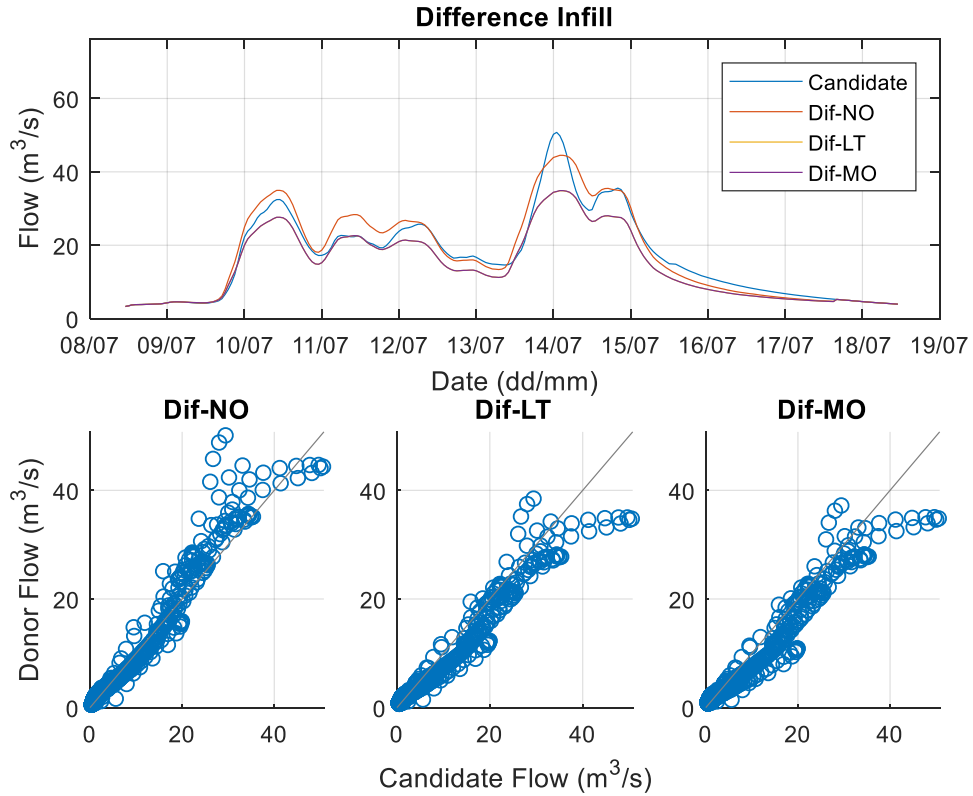
Figure 5.5. Difference infilling method at the Kopuaranga 1 gauge using Kopuaranga 2 donor data, compared to candidate data (blue line). Applied with Long-term (LT), Monthly (MO) and No (NO) multiplier with accompanying scatter plots for each multiplier option.

Infilling methods

The quantitative NSE metric shows both infilling methods performing well, with generally excellent or good values present. With respect to NSE values, the STR method produced 20 values in the excellent bracket compared to the DIF method's 13. Upon observation of graphs however, the DIF method looks to preserve the propagation of flow marginally better than the STR method with no large steps between donor and candidate datasets (Figure 5.6).
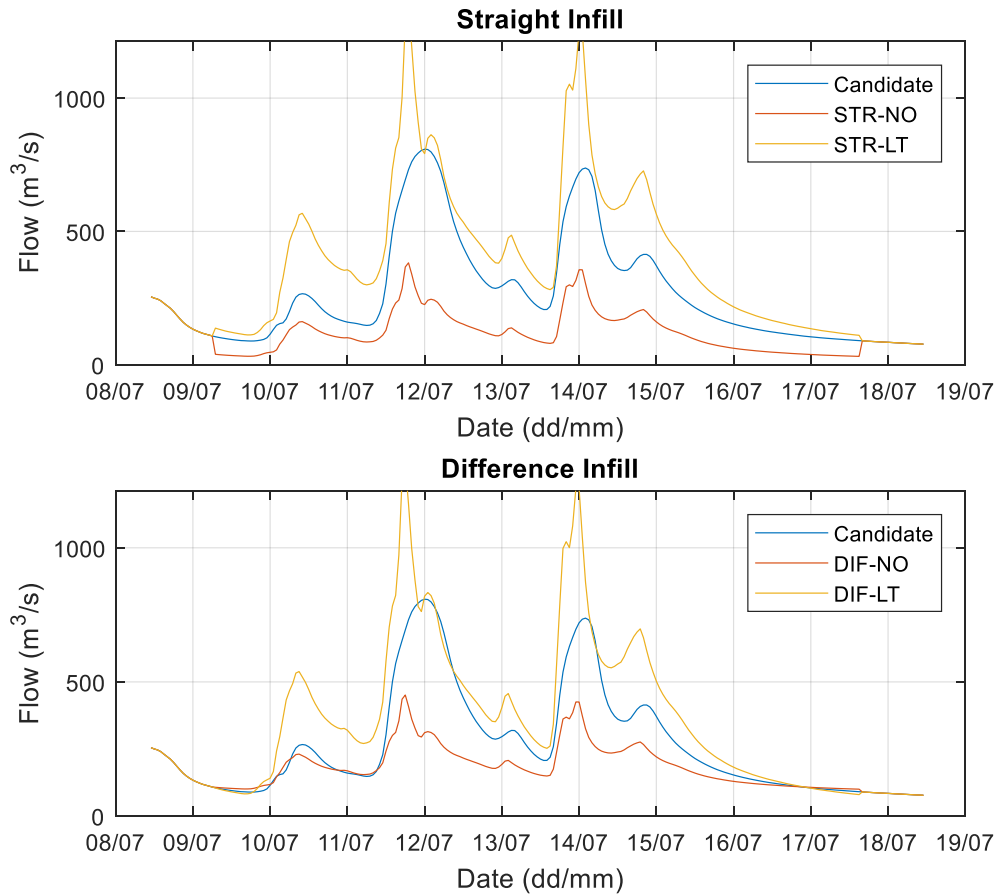
Figure 5.6. Straight (top) and Difference (bottom) infilling methods (yellow line) at the Ruamahanga 3 gauge using Ruamahanga 2 donor data (red line), compared to candidate data (blue line). Applied with long-term (LT) multiplier.

5.4   Flow infilling discussion

Correlations

It is apparent from the infilling results that the initial high correlation of gauges is important. The Otakura infilled dataset demonstrates the worst results across all indices. The Otakura gauge and its donor gauge Ruamahanga 2 are poorly correlated (0.643). The two gauges, while only 6.2 km apart, have drastically different flow totals and catchment areas. It is clear from Figure 5.7 that despite a long-term multiplier of 0.007 the Ruamahanga 2 donor dataset (yellow) is still significantly larger than the Otakura observed flow. The generally good correlations across the other 12 gauges (>0.85) produced mostly excellent results across the measured performance statistics for all of the options investigated (DIF vs STR and LT vs MO). This meant flow infilling from alternate gauges could be carried out with a high level of confidence.
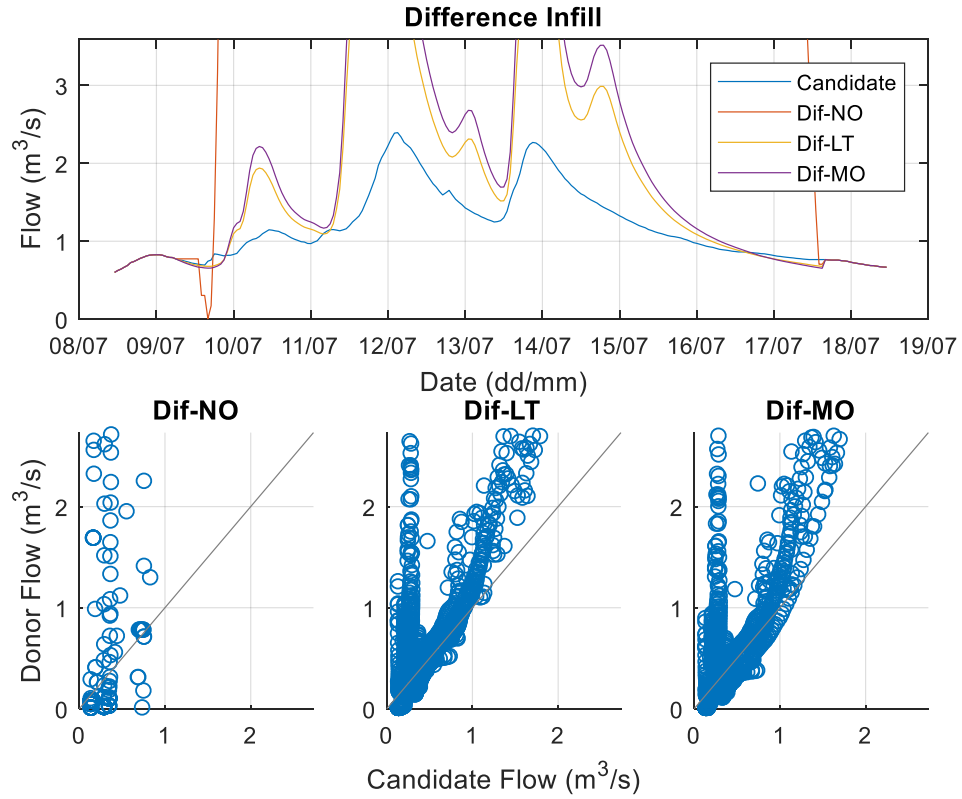
Figure 5.7. Difference infilling method at the Otakura gauge using Ruamahanga 2 donor, compared to candidate data (blue line). Applied with Long-term (LT), Monthly (MO) and No (NO) multiplier with accompanying scatter plots for each multiplier option.

The scatter plots from figure 5.7 show the strong positive bias in the Ruamahanga 2 donor dataset (Dif-NO) and relatively poor correlation at low flows. The limits of the graph are defined by the observed Otakura stream gauge max flow and much of the original Ruamahanga 2 data points are above this limit. The two multiplier datasets are effective at bringing the Ruamahanga dataset down towards a better correlation with the Otakura candidate data, however there remains a significant portion of flow values between 0 and 0.5 m^3/s that do not correlate with the multiplied Ruamahanga datasets. As seen on the hydrograph in Figure 5.7 the multiplied datasets display larger peaks compared with the original Otakura dataset explaining some form of positive bias, however it is difficult to discern where the overestimation of flow by the infilling datasets, below 1 m^3/s is occurring. These artefacts are mostly functions of the infilling method and application of the multiplier, however they are significantly more apparent at the sites that exhibit poor correlations between donor and candidate gauges. This demonstrates the importance of a good initial correlation to reduce errors related to infilling and scaling.

Multiplier

The two multiplier options tested in this investigation produced relatively similar results with respect to the NSE statistics. Further, the hydrograph comparisons show LT and MO options are almost indistinguishable from each other. Values for the monthly multiplier were quite similar to the more representative long-term multiplier. This shows that seasonal flow changes, however significant, occur at both donor and candidate gauges acting to maintain a steady flow ratio throughout the year. In terms of application, the LT multiplier is much simpler to calculate and apply to the donor gauge dataset.

Infilling

The straight infilling and difference infilling methods produced very similar results across the quantitative statistics and showed similar propagations in the hydrographs. The difference method was expected to perform best because it erases steps between donor and candidate flow datasets. However the straight method produced marginally better statistics. This is largely because the donor datasets compared in this research were very similar to their candidate gauges, as such, the correction for stepping in the difference method is less apparent than if the datasets were more different. The large steps between donor and candidate flow are a feature of many hydrological infilling attempts (Harvey *et al.*, 2012), therefore the difference infilling method can be considered superior.

5.5   Flow infilling limitations

The calculation of the scale factors is likely to favour low/medium-flow instead of high flow event peaks. There are more data points at the lower flow volumes and multiplying an extreme by the flow average will result in attenuation of that extreme. This is apparent in the Kopuaranga 1 infilling graph (Figure 5.5), in which the MO and LT infill options under-predict the event peak flow whereas the NO multiplier option performs best. With regards to the accompanied scatter plots, flow at low flows correlates well with observed data and there is a larger spread at higher flows. The Kopuaranga 1 and 2 gauges can be considered a special case because of their extremely high similarity in flow dynamics ($R^2 = 0.975$). However, this application of applying a multiplier to two very similar datasets has shown that the calculated ratio of the two gauges is dominated by the high frequency low flow. This results in attenuation of the

70

extreme values (i.e. peak flows) when applied. A potential solution to this would be to calculate two unique high flow and low flow multipliers, whereby flow recorded at the donor gauge above a specified threshold is scaled with respect to the high flow ratio and similarly for the low-flow measurements.

The Ruamahanga 3 gauge (at Wardell's Bridge) is approximately 30 km upstream of its donor gauge, Ruamahanga 2 (at Waihenga Bridge). As expected the flow recorded at the donor gauge further downstream is significantly larger. This difference in volume is mostly corrected for by the multiplier, however the general shape of the peaks on the hydrograph during event flows is quite different between the two datasets (Figure 5.8). The upstream Ruamahanga 3 gauge (blue line) spikes and recedes quickly, whereas the downstream Ruamahanga 2 gauge (purple), with multiplier applied, takes much longer to recede following the event peak. This is most likely due to input from the Waingawa, Taueru and Waiohine rivers that flow into the Ruamahanga River between where the two gauges are situated, which acts to maintain high flow periods for longer.
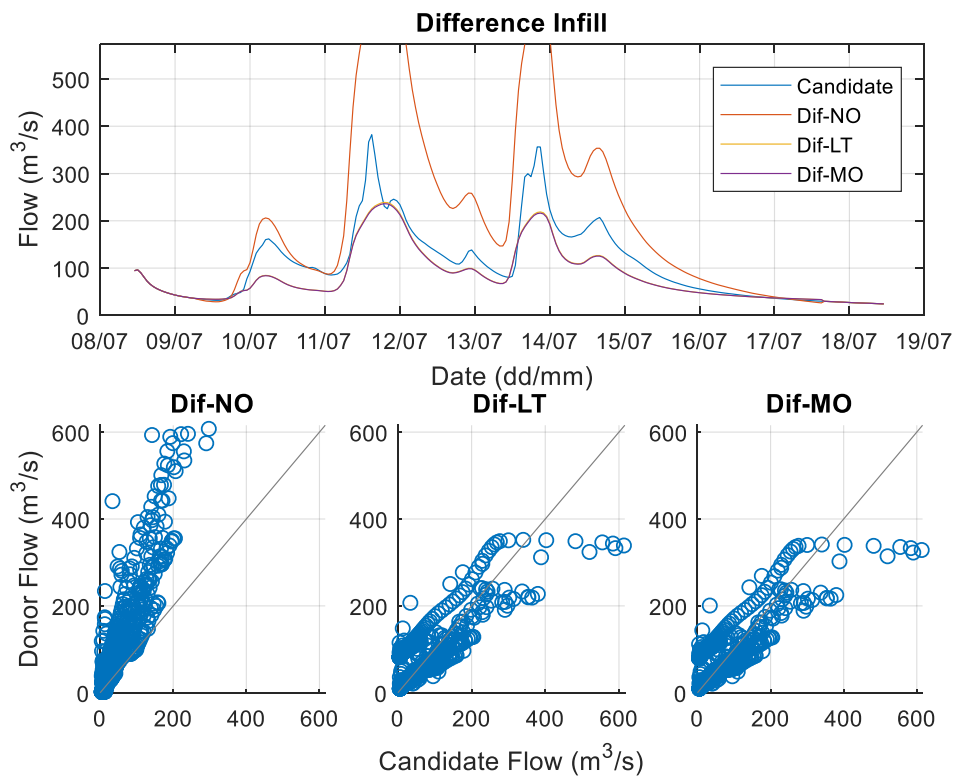
Figure 5.8. Difference infilling method at the Ruamahanga 3 gauge using Ruamahanga 2 donor data, compared to candidate data (blue line). Applied with Long-term (LT), Monthly (MO) and No (NO) multiplier with accompanying scatter plots for each multiplier option.

The multiplier has corrected flow volumes to correlate well with the candidate gauge (Figure 5.8), however the differences in shape of the peak flows is not corrected for by the infilling process. This has significant implications for baseflow investigations as much of this work is based on the shape of streamflow recession curves. The shape of the hydrograph is difficult to correct for. This is the reason for initial correlations to find the donor gauge with the best fit to the candidate gauge. The equipercentile approach of flow infilling with respect to equivalent flow on a flow duration curve could offer solutions to this problem as it matches donor flow to candidate flow with respect to frequency, rather than time. An investigation into this approach to correct for the differences in hydrograph shape is therefore recommended.

A limitation of this method is the inability to extrapolate data. Flow infilling was carried out in gaps of datasets. Specifically for the difference infill method, this is absolutely necessary in order for the calculation to work. This means that flow infilling is limited by the bounds of the start and end points of a dataset, even if a much longer, highly correlated donor gauge is available for infilling. Technically, the straight infilling method could be used to extend a dataset outside of its original temporal bounds, however this is not common (Harvey *et al.*, 2012). Further, something that was not specifically targeted in this investigation was the temporal realms of applicability of the flow infilling method. This investigation removed data from observation datasets at durations of nine and 20 days in order to compare the donor data to the candidate data. The results suggest the flow infilling worked well at these time scales, however no upper limit was established. Further investigation into the time uncertainty limits of the flow infilling processes would be beneficial.

5.6   Flow infilling concluding remarks

The straight and difference infilling methods for transposing donor flow data to gaps of candidate flow data were compared across three unique periods over the 2012 hydrological year – a summer and winter flow event and a spring low-flow period. This investigation was carried out at 13 gauges in the Wairarapa region. Two scaling options were also investigated: long-term and monthly. The NSE performance indicator was used to demonstrate goodness of fit. Visual graphing assessments of hydrographs and scatter plots were also employed.

It is clear from this investigation, that the highly correlated gauges produced the best comparative results. These gauges presented excellent quantitative statistics and graphical comparisons mostly irrespective of the infilling or multiplier options considered. Therefore, this should be the first consideration for future flow infilling applications.

Of the scaling options, the two multipliers improved the dataset comparison with respect to NSE. Where no multiplier worked best, the donor and candidate gauges were already so similar that a multiplier, influenced by more low-flows than high flows, resulted in attenuation of the event peaks. A comparison between the MO and LT methods showed that there was no clear superior option between the two across the 13 gauges. The NSE indicator showed the LT option performed better than the MO option in many catchments, but only marginally. Inspection of the hydrograph and scatter plots shows the two options were very similar. The LT option was chosen as the preferred option because of its comparative simplicity to implement and because the more complex MO option demonstrated no apparent advantages.

The straight and difference infilling methods performed relatively similarly with respect to NSE values. On the hydrograph the straight infill method produced significant artefacts where flow switched from candidate to donor flows and vice versa. The difference infill method performed better because it calculated the flow at time step (t) by adding the difference in flow of the donor gauge to the flow of the candidate gauge from time step (t-1), effectively assimilating the donor gauge data into the candidate gauge data at the point of switch-over. As such, the Difference infill method was deemed to be the method of choice for this investigation moving forward.

The deleted periods in the observed flow were restored and the difference infilling method was applied to actual gaps in the data with the long-term multiplier. The process was applied within the start and end of the candidate datasets. Table 5.9 presents the various infilling that was carried out for each of the 13 gauges, while Figure 5.9 displays the relative change between the original and infilled datasets. Actual infilling was at similar scales to the artificial gaps taken out in this investigation (1 to 32 days) with one exception being the Mangatarere 3 gauge, which had a large gap of approximately 120 days of missing data. The infilling process has allowed for the subsequent investigations in this research to be carried out over longer periods.

Table 5.8. Actual flow infilling at six of the 13 gauges over the investigation period. Listed with their donor gauges, total number of data points infilled and longest continuous gap infilled.

Candidate	Donor	Total no. points	Longest gap
Kopuaranga 1	Kopuaranga 2	28	17 hours
Kopuaranga 2	Kopuaranga 1	862	779 hours (32 days)
Mangatarere 1	Mangatarere 3	1161	394 hours (16 days)
Mangatarere 3	Mangatarere 1	2892	2891 hours (120 days)
Tauherenikau	Waiohine	289	117 hours (5 days)
Waipoua 2	Ruamahanga 3	1036	248 hours (10 days)
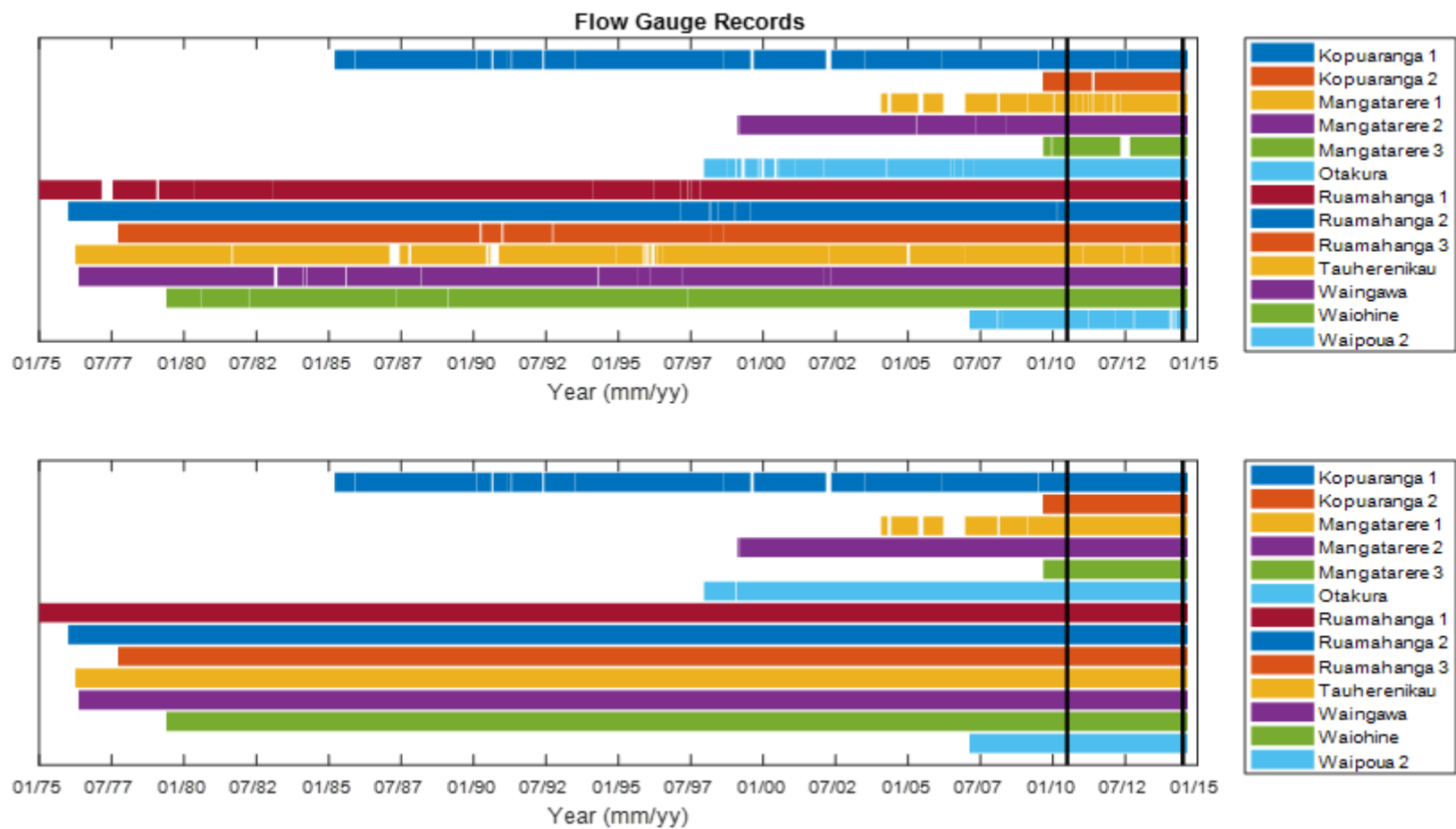
Figure 5.9. Duration of 13 Wairarapa flow gauges showing gaps (top) and after infilling (bottom) and 2011-2014 hydrological years "Investigation Period" (black vertical lines)

The previous chapter describes the flow infilling process carried out on gaps in observed flow datasets. Recursive digital filters calculate flow at each timestep from the value at the previous timestep and therefore, a continuous time series is required. The flow infilling allowed for baseflow separation techniques over the investigation period 1st July 2010 to 30th June 2014. An investigation into baseflow separation techniques is in accordance with the main aims of this research; to inform groundwater-surface water hydrological models.

While difficult to directly measure, baseflow can be inferred using digital frequency analysis of a streamflow dataset. The baseflow contribution to a stream as a portion of total streamflow is known as the baseflow index (BFI). Estimates of baseflow contribution were carried out on 13 flow gauge datasets in the Wairarapa region using Q90/Q50 from flow duration curves, and BFI from the Eckhardt and Bump and Rise recursive digital filters. These methods were chosen because they are relatively simple and are applied to streamflow datasets. Historical records of streamflow are readily available across New Zealand, much more so than chemical tracer data. As such the methods employed in the investigation have potential to be applied by water management organisations quickly and easily.

This chapter presents the results from the baseflow separation investigation applied to the datasets described in Chapter 5 (Section 5.1). It is divided into five sections, the first of which presents the methodologies employed for this baseflow investigation. This is followed by Section two which presents the findings of the various baseflow separation methods and their relation to specific catchment characteristics. Section three presents a discussion of the relative performance of the separation methods with respect to performance indicators identified in the literature, while Section four outlines the limitations surrounding this part of the investigation. The final section presents a summary of baseflow contributions to surface waters in the Wairarapa valley based on results obtained from this investigation.

6.1 Baseflow separation methodology

For the purpose of this investigation, baseflow should be considered to be the portion of total streamflow that is provided by groundwater. Under this definition, baseflow maintains the flow of a stream when quickflow conditions cease (Nathan and McMahon, 1990). Many calculations for baseflow exist that express baseflow as a fraction of total flow, known as the baseflow index (BFI). Estimates of BFI for this investigation were produced by the Eckhardt Digital Filter (EK), the Bump and Rise Digital Filter (BR), and Flow Duration Curves (FDC) (Eckhardt, 2005; Collischonn and Fan, 2013; Stewart, 2015). The baseflow separation methods described below were applied to 13 infilled observed flow gauge datasets described in Chapter 5 over the investigation period from 1 July 2010 to 30 June 2014.

6.1.1 Eckhardt Recursive Digital Filter

The Eckhardt filter is a development of the Chapman filter, which in turn is based on the original Lyne and Hollick recursive digital filter (Nathan and McMahon, 1990; Chapman and Maxwell, 1996). Recursive digital filters separate the low frequency, high magnitude flood flow signal from the high frequency, low magnitude baseflow signal along a time-series. The original Lyne and Hollick recursive digital filter can be written as the following equation:

$$q_{f(t)} = \omega q_{f(t-1)} + \frac{(1 + \omega)}{2} \left(q_{(t)} - q_{(t-1)} \right) \qquad (6.1)$$

where $q_{f(t)}$ is the filtered quickflow response at time step t or the previous time step $t - 1$, $q_{(t)}$ is the total stream flow at time step t, and ω is the filter parameter. In terms of baseflow: $q_{b(t)} = q_{(t)} - q_{f(t)}$. The suggested value for the filter parameter is between 0.9 and 0.95 (Nathan and McMahon, 1990).

The Chapman filter addresses the assumption that outflow from reservoirs remains constant during periods of no direct runoff (Chapman, 1991). In fact, without any recharge groundwater contribution recedes as a function of general catchment drying and easing hydraulic pressure gradients. This assumption was reaffirmed by chemical tracer analysis of streamflow (Chapman and Maxwell, 1996). The Chapman filter replaces the filter parameter (ω) with a master recession parameter to resolve this dynamic response. In terms of baseflow, the Chapman filter can be written as:

$$q_{b(t)} = \left(\frac{\omega}{2 - \omega}\right) q_{b(t-1)} + \left(\frac{1 - \omega}{2 - \omega}\right) q_{(t)} \qquad (6.2)$$

where ω is the master recession constant; calculated as the ratio of flow to preceding flow (i.e. $\sum\left(\frac{q_{(t-1)}}{q_{(t)}}\right)$) for all instances of receding flow (Chapman, 1991).

The Eckhardt filter further builds on its predecessors by assessing the idea that baseflow is proportional to storage (Eckhardt, 2005). The Eckhardt filter incorporates an additional parameter to represent the aquifer storage component, BFI_{max}. In terms of baseflow, the Eckhardt digital filter can be written as:

$$q_{b(t)} = \frac{(1 - BFI_{max})\omega q_{b(t-1)} + (1 - \omega)BFI_{max}q_{(t)}}{(1 - \omega BFI_{max})} \qquad (6.3)$$

where $q_{b(t)} \leq q_{(t)}$. BFI_{max} constrains the maximum value of the baseflow index, which can be changed depending on inferences of the permeability of the catchment.

In a sensitivity analysis investigation of the two parameters, Eckhardt (2012) states the recession constant has a stronger influence on the output BFI. However, this research has found that the BFI_{max} parameter also has a significant impact on the output BFI. Eckhardt (2005) suggests three possible values for the parameter BFI_{max}; 0.8 for perennial streams with porous aquifers, 0.5 for ephemeral streams with permeable aquifers, and 0.25 for perennial streams with hard rock aquifers. These values should be considered as a first approximation (Eckhardt, 2008). Critiques of the BFI_{max} parameter state that it lacks physical basis and the actual value does not relate to natural catchment characteristics (Szilagyi, 2004). Optimising the BFI_{max} parameter with respect to the physical characteristics of the catchment and/or stream is a continual field of investigation. For example, Collischonn and Fan (2013) employ a backward filter method based on the master recession parameter to estimate BFI_{max}, and Zhang *et al.* (2013) calibrated BFI_{max} using a chemical conductivity mass balance method based on baseflow and surface runoff relationships. The Collischonn and Fan (2013) has been employed in this investigation and is described in Section 6.1.4.

Nathan and McMahon (1990) suggest the filters should be passed through three times, forward, backward and forward again. The number of passes affects the degree of smoothing. An experiment by Arnold *et al.* (1995) showed that the value for BFI

reduced by 17% for a second pass and further 10% for a third pass. However, passing a filter over a dataset twice in the forward direction and once in the backward direction could result in a phase distortion along the time axis. As such, Spongberg (2000) suggests passing the filter over a dataset twice. Once in the forward direction and once in reverse. This second pass nullifies any distortion along the time axis from the first pass of the filter.

6.1.2   Bump and Rise recursive digital filter

The bump and rise method for baseflow separation is based on evidence from tracer separation analysis as well as historical literature. Stewart (2015) argues that baseflow response to a rainfall event is temporally divided into two processes: an initial spike resulting from a hysteretic flow response (bump) followed by a steady flow increase due to a rise in overall catchment wetness (rise).

Recent chemical tracer separation analysis demonstrates baseflow on a hydrograph possessing an initial rapid response to a rainfall event (Stewart, 2015). The BR method attempts to recreate this condition by incorporating a 'bump' parameter (f). The second constant 'rise' factor is based off the Hewlett and Hibbert (1967) graphing method, which separates the total flow hydrograph into "quickflow" and "delayed flow" components by arbitrarily projecting a line of constant slope (k) from the beginning of any stream rise until it intersects the falling side of the hydrograph (Figure 6.1 - pink dashed line).
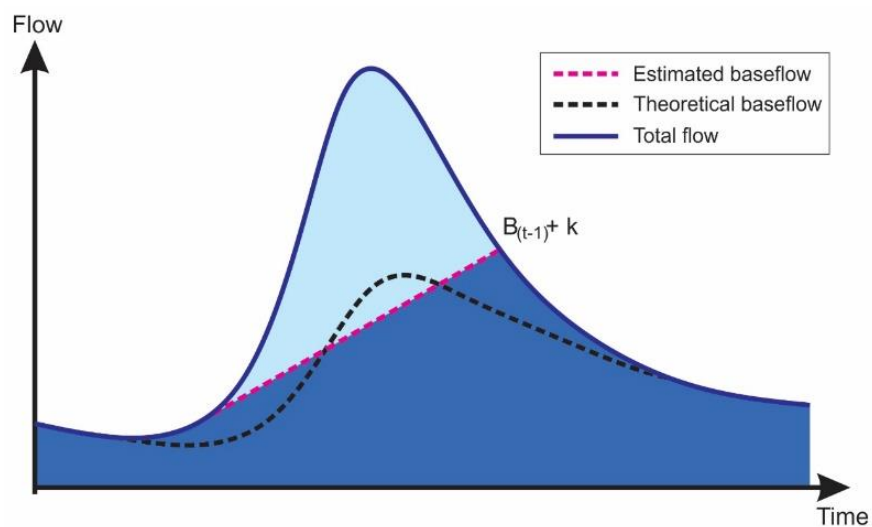
Figure 6.1 Hewlett and Hibbert (1967) constant rise method for baseflow separation (pink dash) and contemporary understanding of baseflow response (black dash). Adapted with changes from Brodie and Hostetler (2005).

The BR hydrograph separation is represented in the following equations:

$$q_{b(t)} = q_{b(t-1)} + k + f\left(q_{(t)} - q_{(t-1)}\right) \text{ for } q_{(t)} > q_{b(t-1)} + k \qquad (6.4)$$

$$q_{b(t)} = q_{(t-1)} \text{ for } q_{(t)} \leq q_{b(t-1)} + k \qquad (6.5)$$

where f is the constant fraction of the increase or decrease of streamflow during an event (bump) and k is the slope of the dividing line (rise). Calibration of the two parameters is by fitting relevant tracer data, or using an optimising process if tracer data is not available (Stewart, 2015). Quickflow and baseflow are generated by significantly different processes and Stewart (2015) asserts that the two processes possess unique recession curves that add to the total flow recession curve. When individual recession analyses are carried out on separated quickflow and baseflow, it is possible to discern unique recession curves that make up the total streamflow recession curve (Figure 6.2) (Stewart, 2015). The late recession is dominated by baseflow recession, whereas early recession is mostly quickflow recession. The "bump" aspect of this filter aims to represent the switch from early to late streamflow recession (pink dash).
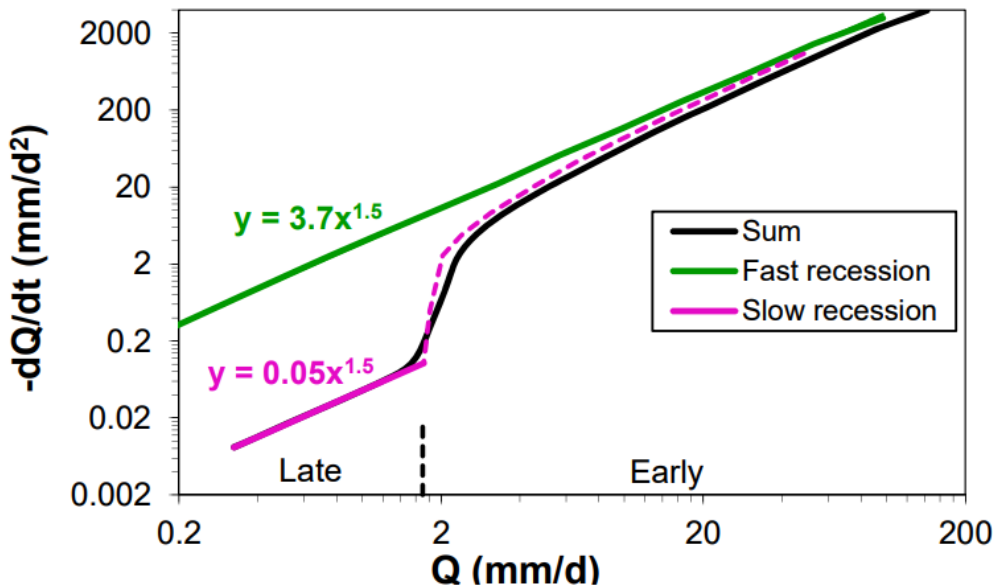
Figure 6.2. Recession plot of stream flow (sum), baseflow and fast recession. The sum is close to slow recession at low flows (baseflow) and moves toward the fast recession curve at high flows. The dashed line represents the "bump" in baseflow response (Stewart, 2015).

6.1.3   Flow Duration Curves

Low-flow and baseflow characteristics can also be investigated with respect to frequency analysis by deriving relationships between the magnitude and frequency of streamflow discharge. Flow duration curves plot the ranked flow of a dataset against a 'percentage exceeded' scale (Figure 6.3). This demonstrates the percentage each flow rate occurs or is exceeded with respect to all other flows in the dataset (Smakhtin, 2001a). FDC analysis is widely used to investigate stream flow characteristics (Welderufael and Woyess, 2010).
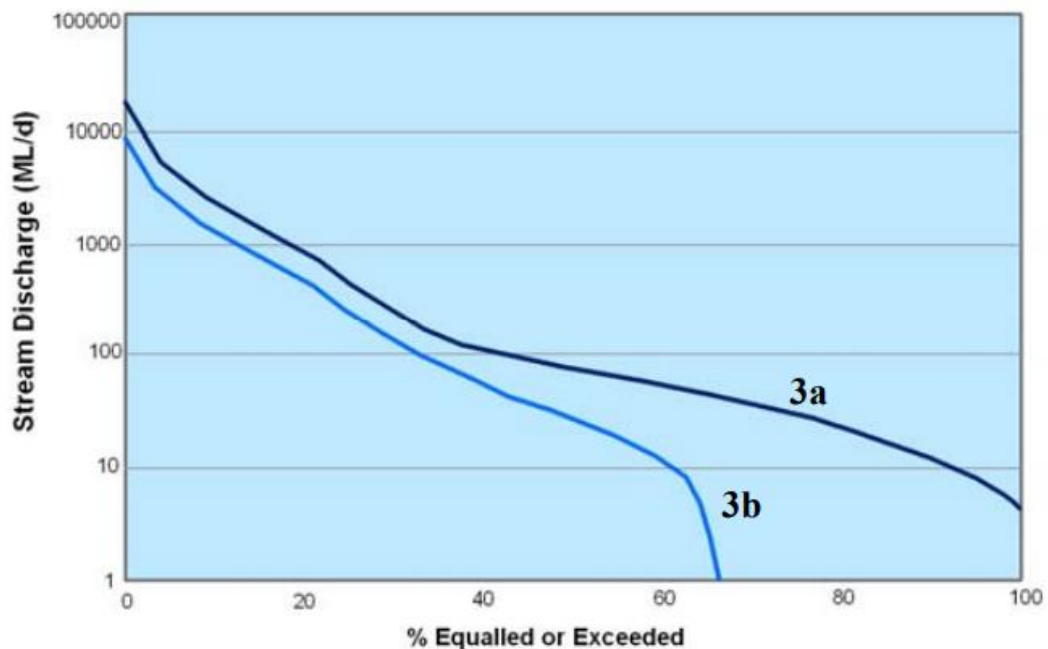
Figure 6.3. Flow duration curves for examples of (3a) high baseflow and (3b) low baseflow streams (Brodie & Hostetler, 2004).

Al-Faraj and Scholz (2014) demonstrate the use of FDC analysis to calculate a low flow index (Q_{90}/Q_{50}). The ratio of Q_{90} (flow value which is equalled or exceeded 90% of the time) to Q_{50} (flow value which is equalled or exceeded 50% of the time, also median flow) can be used as an index for the long-term, low flow characteristics of a stream (Nathan and McMahon, 1990; Smakhtin, 2001b). While useful for initial investigations into catchment characteristics, the Q90/Q50 metric should be

considered somewhat arbitrary as it is not linked directly to groundwater processes (Smakhtin, 2001b; Brodie and Hostetler, 2005).

Collischonn and Fan (2013) use the Q90/Q50 metric as a comparison to estimates of BFI produced by the EK filter. In this investigation, Q90/Q50 was similarly calculated from the FDC of each gauge and used as an initial comparative assessment of low-flow characteristics for each catchment.

6.1.4 *Parameter estimation*

Table 6.1 lists the different baseflow separation methods employed in this investigation, along with the parameters required. In the case of the Eckhardt digital filter, results were initially produced using a BFI_{max} of 0.8 for perennial streams with porous aquifers as suggested by Eckhardt (2008). The backward filter (BF) estimation for BFI_{max} suggested by Collischonn and Fan (2013), described below, was also implemented in order to assess the effects of catchment specific estimation for the BFI_{max} parameter.

Table 6.1. Parameters required for baseflow index estimation of the methods applied to hourly streamflow in the Wairarapa region, New Zealand.

Method	Name	f*	k*	BFI_{max}*	ω*	Ref.
Q90/Q50	Q90/Q50	-	-	-	-	Al-Faraj & Scholz (2014)
Eckhardt (default)	EK(def)	-	-	0.8	Yes	Eckhardt (2007)
Eckhardt (optimised)	EK(opt)	-	-	Yes	Yes	Collischonn & Fan (2013)
Bump & Rise	BR	Yes	Yes	-	-	Stewart (2015)

*f is the bump fraction, k is slope parameter, BFI_{max} is the maximum value of the baseflow index and ω is the recession constant.

Recession constant for EK(def) and EK(opt)

The recession curve is the specific part of the hydrograph after the peak where streamflow reduces. The slope of the recession after event peaks can be calculated as the gradient of flow at each time step compared to the previous time step (Chapman, 1999). To estimate the recession constant, recession values are taken over the entire dataset where flow is receding and averaged. This method has been used to estimate the recession parameter ω for the Eckhardt filter, and also to inform the Collischonn

and Fan (2013) backward filter algorithm for estimating BFI_{max} described below. Further, total flow recession constants were compared to baseflow recession constants in order to optimise the parameters for the BR filter (below).

BFI_{max} parameter for EK(opt)

The first application of the Eckhardt filter uses the recommended values for BFI_{max} (Eckhardt, 2008). All of the catchments were considered to possess perennial streams with porous aquifers and therefore a BFI_{max} of 0.8 was employed at all 13 flow gauges. The Collischonn and Fan (2013) method was used to estimate BFI_{max} for the EK(opt) filter. This technique applies a reverse filter to a hydrograph. By starting the filter at the end of a recession where baseflow is assumed to equal total flow ($b_t = q_t$), the backward filter then multiplies the initial estimate of baseflow by the recession parameter 'ω' to generate baseflow at the previous time step. The backward filter is written as:

$$q_{b(t-1)} = \frac{q_{b(t)}}{\omega} \qquad (6.6)$$

where $q_{b(t)}$ is baseflow at time step t and $q_{b(t)} \leq q_{(t)}$. By applying this filter to the hydrograph record and using the mean value as BFI_{max} in the Eckhardt filter, the BFI could be estimated. BFI_{max} estimates from 15 Brazilian flow gauges were compared to the Q90/Q50 statistic and showed good correlations (Figure 6.4).
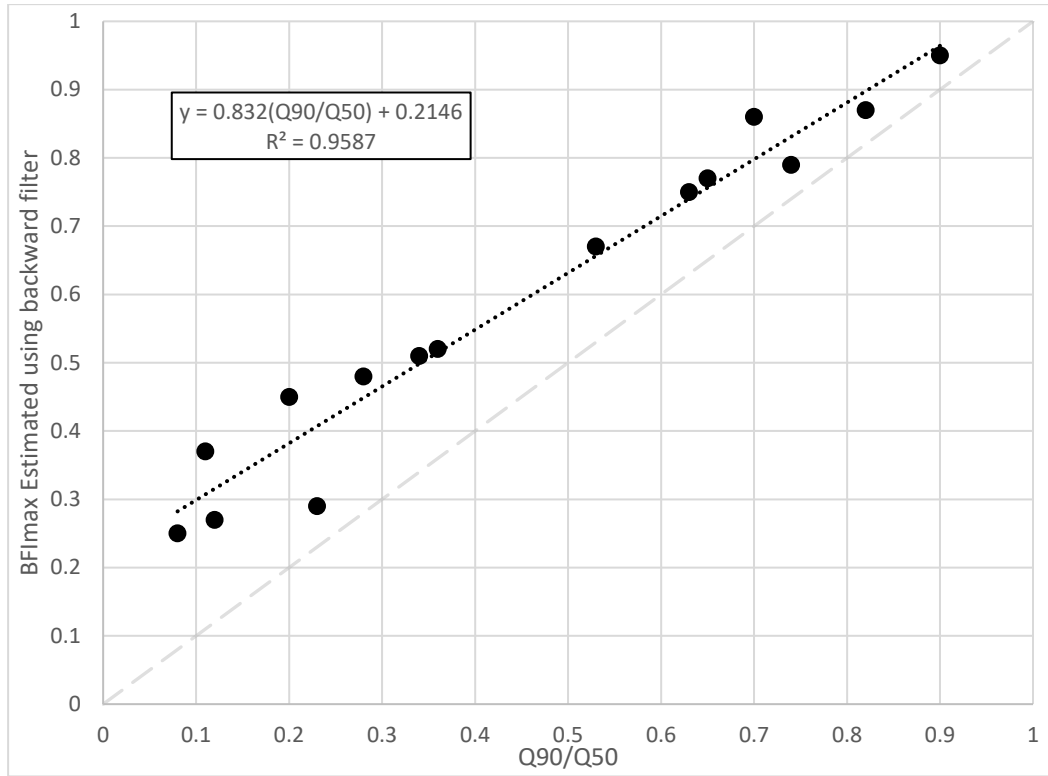
Figure 6.4. Relation of Q90/Q50 and backward filter estimation of BFImax (blackdots) with trend line (black dashes) for 15 gauges in Brazil (Collischonn & Fann, 2013).

The correlations established by Collischonn and Fan (2013) allowed them to simplify calculating BFI_{max} and BFI as a linear relationship to Q90/Q50, shown in the equations below.

$$BFImax = 0.832 \frac{Q_{90}}{Q_{50}} + 0.2146 \qquad (6.7)$$

$$BFI = 0.850 \frac{Q_{90}}{Q_{50}} + 0.163 \qquad (6.8)$$

The three indices – Q90/Q50, BFI_{max} and BFI – are derived from simple flow duration curves in this instance and can be used as catchment specific parameters for digital filtering algorithms and hydrological models (Collischonn and Fan, 2013; Zhang *et al.*, 2013). This investigation was based on results from 15 gauging stations in South and Central Brazil (Collischonn and Fan, 2013). Gauging stations were selected to represent a wide range of catchments, from groundwater dominated rivers to surface runoff dominated rivers. They are generally not associated with each other as with the Wairarapa gauges investigated here. With regard to the geoclimatic context, the setting can be considered relatively similar to conditions present in the Wairarapa valley. The

geology of South and Central Brazil is predominantly made up of porous sedimentary basins with more impermeable volcanic and carbonaceous formations at higher elevations (Negreiros *et al.*, 2009). There are slight differences in climate, with generally higher rainfall totals and higher average temperatures in the Brazil region, between 18 and 26 °C across the investigation areas (Alvares *et al.*, 2013). Results between the investigations are comparable, although more variability in hydrological and physiographic settings is present in the Collischonn and Fan (2013) investigation.

f and *k* parameters for Bump and Rise filter

The two parameters for the BR digital filter (*f* and *k*) were estimated for each flow gauge using Monte Carlo analysis. The filter was applied to the flow gauge dataset using random values that were within realistic bounds for each of the two parameters (0.01-0.5 and 0.001-0.09 for *f* and *k* respectively - *pers. comm.* Stewart, 2018). Results were ignored if the BR output value for BFI was larger than the BFI_{max} calculated from the FDC using equation 6.7. The baseflow recession constant for each iteration was calculated from the baseflow dataset and then compared to the total flow recession constant (Chapman, 1999). The combination of values for *f* and *k* that produced the best fit to the total flow recession were chosen as the parameters for that gauge. An exception was made for the Otakura gauge, whereby *f* values larger than 0.01 resulted in BFI estimates being greater than the BFI_{max}. As such the analysis was rerun for this gauge using lower limits (0.001 – 0.01 and 0.0001 – 0.01 for *f* and *k* respectively).

Summary

Baseflow separation techniques employed in this investigation include the Eckhardt digital filter (EK) and the Bump and Rise digital filter (BR). The various methods were applied to 13 continuous flow gauge datasets situated throughout the Wairarapa valley. Outputs were compared with each other and also with Q90/Q50, area, mean flow, specific discharge and predominant geology. The EK filter was applied using a defualt BFI_{max} value of 0.8 (EK(def)) (Eckhardt, 2008) and again using a catchment specific value for BFI_{max} (EK(opt)), estimated from a backward filter algorithm (Collischonn and Fan, 2013). The BR filter was applied using catchment specific parameters that dictate the baseflow response to an event (Stewart, 2015). The bump (*f*) and rise (*k*) parameters were calculated using Monte Carlo analysis and optimised to fit the total flow master recession curve at each gauge. Catchment specific parameters are

presented in Table 6.2. Results of the baseflow estimation comparison are presented in Section 6.2.

Table 6.2. Catchment specific parameters for baseflow estimation include: recession constant (w), BFImax from backward filter (BF) algorithm and default (def) for Eckhardt filter, bump (f) and rise (k) parameters for BR filter and FDC low flow ratio (Q90/Q50).

No.	Candidate	ω	BF BFI$_{max}$	def BFI$_{max}$	f	k	Q90/Q50
3	Kopuaranga 1	0.997	0.48	0.8	0.290	0.0100	0.30
4	Kopuaranga 2	0.997	0.61	0.8	0.257	0.0100	0.36
5	Mangatarere 1	0.983	0.83	0.8	0.011	0.0042	0.09
6	Mangatarere 2	0.985	0.79	0.8	0.092	0.0099	0.28
7	Mangatarere 3	0.988	0.77	0.8	0.073	0.0100	0.22
8	Otakura	0.998	0.73	0.8	0.001	0.0008	0.22
12	Ruamahanga 1	0.949	0.80	0.8	0.500	0.0100	0.39
13	Ruamahanga 2	0.996	0.65	0.8	0.499	0.0099	0.29
14	Ruamahanga 3	0.985	0.81	0.8	0.500	0.0100	0.31
16	Tauherenikau	0.966	0.85	0.8	0.492	0.0100	0.34
17	Waingawa	0.962	0.84	0.8	0.500	0.0100	0.38
18	Waiohine	0.963	0.82	0.8	0.499	0.0100	0.39
20	Waipoua 2	0.974	0.87	0.8	0.387	0.0100	0.21

6.2   Baseflow separation results

Table 6.3 presents estimates of BFI from the two Eckhardt filter options, the Bump and Rise filter and the Q90/Q50 statistic from a flow duration curve at each of the 13 gauges analysed in this investigation.

Table 6.3. Baseflow indices generated by Eckhardt recursive digital filter using two different values for BFImax parameter: backward filter EK(opt) and default EK(def) as well as estimates from Bump & Rise filter (BR) and Q90/Q50 metric for 13 gauges in the Wairarapa region.

No.	Candidate	EK(opt)	EK(def)	BR	Q90/Q50
3	Kopuaranga 1	0.21	0.51	0.61	0.30
4	Kopuaranga 2	0.35	0.57	0.67	0.36
5	Mangatarere 1	0.60	0.56	0.44	0.09
6	Mangatarere 2	0.52	0.53	0.53	0.28
7	Mangatarere 3	0.50	0.53	0.50	0.22
8	Otakura	0.51	0.61	0.64	0.22
12	Ruamahanga 1	0.48	0.48	0.58	0.39
13	Ruamahanga 2	0.39	0.57	0.54	0.29
14	Ruamahanga 3	0.56	0.55	0.60	0.31
16	Tauherenikau	0.60	0.53	0.62	0.34
17	Waingawa	0.57	0.52	0.61	0.38
18	Waiohine	0.52	0.50	0.57	0.39
20	Waipoua 2	0.65	0.55	0.61	0.21

Figure 6.5 presents baseflow estimates from the various methods at each gauge. The graph is set out with respect to catchment mean flow from low (left) to high (right). This graph displays an apparent narrowing of results across all estimates at the higher mean flow gauges compared to a larger spread at the lower mean flow gauges. This could suggest that the baseflow separation filters are more precise at gauging sites where more discharge is measured. The relative spread on graph indicates there is no significant correlation between baseflow ratio and mean flow volume across the 13 gauges investigated.

The Mangatarere 2 gauge showed the best agreement across the 3 methods, however the BFI values of 0.52, 0.53 and 0.53 were almost double that of the Q90/Q50 reference value. In fact, BFI estimates generated from the three methods were larger than the Q90/Q50 metric at all gauges except the Kopuaranga 1 and Kopuaranga 2 gauges (Figure 6.5). This highlights the temperamental nature of the Q90/Q50 metric and questions its effectiveness for comparisons in the Wairarapa valley.
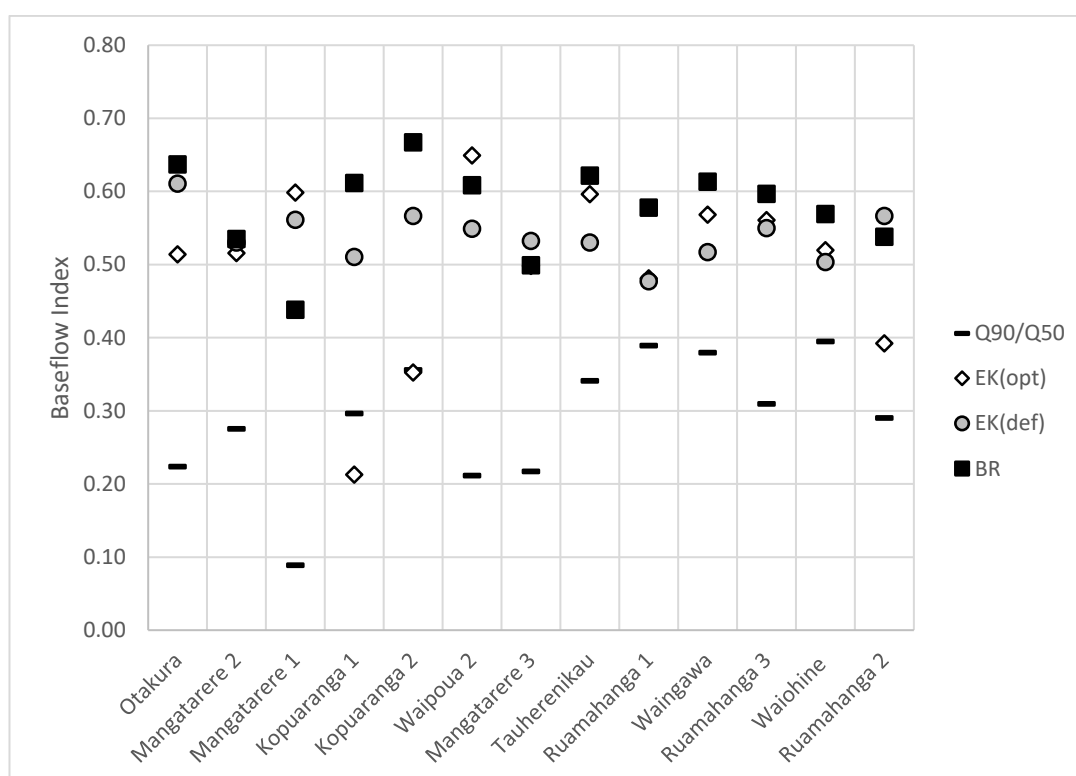
Figure 6.5. Baseflow indices generated by the Eckhardt filter using two different values for BFImax (Backward filter method (hollow diamonds) and prescribed 0.8 (grey circles)) and the Bump & Rise filter (black squares) plotted along with Q90/Q50 metric (black dashes) at 13 gauges in the Wairarapa.

The baseflow estimates generated by digital filters have been compared to Q90/Q50 statistics and specific discharge across the 13 Wairarapa gauges. These comparisons are presented in Figures 6.6 and 6.7. The EK(opt) filter shows no correlation to the Q90/Q50 metric, while the EK(def) filter exhibits an inverse relationship (negative R). The best positive correlation of BFI estimates to the Q90/Q50 metric comes from the BR filter, despite an over estimation bias at all gauges. All three sets of BFI estimates show no correlation to catchment specific discharge compared in Figure 6.7.

Figure 6.6 shows that there is a relatively small spread of Q90/Q50 values for the gauges analysed in the Wairarapa valley. All but the Mangatarere 1 gauge exhibit Q90/Q50 statistics between 0.20 and 0.40. The comparative BFI estimates exhibit similar clustering, with majority of estimates from all methods being between 0.48 and 0.68. Figure 6.6 shows the two Kopuaranga gauges and the largest mean flow Ruamahanga 2 gauge share a distinct cluster of EK(opt) BFI values that could be considered correlative to their respective Q90/Q50 statistics.
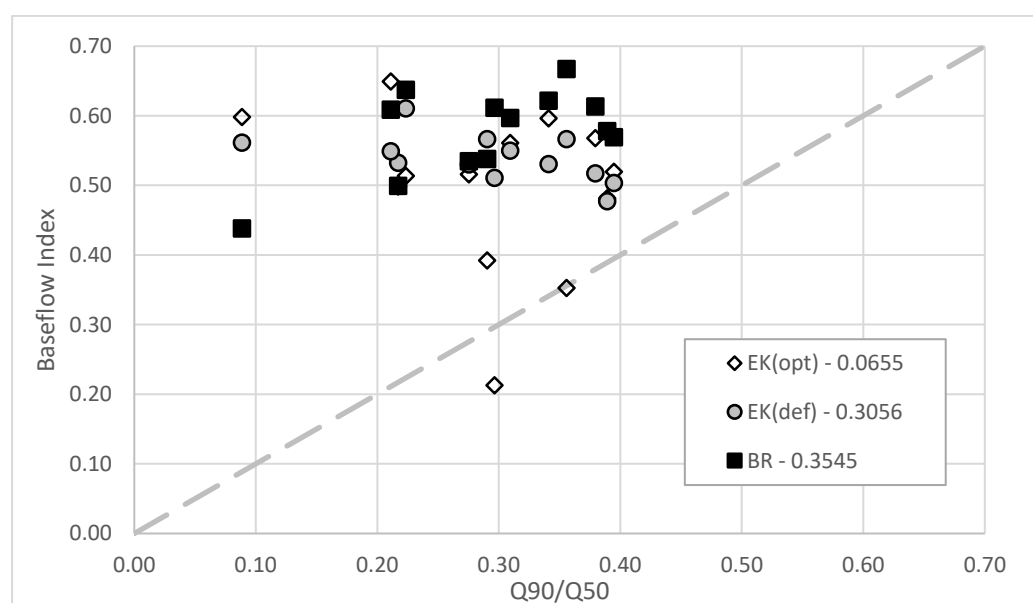
Figure 6.6. Baseflow indices generated by EK(opt) (hollow diamonds) EK(def) (grey circles) and BR filters against Q90/Q50 metric at 13 gauges in the Wairarapa valley, New Zealand. Legend shows R² values.
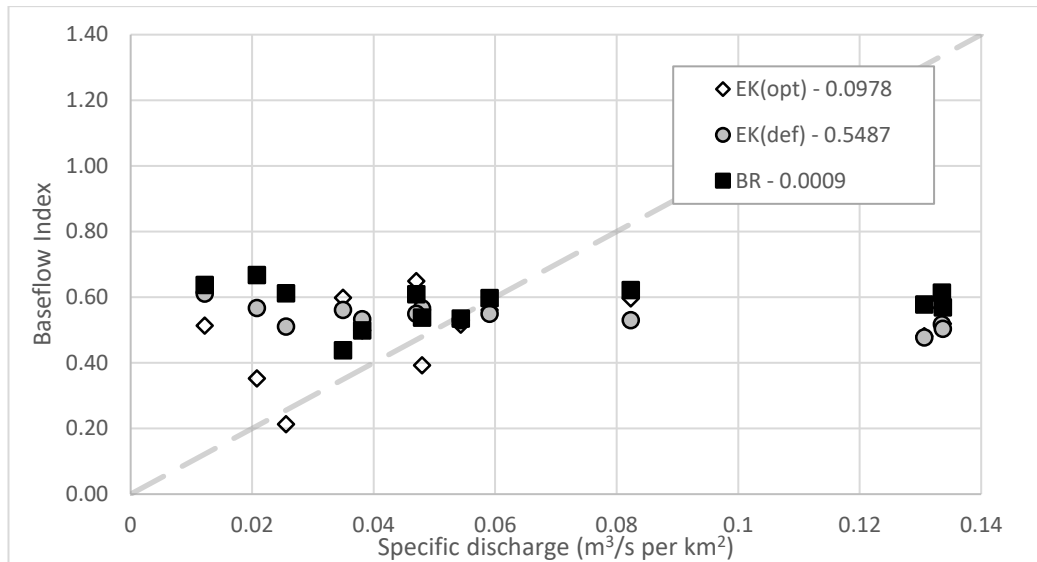
Figure 6.7. Baseflow indices generated by EK(opt) (hollow diamonds) EK(def) (grey circles) and BR filters against specific discharge at 13 gauges in the Wairarapa valley, New Zealand. Legend shows R^2 values.

Quantitative comparisons show the BR filter produces BFI estimates at each gauge with the highest correlation to Q90/Q50. However, in terms of graphical comparisons, performance of the filter techniques is varied and can be categorised with respect to distinct high, medium and low mean flow volumes. Figure 6.8 presents the total flow hydrograph at the Ruamahanga 3 gauge accompanied by the three estimated baseflow separated hydrographs. This Ruamahanga 3 graph can be considered representative of the higher flow gauge hydrographs: Ruamahanga 1, Ruamahanga 2, Tauherenikau, Waiohine and Waingawa. These 6 gauges have very similar hydrograph plots and outputs. They also have the 6 highest average flow rates out of the 13 gauges analysed (> 9 m³/s). Average BFI estimates for the various methods at these 6 gauges are presented in Table 6.4. The hydrographs presented herein show a two week period in October. This was done to view the shape of the baseflow response to an individual peak flow event. Parameters for the filter algorithms are constant and so this snapshot is representative of the baseflow response to peak flow at all times of the year.
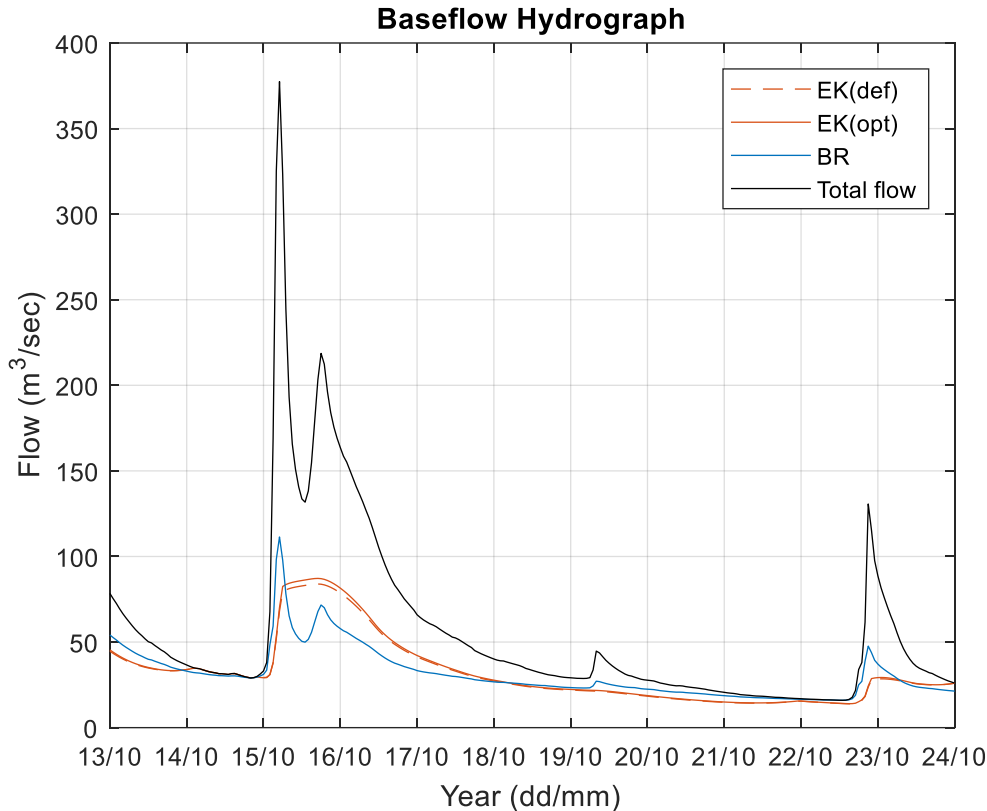
Figure 6.8. Ruamahanga 3 total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).

Table 6.4. Average BFI for each method and standard deviation of the six largest flow datasets: Ruamahanga 1, Ruamahanga 2, Ruamahanga 3, Tauherenikau, Waiohine and Waingawa. Average parameter value for the gauges also provided where applicable.

Method	BFImax	f	k	Avg. BFI	σ
Eckhardt (def)	0.8	-	-	0.52	0.03
Eckhardt (opt)	0.8	-	-	0.52	0.07
Bump & Rise	-	0.50	0.01	0.59	0.03
Q90/Q50	-	-	-	0.35	0.04

Figure 6.8 shows the BR method is responsive to total flow peaks. In contrast, both EK filter methods exhibit heavily attenuated baseflow responses. Further comparisons show that the BR receding limb is very abrupt, whereas the two EK methods possess longer recessions that eventually drop below the BR output. At low flows total flow and BR baseflow are equal, whereas the EK baseflow hydrographs are below the total flow.

Baseflow separation at the Mangatarere gauges exhibits very similar outputs and can be represented by the Mangatarere 3 hydrograph Figure 6.9. It shows a very limited

response of baseflow to event flow by the BR filter compared to the high flow gauges represented in Figure 6.8. The Mangatarere gauges had the lowest BR BFI values. This is because the estimated values for parameter f at these gauges were an order of magnitude smaller than their counterparts at the "high flow" gauges (Table 6.5).
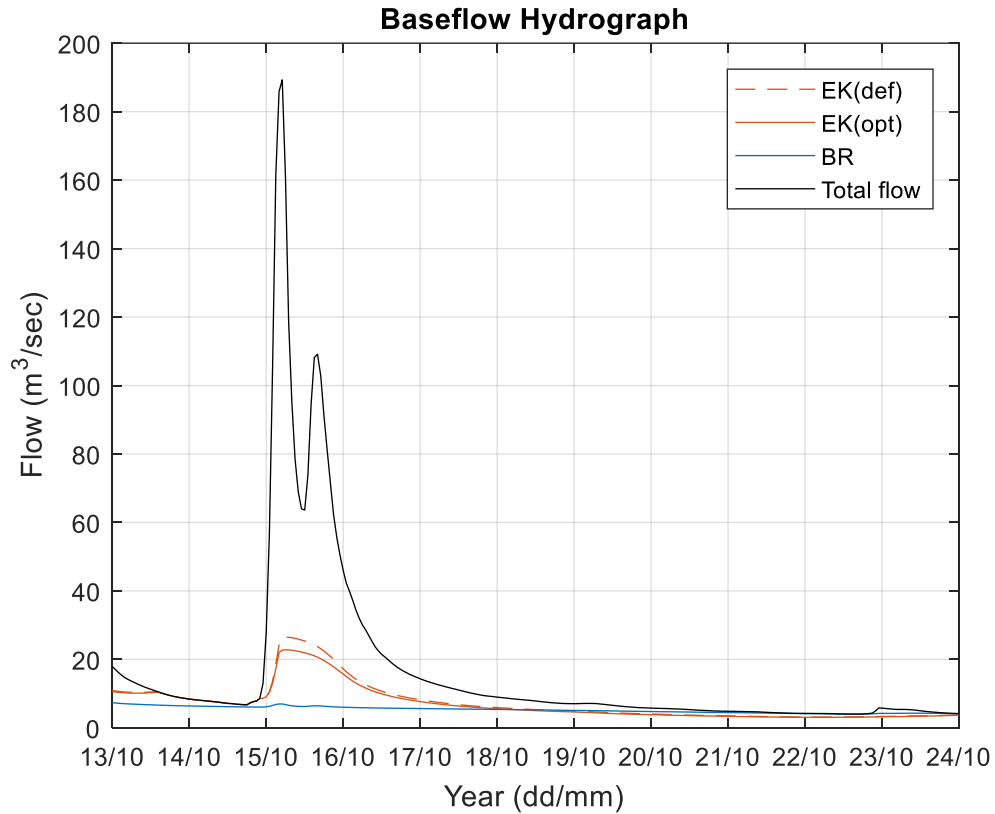
Figure 6.9. Mangatarere 3 total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).

Table 6.5. Average BFI for each method and standard deviation of the three Mangatarere flow gauge datasets. Average parameter value for the gauges also provided where applicable.

Method	BFImax	f	k	BFI	σ
Eckhardt (def)	0.8	-	-	0.54	0.01
Eckhardt (opt)	0.8	-	-	0.54	0.04
Bump & Rise	-	0.06	0.01	0.49	0.04
Q90/Q50	-	-	-	0.19	0.08

The two Kopuaranga gauges show similar baseflow separation outputs and are represented by the Kopuaranga 1 total flow/baseflow hydrograph in Figure 6.10. At these generally lower mean flows, it appears the EK(opt) filter significantly

91

underestimates flow across the two week period as it never matches up with the total flow hydrograph. Another observation is the relative responsiveness of the EK(def) filter to peaks in comparison to the BR filter despite the BR filter recording higher BFI values at the two flow gauges (Table 6.6).
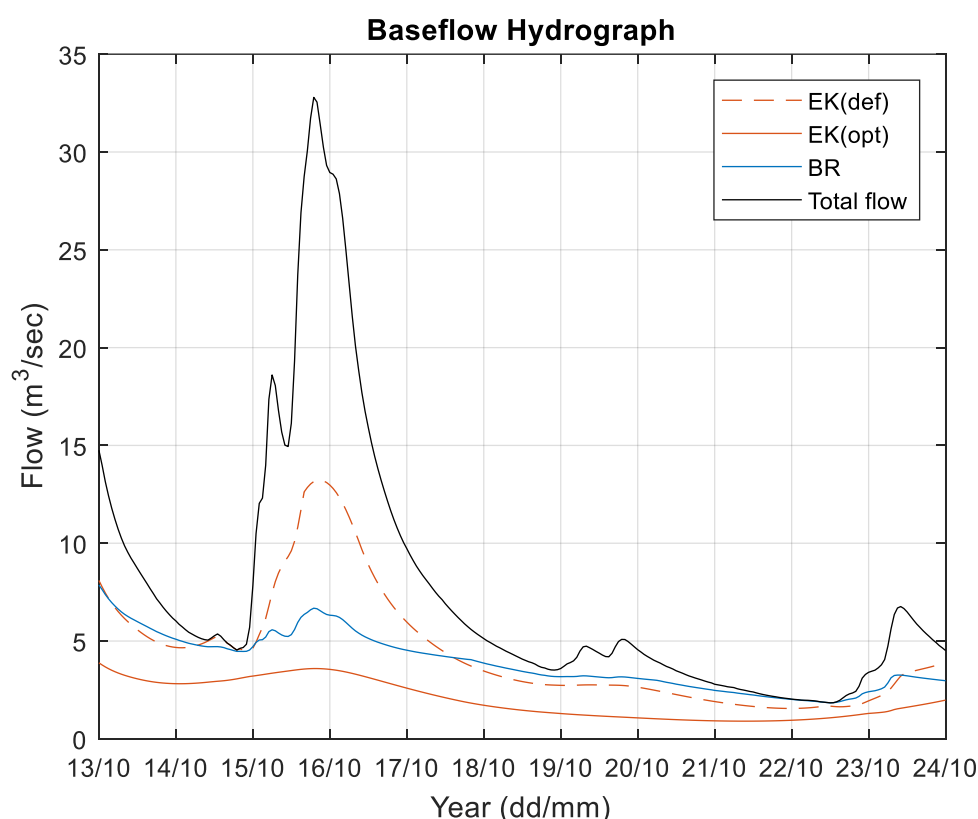
Figure 6.10. Kopuaranga 1 total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).

Table 6.6. Average BFI for each method and standard deviation of the two Kopuaranaga flow gauge datasets. Average parameter value for the gauges also provided where applicable.

Method	BFImax	f	k	BFI	σ
Eckhardt (def)	0.8	-	-	0.54	0.02
Eckhardt (opt)	0.55	-	-	0.28	0.07
Bump & Rise	-	0.27	0.01	0.64	0.03
Q90/Q50	-	-	-	0.33	0.03

The Otakura baseflow separation is presented in Figure 6.11. This site exhibits the lowest mean flow (0.56 m^3/s). Like the Mangatarere gauges, BR baseflow is unresponsive to total flow peaks but meets up with total flow when quickflow is negligible. The EK filters continue to under predict these sections of the hydrograph.

92

The *f* and *k* parameters for the BR filter at the Otakura gauge are significantly lower than the other gauges by an order of two magnitudes (0.001 and 0.0008 respectively), which explains the muted response to total flow. The BR filter maintains a steady level of non-event period baseflow – more than EK – and therefore produces the highest BFI of 0.64. EK(opt) and EK(def) BFI values are 0.51 and 0.61 respectively.
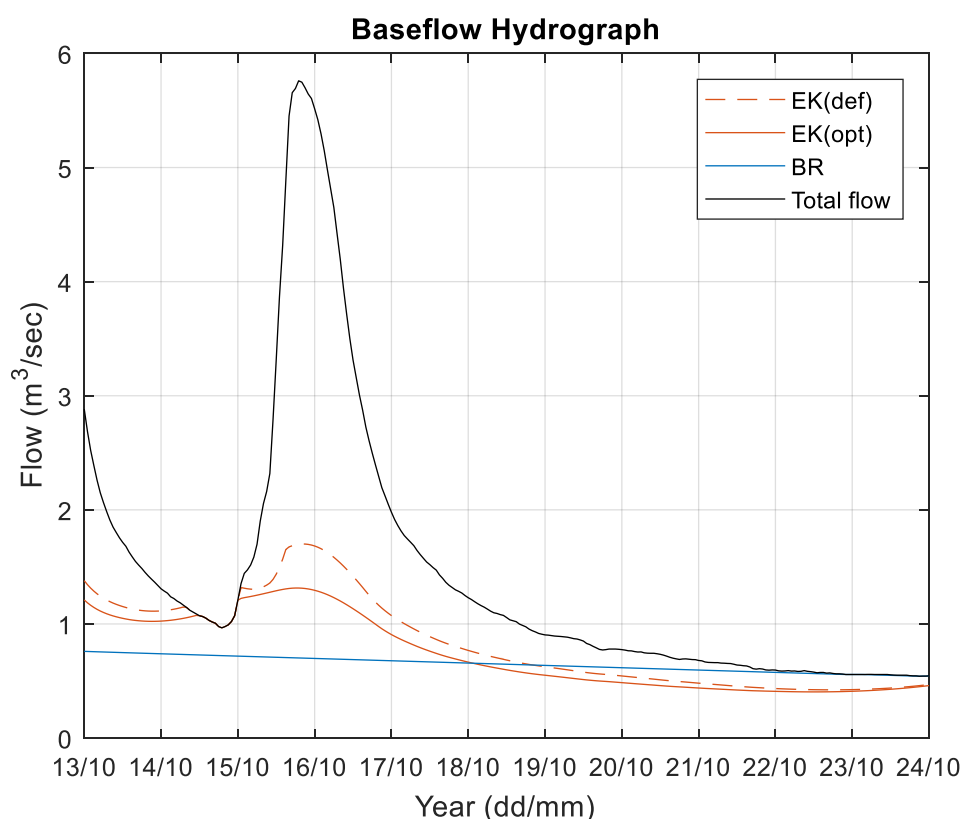
Figure 6.11. Otakura total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFI$_{max}$ (red line) and Bump and Rise filter (blue line).

6.3 Baseflow separation discussion

The BFI estimates produced from various techniques in this investigation suggest there is a relatively high connectivity between groundwater and surface water in the Wairarapa region. The average BFI values for each of the three filter applications across all 13 gauges is presented in Table 6.7 along with their respective averaged parameters.

Table 6.7. Average BFI for each method and standard deviation of 13 flow gauge datasets in the Wairarapa. Average parameter values also provided where relevant.

Method	BFImax	*f*	*k*	BFI	σ
Eckhardt (def)	0.8	-	-	0.54	0.03

Eckhardt (opt)	0.76	-	-	0.50	0.11
Bump & Rise	-	0.32	0.01	0.58	0.06
Q90/Q50	-	-	-	0.29	0.08

The three filters will be discussed below in the context of their relation to the Q90/Q50 statistics, individual response to rainfall events, ability to match total flow in low-flow/recession periods and the overarching influence of the filter parameters. The applicability of the methods also appears to vary depending on average flow volumes at each gauge, this aspect will be discussed throughout.

BFI comparison to Q90/Q50

The EK(def), EK(opt) and BR filter methods indicate the average baseflow contribution to streams in the Wairarapa region is just above 50% of total streamflow over the investigation period. This suggests a high connectivity between ground and surface water bodies. In contrast, the average Q90/Q50 statistic for the same 13 gauges is just less than 30%.

The two applications of the EK filter, using different values for the BFI_{max} parameter produced results that had no apparent relation to the Q90/Q50 metric. The BR filter, using two parameters (*f* and *k*) to match the total flow recession curve (Stewart, 2015) showed the best correlation to Q90/Q50, albeit not a strong one ($R^2 = 0.35$). All of the BFI estimates were larger than the Q90/Q50 statistics at the respective gauges. These results are significantly different to the conclusions from Collischonn and Fan (2013), who reproduced good correlations between EK(opt) and Q90/Q50. Figure 6.12 presents the backward filter BFI_{max} results from this investigation on the same graph as the Collischonn and Fan (2013) results in relation to Q90/Q50. Their results show a much larger spread of Q90/Q50 values and a relatively good correlation with backward filter BFI_{max} estimates. The Wairarapa backward filter BFI_{max} estimates show an over-estimation bias and poor correlation to Q90/Q50, however both BFI_{max} and Q90/Q50 estimates are over a much smaller range compared with the Collischonn and Fan (2013) dataset.
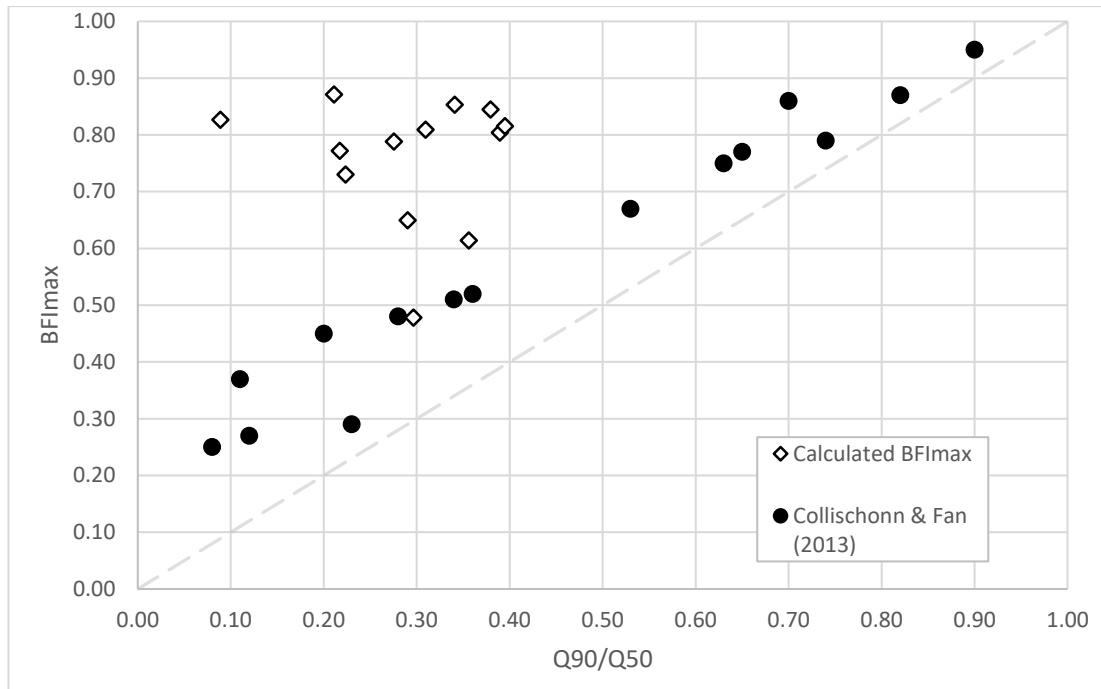
Figure 6.12. Backward filter BFImax values for 13 gauges in the Wairarapa region, New Zealand (hollow diamonds) and backward filter BFImax values for 15 gauges in Brazil (black circles) (Collischonn & Fann, 2013) in relation to catchment specific Q90/Q50 metrics.

Flow records in the Wairarapa region demonstrate acute responses to event flow, whereby rivers recede to low-flow volumes very quickly after rainfall events. This means the record consists of a relatively large number of low and high flow periods, with less medium flow values. As such, the ratio of Q90/Q50 is likely to be a small number, relative to a catchment with more consistent flow characteristics. Therefore, the Q90/Q50 metric could potentially be underestimating the low-flow proportion due to the responsive nature of streams in the Wairarapa. Despite being a significant feature of the Collischonn and Fan (2013) investigation, the arbitrary nature of the Q90/Q50 statistic as an estimate for baseflow contribution has been restated within this research in agreement with other bodies of work (Smakhtin, 2001b; Brodie and Hostetler, 2005).

The poor correlations of the BFI results to Q90/Q50 and specific discharge suggest that area, average flow volumes and the frequency of medium to low flows are not the most significant drivers in terms of the baseflow contribution to a stream. This statement potentially challenges results presented by Collischonn and Fan (2013), however it supports the idea that baseflow is largely dependent on the porosity of the underlying geology and the interaction boundary between stream and aquifer

(Sophocleous, 2002; Eckhardt, 2005). The Q90/Q50 statistic and specific discharge are mostly unrelated to subsurface processes (Brodie and Hostetler, 2005). As such an alternative statistic that is comparable to BFI outputs from digital filters, and that is representative of groundwater-surface water exchange would be informative.

Baseflow comparisons to total flow hydrographs

The BR filter was very responsive to event flow at the flow gauges that had larger average flow rates (> 9 m^3/s). The baseflow response of the two EK filter outputs at these gauges was more subdued compared to the BR outputs. With respect to the theoretical understanding of baseflow response to event flow (Brodie and Hostetler, 2005), it appears the EK filter was successful at recreating an attenuated baseflow response to event flow. In contrast, the BR filter produced a baseflow hydrograph that could be considered overly responsive. On most occasions, the EK filter baseflow hydrographs reproduced the main peak of an event flow, but did not capture the minor total flow fluctuations in the same way that the BR filter did.

At gauges that had lower average flow rates (< 9 m^3/s), the BR filter was generally unresponsive to total flow event peaks. This highlights the uncertainties based on the experimental parameter estimation for the BR filter, discussed below. In contrast, the EK filter baseflow hydrographs remained consistent with their subdued yet significant baseflow estimations at all levels of flow across the two parameter estimation techniques. This demonstrates a structural robustness in the EK filter algorithms because the general response to event flow dynamics is consistent, irrespective of the BFI$_{max}$ parameter.

The EK filter estimates were mostly underestimating baseflow at low-flow periods compared to the BR filter, which consistently joined with the total flow hydrograph at low flows. Total flow is largely made up of baseflow during low-flow periods when quickflow is considered to be negligible (Brodie and Hostetler, 2005). The baseflow hydrograph is therefore expected to equal the total flow hydrograph at this time. The BR filter algorithm generates outputs that are consistent with this theoretical understanding of baseflow during low-flow periods. It does this independently of the parameter values, which means the structure of the algorithm should be considered favourable during low-flow periods.

In summary, the structure of the two algorithms simulate baseflow hydrographs that exhibit unique strengths and weaknesses when compared to the Brodie and Hostetler (2005) defined theoretical baseflow response to rainfall events. The EK filters simulated an attenuated, yet significant baseflow response to an event, however it largely underestimated baseflow contribution to streams during low-flow periods. In contrast, the BR filter generated a baseflow peak that could be considered overly responsive, while consistently made baseflow contribution equal total flow during low-flow periods. These features are functions of the algorithm structure and are mostly independent of the parameter values.

Geological influence

Baseflow estimates from within the Mangatarere catchment have been assessed here with respect to the geological setting of the catchment. This catchment was selected because it is the area where the model validation investigation takes place. The Mangatarere catchment is representative of the Wairarapa region, with headwaters that flow out of the less permeable Tararua ranges at high elevation before reaching the more permeable Q1 gravels on the plains. The analogous nature of the Mangatarere catchment with respect to the wider region means that the conclusions and outcomes of this investigation can be applied to the Ruamahanga catchment and other similar regions (Canterbury plains, Otago plains, etc).

The parameters employed for baseflow separation attempt to represent physical catchments in various ways. The EK(def) BFI_{max} assumes a hard rock/porous rock definition to capture groundwater influence (Eckhardt, 2008), whereas the backward filter BFI_{max} method for EK(opt) assumes geological representation as exhibited by the total flow recession curve (Collischonn and Fan, 2013). Similarly, the BR filter parameters f and k relate to quickflow and baseflow recession curves which are considered to be functions of physical catchment characteristics (Stewart, 2015).

Figure 6.13 displays the general geology of the Mangatarere catchment upstream of the Mangatarere 3 gauge. From this figure, it is clear to see that the Mangatarere catchment is dominated by relatively permeable sedimentary rock. This observation agrees with the baseflow separation estimates of a relatively high level of interaction between groundwater and surface water bodies. Figure 6.13 shows that the catchment above the most upstream Mangatarere 2 gauge is dominated by less permeable

sandstone basement rock and one might assume a lower BFI might be calculated at this gauge compared to the gauges whose catchments drain out of more porous gravels further downstream. In fact the BFI estimates of these three gauges are relatively similar, with the upstream Mangatarere 2 gauge recording almost the same BFI as the farthest downstream Mangatarere 3 gauge. While this initially suggests poor sensitivity from the baseflow separation methods, it is apparent that the areas directly adjacent to the stream above the Mangatarere 2 gauge are in fact predominantly gravels, associated with high connectivity. Such an observation calls into question the Eckhardt (2008) definition of a catchment with a "porous aquifer" because the geology directly associated with the stream is significantly more porous than the average porosity of the wider catchment geology.

A further consideration is the influence of active faults in this area. The Mangatarere catchment is dissected by three significant faults (Figure 6.13) and groundwater flow should be expected to change course drastically as a result of these significant features (Freeze and Cherry, 1979). Accordingly, Jones and Gyopari (2006) record a switch to gaining stream reaches as the Mangatarere river flows over the Carterton fault. The positioning of the gauges is such that Mangatarere 2 (highest upstream) is situated above all fault intersections and Mangatarere 1 (midstream) is approximately one kilometre downstream of the Carterton fault. This stream reach would be expected to possess a higher BFI. From the results of this investigation, the BR filter registers an increase of BFI from 0.44 to 0.53 between the Mangatarere 2 and Mangatarere 1 gauges before dropping to a BFI of 0.50 at the downstream Mangatarere 3 gauge. The Q90/Q50 statistic demonstrates a similar spike in baseflow contribution at the middle gauge. These observations could relate to the influx of groundwater due to the Carterton fault. In support, the two gauges downstream of the fault intersection exhibit generally similar BFI estimates across all methods. Further, the two EK filter outputs record the midstream gauge as possessing the highest portion of baseflow of the three gauges. Despite these observations, it is difficult to be certain whether the midstream spike in BFI is due to the influence of the Carterton fault intersection, which highlights the inability of baseflow separation filters to capture high-resolution differences in a dynamic hydrological setting. The Mangatarere catchment and wider Wairarapa region exhibits high groundwater-surface water connectivity. This conclusion can

been inferred by the predominantly porous geology in the region, especially the plains, and outputs from digital baseflow filters.
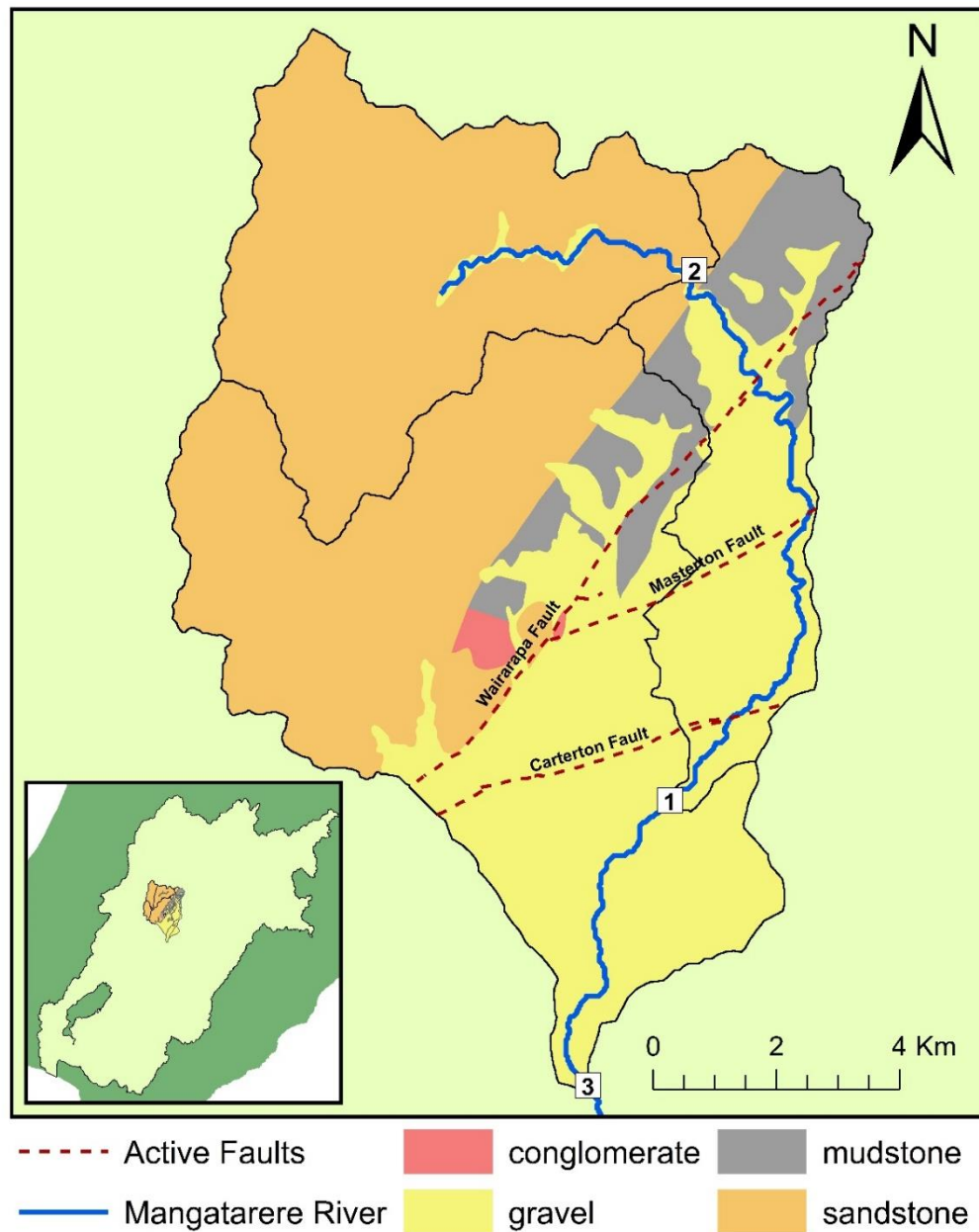
Figure 6.13. Geological map of the catchment upstream of the Mangatarere 3 gauge, Wairarapa valley, New Zealand showing dominant surficial geology, active tectonic faults and location of the Mangatarere gauging locations (1, 2 and 3 in boxes) – Inset: The location of the catchment in regards to the Wairarapa valley.

Parameter estimation

The parameters and outputs of baseflow separation investigations are often legitimised by an "expert eye" and it is difficult to unequivocally prove correctness. The EK and BR digital filters claim to stand apart from previous hydrograph separation techniques as they address the need for parameters that reflect physical characteristics of a

catchment (Eckhardt, 2005; Stewart, 2015). These methods of baseflow estimation are informed by quantitative chemical tracer analysis. It is crucial that parameters are well understood and quantified, as they have a significant effect on the digital filter BFI estimations. This is apparent in a comparison between the EK(def) and EK(opt). The EK(def) filter exhibits the smallest range of BFI values across the 13 gauges despite varied levels of average flow. The EK(opt) filter produces a wider range of results which is due to the optimised values of BFI_{max} calculated for each gauge.

The BFI_{max} values used for the EK filter in this investigation produced BFI estimates that showed no correlation to total discharge, catchment area or the Q90/Q50 discharge frequency statistic. A relationship was expected in accordance with Collischonn and Fan (2013) paper. The BFI_{max} parameter is changed depending on an assumption of the porosity of the aquifer in the catchment and these subsurface characteristics are likely to have a minimal relationship to the surface catchment characteristics analysed in this investigation. While it is understood that baseflow contribution is largely dependent on geological characteristics, it is difficult to quantify this influence and Eckhardt (2008) definitions are ambiguous. For example, the Mangatarere 2 catchment could be characterised as a relatively hard rock catchment when looking at the entire catchment, however the directly adjacent stream geology consists of porous Q1 gravels (Figure 6.13). As such, the average porosity of a catchment may not be a good indicator to inform on parameters like BFI_{max}. Bloomfield *et al.* (2009) present an investigation into the geological controls on baseflow index. They consider the areal extents of numerous geological classes based on lithostratigraphic and hydrogeological classification schemes in a linear regression method to correlate to BFI. This work has potential to inform BFI_{max}, thus giving physical meaning to the eventual output of BFI from digital separation filters.

Using $BFI_{max} = 0.8$ for the Eckhardt filter should be considered relatively arbitrary (Szilagyi, 2004), so EK(opt) was applied as a comparison. The Collischonn and Fan (2013) method for optimising BFI_{max} to the particular flow regime relies on using the recession constant, which is already a separate parameter employed in the Eckhardt recursive digital filter. By using the recession constant as both a parameter and a means for estimating a second parameter (BFI_{max}), the two are likely to be intercorrelated. Further, using total flow recession will result in uneven weighting of the filter to total

streamflow processes as opposed to groundwater characteristics identified by recession analysis on baseflow separated datasets (Stewart, 2015). Establishing a value for BFI_{max} that is a function of the subsurface and boundary characteristics would elevate the effectiveness of the Eckhardt filter and likely produce more accurate and catchment specific results.

The poor performance of the BR filter at lower flows is likely to be due to the initial estimation of the f and k parameters. With the constant rise parameter (k) being so low the estimates for baseflow were effectively normalised at every point by a ratio of close to 0. Hence the representation of baseflow as a flat line. The BR filter showed excellent promise in this investigation at higher flow gauges, producing similar results to the EK filter despite using significantly different algorithms and parameter sets. The advantage of the BR filter is that the parameters are specifically related to recessional baseflow processes and these can actually be quantified by chemical tracer analysis. However chemical tracer analysis is expensive and time-consuming, so improvement on the optimisation process should be considered an important avenue for research.

6.4   Baseflow separation limitations

One potential caveat with this investigation is the relatively small spread of the Q90/Q50 statistics for the 13 gauges analysed. Collischonn and Fan (2013) mention the fact that their dataset of Q90/Q50, estimate from 15 gauges, spanned a large range (0.08 – 0.9) and they were able to generate a good correlation between backward filter estimated BFI_{max} values and Q90/Q50. The small distribution of the Q90/Q50 from 12 gauges in this investigation (0.2 – 0.4) is likely to be a barrier to generating such correlations. The range might have been more significant if Q90/Q50 and BFI_{max} had been calculated at all 20 gauges supplied by GWRC, however the quality of these flows was uncertain. Further, as the gauges are positioned in relatively similar catchments and considered sub-catchments of the regional Ruamahanga flow catchment, it is likely that estimates would have been quite similar to those already produced in this investigation. Further investigations in the same vein as Collischonn and Fan (2013) across New Zealand with a focus on a variety of hydrological settings would help to address this uncertainty.

From this investigation, it is clear that the parameters for both the EK and BR filters have a significant effect on the calculated BFI estimation. While the parameters are

ideally related to groundwater-surface water interaction processes, they are mostly estimated using datasets related to streamflow. Eckhardt (2008) even states that the initial suggestions for BFI_{max} should be considered first order approximations. There is a growing body of work dedicated to developing more finely tuned estimates for these filter parameters. This is important because accurate parametrisation is paramount to trustworthy estimations of BFI by digital filter separation. More work should be carried out to relate these parameters to measurable catchment characteristics with a specific focus on subsurface, geological features.

Another issue with baseflow estimation is that baseflow contribution to a stream cannot actually be directly measured. This investigation has highlighted the arbitrary nature of comparing estimates of BFI from digital filters, to estimates of BFI from Q90/Q50. None of these metrics can be considered the "correct" value for baseflow and therefore true correctness cannot be proved. Despite this subjectivity, investigations of this type will remain relevant because they are simple, cheap and only require flow data. For these reasons, it should be considered pertinent to further inform and develop baseflow separation digital filters in the future by using quantitative chemical tracer data to produce objective and groundwater specific parameters.

The value for BFI represents an average groundwater contribution for the entire length of stream above a gauge over a certain time period. As such it does not have the capability to delineate gaining or losing reaches along a stream or identify the temporal variability of groundwater – surface water interaction. More intensive field and chemical tracer investigations such as those carried out by Guggenmos (2010) have found such variability exists in streams in the Wairarapa valley, however these investigations are expensive, time-intensive and they also infer catchment conclusions based on point measurement results. A larger number of gauging stations would increase the resolution of these investigations, while applying the filters to specific time periods (i.e. summer months) would be beneficial for gathering temporal information about groundwater-surface water interactions in the Wairarapa valley.

6.5 Baseflow separation concluding remarks

Recursive digital filter estimates of baseflow index were calculated and compared at 13 flow gauges in the Wairarapa valley, New Zealand. The Eckhardt digital filter produced average BFI estimates of 0.54 and 0.50 for the respective (default) and

102

(optimised) applications. These estimates were significantly larger than the Q90/Q50 statistics, however graphical comparisons showed the filter ably simulated propagation of baseflow in response to total flow event peaks. The baseflow hydrograph was represented as a dampened version of the total flow peaks. At low-flows, the EK filter options failed to intersect with the low-flow portions of the hydrograph at most sites. This suggests that either the catchment never dropped to solely baseflow contribution conditions, or that the digital filter was under predicting at these low-flow volumes.

The bump and rise filter showed an excessive response to event flow and was most effective at matching up with low-flow non-event periods. However, it is heavily dependent on the two parameters f and k. The Monte-Carlo analysis used in this investigation for estimating the two parameters was done on the basis that the two parameters concatenated to match the total flow recession constant. However, the estimated parameters were not effective at reproducing a responsive baseflow hydrograph at gauges with lower average flow ($< 9 \text{ m}^3$/s). This uncertainty led to the BR filter being omitted from the model comparison in Chapter seven as the Mangatarere gauges exhibit relatively low mean flow values ($< 5 \text{ m}^3$/s). The primary method for estimating the two parameters is with quantitative chemical tracer analysis (Stewart, 2015). If more chemical separation analysis work was carried out at a variety of rivers a robust relationship is likely to be established with respect to some measurable streamflow characteristic and Monte-Carlo analysis would not be necessary. This work would be similar to the efforts of Collischonn and Fan (2013) who developed mathematical relationships for estimating BFI_{max} for the Eckhardt filter.

Baseflow separation investigations show a high connectivity between groundwater and surface water in the Wairarapa valley with, on average, approximately 52% of streamflow being attributed to baseflow contribution. The shallow surface geology of the region is predominantly permeable fluvial gravel deposits. Such physiographic settings support the relatively large estimates of BFI calculated in this investigation. With this conclusion, it can be inferred that other regions in New Zealand that have similar geological settings will also exhibit a significant groundwater-surface water interaction. This highlights the importance for hydrological models to incorporate groundwater flow packages when investigating hydrological systems over large spatial

scales. For water management decisions, it stresses the importance of considering a stream and the associated unconfined aquifer in these regions as a single resource in order to avoid double allocation.

This investigation uses model validation techniques to assess the performance of two hydrological models, TopNet-0 and TopNet-GW, to simulate streamflow in the Mangatarere catchment, Wairarapa. Baseflow separation techniques investigated in Chapter 6 are also utilised in order to evaluate the models ability to match low-flow discharges. Model validation is an important stage in the process of model development, whereby independently recorded flow data and their hydrologic indices are commonly used to confirm hydrologic models (Refsgaard and Henriksen, 2004). Methods of validation include a combination of quantitative and graphical dataset comparisons of observed flow and modelled flow (Moriasi *et al.*, 2007). The validation process allows for inter-model comparisons between TopNet-0 and TopNet-GW. The primary difference between the two models is the additional deep groundwater reservoir (S_D) in the TopNet-GW model. TopNet-0 only has one subsurface groundwater reservoir. It is this significant structural modification between the models that this investigation aimed to assess in the context of simulating streamflow and baseflow in the Mangatarere catchment

This chapter is separated into six sections. The first section presents information about the output data produced by the two hydrological models. This is followed by a description of the methodology carried out for this model validation investigation. Section three presents the results of the model validation for each of the three Mangatarere gauges. These results are critically discussed in section four, followed by a consideration of any limitations of the investigation in section five. Finally, concluding remarks on the model validation investigation are presented in section six.

7.1 Modelled output data

Datasets produced by TopNet-0 and TopNet-GW were provided by the National Institute of Water and Atmospheric Sciences (NIWA). TopNet-0 outputs were available at stream reaches that matched the positioning of all 20 observation flow gauges in the Wairarapa region, whereas TopNet-GW has only been applied in the Mangatarere catchment. Model validation results from this investigation can be considered representative of the wider Wairarapa valley as the physiographical characteristics are similar.

Strahler one sub-catchments used for TopNet application in the Mangatarere catchment are presented in Figure 7.1. Outputs from the models are presented as hourly flow in cumecs (m³/s), 24 hours per day. The modelled estimates of flow have no gaps and span the duration of the investigation period (01/07/2010 – 30/06/2014). Parameter sets for the TopNet-GW model run are presented in Table 7.1 – one for above the gorge and one for below the gorge as a function of the changes in geology, land cover and topography – TopNet-0 parameter sets were unavailable.
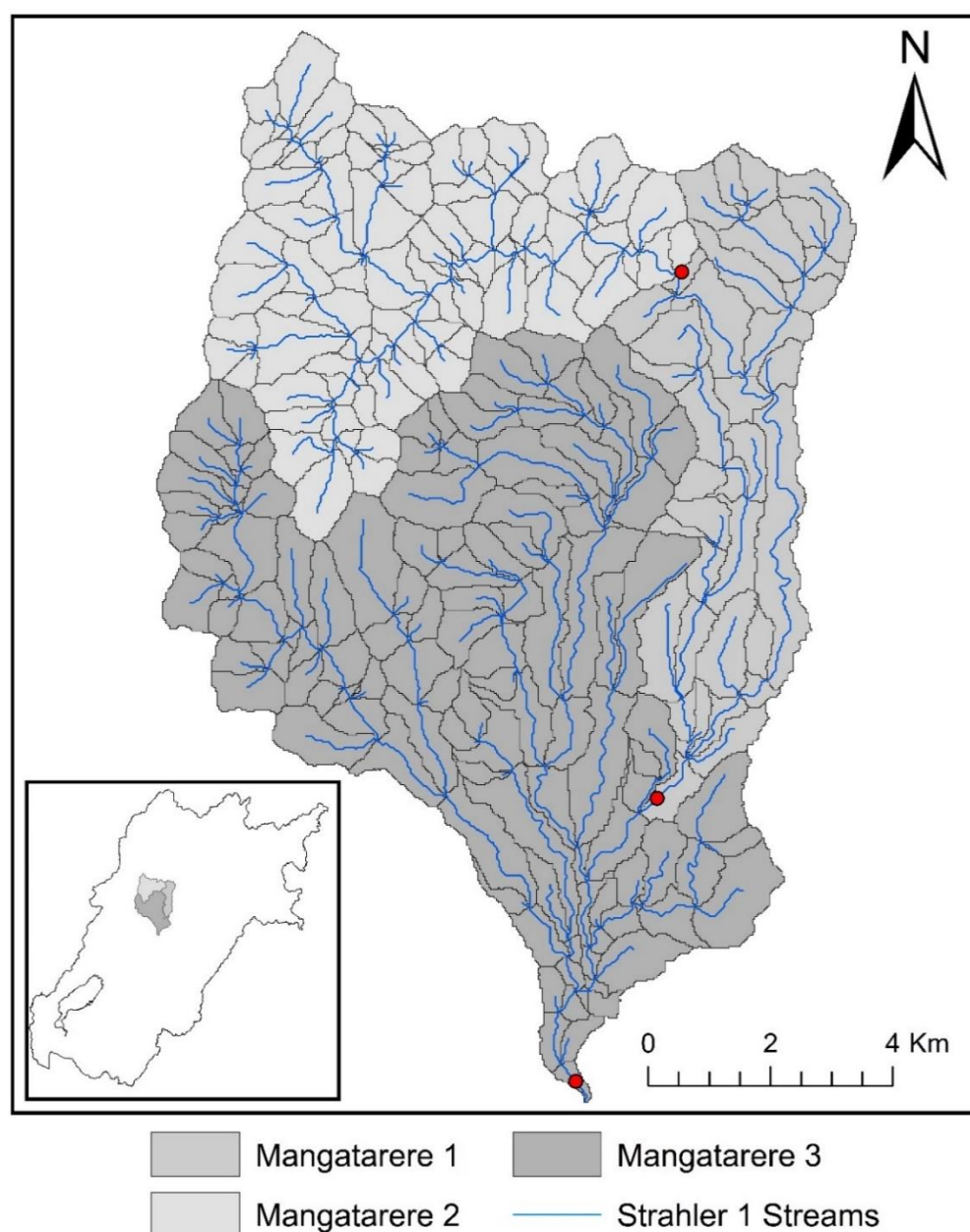
Figure 7.1. Watersheds for Mangatarere 1 (medium grey), Mangatarere 2 (light grey) and Mangatarere 3 (dark grey) flow gauge sites. Watersheds are delineated into strahler 1 sub-catchments demonstrating the scale of TopNet model application. Exact flow gauge locations are shown as red dots. Inset - Wairarapa region with Mangatarere gauge watersheds in grey.

Table 7.1. Parameter sets for the Mangatarere TopNet-GW model run. One for above the gorge opening and one for below.

Parameter Name	Above Gorge	Below Gorge	Parameter Name	Above Gorge	Below Gorge
topmodf	0.113*default	0.717*default	overvel	2.2*default	0.905*default
hydcon0	7.4*default	504*default	canscap	3.49*default	1.32*default
swater1	0.122*default	0.21*default	canenhf	0.45*default	0.393*default
swater2	0.401*default	0.873*default	salbedo	0.634*default	0.828*default
dthetat	0.408*default	1.67*default	atmlaps	0.564*default	-
ch_cexp	2.22*default	1.65*default	gucatch	0.798*default	0.99*default
ga_psif	1.83*default	3.33*default	r_man_n	2.99*default	3.24*default

7.2 Model validation methodology

Outputs from the TopNet-0 and TopNet-GW catchment scale hydrological models were compared to observed flow data at the Mangatarere 1, Mangatarere 2 and Mangatarere 3 gauges. The validation process was carried out with respect to two components: (1) qualitative, or graphical, evaluation of model performance and (2) quantitative evaluation of model performance. Qualitative comparisons include analysis of time series, cumulative discharge, frequency distribution of the time series (flow duration curves), scatter plots and average monthly flow values. Quantitative statistics include correlation coefficients, NSE, PBIAS and RSR performance metrics (see Chapter 3). This suite of comparison techniques satisfies the standard procedure of model validation as suggested by Biondi *et al.* (2012). Quantitative performance metrics used to indicate good or bad agreement between datasets are presented below in Table 7.2 along with the levels of acceptance for each statistic. Performance limits were selected following a review of relevant literature and employed to ensure conditional validation (Refsgaard and Henriksen, 2004).

Table 7.2. Quantitative metrics used to evaluate model performance (Biondi, 2012; Moriasi et al., 2007).

Metric	Name	Excellent	Good	Adequate	Poor	Ref.
NSE	Nash-Sutcliffe Efficiency coefficient	> 0.8	> 0.6	> 0.4	< 0.4	Zammit & Yang (2017)
PBIAS	Percent Bias	<± 1%	<± 5%	<± 10%	>± 10%	Moriasi et al., (2007)
RSR	RMSE-observations standard deviation ratio	< 1.5	< 2.5	< 3.5	> 3.5	Moriasi et al., (2007)

In order to emphasise the baseflow component, comparisons using log flow were also carried out in accordance with the report by Zammit and Yang (2017). By transforming flow data into the log domain, the relative differences at low-flows are easier to observe and quantitative comparisons are less influenced by extreme high flows (Zammit and Yang, 2017). Baseflow estimates using the Eckhardt filter were computed for the model outputs, firstly using a BFI_{max} of 0.8 as suggested by Eckhardt (2008), followed by an optimised BFI_{max} calculated by the backward filter method from Collischonn and Fan (2013). Parameters for this baseflow investigation are presented in Table 7.3. The BR separation method was left out of this investigation due to uncertainty based around the Monte-Carlo estimation of the f and k parameters at low-flow gauges like the Mangatarere sites and because the average BFI estimates produced by the BR method were relatively similar to those produced by the EK filter. Values for BFI produced from the modelled outputs are compared to BFI values produced from the observed flow in order to assess the model's ability to represent baseflow. The entire comparison investigation was carried out on the three flow gauges over the investigation period from 01/07/2010 to 30/06/2014.

Table 7.3. Catchment specific parameters used for baseflow separation and Q90/Q50 for three Mangatarere gauges. Backward filter BFImax calculated from recession constant (ω).

Gauge	Dataset	ω	BF BFI_{max}	ECK BFI_{max}	Q90/Q50
Mangatarere 1	Observed	0.983	0.83	0.8	0.09
	Topnet-0	0.804	0.56	0.8	0.60
	TopNet-GW	0.985	0.75	0.8	0.22
Mangatarere 2	Observed	0.985	0.79	0.8	0.28
	TopNet-0	0.955	0.35	0.8	0.40
	TopNet-GW	0.990	0.76	0.8	0.28
Mangatarere 3	Observed	0.988	0.77	0.8	0.22
	TopNet-0	0.986	0.27	0.8	0.44
	TopNet-GW	0.987	0.75	0.8	0.23

7.3   Model validation results

As stated in the beginning of this chapter, this investigation hoped to compare the two unique model structures with respect to total flow and baseflow simulations. However it became clear throughout the investigation that the TopNet-0 model displayed significant bias in simulation of flow totals. This suggests that the particular application of the TopNet-0 model was significantly lacking in calibration and

parametrisation effort compared to the TopNet-GW model. The TopNet-GW model is more recent and because of its potential application to the SAM project (see Chapter 1), model development efforts have been focussed on this version, rather than the older TopNet-0 model (GNS, 2018). As such, a fair structural comparison between the two simulations is not possible. The following presentation of results and latter discussion, while describing the performance of TopNet-0, is generally focussed in TopNet-GW comparisons to observed flow datasets.

Model validations are presented at each of the three Mangatarere gauges in order of their successive positioning along the Mangatarere stream starting with the upstream gauge. They have been numbered with respect to alphabetical order. As such, Mangatarere at Gorge (Mangatarere 2) is upstream of the gauge at Belvedere Bridge (Mangatarere 1). Mangatarere at SH2 (Mangatarere 3) is the furthest downstream of the three gauges.

7.3.1   Mangatarere at Gorge

The quantitative performance metrics of the TopNet-0 and TopNet-GW outputs in relation to observed streamflow are presented in Table 7.4 for the investigation period. Figure 7.2 presents the comparison between observed and simulated flow at the Mangatarere 2 flow gauge location as standard and cumulative hydrographs along with a flow duration curve.

A TopNet-0 NSE score of -0.52, as well as generally poor PBIAS results indicates inadequate performance. The positive PBIAS of 27.50 indicates model under estimation. This can be observed on the cumulative graph in Figure 7.2, where the TopNet-0 outputs track below the other two datasets. TopNet-GW outputs indicate a much better simulation than TopNet-0, with a good NSE score of 0.35. A good PBIAS (4.04) and correlation coefficient of 0.64 at normal flow comparisons supports this conclusion. The low positive PBIAS indicates a small under-estimation by the model, this can be seen on the cumulative graph in Figure 7.2. With regards to the log domain, NSE is 0.78 – just outside the conditions for an excellent score – and the correlation coefficient improves to 0.88. The flow duration curve shows slightly higher quantity of flow at low flows ($> 50\%$) compared to high flows ($< 50\%$) by TopNet-GW compared to the observed data.

Table 7.4. Simulated streamflow validation statistics for Mangatarere 2 gauge over the validation period.

Gauge	NSE	logNSE	RSR	logRSR	PBIAS	R	logR^2
TopNet-0	-0.52	-0.24	1.23	1.11	27.50	0.25	0.64
TopNet-GW	0.35	0.78	0.81	0.47	4.04	0.64	0.88

Figure 7.2. Observed and simulated cumulative hydrographs and flow duration curves at the Mangatarere 2 gauge over the investigation period.

Figure 7.3 presents scatter plot comparisons between simulated and observed flows. It typically shows a larger spread at higher flows for TopNet-0, which also shows a large band between the 10th and 90th error percentiles (black dots). In contrast, the TopNet-GW outputs exhibit a good fit with the observed data at all flows. The 80% error statistic is narrow which implies the relative error is low. Both models exhibit larger uncertainties with increasing flow volumes.
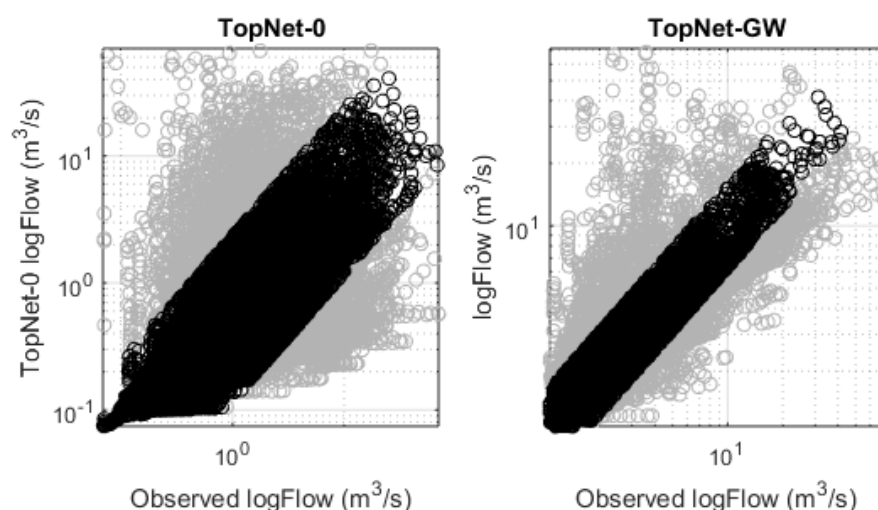
Figure 7.3. Scatter plots of observed flow against TopNet-0 (left) and TopNet-GW (right) at the Mangatarere 2 gauge. Data points in black are within the 80th percentile with regards to an error deviation statistic.

Table 7.5 presents the observed and simulated monthly average flows, calculated from the respective datasets over the four-year period, these averages are also presented as a graph in Figure 7.4. From the graph, it is apparent that the two modelled datasets predict similar values to the observed flow during the dry summer months (Dec - Mar). TopNet-0 drastically under predicts flow throughout most of the year. TopNet-GW tracks well with the observed dataset, despite a small underestimation of average flow in winter and spring, from July to November.

Table 7.5. Simulated and observed monthly average flows (m^3s) for the Mangatarere 2 gauge over the validation and the historic flow record (GWRC, 2017)

Months	TopNet-0	TopNet-GW	Observed
January	0.89	1.44	1.25
February	0.39	0.48	0.43
March	1.17	1.26	1.18
April	1.41	2.31	2.20
May	0.86	2.03	1.97
June	1.76	2.03	2.18
July	1.76	2.79	3.10
August	1.38	1.73	2.16
September	1.06	3.08	3.42
October	1.00	2.17	2.45
November	0.60	1.10	1.33
December	0.61	0.51	0.64
Annual	**1.08**	**1.75**	**1.87**
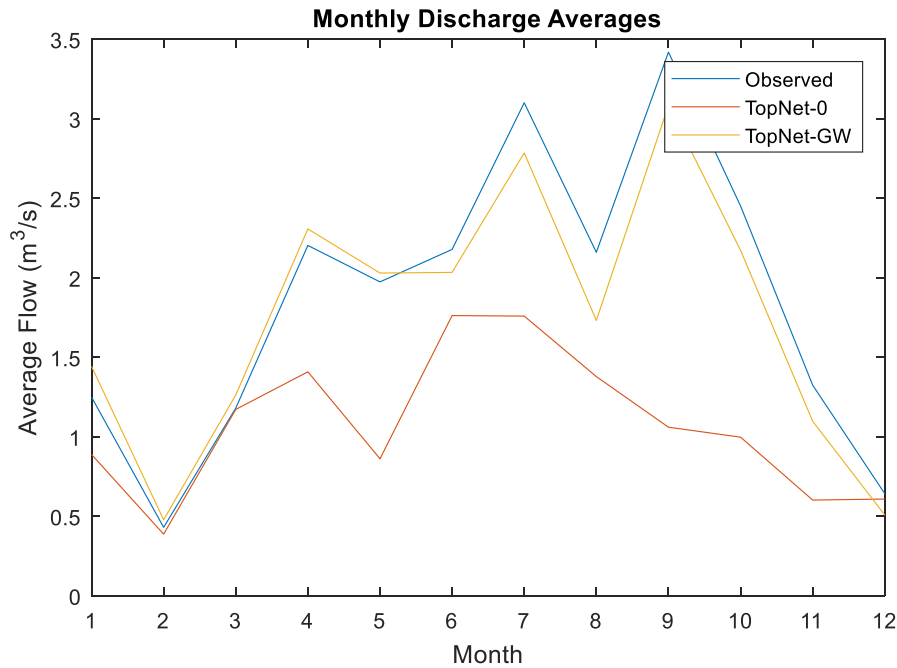
Figure 7.4. Observed and simulated monthly average flow at Mangatarere 2 over the investigation period 2011-2014.

Baseflow estimates for the Mangatarere 2 gauge are presented in Table 7.6. Figure 7.5 presents the simulated and observed hourly low flow hydrographs (i.e., for discharge below observed mean flow). BFI estimates for TopNet-0 are significantly lower than the observed and TopNet-GW BFI estimates even when BFI_{max} was the same value (EK(def)). The optimised TopNet-0 BFI_{max} parameter value of 0.35 was considerably smaller than those calculated for the other datasets (BFI_{max} = 0.79 and 0.76 for observed and TopNet-GW respectively - Table 7.3).

TopNet-GW exhibits shallow recession periods that maintain relatively higher low flows than the observed flow hydrograph in Figure 7.5. In contrast, TopNet-0 flow (red line) can be seen dropping quickly to low flow levels directly after an event peak which is indicative of limited baseflow contribution to a stream. Despite the overestimation at low-flow by TopNet-GW, the general shape and timing of the recession curve on the graph is relatively similar to that of the observed flow dataset. For this reason TopNet-GW BFI values generally match BFI estimates produced by the observed flow.

Table 7.6. Baseflow estimates from digital filter and Q90/Q50 methods from observed flow, TopNet-0 and TopNet-GW outputs over the investigation period.

Mangatarere 2	Dataset	EKF(opt)	EKF(def)	Q90/Q50
	Observed	0.52	0.53	0.28
	TopNet-0	0.10	0.37	0.40
	TopNet-GW	0.50	0.56	0.28
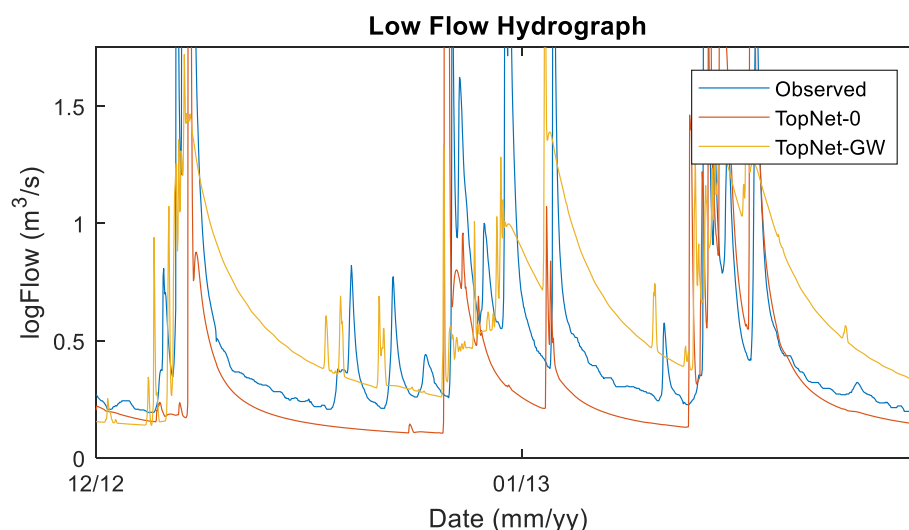
Figure 7.5. Observed and simulated hourly low flow hydrograph (i.e., flow below observed mean flow) over a 2012/13 summer period at Mangatarere 2 gauge.

7.3.2   *Mangatarere at Belvedere Bridge*

The quantitative performance metrics of the TopNet-0 and TopNet-GW outputs in relation to observed streamflow are presented in Table 7.7 for the investigation period (01/07/2010-30/06/2014). Figure 7.6 presents the comparison between observed and simulated flow at the Mangatarere 1 flow gauge location as standard and cumulative hydrographs and flow duration curves.

Table 7.7. Simulated streamflow validation statistics for Booths gauge over the validation period.

Gauge	NSE	logNSE	RSR	logRSR	PBIAS	R	logR^2
TopNet-0	-18.36	-1.18	4.40	1.48	-163.22	0.04	0.25
TopNet-GW	0.15	0.65	0.92	0.59	4.46	0.56	0.81

A negative NSE score (-18.36) indicates model failure by the TopNet-0 outputs at the Mangatarere 1 gauge. The large negative value for PBIAS (-163.22) states the model is over-estimating the flow volume at most of the data points. This is especially apparent in the two graphs in Figure 7.6, with the TopNet-0 line (red) well above the

observed flow dataset on both graphs. With respect to the log domain, quantitative model performance metrics improve, however an NSE of -1.18 indicates the mean of observed flow is a better predictor than the model output.

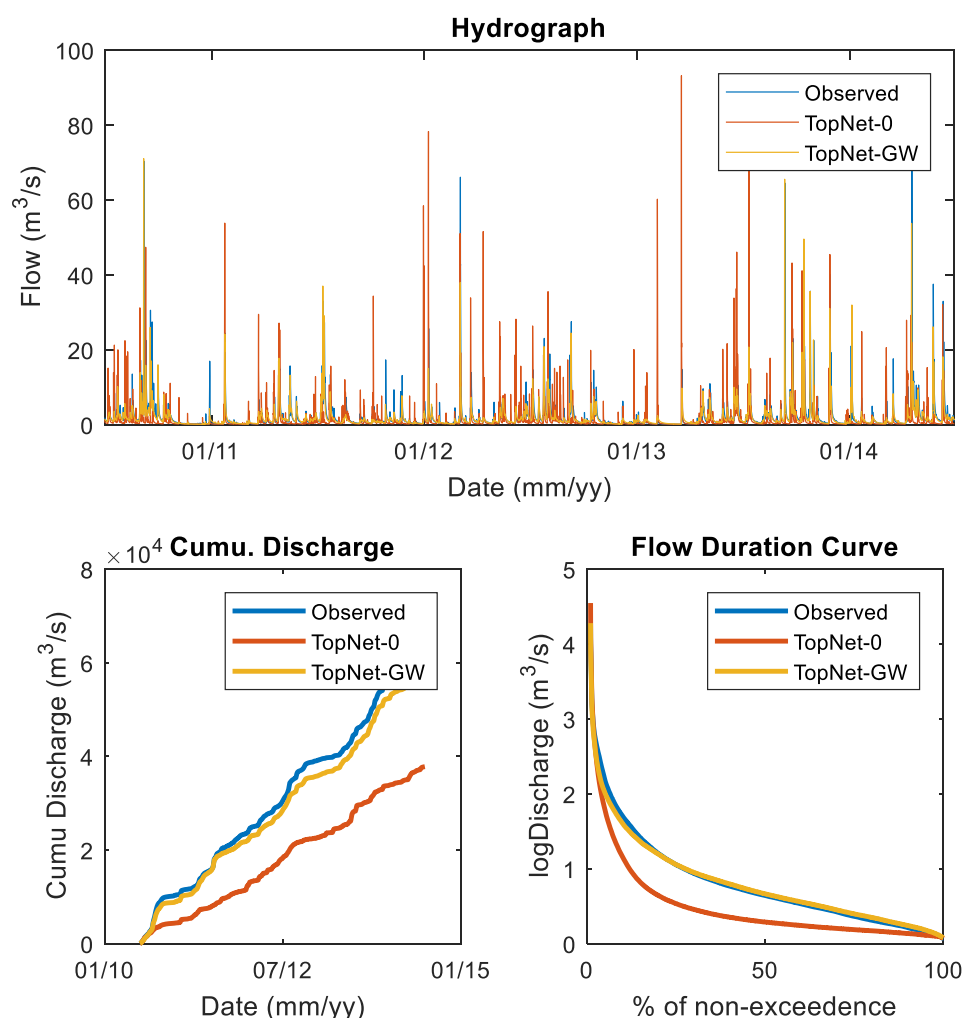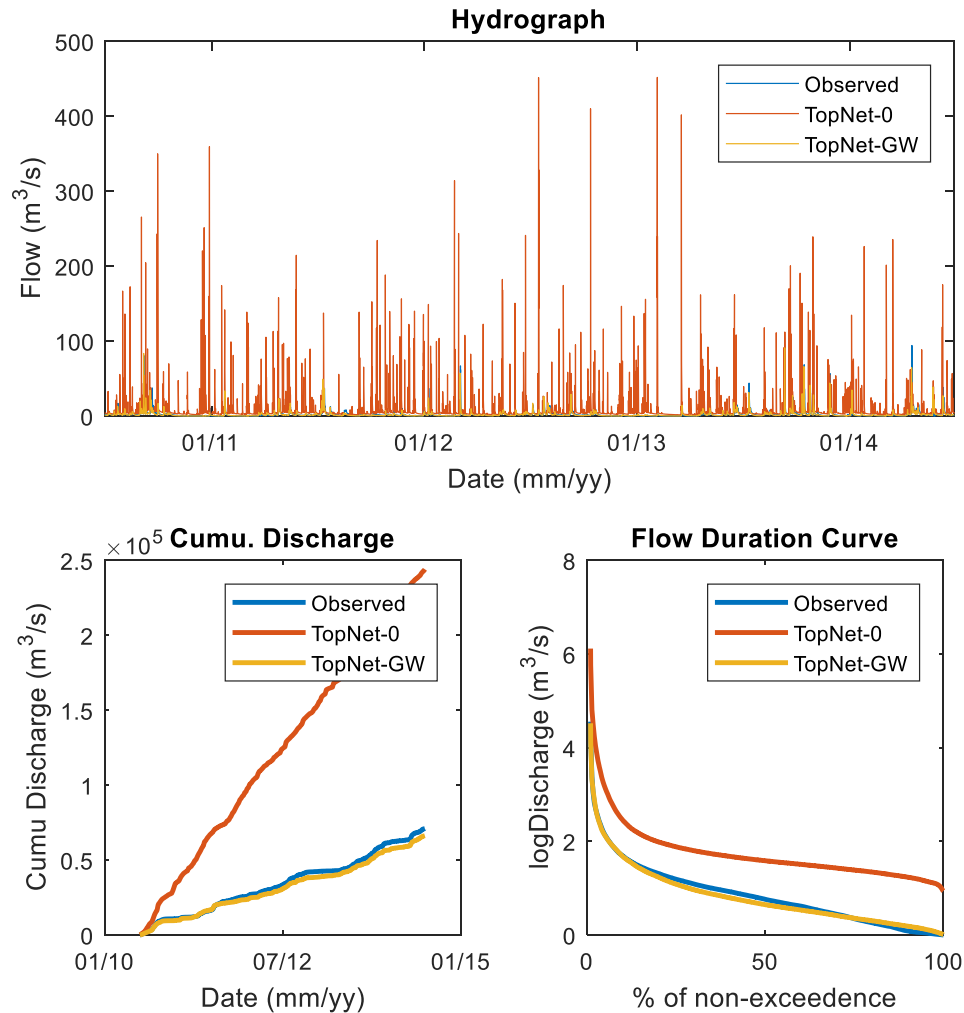*Figure 7.6. Observed and simulated cumulative hydrographs and flow duration curves at the Mangatarere 1 gauge over the investigation period.*

TopNet-GW exhibits a positive, albeit low NSE score (0.15) which suggests less than adequate model performance. The good PBIAS statistic further suggests that TopNet-GW does a good job at simulating observed flow levels despite a slight under estimation of flow as seen on the cumulative graph in Figure 7.6. The flow duration curve shows that TopNet-GW has a higher number of low-flows above the 75th percentile with fewer medium flows occurring below the 50th percentile. The low-flow

comparison (log domain) shows a good NSE score (0.65) and general correlation coefficient of 0.81.

Figure 7.7 presents scatter plot comparisons between simulated and observed flows over the validation period. The TopNet-GW outputs (right) appear more correlated to observed flow compared to TopNet-0 (left). Further, the spread between the 10th and 90th percentiles (black dots) is much narrower for TopNet-GW which shows good general agreement between the two datasets. The spread of data increases at higher flows for TopNet-GW. The TopNet-0 results are generally poor at all levels of flow.
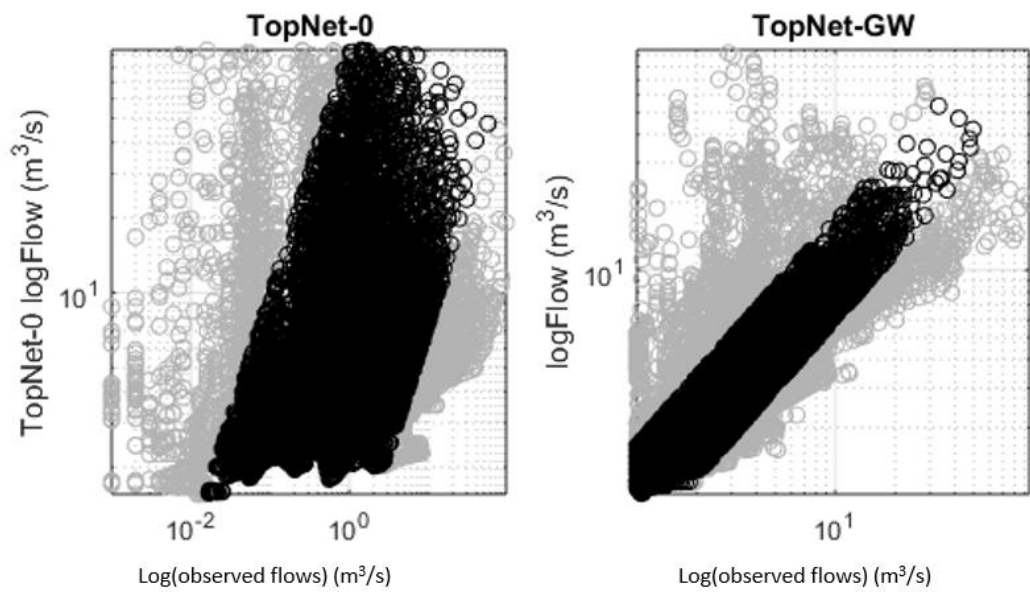
Figure 7.7. Scatter plots of observed flow against TopNet-0 (left) and TopNet-GW (right) at the Mangatarere 1 gauge. Data points in black are within the 80th percentile with regards to an error deviation statistic.

Table 7.8 presents the observed and simulated monthly average flows, these averages are also presented as a graph in Figure 7.8. TopNet-0 produces a higher flow regime over all seasons compared to the other datasets. TopNet-GW tracks well with observed data, especially over the summer period from December to May, however there is apparent underestimation by TopNet-GW between July and October. Regarding annual means, TopNet-GW simulations and observed flow dataset are relatively similar at 1.89 and 2.03 m^3/s respectively, while mean annual flow from TopNet-0 is much larger at 6.97 m^3/s.

Table 7.8. Simulated and observed monthly average flows (m³s) for the Mangatarere 1 gauge over the validation and the historic flow record (GWRC, 2017)

Months	TopNet-0	TopNet-GW	Observed
January	7.67	1.52	1.15
February	5.71	0.42	0.33
March	5.96	1.33	1.04
April	6.58	2.55	2.46
May	7.10	2.20	1.98
June	7.03	2.25	2.51
July	5.50	3.14	3.56
August	5.39	1.89	2.64
September	9.19	3.39	3.73
October	9.01	2.32	2.79
November	6.11	1.15	1.46
December	8.26	0.46	0.57
Annual	**6.97**	**1.89**	**2.03**
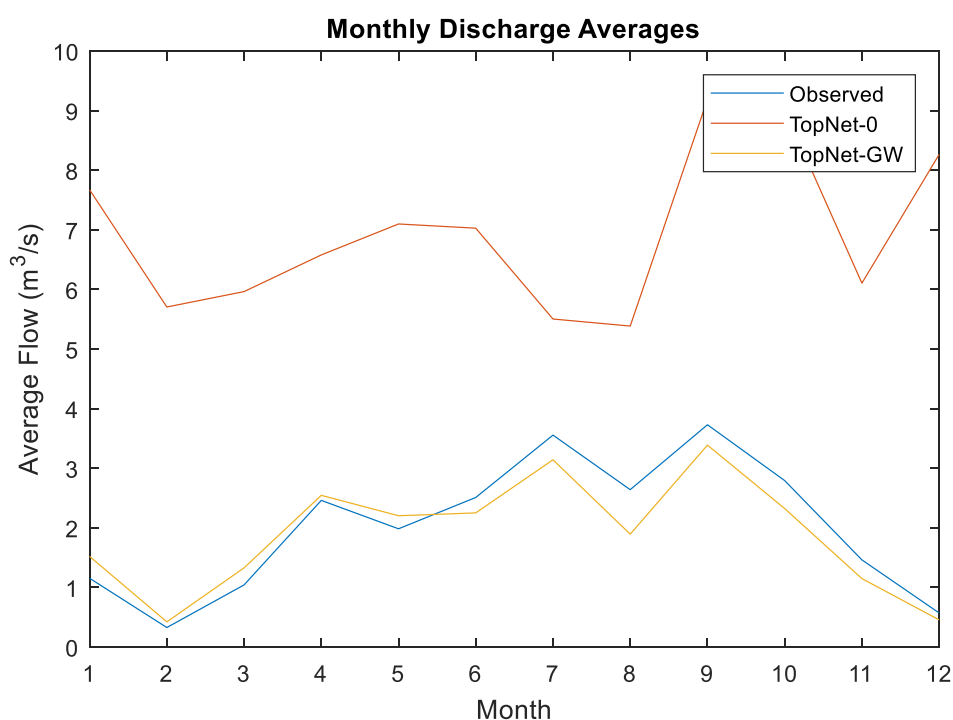
Figure 7.8. Observed and simulated monthly average flow at Mangatarere 1 over the investigation period 2011-2014.

Baseflow estimates for the Mangatarere 1 gauge are presented in Table 7.9. The BFI's produced from the TopNet-GW outputs are similar to the observed flow BFI's and despite the significant overestimation of total flow volumes, the BFI's produced for TopNet-0 are low compared to the other two datasets. Although the flow duration

curve in Figure 7.6 shows a higher representation of low-flows by TopNet-GW compared to observed flows, the BFI estimates are lower than the observed flow estimates. Interestingly, the observed flow Q90/Q50 statistic is significantly lower than other estimates. This is not represented on the hydrographs and in a catchment that is characterised by permeable gravels, a BFI of 0.09 is considered highly unlikely.

Table 7.9. BFI estimates from digital filter and Q90/Q50 methods from observed flow, TopNet-0 and TopNet-GW outputs over the investigation period.

Mangatarere 1.	Candidate	EKF(opt)	EKF(def)	Q90/Q50
	Observed	0.60	0.56	0.09
	TopNet-0	0.28	0.47	0.60
	TopNet-GW	0.49	0.55	0.22

Figure 7.9 presents the simulated and observed hourly low flow hydrographs (i.e., for discharge below observed mean flow). This magnified low-flow perspective presents multiple insights into propagation of flow for the two models. Firstly, the over-estimation of TopNet-0 outputs means there are no data points for TopNet-0 over this period that are below the mean of the observed flow data and so it does not appear on the low-flow graph. Secondly, this graph shows there are recessional discrepancies between observed and TopNet-GW flow hydrographs, despite good quantitative scores. With respect to the receding limbs, TopNet-GW exhibits a slower recession with the curve being less steep than the observed flow hydrograph.
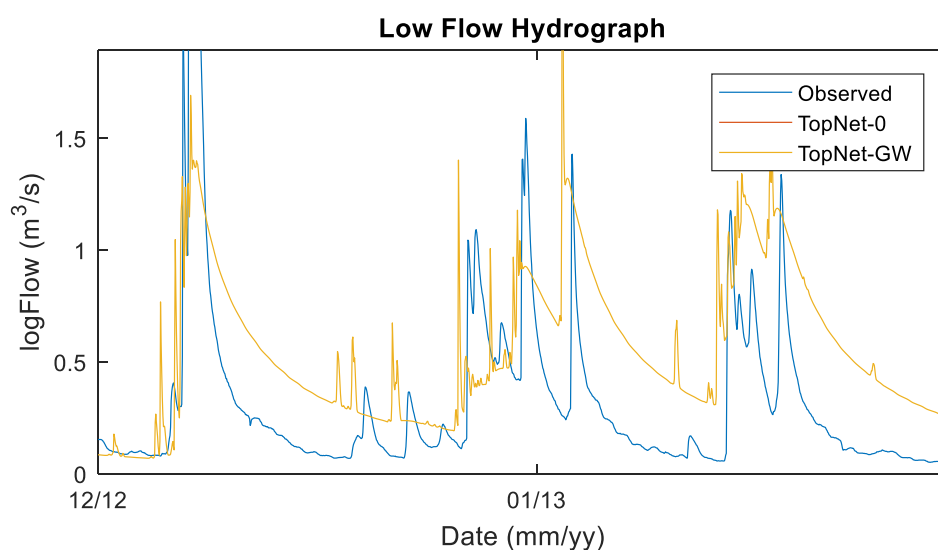
Figure 7.9. Observed and simulated hourly low flow hydrograph (i.e., flow below observed mean flow) over a 2012/13 summer period at Mangatarere 1 gauge.

7.3.3   Mangatarere at SH2

Quantitative performance metric for Mangatarere 3 gauge are presented in Table 7.10. Figure 7.10 presents the comparisons at the Mangatarere 3 flow gauge location as standard and cumulative hydrographs and flow duration curves.

Table 7.10. Simulated streamflow validation statistics for Booths gauge over the validation period.

Gauge	NSE	logNSE	RSR	logRSR	PBIAS	R	logR^2
TopNet-0	-1.61	-0.22	1.62	1.11	-57.78	0.26	0.46
TopNet-GW	0.24	0.82	0.87	0.42	8.27	0.52	0.91
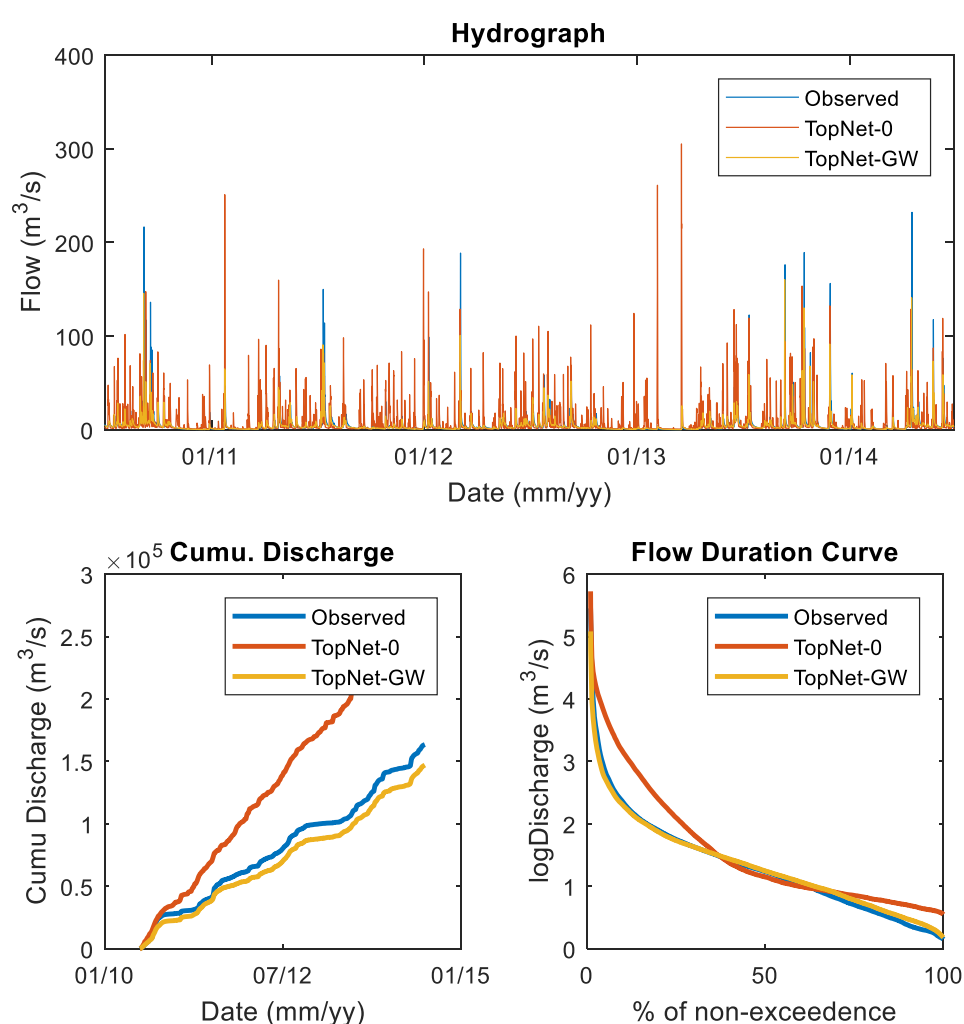
Figure 7.10. Observed and simulated cumulative hydrographs and flow duration curves at the Mangatarere 3 gauge over the investigation period.

118

TopNet-0 outputs exhibit a poor model performance across the range of quantitative statistics, particularly NSE (-1.61). The significant negative PBIAS score of -57.78 indicates model overestimation, which is apparent on the cumulative graph in Figure 7.10. There is only a slight improvement in statistics in the log domain.

TopNet-GW shows an adequate NSE score (0.24), which improves to an excellent value in the log domain (0.82). An adequate positive PBIAS (8.27) indicates a slight underestimation by the model in comparison to the observed streamflow dataset. Further, a correlation coefficient in the log domain of 0.91 indicates a good representation of flow produced by the TopNet-GW model at low flow regimes. These statistical conclusions are supported by general agreement of the TopNet-GW and observed data flow duration curves in Figure 7.10. As with the other two gauges, TopNet-GW exhibits a higher frequency of low-flows and lower frequency of high-flows. This partly explains the underestimation bias because if there is relatively more low flow data points and relatively less high flow data points, then the cumulative sum of those data points will be lower than the sum of the observed flow data.

Scatter plot comparisons between simulated and observed flows exhibit a wider spread toward higher flows for the TopNet-0 graph (Figure 7.11), which also shows a wide band between the 10th and 90th error percentiles (black dots). The TopNet-GW outputs also exhibit an increased spread compared to observed flow at higher flows. However, the 80% error statistic is narrow which implies the relative error is low.
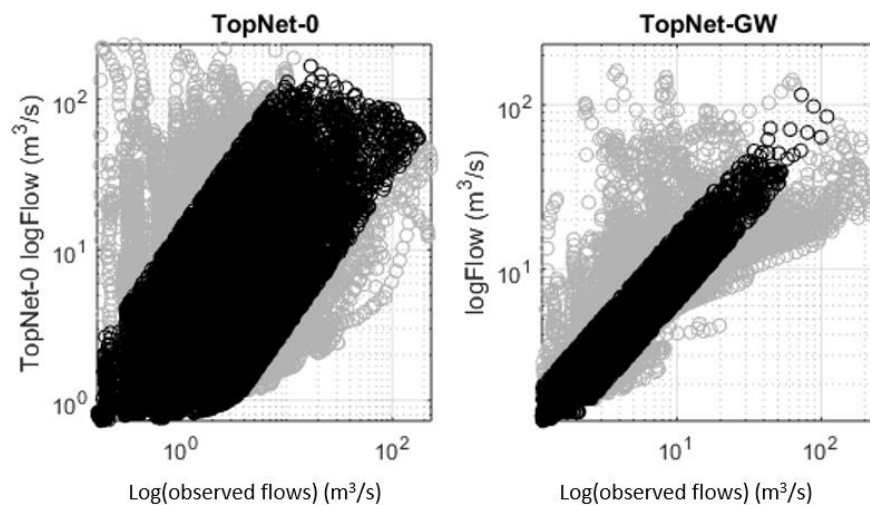
Figure 7.11. Scatter plots of observed flow against TopNet-0 (left) and TopNet-GW (right) at the Mangatarere 3 gauge. Data points in black are within the 80th percentile with regards to an error deviation statistic.

Table 7.11 presents the observed and simulated monthly average flows over the four year period, these averages are also presented as a graph in Figure 7.12. With respect to annual means, the TopNet-0 mean is almost double that of the TopNet-GW and observed flow annual means. The TopNet-GW annual mean is within a 0.5 cumec range of the observed flow mean. It appears to produce very similar mean flow values for the summer and autumn months, before generally under-estimating flows throughout the wetter winter and spring periods.

Table 7.11. Simulated and observed monthly average flows (m^3s) for the Mangatarere 1 gauge over the validation and the historic flow record (GWRC, 2017)

Months	TopNet-0	TopNet-GW	Observed
January	6.68	2.87	2.72
February	3.59	1.00	0.83
March	7.06	2.49	2.57
April	9.05	5.25	5.56
May	7.49	5.06	4.69
June	10.72	5.39	5.45
July	10.31	7.30	8.46
August	8.44	4.89	5.82
September	10.10	7.12	9.16
October	9.20	5.04	5.94
November	6.55	2.67	3.35
December	5.87	1.14	1.26
Annual	**7.94**	**4.20**	**4.67**
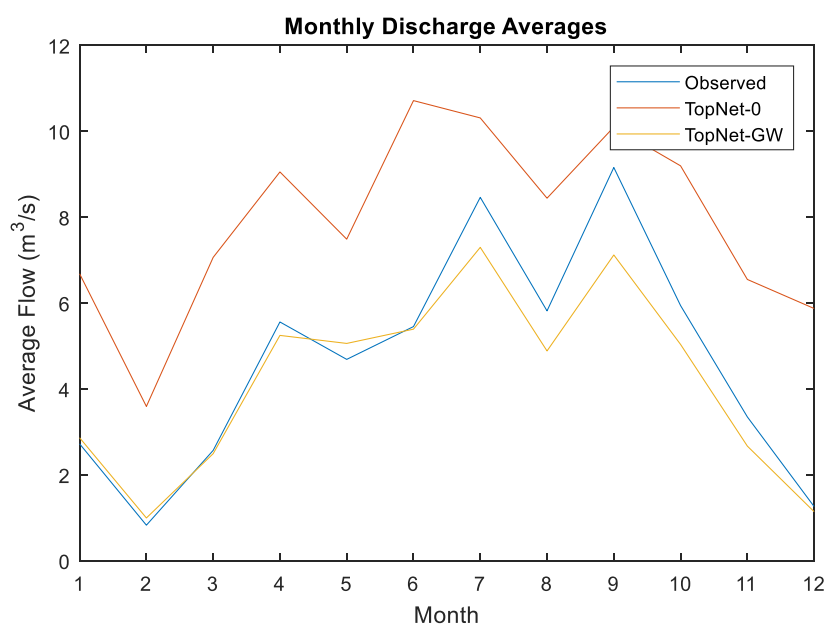
Figure 7.12. Observed and simulated monthly average flow at Mangatarere 3 over the investigation period 2011-2014.

Baseflow estimates for the Mangatarere 3 gauge are presented in Table 7.12 while Figure 7.13 presents the simulated and observed hourly low flow hydrographs (i.e., for discharge below observed mean flow). Baseflow estimates for TopNet-0 are much lower compared to the other two datasets, and yet, similar to the other gauges, the Q90/Q50 statistic is twice as large. TopNet-GW exhibits very similar BFI estimates in comparison to the observed flow BFIs. The likeness of low-flow dynamics between TopNet-GW and the observed dataset is apparent in Figure 7.13, where the two hydrographs exhibit very similar recession curves. In comparison, the recession of TopNet-0 is extremely abrupt, indicative of a more disconnected hydrological system.

Table 7.12. Baseflow estimates from digital filter and Q90/Q50 methods from observed flow, TopNet-0 and TopNet-GW outputs over the investigation period.

Mangatarere 3	Candidate	EKF(opt)	EKF(def)	BRF	Q90/Q50
	Observed	0.50	0.53	0.50	0.22
	TopNet-0	0.06	0.39	0.22	0.44
	TopNet-GW	0.50	0.55	0.56	0.23
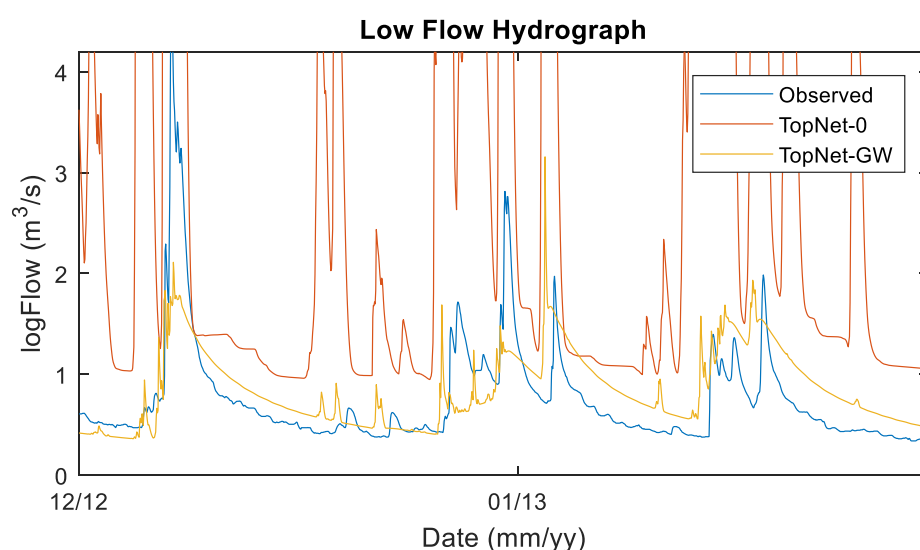
Figure 7.13. Observed and simulated hourly low flow hydrograph (i.e., flow below observed mean flow) over a 2012/13 summer period at Mangatarere 3 gauge.

7.3.4   Summary

Model validation methods carried out for TopNet-0 and TopNet-GW demonstrated that TopNet-GW outputs matched the observed gauge streamflow datasets generally well. In comparison, TopNet-0 outputs produced consistently poor comparisons to the observed streamflow data at all gauges. General discrepancies in flow simulation by

TopNet-GW compared to observed flow included: (a) higher frequency of low-flows that was balanced by a lower frequency of high-flows, (b) larger uncertainty at high flows with regards to the scatter plots, (c) shallower recession periods that led to higher low-flow volumes and (d) lower average monthly flows during winter and spring months

The best simulations of flow by TopNet-GW were at the Mangatarere 1 and Mangatarere 3 gauge sites. These are the two downstream gauges that are on the Wairarapa plains. In contrast the Mangatarere 2 gauge is stationed at the edge of the more complex interface between plains and mountainous Tararua ranges.

7.4 Model validation discussion

Quantitative and graphical comparisons of TopNet-0 and TopNet-GW against observed hourly flow showed that TopNet-GW outputs produced acceptable, and in many cases excellent, simulations of flow at the three flow Mangatarere river flow gauge sites. TopNet-0 outputs were consistently poor, producing negative NSE values at every gauge and not tracking well with observed flow in any of the graphical comparisons. In terms of a structural comparison of the two hydrological models, the significant bias exhibited by TopNet-0 at all gauges indicated poor calibration and parametrisation more than necessarily indicating inadequate model structure. Conversely, the small bias of the TopNet-GW results suggested excellent calibration and parametrisation. An assessment of the influence of the two model structures on simulated outputs is difficult to assess when these aspects of the two models are not consistent and therefore it is not extensively discussed. Instead, this discussion focusses on the ability of TopNet-GW to simulate total streamflow, specifically at low-flow periods, with a strong representation of groundwater dynamics estimated by baseflow separation.

With regards to low-flow dynamics, BFI's produced from the TopNet-GW outputs showed excellent agreement with BFI's calculated from the observed streamflow data, even when unique parameters were estimated for the model data (EK(opt)). The high BFI values were demonstrated on the hydrographs by relatively shallow recession curves. In contrast, the TopNet-0 BFI's were considerably lower and the TopNet-0 hydrograph displayed very abrupt recession curves. The agreement of BFI values and the good match to the shape of the observed flow recession curve indicate that not only

122

is TopNet-GW simulating total flow to a high standard, it is doing so in a way that is consistent with respect to baseflow contribution.

Seasonally, TopNet-GW produced excellent simulations throughout the summer and autumn months. However there was a general trend across all three gauges of under-estimation of simulations at some of the higher flow regimes through winter and early-spring. The scatter plots at each of the three gauges showed a tendency for model predictions to drift at the higher flows, this is likely to be the reason flow simulation is more uncertain during the winter months when high flows occur most frequently. A similar TopNet-0 investigation by Zammit and Yang (2017) in the Wairarapa region made similar conclusions, attributing high flow underestimation to a lack of rain gauges in the Tararua Ranges. This leads to poor spatial averaging of precipitation inputs (Zammit and Yang, 2017). Correcting for simulation at high flows is a significant area of investigation moving forward. However the model was implemented to enable low-flow forecasting for better management during drought and summer pressure months and therefore the issue can be considered relatively minor in this context.

At all three sites, TopNet-GW tracks below observed streamflow on the cumulative flow graphs. The general underestimation of flow, according to the flow duration curve graphs, is due to fewer medium to high flow values (above 60% occurrence) and more low flow values (below 10% occurrence) relative to the observed dataset. This discrepancy in the model is not necessarily represented by the BFI comparisons, because the calculation of BFI is as a proportion of total flow. Which means that if the timing and shape of the simulation matches the observed flow dataset, then BFI estimates will be similar. The relative dominance of low-flow values by TopNet-GW can also be observed in the low-flow graphs, where a slow recession is maintained at low flows and peak flows are poorly simulated. Further, the high-flow winter months are predominantly underestimated by TopNet-GW. It is possible that the model has diverted a larger portion of the water budget to the additional groundwater store rather than directly to the river, thus resulting in underestimation of event flow peaks and longer recessional periods as groundwater processes (slow recession) are over-represented. This highlights a need for further investigation into flow pathway parametrisation of the TopNet-GW model.

7.5 Model validation limitations

The TopNet-0 model showed significantly poor simulation results. This is most likely due to parametrisation and calibration issues rather than the structural performance. While TopNet-GW did simulate much better results, it is difficult to unequivocally state whether this improvement is due to the additional groundwater store (S_D) or just superior parametrisation and calibration compared with TopNet-0. Supporting information describing the relative differences of parametrisation and calibration of the two model applications was not available at the time of this research. An assessment of this information would provide significant insights into such questions.

Uncertainty analysis of the two models was not employed in this investigation. This is a significant step in the model development process and common to many hydrological investigations. Uncertainties in models arise at all stages, specifically with respect to input data, model structure, parametrisation and calibration errors (McMillan *et al.*, 2010). Zammit and Yang (2017) applied a Generalised Likelihood Uncertainty Estimation (GLUE), which is widely used in hydrological modelling (Beven and Binley, 1992). The GLUE approach includes obtaining a large number of randomly sampled behavioural datasets to produce behavioural simulations, which can then be used to derive uncertainties in the simulated flow (Beven and Binley, 1992).

Further uncertainty analysis of the observed streamflow data was also not considered in this investigation. Standard streamflow uncertainty analysis assesses random or systematic errors in the recorded data (river water levels/discharge data) used for comparison with the model outputs (Refsgaard and Storm, 1990). In many cases, an arbitrary relative uncertainty value is assigned for all discharge data (McMillan *et al.*, 2010). Léonard *et al.* (2000) and Schmidt (2002) approximate errors up to approximately 6% of the flow value provided by the meter, whereas Pelletier (1988) found the uncertainty of discharge measurements might be as high as 20% of the observed value. For this investigation, confidence was placed in the GWRC's assessment of measured streamflow quality and their reliably maintained flow rating curves.

Comparisons of modelled and observed estimates of BFI were found to be informative, however the uncertainty of baseflow separation outputs is an area that would benefit from further investigation. The root of such uncertainty is based around the parameter

124

estimation of the baseflow separation filters. Decisions not to use the observed flow backward filter BFI_{max} on the simulated flow outputs and other variations of parameters were considered but not quantitatively assessed or discussed. The BR method, which showed promise for baseflow separation at gauges with high average flow was not employed in the model comparison due to parameter estimation uncertainty at low flow gauges such as those in the Mangatarere catchment. Finally, a full dataset quantitative comparison could have been applied to the baseflow datasets generated by the digital filters instead of just comparing the averages of the modelled and observed baseflow datasets.

7.6 Model validation concluding remarks

The model validation procedure demonstrated a significant improvement of modelled results between TopNet-0 and TopNet-GW hydrological models. TopNet-0 produced generally poor simulations of flow at three sites in the Mangatarere catchment, whereas TopNet-GW modelled streamflow at a high standard. This emphasises the significance of representing groundwater in catchment scale models in areas with significant connectivity between groundwater and surface water. However, inferred discrepancies in parametrisation and calibration between the two models means the relative improvement of flow simulations is difficult to fully attribute to the additional TopNet-GW groundwater reservoir (S_D).

TopNet-GW exhibited BFI indices of roughly 0.5 at all three gauges, which are similar to the observed flow BFI estimates generated from the baseflow separation investigation in Chapter six. The good match of BFI estimates indicates that the TopNet-GW model simulates total streamflow with a high regard to the groundwater sourced baseflow component in the Mangatarere catchment. This conclusion is supported by the graphical observations, whereby the modelled flow on the low-flow hydrographs exhibit similar timing and shape to their observed streamflow datasets, especially with respect to recession periods. However, an apparent underestimation of total cumulative flow and monthly average flow during winter by TopNet-GW is potentially due to over-representation of the groundwater component over the four year investigation period.

The addition of baseflow index comparisons to a standard procedure of model validation has shown to be useful in order to assess whether the hydrological models

are representing groundwater flow dynamics in total flow simulations. Results showed that the TopNet-GW model produced BFI's similar to those estimated from observed streamflow and also simulated a good match to total flow of the two models. Conversely, the TopNet-0 outputs for BFI were significantly less than the observed streamflow estimates, this was observed to be function of a very quick recession in the TopNet-0 model simulations. The TopNet-0 simulations of total streamflow were also generally poor. This suggests that a good match between modelled and observed flow BFI estimates coincides with a good simulation of total streamflow at the three sites analysed in this investigation. With regards to the wider Wairarapa valley, similar BFI estimates of approximately 0.52 indicate that baseflow contribution is an equally important regional factor. Thus, stressing the importance of representative groundwater packages in catchment-scale hydrological models. The model validation procedure has demonstrated that a relatively efficient, conceptual model is able to simulate streamflow in areas, such as low-relief plains, that were once problematic (Yang *et al.*, 2017). Furthermore, the able simulation of low-flow dynamics, as investigated by BFI comparisons, means model predictions in these areas, which deal with significant low-flow water stresses during summer months, will be more reliable.

This particular investigation uses relatively subjective comparisons between modelled and observed baseflow estimates. Further understanding and estimation of robust baseflow indices from digital separation filters, along with comprehensive uncertainty analysis of observed streamflow and model outputs would result in more robust results that possessed a high level of reproducibility.

Chapter 8
Conclusions and recommendations

The overall aim of this research was to determine if existing baseflow separation methods could be used to inform on hydrological systems and partially validate catchment scale hydrological models in the Wairarapa valley, New Zealand. In order to achieve this, modelled flow outputs were compared to observed streamflow data in the Mangatarere catchment using standard model performance metrics and baseflow estimates from digital separation filters. An initial process of flow infilling was carried out to complete the observed streamflow records and allow for comparisons to be applied over four consecutive hydrological years. The objectives specified in Chapter 1 to meet the overall aim are addressed below.

8.1   Overall conclusions

The initial flow infilling investigation was successful largely due to the high correlation between donor and candidate gauges. The straight and difference infilling methods performed similarly across a range of statistical and graphical comparisons, demonstrating mostly excellent correlations and NSE values. The difference infilling method was deemed to produce the least number of artefacts compared to the straight infilling method, while the long-term multiplier was favoured over the monthly multiplier because of its ease of calculation and application. Results from the baseflow separation investigation demonstrate significant ground and surface water interaction throughout the Wairarapa valley, New Zealand. Baseflow separation methods employed at 13 Wairarapa gauges indicated an average proportion of baseflow contribution of approximately 52% of total streamflow over the four hydrological year period. Thus, highlighting the importance of groundwater representation in catchment-scale, hydrological modelling investigations. Conditional model validations at three flow gauges along the Mangatarere stream demonstrated that TopNet-GW simulations perform well at estimating total streamflow while representing baseflow contribution well. Simulations were compared to observed streamflow records using quantitative statistics and objective graphical comparisons. The TopNet-GW model outputs had consistently high NSE scores (> 0.78) and tracked well with the observed flow dataset at low flows in particular. BFI estimates calculated from the simulated flow datasets were found to be similar to the observed flow BFI estimates.

Flow Infilling

Flow infilling investigations in Chapter five assessed a standard, straight infilling method and a novel, difference infilling method using a suite of quantitative dataset comparison metrics. It was expected that the straight infilling method could exhibit significant steps when candidate data switched to donor data and vice versa. However, both the straight and difference infilling methods worked well. This was mostly due to the fact that the datasets investigated possessed such high correlations (> 0.85). When applied to datasets with worse correlations the straight infill method could exhibit these artefacts because of the direct placement of scaled donor data into the candidate dataset with no regard for the adjacent candidate data values. The difference infill method takes the scaled donor data and adds the flow difference (positive or negative) to the previous flow in the dataset to smooth the transition from candidate to donor data and is therefore the recommended option. Different multipliers were also assessed: a ratio of donor to candidate flow along the entire dataset, and a ratio of donor to candidate flow calculated for each month of the year. With respect to their performance, the long-term multiplier worked just as well as the monthly multiplier because the two values turned out to be very similar ratios in most cases. Due to the fact that the long-term multiplier is simpler to calculate and implement, it is recommended over the monthly calculated multiplier.

Baseflow Separation

The baseflow investigation, presented in Chapter six, utilised infilled streamflow datasets from the Wairarapa region in an attempt to estimate baseflow index from the Eckhardt and Bump & Rise recursive digital filters. The two filters demonstrated a high degree of connectivity between groundwater and surface water in the Wairarapa valley, New Zealand. According to the results of these filters, baseflow contribution to streams across the Wairarapa plains accounts for approximately half of the total flow. The indices produced in this investigation are from an average of four consecutive hydrological years, thus results can be considered representative of average annual baseflow contribution. This means they do not account for the significant variability of groundwater-surface water interaction at seasonal and individual rainfall event time scales. Digital baseflow separation is fast, simple and only requires streamflow datasets to generate estimates of the baseflow contribution to a stream. However, the methods remain relatively arbitrary and it is difficult to prove

correctness because of the theoretical nature of baseflow and relative sparsity of chemical tracer separation analysis. Further, the parameters employed in digital baseflow separation filters exhibit a significant influence on the final output, yet parameter estimation appears to be rather uncertain.

The porous sedimentary gravels that span the Wairarapa valley are likely to have a significant role on the high degree of groundwater-surface water connectivity inferred by the digital separation filters in this investigation. Expansive sedimentary basins similar to the Wairarapa plains are common throughout New Zealand, most of which possess a significant demand for freshwater due to intensive agricultural practices (Moreau and Bekele, 2015). The results of this baseflow separation investigation highlight the importance of considering groundwater-surface water interaction in these regions when addressing consents and water management issues.

Model validation and baseflow comparison

Chapter seven presented a methodology for model validation which can be applied to other catchments of interest. The two model output datasets investigated in the Mangatarere catchment were generated by the TopNet-0 and TopNet-GW conceptual hydrological models. Conceptual models are considered to be more computationally efficient than alternative model structures and represent hydrological reservoirs well. Their potential due to these aspects is often compromised by poorer accuracy and precision of outputs. However, continued efforts to assess the model outputs against observed data helps to reduce uncertainty, thus producing more accurate and precise outputs from conceptual models. Modelled outputs in this investigation were compared to observed flow at three sites along the Mangatarere River. Negative NSE scores and a large PBIAS (> 50%) indicated that TopNet-0 produced relatively poor simulations. This is most likely due to poor initial parametrisation and calibration as inferred by the significant bias in the outputs. This inequality of parametrisation and calibration meant a robust structural comparison between the two models was not possible.

TopNet-GW was generally successful at simulating streamflow at all three sites in the Mangatarere catchment. With a specific interest in low-flows, TopNet-GW performed well with 'good' annual mean NSE scores at two of the sites and an excellent score at the third when flows were compared in the log transformed domain. Eckhardt filter

baseflow estimates from the TopNet-GW simulated flow, using optimised values for the BFI_{max} parameter, demonstrated sound agreement with BFI estimates produced from the observed streamflow datasets at all three sites. It is important to note monthly calculations of NSE would have been beneficial, owing to the evident seasonality of discharge in the catchment.

An underestimation of total cumulative flow and of average flow during winter high-flow periods, along with a higher frequency of low-flow values compared to the observed dataset suggests that the TopNet-GW model was diverting too much water to the additional groundwater reservoir. Flow is therefore is over-represented at low-flows and under-represented at high-flows. This observation indicates the deep TopNet-GW additional groundwater reservoir performs well, but further parametrisation of relative flow pathways could improve prediction of streamflow by TopNet-GW. Conclusions drawn from the Mangatarere catchment investigation can be considered relevant to the wider Wairarapa region as the climate and geology characteristics are generally analogous. Future model investigations in this region will benefit greatly from adequately representing groundwater-surface water interaction through the TopNet-GW additional groundwater reservoir.

8.2 Avenues for future research

Although this research provided some insight into the connectivity between ground and surface water bodies and presented a potential method for evaluating groundwater representation in catchment scale hydrological models through recursive digital filters, there are still a number of avenues for future research and recommendations for improvement of hydrological investigations in New Zealand. These are presented in the following two sections. This research investigated the contribution of groundwater to streams as an average over four consecutive hydrological years and used this information to partially validate two catchment-scale hydrological models. However the dominant processes surrounding this interaction and the spatial and temporal scale at which this interaction occurs is not fully understood.

A number of research topics and questions that were only briefly touched on in this research and that require further attention include:
- With respect to the initial flow infilling, there appears to be no suggested upper limit in terms of the size of the missing data gap. This investigation used

infilling on gaps generally less than one month, however one gauge was infilled for almost one third of an entire year. The impacts that this extent of infilling had on the final comparison output is unknown. Future work would be possible by testing the infilling methods on observed flow at incremental temporal units using similar dataset comparison methods employed in this investigation.

- Baseflow separation filters are heavily influenced by parameters, and parameter estimation is still uncertain. Future local scale work into recession analysis, chemical tracer analysis, and catchment physiographical influences will help to inform regional scale baseflow separation filter parameters to ensure more robust estimates of BFI. A list of potential influential physiographical factors include, but is not limited to; catchment slope, area, geology, soil layers, stream – aquifer boundary etc.

- The temporal effects on baseflow estimates by digital separation filters is also not well quantified. This investigation utilised averages over four consecutive hydrological years. Investigation into baseflow changes over time, with a particular focus on dry summer periods when abstraction is at maximum, could be insightful. Further, winter and spring months appear to be a period when TopNet-GW underestimates flows. Temporal baseflow investigations could potentially help to resolve this issue.

- This investigation assessed the use of streamflow datasets for baseflow separation and model validation applications. GWRC also administer many groundwater stage monitoring sites that could provide insights into groundwater flow dynamics. Research into the relationships between groundwater gauges, digital baseflow parameters and streamflow gauges could prove insightful.

- It is evident from this investigation that the TopNet-GW models underestimate high flow rates, this statement is supported by findings in Zammit and Yang (2017). As such, their application over winter months is less certain. Further investigation into the interpolation of rainfall from rain gauges would help to inform modelling attempts. This is particularly important for catchments like the Mangatarere that initiate at elevation in the Tararua ranges (high rainfall)

and terminate on the Wairarapa plains (low rainfall). Furthermore, monthly comparisons of modelled versus observed data with respect to NSE, RSR and PBIAS would prove insightful in catchments that possess evident seasonality of discharge.

- Though initially postulated, this investigation did not explore the influence of large-scale climatic modes on flow dynamics in the Wairarapa valley. Frequency analysis of streamflow with respect to annual and decadal climate cycles could present insights into the relative influence of ENSO, PDO and SAM on streamflow variability. Meanwhile, examination of summer baseflow estimates over the last 30 years could inform on the influence of anthropogenic abstraction to stream and groundwater flow variability.

8.3   Recommendations

Results from the baseflow separation investigation (Chapter 6) suggest the contribution of groundwater to surface water in the Mangatarere River and wider Wairarapa valley surface waters is significant. The model validation investigation (Chapter 7) shows that when groundwater-surface interaction is accounted for in the model (i.e. TopNet-GW's 'S_D' reservoir) total streamflow is more accurately simulated. The establishment of this connectivity has noteworthy implications for water quality and quantity issues as agricultural practices continue to intensify in the region. Demand for groundwater as a resource is increasing and a reduction in quantity will lead to poorer water quality in both surface and groundwater reservoirs. Therefore, environmental management decisions, with regards to agricultural demands for freshwater, need to account for the relatively high connectivity between surface water and shallow groundwater bodies and the implication this may have on long term water supply and quality issues.

The flow infilling in this research turned out to be extremely successful, however it is likely that other catchments will not have gauges that exhibit such high correlations for use as donor flow gauges. This investigation has highlighted the substantial potential of historical flow datasets for estimating catchment parameters and validating hydrological models. As such, it is recommended that flow gauges should be reliably maintained with all efforts made to reduce the amount of error or malfunction in order to retain reliable and continuous flow records for future use. Further, a higher density

132

of flow gauges along streams would provide for spatial comparisons at finer resolutions. These recommendations would improve the approach, implementation and results of future environmental monitoring programmes and facilitate more informed environmental management decisions.

8.4 Summary

Guggenmos (2010) and Jones and Gyopari (2006) have inferred both gaining and losing reaches in streams throughout the Wairarapa valley from groundwater bore measurements and more intensive chemical tracer analysis. While investigations of this type are useful, they tend to be applied at smaller spatial scales compared to the spatial extents of hydrological models and results may refer to a very small region of the catchment they are trying to represent. Baseflow separation by digital filters estimates baseflow along a flow dataset by assessing the non-event flow periods of a river that are attributed to groundwater inputs. The specific application of this method means the results are representative averages of the entire river upstream of that particular flow gauge. Therefore, baseflow separation techniques can act as a bridge between the small-scale chemical tracer investigations and regional scale modelling investigations. As baseflow separation methods continue to be improved by chemical tracer studies, they can in turn be utilised to inform catchment-scale hydrological models. The holistic suite of methods considered in this research will advance the field of hydrology by furthering understanding of the significant interaction between shallow groundwater and surface water reservoirs at multiple spatial and temporal scales. These insights can be applied to other regions in New Zealand that are known to exhibit a high connectivity between groundwater and surface water such as the Otago and Canterbury plains. These areas are also under pressure with respect to abstraction from shallow groundwater reservoirs during summer dry periods.

Hydrological modelling of low-flow stream characteristics and stream-aquifer interactions is increasingly important in many agricultural basins throughout New Zealand, due to the increasing demand for groundwater abstraction that coincides with peak drought conditions every year (Fenwick *et al.*, 2004). These issues are likely to be exacerbated as climate change induced droughts become more prevalent (Tait *et al.*, 2002). By understanding the interaction between these reservoirs and incorporating this knowledge into conceptual hydrological models, water management options can

set realistic limits such that abstraction does not exceed the recharge rate of those reservoirs. Conceptual models have a huge potential compared with traditional model attempts because they are generally more efficient and require less catchment specific data. However model validation and comparisons against low-flow indices should be considered crucial to ensure accuracy and precision of the modelled outputs.

In many catchments, intensive chemical tracer and temperature investigations of groundwater-surface water connectivity are not available. Therefore, easily accessible historic flow records offer an alternative source of data. Investigations utilising streamflow have potential to provide information on surface and sub-surface catchment characteristics as well as ably inform catchment-scale hydrologic models. The data conditioning, baseflow investigation and model validation processes presented here can be considered a case-study moving forward for the investigation of groundwater characteristics, for improved model performance and water management, in catchments that are based on low-relief plains and that face water shortages around New Zealand.

References

Al-Faraj, F.A.M., Scholz, M., 2014. Incorporation of the Flow Duration Curve Method Within Digital Filtering Algorithms to Estimate the Base Flow Contribution to Total Runoff. Water Resources Management 28, 5477-5489.

Alvares, C.A., Stape, J.L., Sentelhas, P.C., de Moraes, G., Leonardo, J., Sparovek, G., 2013. Köppen's climate classification map for Brazil. Meteorologische Zeitschrift 22, 711-728.

Arnold, J., Allen, P., Muttiah, R., Bernhardt, G., 1995. Automated base flow separation and recession analysis techniques. Groundwater 33, 1010-1018.

Bandaragoda, C., Tarboton, D.G., Woods, R., 2004. Application of TOPNET in the distributed model intercomparison project. Journal of Hydrology 298, 178-201.

Beadel, S., Perfect, A., Rebergen, A., Sawyer, J., 2000. Wairarapa Plains ecological district. Survey report for the Protected Natural Areas Programme. Wellington Conservancy, Department of Conservation, Wellington.

Bear, J., 1979. Groundwater hydraulics. McGraw, New York.

Begg, J., Brown, L., Gyopari, M., Jones, A., 2005. A review of Wairarapa geology– with a groundwater bias. Institute of Geological and Nuclear Sciences Client Report 159.

Begg, J., Johnston, M., 2000. Geology of the Wellington are: scale 1:250,000. Institute of Geological & Nuclear Sciences, Lower Hutt.

Beven, K., 1989. Interflow. Unsaturated Flow in Hydrologic Modeling. Springer, pp. 191-219.

Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrological processes 6, 279-298.

Beven, K., Freer, J., 2001. A dynamic topmodel. Hydrological processes 15, 1993-2011.

Beven, K., Kirkby, M., 1979. A physically based, variable contributing area model of basin hydrology. Hydrological Sciences Journal 24, 43-69.

Beven, K., Lamb, R., Quinn, P., Romanowicz, R., Freer, J., Singh, V., 1995. Computer models of watershed hydrology. Littleton: Water Resources Publications 627, 668.

Beven, K., Young, P., 2013. A guide to good practice in modeling semantics for authors and referees. Water Resources Research 49, 5092-5098.

Biondi, D., Freni, G., Iacobellis, V., Mascaro, G., Montanari, A., 2012. Validation of hydrological models: Conceptual basis, methodological approaches and a proposal for a code of practice. Physics and Chemistry of the Earth, Parts A/B/C 42, 70-76.

Bloomfield, J., Allen, D., Griffiths, K., 2009. Examining geological controls on baseflow index (BFI) using regression analysis: An illustration from the Thames Basin, UK. Journal of Hydrology 373, 164-176.

Blöschl, G., Sivapalan, M., 1995. Scale issues in hydrological modelling: a review. Hydrological processes 9, 251-290.

Brodie, R., Hostetler, S., 2005. A review of techniques for analysing baseflow from stream hydrographs. Proceedings of the NZHS-IAH-NZSSS 2005 conference. Auckland New Zealand.

Brunke, M., Gonser, T., 1997. The ecological significance of exchange processes between rivers and groundwater. Freshwater biology 37, 1-33.

Bunny, T., Thompson, M., Keenan, L., 2014. Water allocation in the Ruamahanga whaitua. Greater Wellington Regional Council, Masterton.

Cai, W., Borlace, S., Lengaigne, M., Van Rensch, P., Collins, M., Vecchi, G., Timmermann, A., Santoso, A., McPhaden, M.J., Wu, L., 2014. Increasing frequency of extreme El Niño events due to greenhouse warming. Nature climate change 4, 111.

Chapman, T., 1991. Comment on "Evaluation of automated techniques for base flow and recession analyses" by RJ Nathan and TA McMahon. Water Resources Research 27, 1783-1784.

Chapman, T., 1999. A comparison of algorithms for stream flow recession and baseflow separation. Hydrological Processes 13, 701-714.

Chapman, T., Maxwell, A., 1996. Baseflow separation-comparison of numerical methods with tracer experiments. Hydrology and Water Resources Symposium 1996: Water and the Environment; Preprints of Papers. Institution of Engineers, Australia, p. 539.

Clark, M.P., Rupp, D.E., Woods, R.A., Zheng, X., Ibbitt, R.P., Slater, A.G., Schmidt, J., 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, 1309-1324.

Collischonn, W., Fan, F.M., 2013. Defining parameters for Eckhardt's digital baseflow filter. Hydrological Processes 27, 2614-2622.

Dahm, C.N., Grimm, N.B., Marmonier, P., Valett, H.M., Vervier, P., 1998. Nutrient dynamics at the interface between surface waters and groundwaters. Freshwater Biology 40, 427-451.

de Joux, R., Scarf, F., 2010. Review of the Hydrology, Water Allocation and Management of the Pareora Catchment water resources. Environment Canterbury.

Eckhardt, K., 2005. How to construct recursive digital filters for baseflow separation. Hydrological Processes 19, 507-515.

Eckhardt, K., 2008. A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. Journal of Hydrology 352, 168-173.

Eckhardt, K., 2012. Analytical sensitivity analysis of a two parameter recursive digital baseflow separation filter. Hydrology and Earth System Sciences 16, 451-455.

Fenwick, G., Thorpe, H., White, P., 2004. Groundwater systems. Freshwaters of New Zealand, 29.18.

Freeze, R.A., Cherry, J.A., 1979. Groundwater., 1979. Printice-Hall Inc, Englewood Cliffs.

GNS, 2018. Smart Models for Aquifer Management (SAM). GNS Science, https://www.gns.cri.nz/Home/Our-Science/Environment-and-Materials/Groundwater/Research-Programmes/Smart-Models-for-Aquifer-Management-SAM.

Gonzales, A., Nonner, J., Heijkers, J., Uhlenbrook, S., 2009. Comparison of different base flow separation methods in a lowland catchment. Hydrology and Earth System Sciences 13, 2055-2068.

Gordon, N.D., 1986. The southern oscillation and New Zealand weather. Monthly weather review 114, 371-387.

Goring, D.G., 1994. Kinematic shocks and monoclinal waves in the Waimakariri, a steep, braided, gravel-bed river. Proceedings of the International Symposium on

waves: Physical and numerical modelling, University of British Columbia, Vancouver, Canada.

Guggenmos, M., 2010. Groundwater and surface water interaction, Wairarapa valley, New Zealand.

GWRC, 2014. The Climate of the Ruamahanga Catchment - Summary Report. Greater Wellington Regional Council, http://www.gw.govt.nz.

GWRC, 2015. Proposed Natural Resources Plan. Greater Wellington Regional Council, Wellington.

GWRC, 2017. Environmental Monitoring & Research. Greater Wellington Regional Council, http://graphs.gw.govt.nz/.

GWRC, 2018. Restrictions - Ruamahanga River and Streams in Wairarapa valley. Greater Wellington Regional Council, http://www.gw.govt.nz/restrictions-ruamahanga-river-and-streams-in-wairarapa-valley/.

Haitjema, H.M., Mitchell-Bruker, S., 2005. Are water tables a subdued replica of the topography? Ground Water 43, 781-786.

Hall, F.R., 1968. Base-flow recessions—a review. Water Resources Research 4, 973-983.

Harvey, C.L., Dixon, H., Hannaford, J., 2012. An appraisal of the performance of data-infilling methods for application to daily mean river flow records in the UK. Hydrology Research 43, 618-636.

Herschy, R., 2002. The uncertainty in a current meter measurement. Flow Measurement and Instrumentation 13, 281-284.

Hewlett, J.D., Hibbert, A.R., 1967. Factors affecting the response of small watersheds to precipitation in humid areas. Forest hydrology 1, 275-290.

Hug, T., Benedetti, L., Hall, E.R., Johnson, B.R., Morgenroth, E., Nopens, I., Rieger, L., Shaw, A., Vanrolleghem, P.A., 2009. Wastewater treatment models in teaching and training: the mismatch between education and requirements for jobs. Water Science and Technology 59, 745-753.

Hughes, Gyopari, M., 2014. Wairarapa Valley groundwater resource investigation: Framework for conjunctive water management. Greater Wellington Regional Council, Wellington.

Hughes, Smakhtin, V., 1996. Daily flow time series patching or extension: a spatial interpolation approach based on flow duration curves. Hydrological Sciences Journal 41, 851-871.

Hunt, B., 1999. Unsteady stream depletion from ground water pumping. Groundwater 37, 98-102.

Jiang, N., Griffiths, G., Lorrey, A., 2013. Influence of large-scale climate modes on daily synoptic weather types over New Zealand. International Journal of Climatology 33, 499-519.

Jones, A.G., Baker, T., 2005. Groundwater monitoring technical report. Greater Wellington Regional Council.

Jones, A.G., Gyopari, M., 2006. Regional conceptual and numerical modelling of the Wairarapa groundwater basin. Greater Wellington Regional Council.

Kalbus, E., Reinstorf, F., Schirmer, M., 2006. Measuring methods for groundwater, surface water and their interactions: a review. Hydrology and Earth System Sciences Discussions 3, 1809-1850.

Kamp, P., 1992. Landforms of Wairarapa: a geological perspective. Landforms of New Zealand. Longman Paul, Auckland, NZ, 367-381.

Keenan, L., Thompson, M., Mzila, D., 2012. Freshwater allocation and availability in the Wellington region: State and trends. Greater Wellington Regional Council, Wellington.

Kidston, J., Renwick, J., McGregor, J., 2009. Hemispheric-scale seasonality of the Southern Annular Mode and impacts on the climate of New Zealand. Journal of Climate 22, 4759-4770.

Kottegoda, N., Elgy, J., 1977. Infilling missing flow data. Modeling Hydrologic Processes, edited by: Morel-Seytoux, HJ, Salas, JD, Sanders, TG, and Smith, RE, Proc. Fort Collins 3rd Int. Hydrol. Symp., On Theoretical and Applied Hydrology, Colorado State University, Fort Collins, Colorado, USA, pp. 27-29.

Le Coz, J., 2012. A literature review of methods for estimating the uncertainty associated with stage-discharge relations. WMO Rep. PO6a.

Lee, J., Begg, J., 2002. Geology of the Wairarapa area: scale 1:150,000. Institute of Geological & Nuclear Sciences, Lower Hutt.

Léonard, J., Mietton, M., Najib, H., Gourbesville, P., 2000. Rating curve modelling with Manning's equation to manage instability and improve extrapolation. Hydrological sciences journal 45, 739-750.

LINZ, 2016. NZ 8m Digital Elevation Model (2012). Land Information New Zealand, https://data.linz.govt.nz/layer/1768-nz-8m-digital-elevation-model-2012/metadata/.

Marshall, G.J., 2003. Trends in the Southern Annular Mode from observations and reanalyses. Journal of Climate 16, 4134-4143.

McConchie, J., 2000. From shaky beginnings. Dynamic Wellington. Eds McConchie, J., Winchester, D., Willis, R. Institute of Geography, Victoria University of Wellington. Wellington.

McCuen, R.H., Knight, Z., Cutter, A.G., 2006. Evaluation of the Nash–Sutcliffe efficiency index. Journal of Hydrologic Engineering 11, 597-602.

McDonnell, J., Bonell, M., Stewart, M., Pearce, A., 1990. Deuterium variations in storm rainfall: Implications for stream hydrograph separation. Water Resources Research 26, 455-458.

McMillan, H., Booker, D., Cattoën, C., 2016. Validation of a national hydrological model. Journal of Hydrology 541, 800-815.

McMillan, H., Freer, J., Pappenberger, F., Krueger, T., Clark, M., 2010. Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions. Hydrological Processes 24, 1270-1284.

McMillan, H., Hreinsson, E., Clark, M., Singh, S., Zammit, C., Uddstrom, M., 2013. Operational hydrological data assimilation with the recursive ensemble Kalman filter. Hydrology and Earth System Sciences 17, 21.

McMillan, H., Westerberg, I., 2015. Rating curve estimation under epistemic uncertainty. Hydrological processes 29, 1873-1882.

MfE, 2018. Climate change projections for the Wellington and Wairarapa region. Ministry for the Environment, http://www.mfe.govt.nz/climate-change/how-climate-change-affects-nz/how-might-climate-change-affect-my-region/wellington.

Moreau, M., Bekele, M., 2015. Groundwater component of the Water Physical Stock Account (WPSA). GNS Science.

Morgan, M., Hughes, B., 2001. Wellington. Groundwaters of New Zealand MR Rosen and PA White (eds.). New Zealand Hydrological Society Inc., Wellington, 397-410.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE 50, 885-900.

Mullan, A., 1995. On the linearity and stability of Southern Oscillation-climate relationships for New Zealand. International Journal of Climatology 15, 1365-1386.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. Journal of hydrology 10, 282-290.

Nathan, R., McMahon, T., 1990. Evaluation of automated techniques for base flow and recession analyses. Water Resources Research 26, 1465-1473.

Negreiros, G.d., Alencar, A., Schlesinger, P., Nepstad, D., Lefebvre, P., 2009. Pre-LBA RADAMBRASIL Project Data. ORNL DAAC.

NIWA, 2018. The National Climate Database. National Institute of Water and Atmospheric Sciences, https://www.niwa.co.nz/our-services/online-services.

Pelletier, P.M., 1988. Uncertainties in the single determination of river discharge: a literature review. Canadian Journal of Civil Engineering 15, 834-850.

Perrie, A., 2005. Lake Wairarapa water quality monitoring technical report. Greater Wellington Regional Council, Wellington.

Philander, S.G.H., 1983. El Nino southern oscillation phenomena. Nature 302, 295.

Pinder, G.F., Jones, J.F., 1969. Determination of the ground-water component of peak discharge from the chemistry of total runoff. Water Resources Research 5, 438-445.

Rantz, S.E., 1982. Measurement and computation of streamflow: volume 2, computation of discharge. USGPO.

Rees, G., 2008. Hydrological data. Manual on Low-flow Estimation and Prediction. Operational Hydrology Report.

Refsgaard, J.C., 1997. Parameterisation, calibration and validation of distributed hydrological models. Journal of hydrology 198, 69-97.

Refsgaard, J.C., Henriksen, H.J., 2004. Modelling guidelines—terminology and guiding principles. Advances in Water Resources 27, 71-82.

Refsgaard, J.C., Storm, B., 1990. Construction, calibration and validation of hydrological models. Distributed hydrological modelling. Springer, pp. 41-54.

Robinson, S., 2004. Simulation: the practice of model development and use. Wiley Chichester.

Rykiel, E., 1996. Testing ecological models: the meaning of validation. Ecological modelling 90, 229-244.

Salinger, M., Mullan, A., 1999. New Zealand climate: temperature and precipitation variations and their links with atmospheric circulation 1930–1994. International Journal of Climatology: A Journal of the Royal Meteorological Society 19, 1049-1071.

Salinger, M., Renwick, J., Mullan, A., 2001. Interdecadal Pacific oscillation and south Pacific climate. International Journal of Climatology 21, 1705-1721.

Schmidt, A.R., 2002. Analysis of stage-discharge relations for open-channel flows and their associated uncertainties. University of Illinois at Urbana-Champaign.

Schwartz, F.W., Zhang, H., 2003. Fundamentals of Groundwater John Wiley & Sons. New York 583.

Shiklomanov, I., 1993. Water in Crisis: A Guide to the World's Fresh Water Resources, PH Gleick, ed. Oxford Univ. Press, New York 1, 13-24.

Singh, J., Knapp, H., Demissie, M., 2004. Hydrologic modeling of the Iroquois River watershed using HSPF and SWAT. ISWS CR 2004–08. Illinois State Water Survey, Champaign. CR/ISWSCR2004-08. pdf.

Sklash, M.G., Farvolden, R.N., 1979. The role of groundwater in storm runoff. Developments in Water Science. Elsevier, pp. 45-65.

Sloto, R.A., Crouse, M.Y., 1996. HYSEP, a computer program for streamflow hydrograph separation and analysis.

Smakhtin, V., 2001a. Estimating continuous monthly baseflow time series and their possible applications in the context of the ecological reserve. Water Sa 27, 213-218.

Smakhtin, V., 2001b. Low flow hydrology: a review. Journal of hydrology 240, 147-186.

142

Sophocleous, M., 2002. Interactions between groundwater and surface water: the state of the science. Hydrogeology Journal 10, 52-67.

Sophocleous, M., Koussis, A., Martin, J., Perkins, S., 1995. Evaluation of simplified stream-aquifer depletion models for water rights administration. Groundwater 33, 579-588.

Spongberg, M., 2000. Spectral analysis of base flow separation with digital filters. Water Resources Research 36, 745-752.

Stewart, M., 2015. Promising new baseflow separation and recession analysis methods applied to streamflow at Glendhu Catchment, New Zealand. Hydrology and Earth System Sciences 19, 2587-2603.

Stewart, M., Mehlhorn, J., Elliott, S., 2007. Hydrometric and natural tracer (oxygen-18, silica, tritium and sulphur hexafluoride) evidence for a dominant groundwater contribution to Pukemanga Stream, New Zealand. Hydrological Processes: An International Journal 21, 3340-3356.

Stoelzle, M., Stahl, K., Weiler, M., 2013. Are streamflow recession characteristics really characteristic? Hydrology and Earth System Sciences 17, 817-828.

Szilagyi, J., 2004. Vadose zone influences on aquifer parameter estimates of saturated-zone hydraulic theory. Journal of hydrology 286, 78-86.

Tait, A., Bell, R., Burgess, S., Gorman, R., Gray, W., Larsen, H., Mullan, B., Reid, S., Sansom, J., Thompson, C., Wratt, D., 2002. Meteorological Hazards and the Potential Impacts of Climate Change in Wellington Region A Scoping Study. Wellington Regional Council, Wellington.

Tallaksen, L., 1995. A review of baseflow recession analysis. Journal of hydrology 165, 349-370.

Thompson, D.W., Solomon, S., 2002. Interpretation of recent Southern Hemisphere climate change. Science 296, 895-899.

Todd, D.K., Mays, L.W., 2005. Groundwater hydrology. John Wiley & Son, Inc., New Jersey, 3.

Tomkins, K.M., 2014. Uncertainty in streamflow rating curves: methods, controls and consequences. Hydrological Processes 28, 464-481.

Watts, L., Perrie, A., 2007. Lower Ruamahanga River instream flow assessment. Greater Wellington Regional Council.

Welderufael, W.A., Woyess, Y.E., 2010. Stream flow analysis and comparison of base flow separation methods: Case study of the Modder River Basin in Central South Africa. European Water 31, 3-12.

Winter, T.C., 1998. Ground water and surface water: a single resource. DIANE Publishing Inc.

Winter, T.C., 1999. Relation of streams, lakes, and wetlands to groundwater flow systems. Hydrogeology Journal 7, 28-45.

Wittenberg, H., Sivapalan, M., 1999. Watershed groundwater balance estimation using streamflow recession analysis and baseflow separation. Journal of hydrology 219, 20-33.

Woessner, W.W., 2000. Stream and fluvial plain ground water interactions: rescaling hydrogeologic thought. Groundwater 38, 423-429.

Yang, J., McMillan, H., Zammit, C., 2017. Modeling surface water-groundwater interaction in New Zealand: Model development and application. Hydrological Processes 31, 925-934.

Yeh, S.-W., Kug, J.-S., Dewitte, B., Kwon, M.-H., Kirtman, B.P., Jin, F.-F., 2009. El Niño in a changing climate. Nature 461, 511.

Younger, P.L., 2009. Groundwater in the environment: an introduction. John Wiley & Sons.

Zammit, C., Yang, J., 2017. Surface Hydrology in the Ruamahanga Hill Country catchment: model calibration and validation. NIWA, Christchurch.

Zhang, R., Li, Q., Chow, T.L., Li, S., Danielescu, S., 2013. Baseflow separation in a small watershed in New Brunswick, Canada, using a recursive digital filter calibrated with the conductivity mass balance method. Hydrological processes 27, 2659-2665.

Appendix A
Flow infilling multipliers and graphs

Table A.1. January to June monthly multipliers used for monthly flow infilling, calculated from the Candidate mean flows divided by the Donor mean flows over each month.

Candidate	Jan	Feb	Mar	Apr	May	Jun
Kopuaranga 1	0.775	0.656	0.715	0.801	0.810	0.762
Kopuaranga 2	1.288	1.525	1.399	1.249	1.234	1.312
Mangatarere 1	0.432	0.415	0.402	0.439	0.418	0.451
Mangatarere 2	1.116	1.132	1.108	0.946	0.994	0.889
Mangatarere 3	2.316	2.408	2.489	2.281	2.391	2.216
Otakura	0.005	0.005	0.005	0.006	0.006	0.008
Ruamahanga 1	1.026	1.014	0.979	0.985	0.948	0.976
Ruamahanga 2	3.568	3.337	3.571	3.646	3.528	3.437
Ruamahanga 3	0.280	0.300	0.280	0.274	0.284	0.291
Tauherenikau	0.320	0.313	0.344	0.369	0.393	0.423
Waingawa	0.974	0.987	1.022	1.016	1.055	1.025
Waiohine	2.575	2.462	2.421	2.421	2.336	2.340
Waipoua 2	0.159	0.140	0.162	0.174	0.184	0.161

Table A.2. July to December monthly multipliers used for monthly flow infilling, calculated from the Candidate mean flows divided by the Donor mean flows over each month.

| Candidate	Jul	Aug	Sep	Oct	Nov	Dec
Kopuaranga 1	0.777	0.749	0.812	0.803	0.698	0.647
Kopuaranga 2	1.285	1.337	1.233	1.245	1.433	1.548
Mangatarere 1	0.429	0.447	0.402	0.461	0.444	0.463
Mangatarere 2	0.830	0.821	0.879	0.895	0.963	1.061
Mangatarere 3	2.330	2.237	2.489	2.167	2.253	2.157
Otakura	0.009	0.009	0.008	0.006	0.006	0.006
Ruamahanga 1	0.938	0.941	0.996	1.006	1.017	1.016
Ruamahanga 2	3.529	3.520	3.446	3.363	3.530	3.703
Ruamahanga 3	0.283	0.284	0.291	0.297	0.283	0.270
Tauherenikau	0.435	0.414	0.360	0.359	0.331	0.308
Waingawa	1.066	1.063	1.004	0.994	0.983	0.984
Waiohine	2.354	2.265	2.330	2.480	2.554	2.586
Waipoua 2	0.167	0.163	0.165	0.186	0.161	0.134
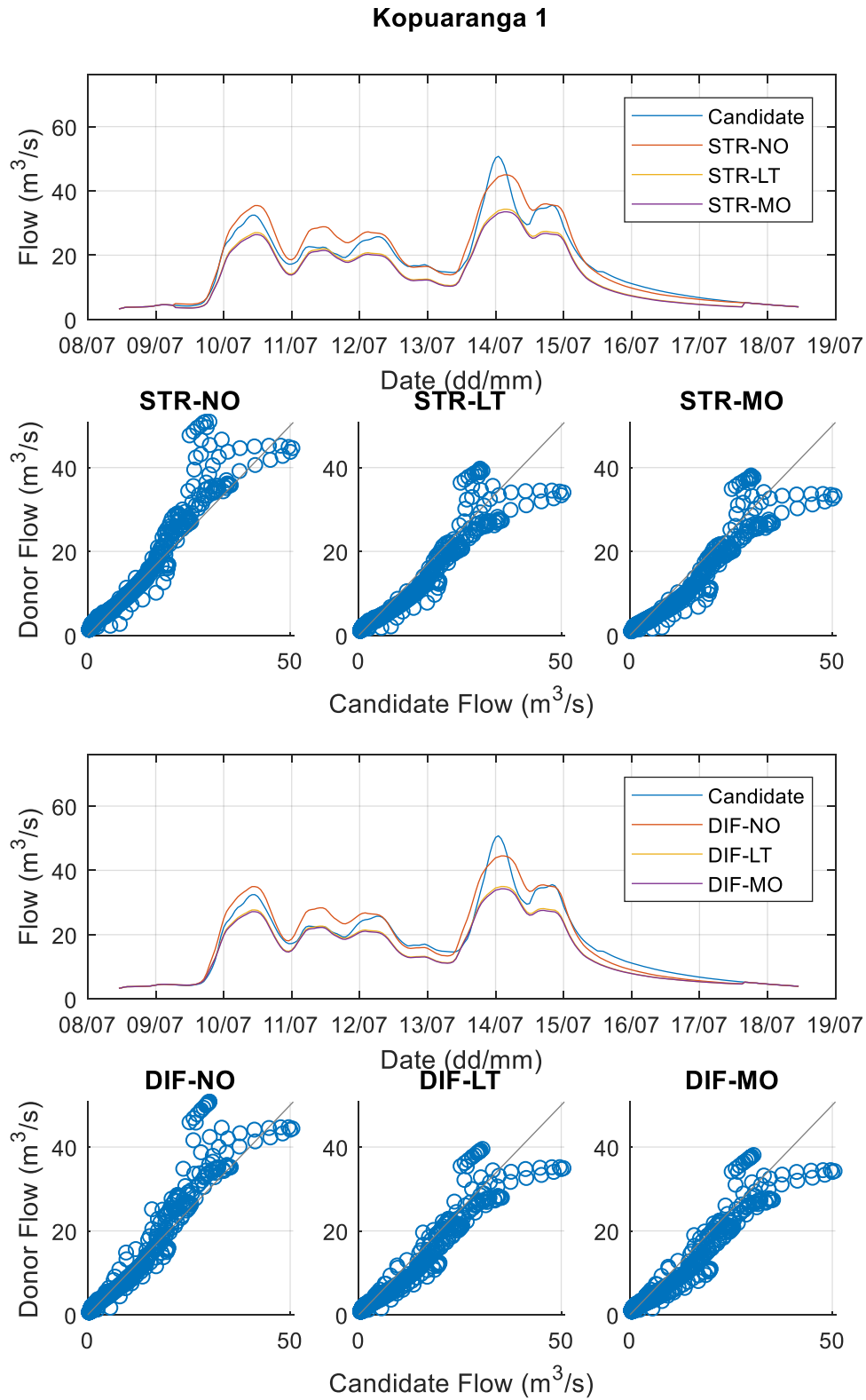
Figure A.1. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Kopuaranga 1.
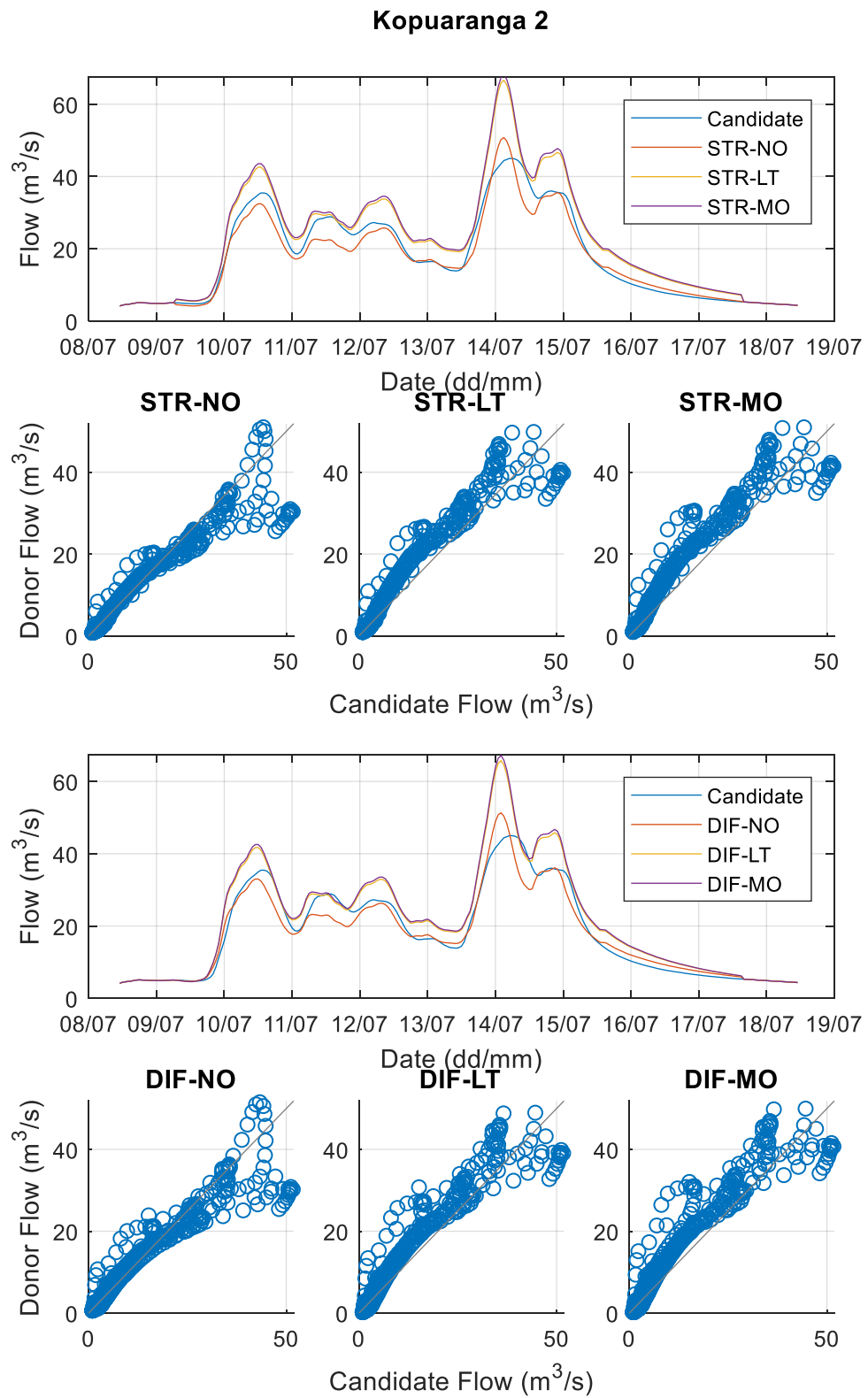
Figure A.2. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Kopuaranga 2.
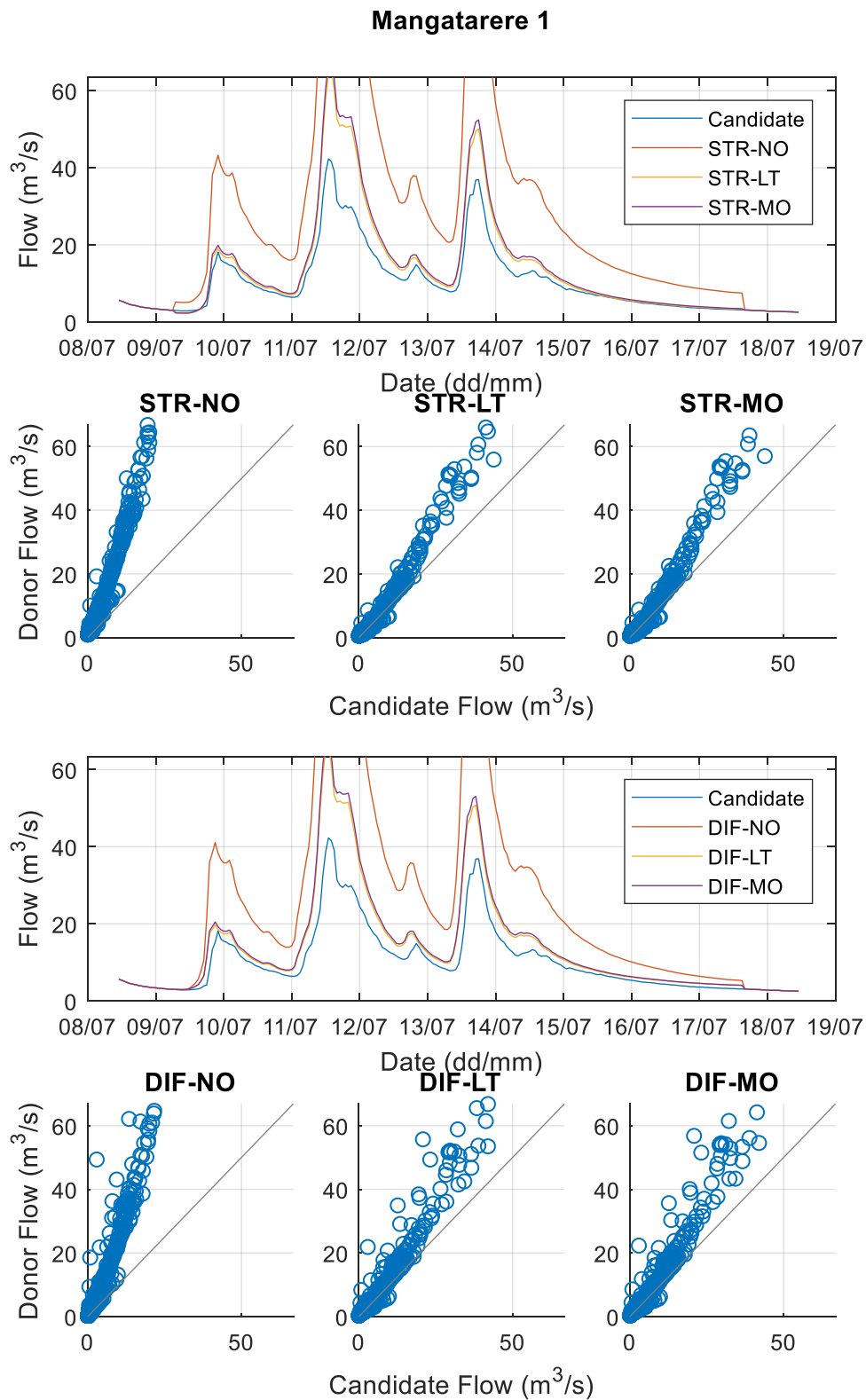
Figure A.3. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Mangatarere 1.
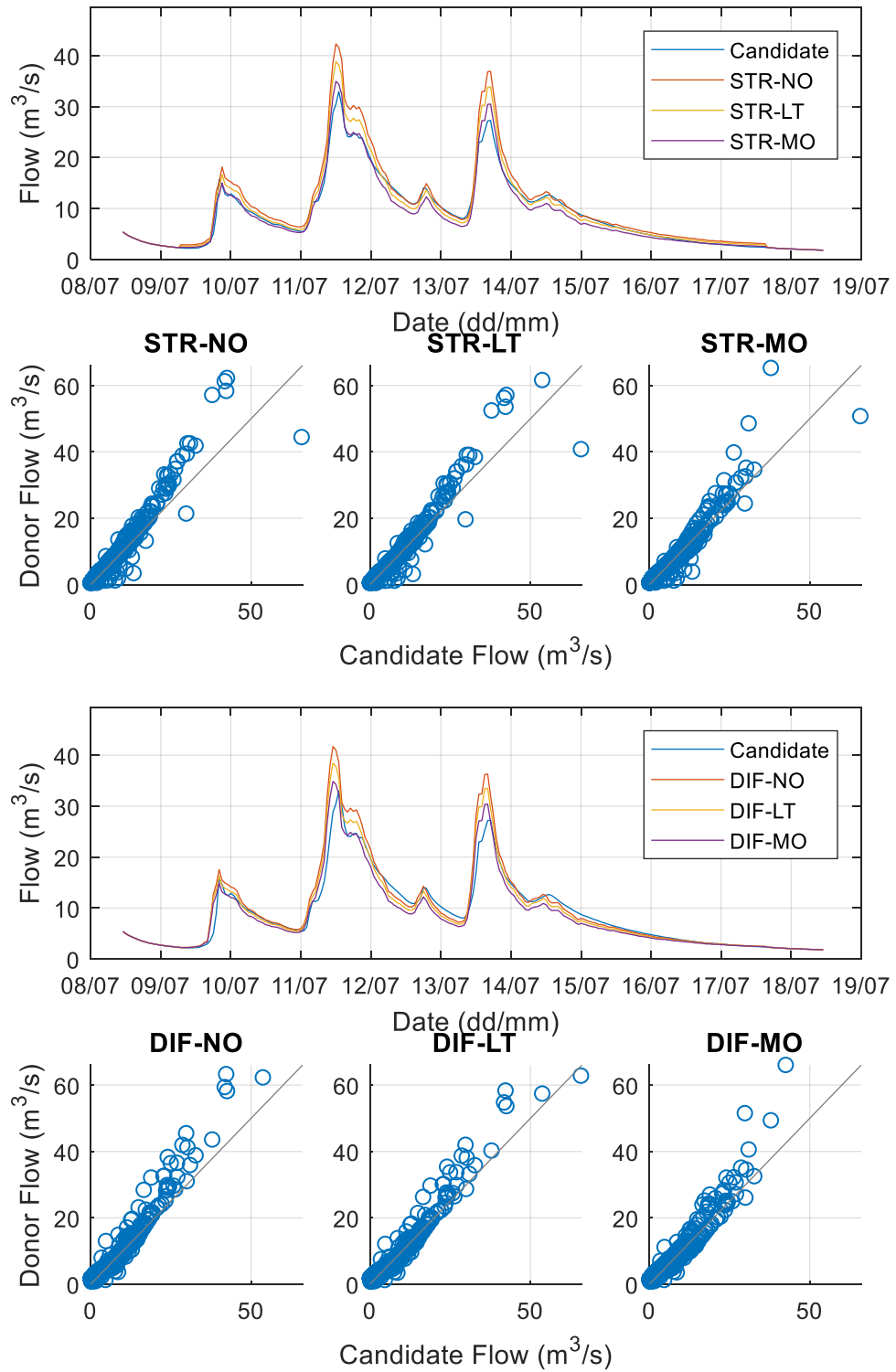
Figure A.4. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Mangatarere 2.
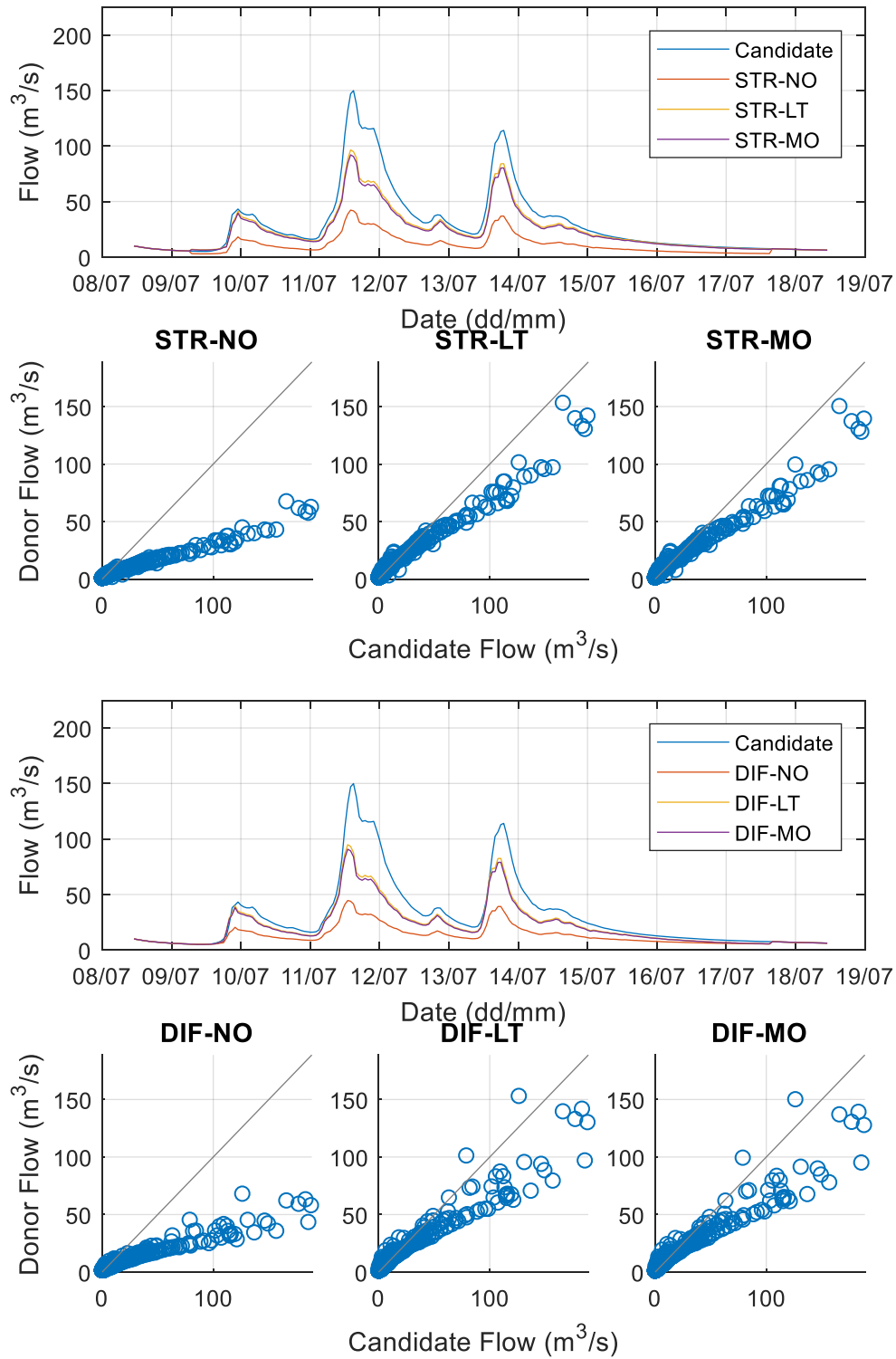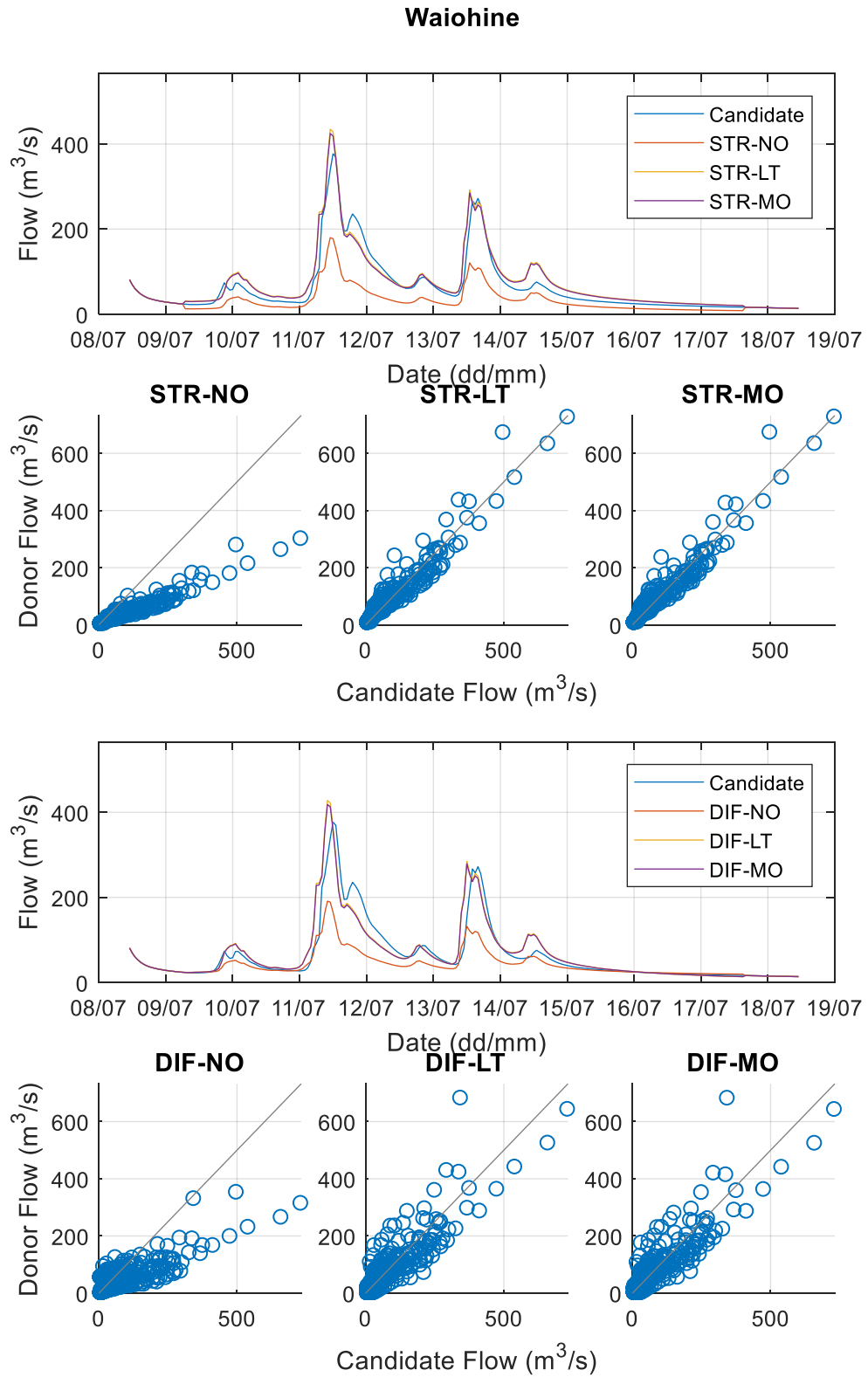
Figure A.5. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Mangatarere 3.
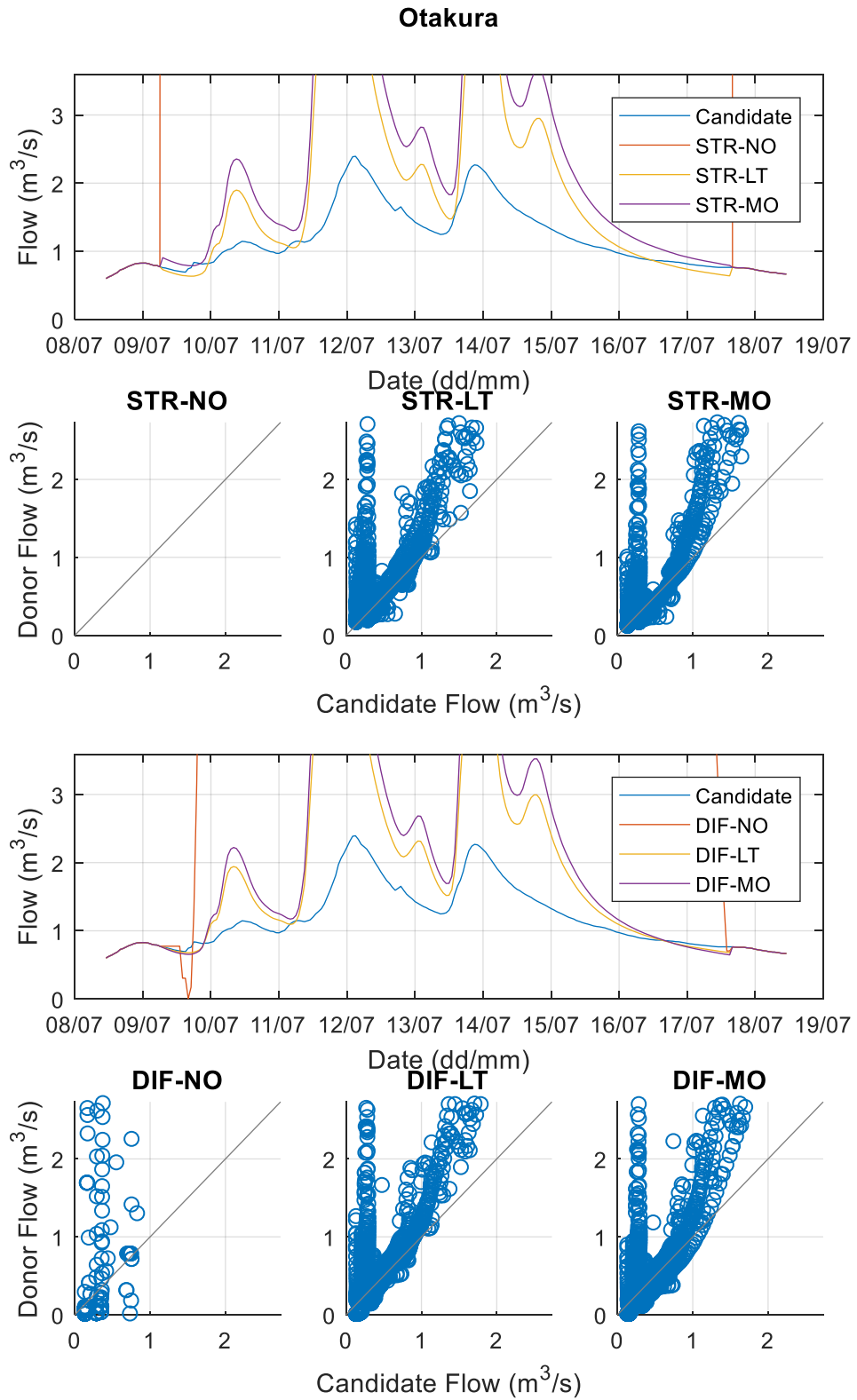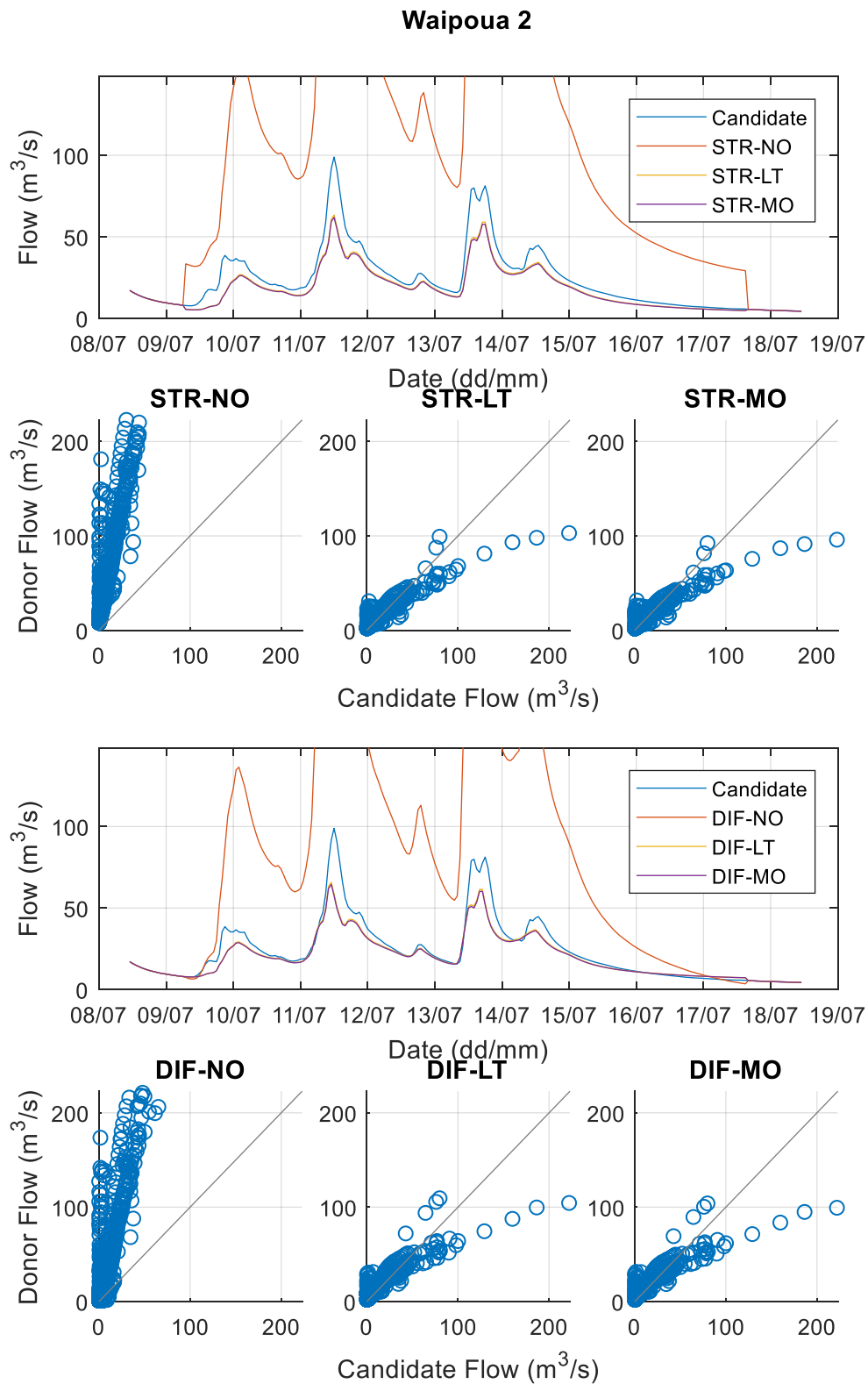
Figure A.6. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Otakura.
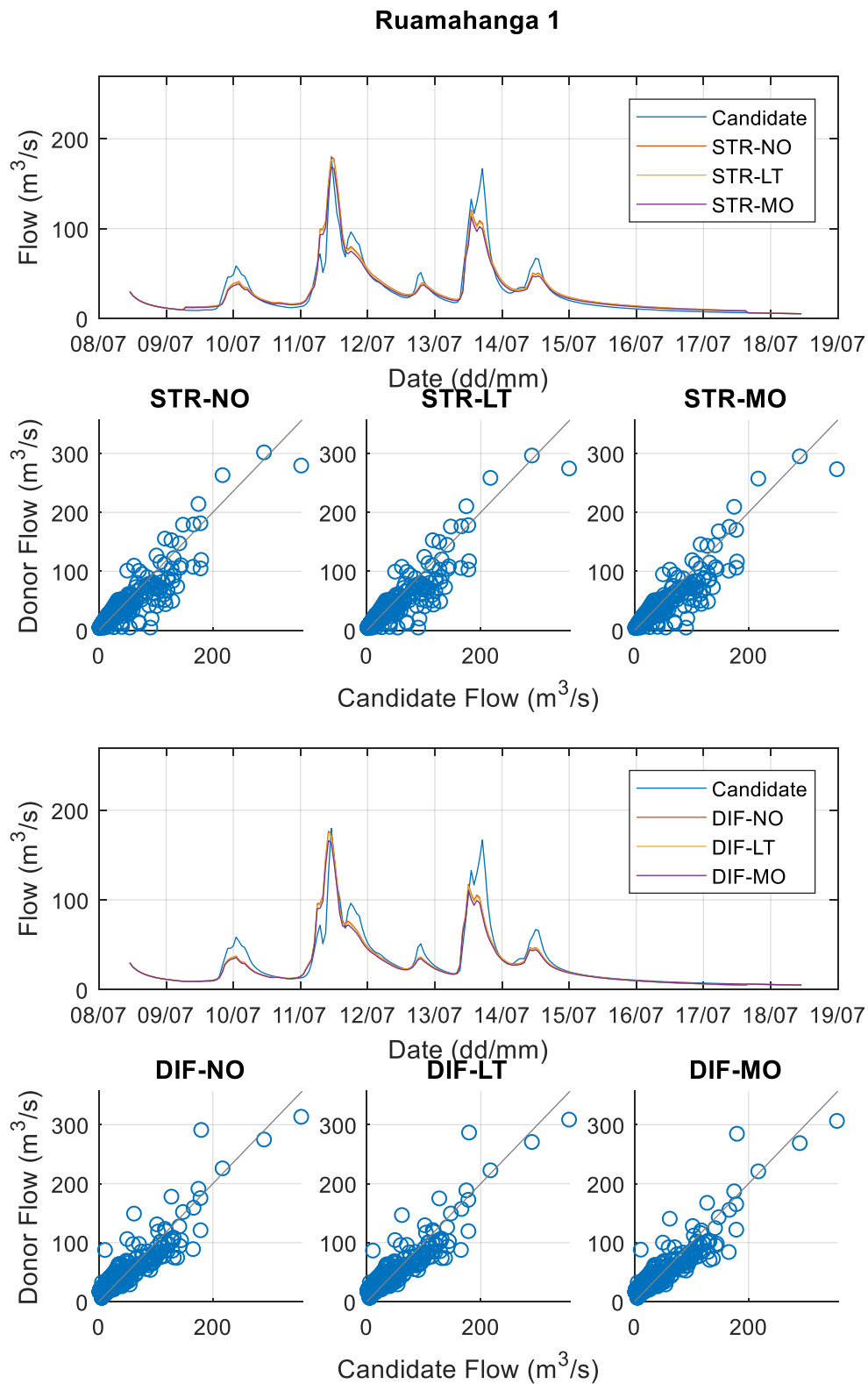
Figure A.7. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Ruamahanga 1.
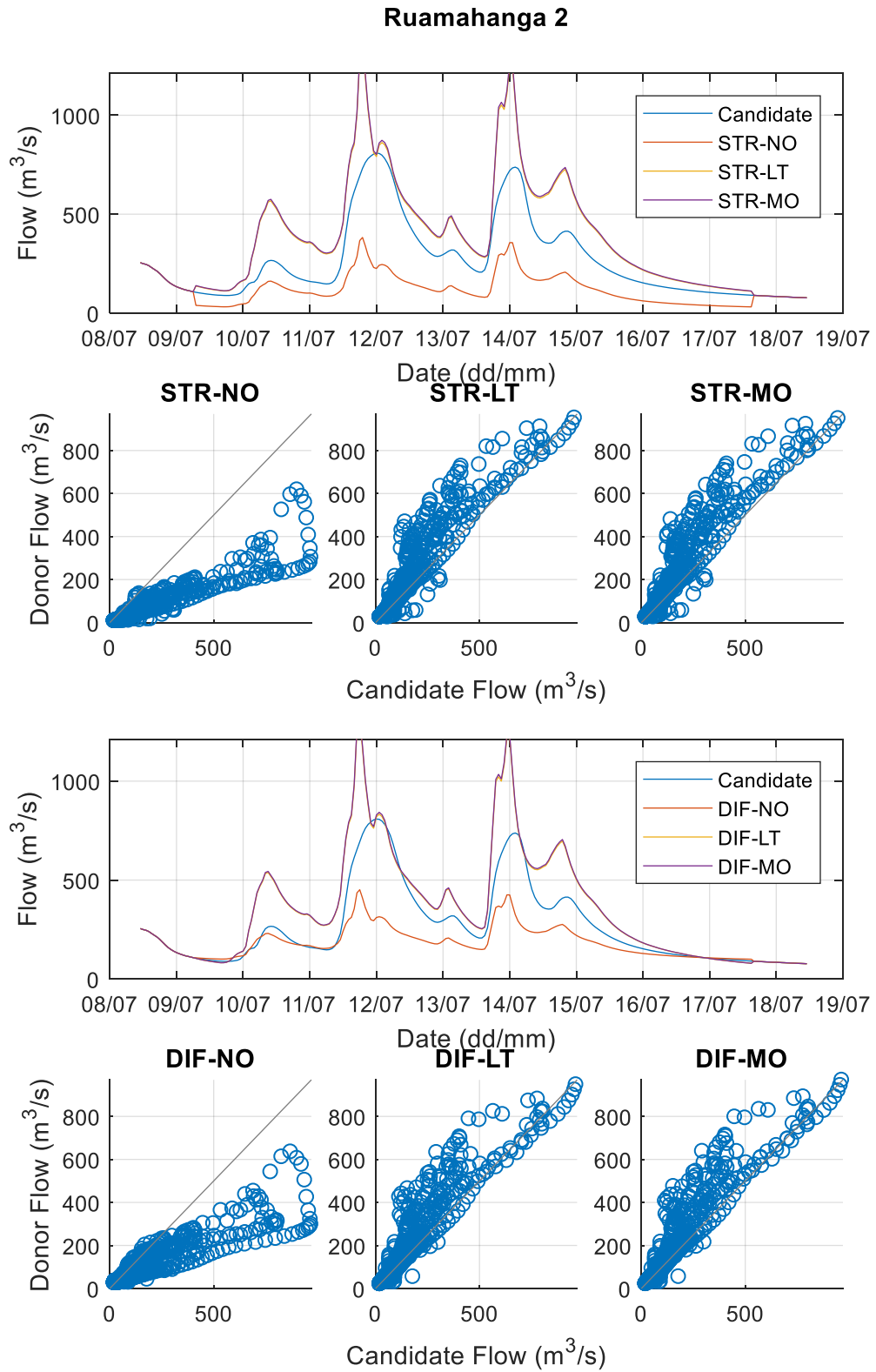
Figure A.8. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Ruamahanga 2.
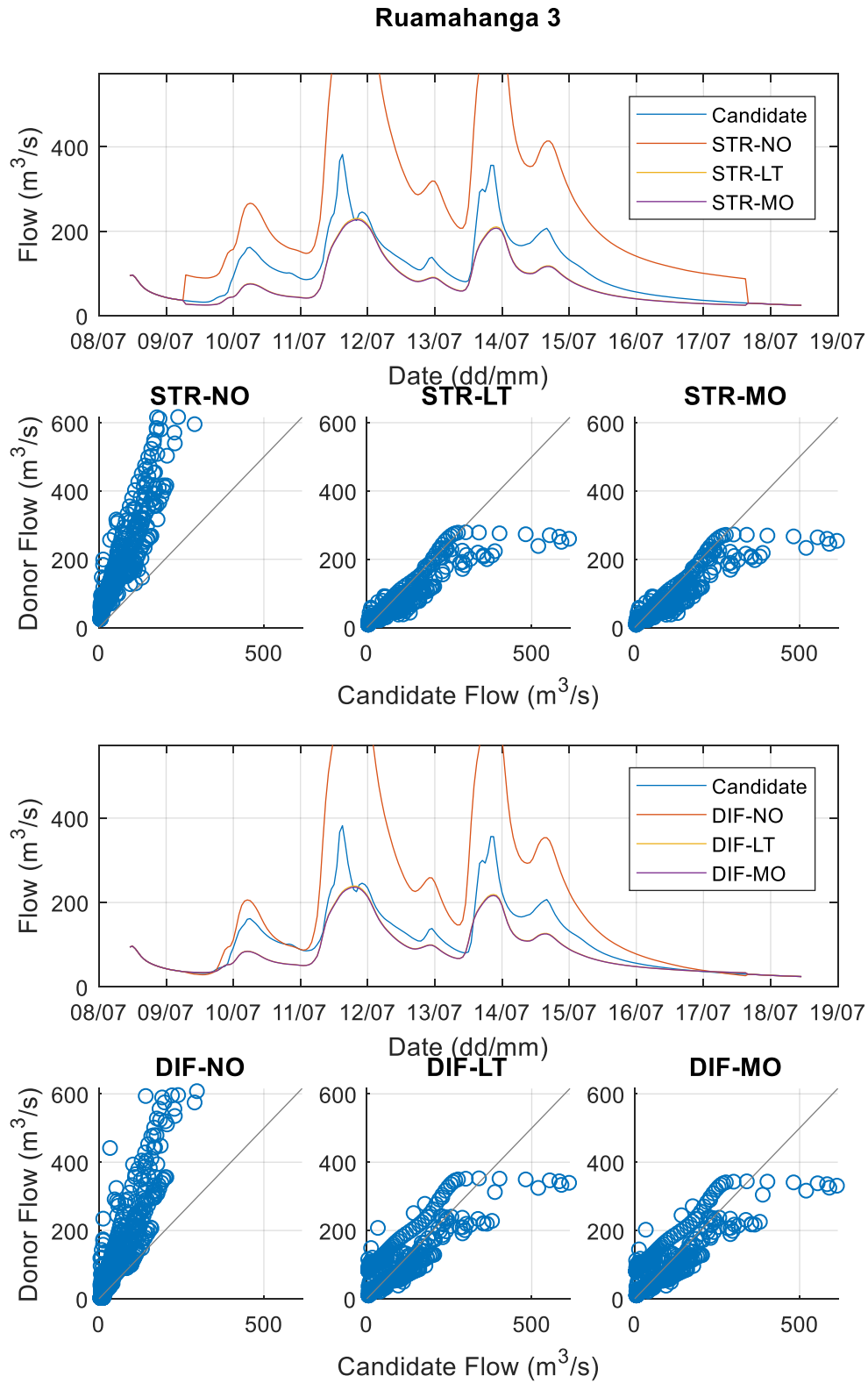
Figure A.9. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Ruamahanga 3.
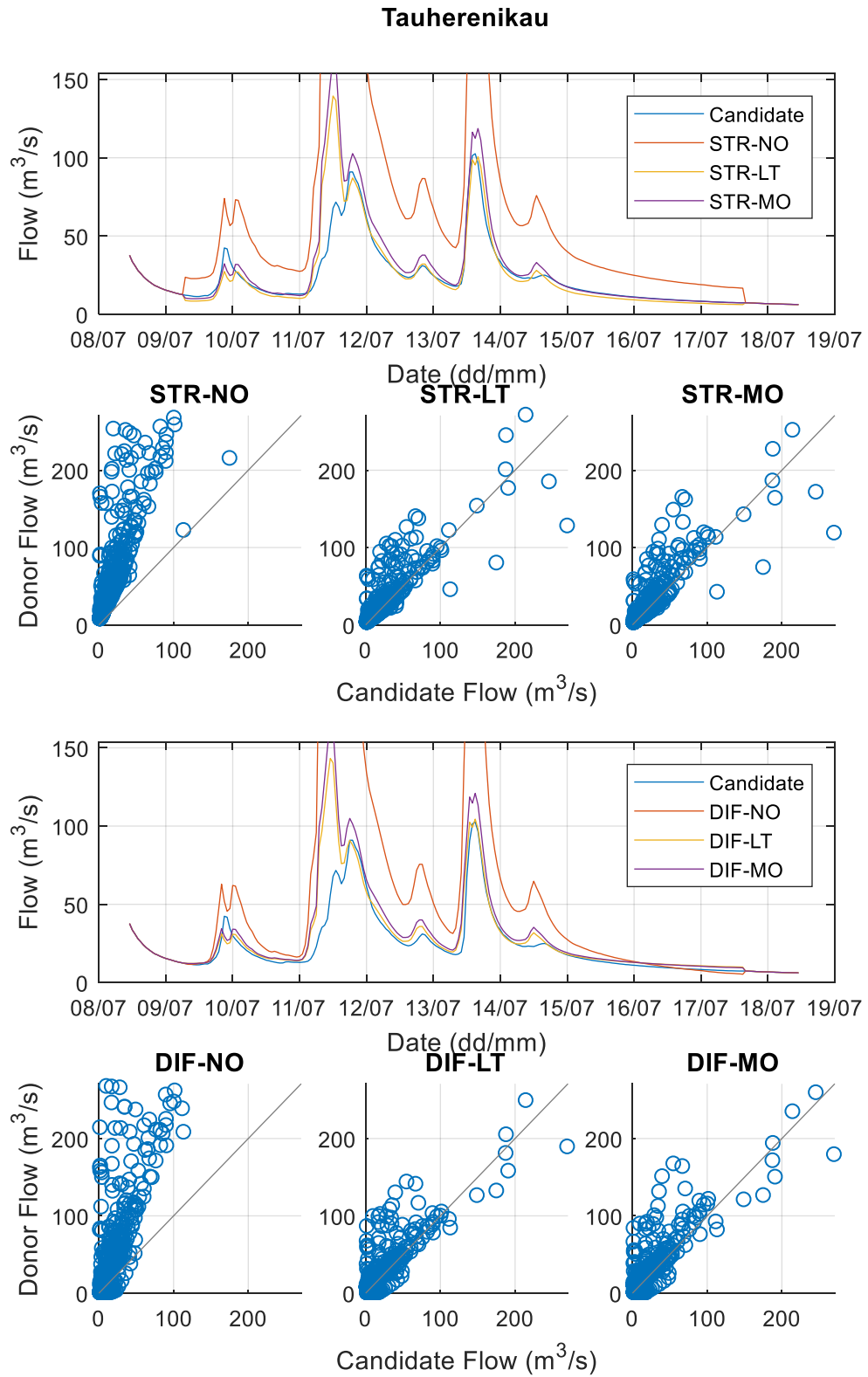
Figure A.10. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Tauherenikau.
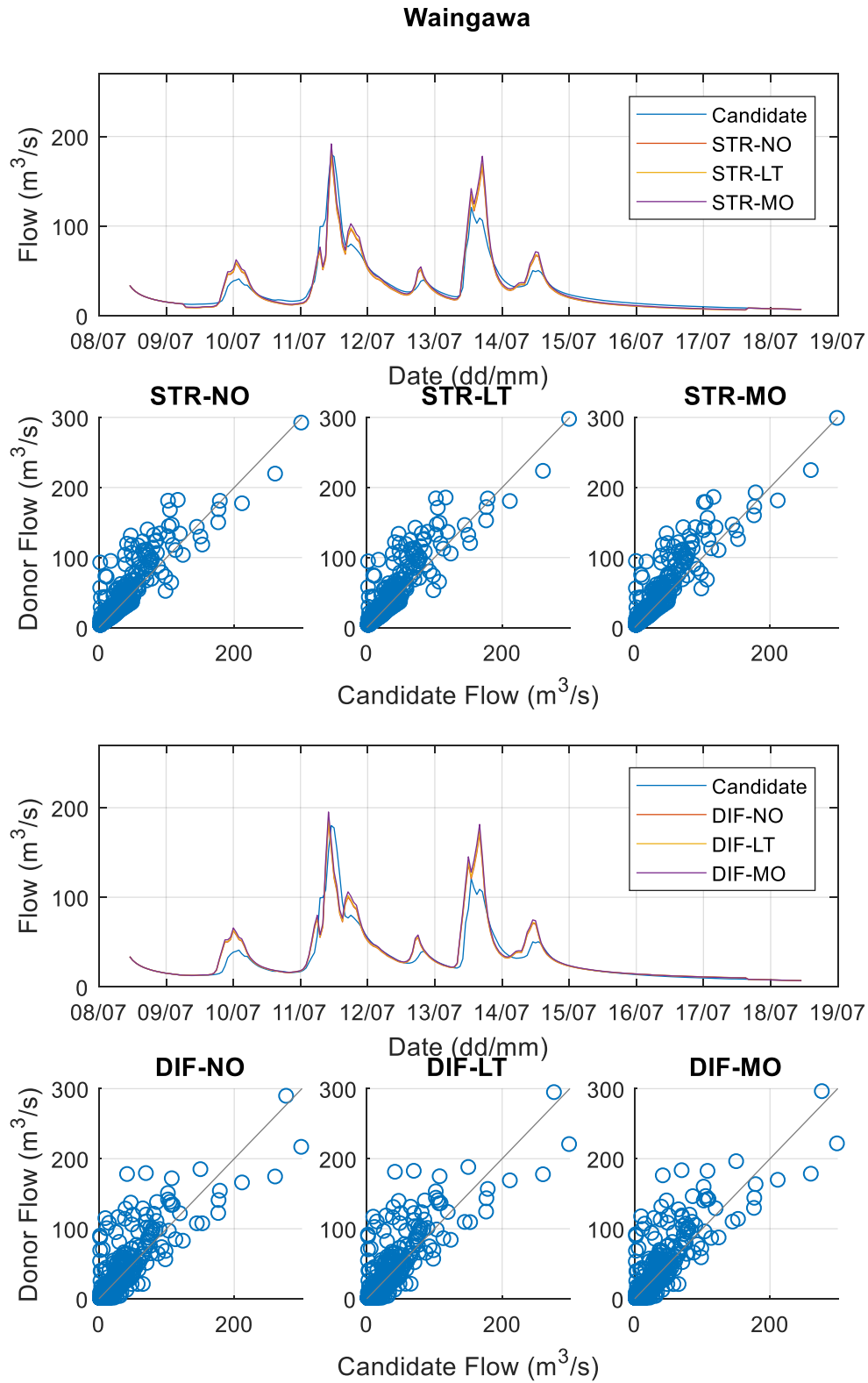
Figure A.11. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Waingawa.

Figure A.12. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Waiohine.

157

Figure A.13. LT (yellow), MO (purple) and NO (red) multipliers applied to straight (top) and difference (bottom) infilling method comparisons at Waipoua 2.

Table B.1. Baseflow indices generated by Eckhardt recursive digital filter using two different values for BFImax parameter: backward filter EK(opt) and default EK(def) as well as estimates from Bump & Rise filter (BR) and Q90/Q50 metric for 13 gauges in the Wairarapa region.

No.	Candidate	EK(opt)	EK(def)	BR	Q90/Q50
3	Kopuaranga 1	0.21	0.51	0.61	0.30
4	Kopuaranga 2	0.35	0.57	0.67	0.36
5	Mangatarere 1	0.60	0.56	0.44	0.09
6	Mangatarere 2	0.52	0.53	0.53	0.28
7	Mangatarere 3	0.50	0.53	0.50	0.22
8	Otakura	0.51	0.61	0.64	0.22
12	Ruamahanga 1	0.48	0.48	0.58	0.39
13	Ruamahanga 2	0.39	0.57	0.54	0.29
14	Ruamahanga 3	0.56	0.55	0.60	0.31
16	Tauherenikau	0.60	0.53	0.62	0.34
17	Waingawa	0.57	0.52	0.61	0.38
18	Waiohine	0.52	0.50	0.57	0.39
20	Waipoua 2	0.65	0.55	0.61	0.21

Figure B.1. Kopuaranga 1 total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).
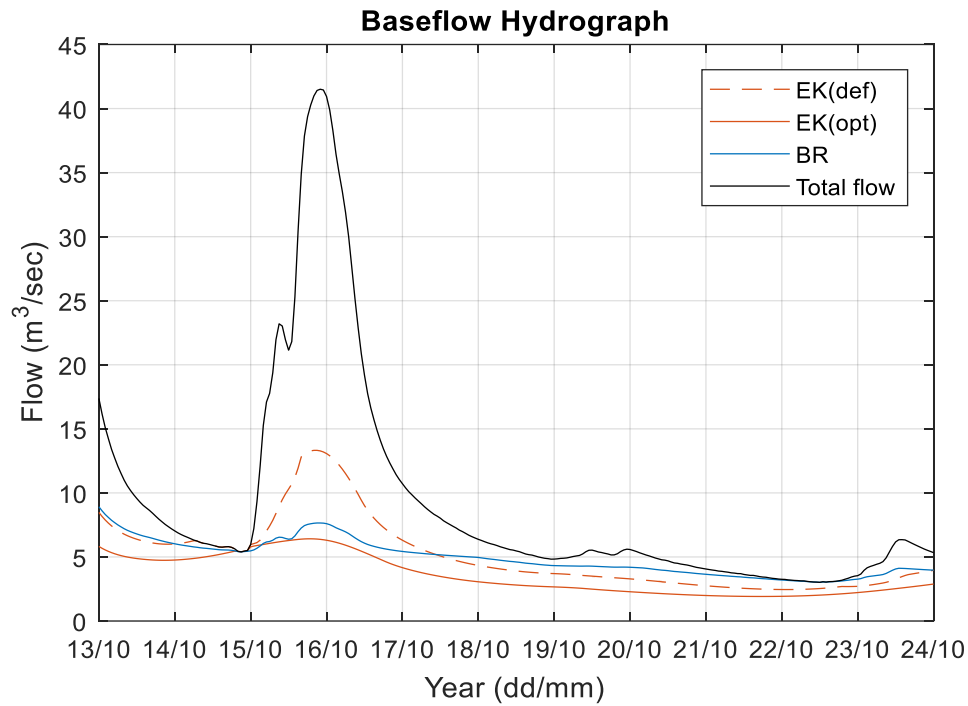
Figure B.2. Kopuaranga 2 total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).
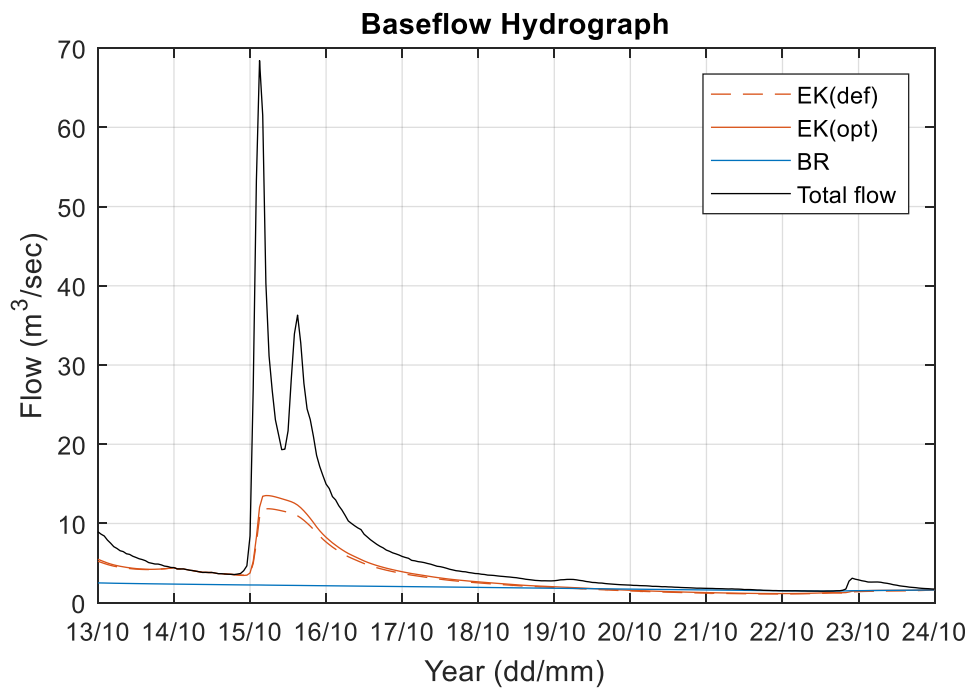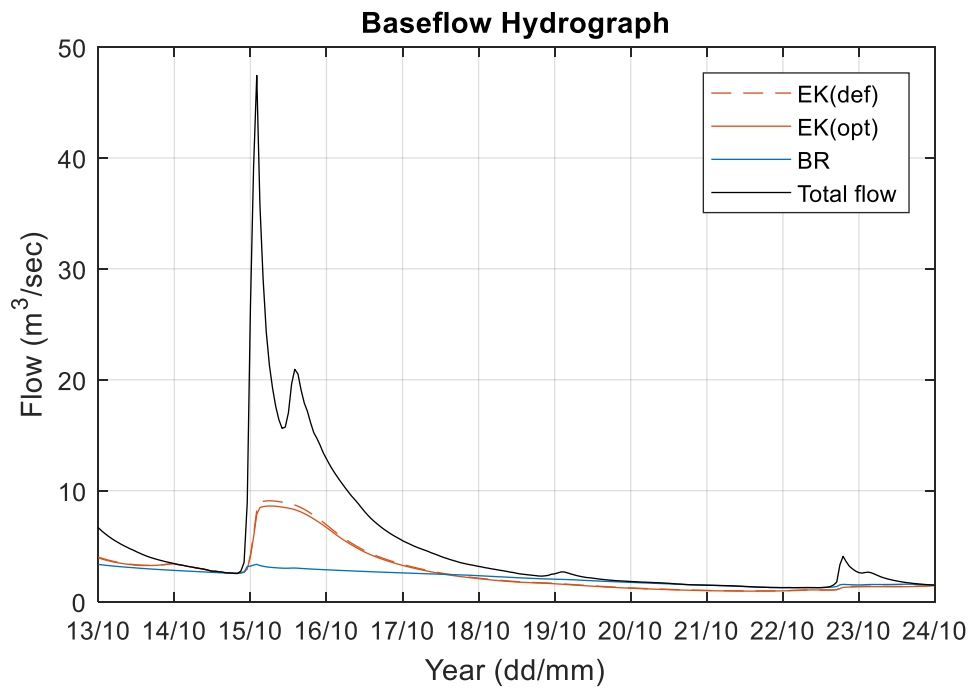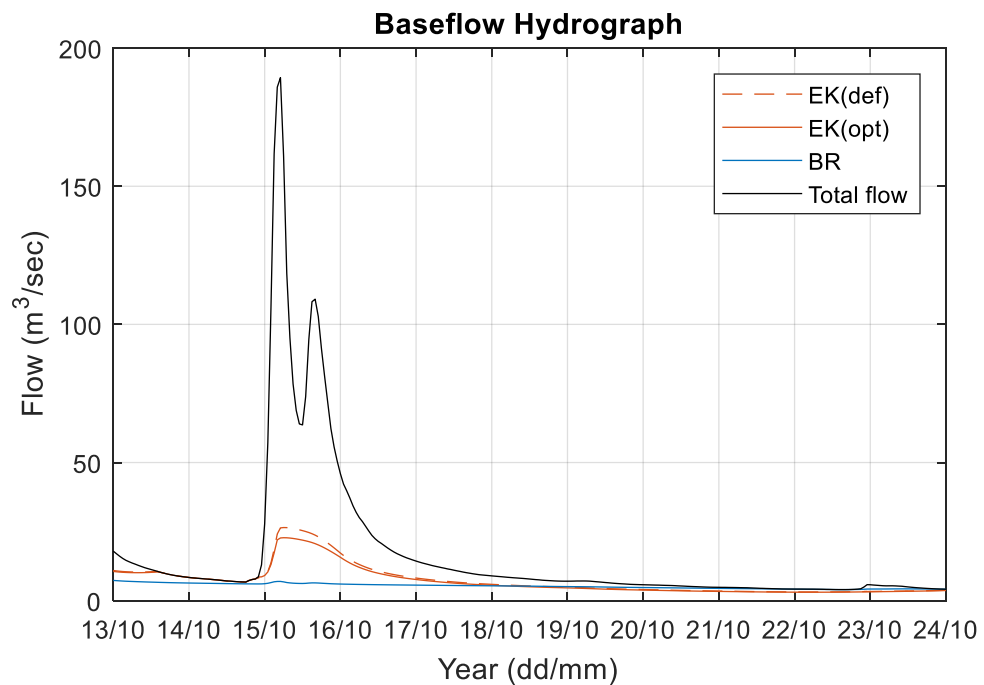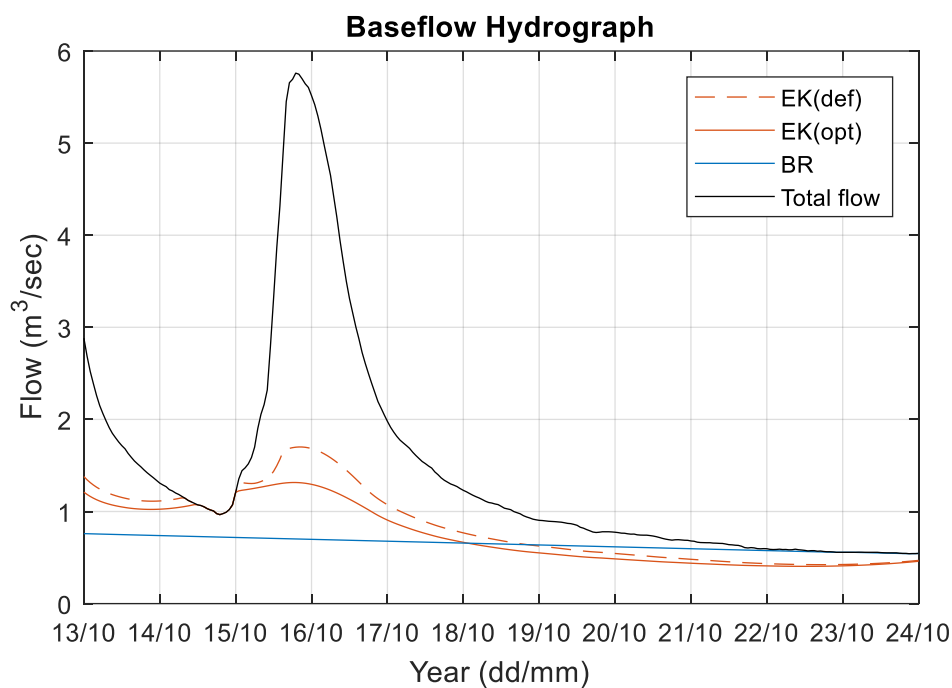

Figure B.3. Mangatarere 1 total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).
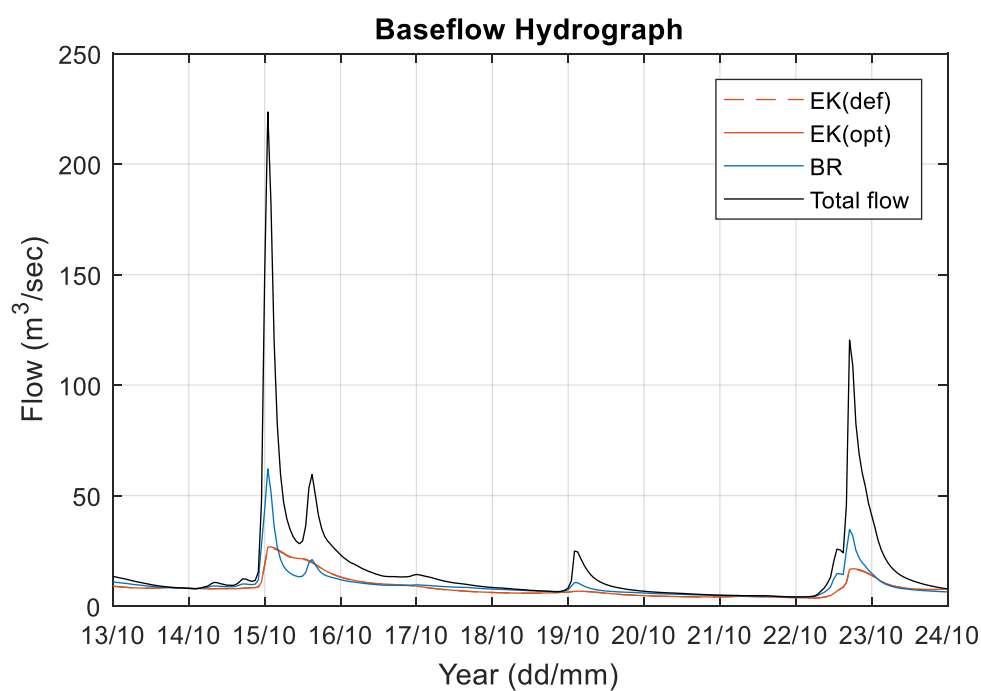
160

Figure B.4. Mangatarere 2 total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).

Figure B.5. Mangatarere 3 total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).
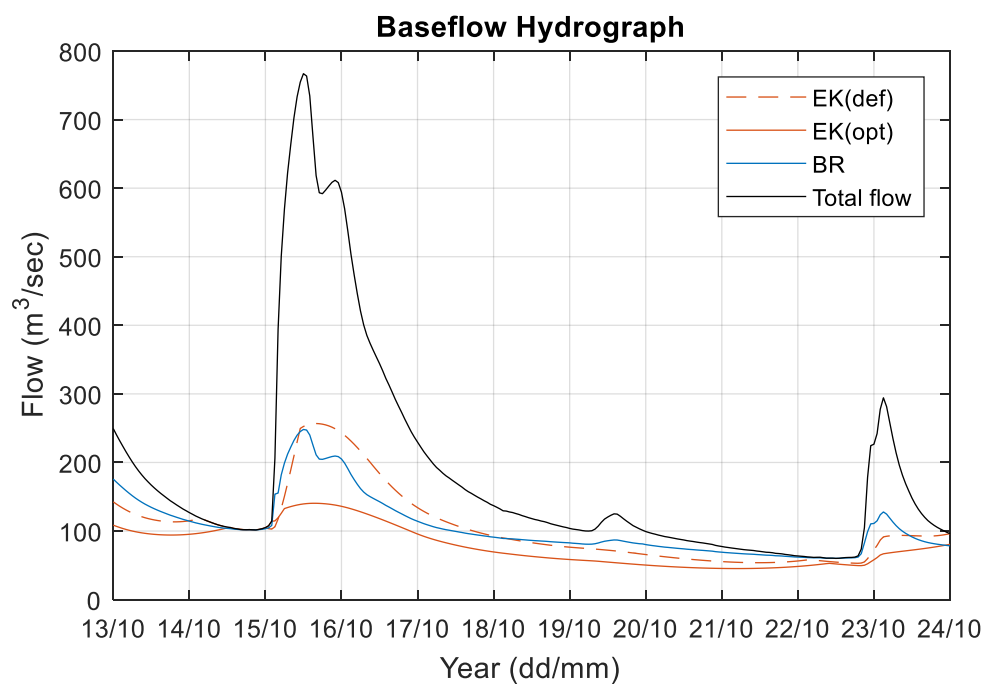
Figure B.6. Otakura total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).
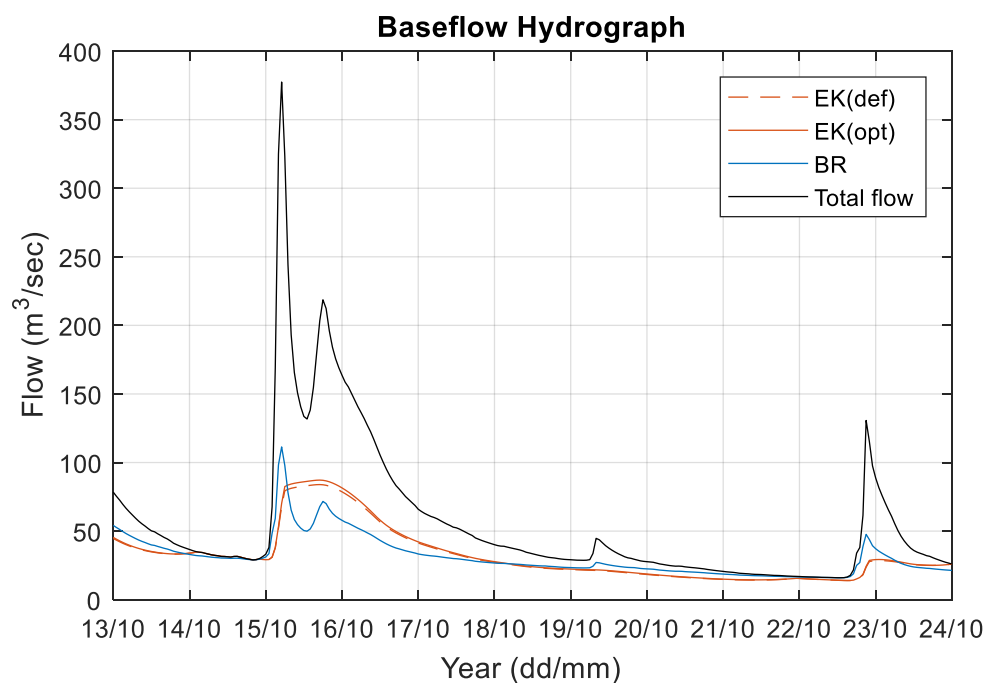

Figure B.7. Ruamahanga 1 total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).

Figure B.8. Ruamahanga 2 total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).
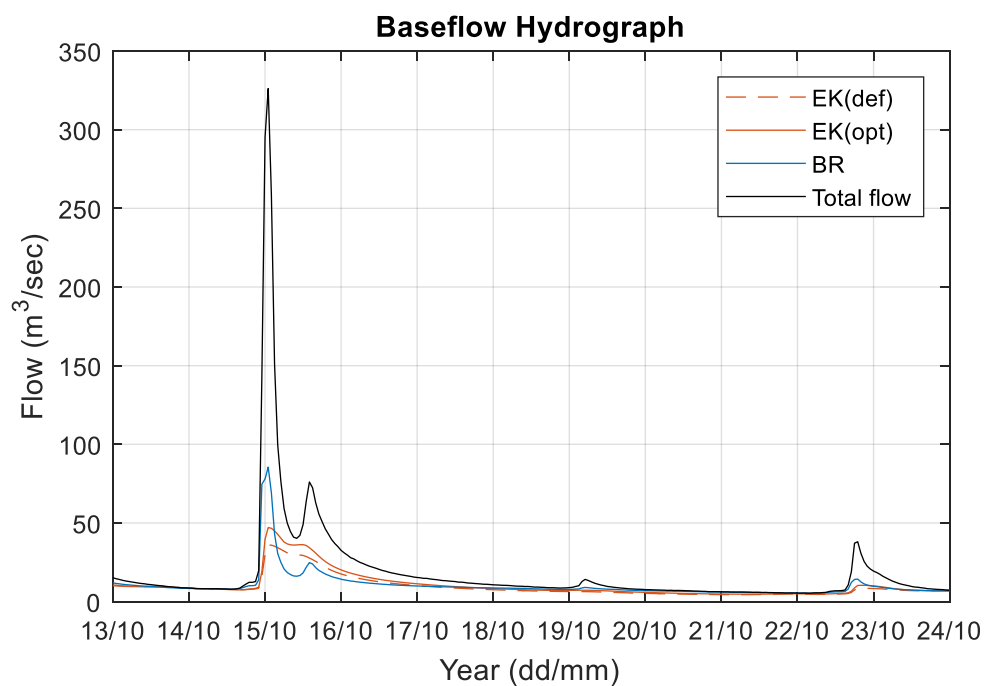

Figure B.9. Ruamahanga 3 total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).

163

Figure B.10. Tauherenikau total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).
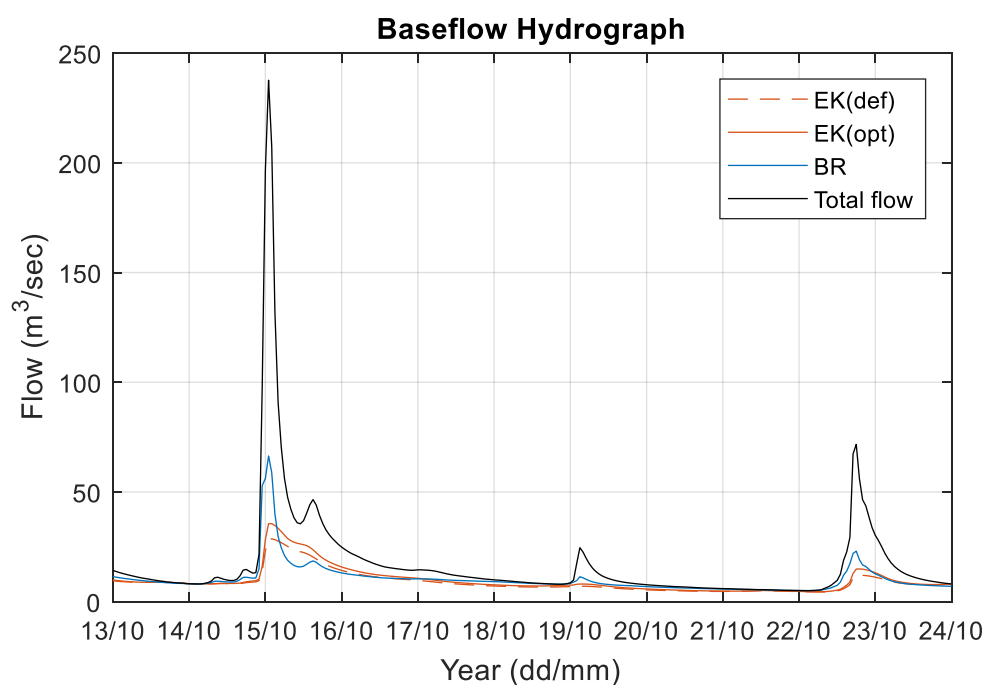

Figure B.11. Waingawa total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).
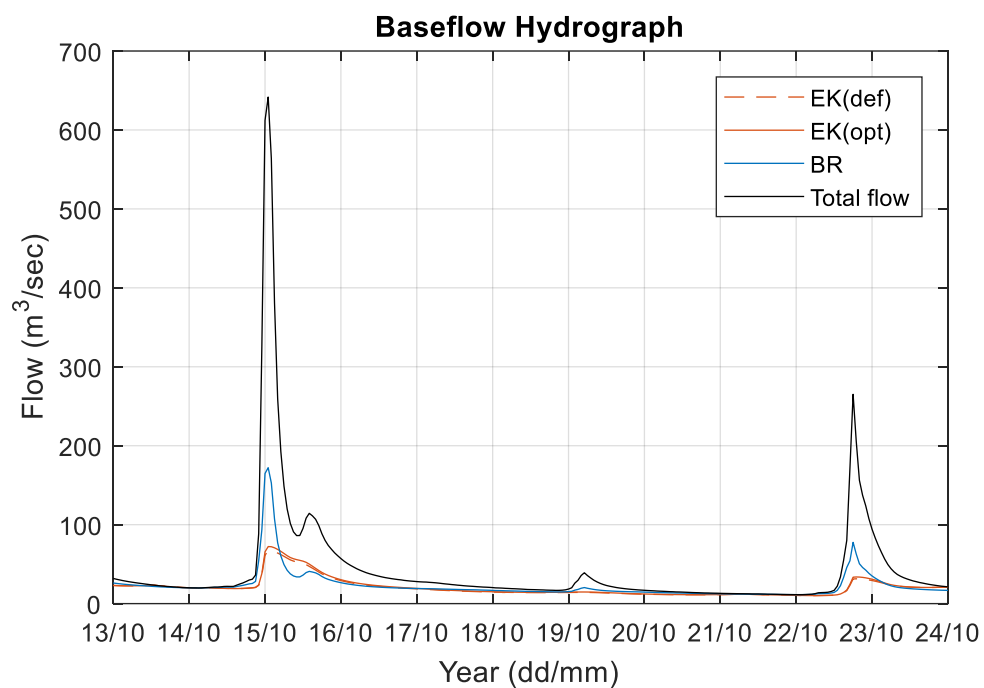
Figure B.12. Waiohine total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).
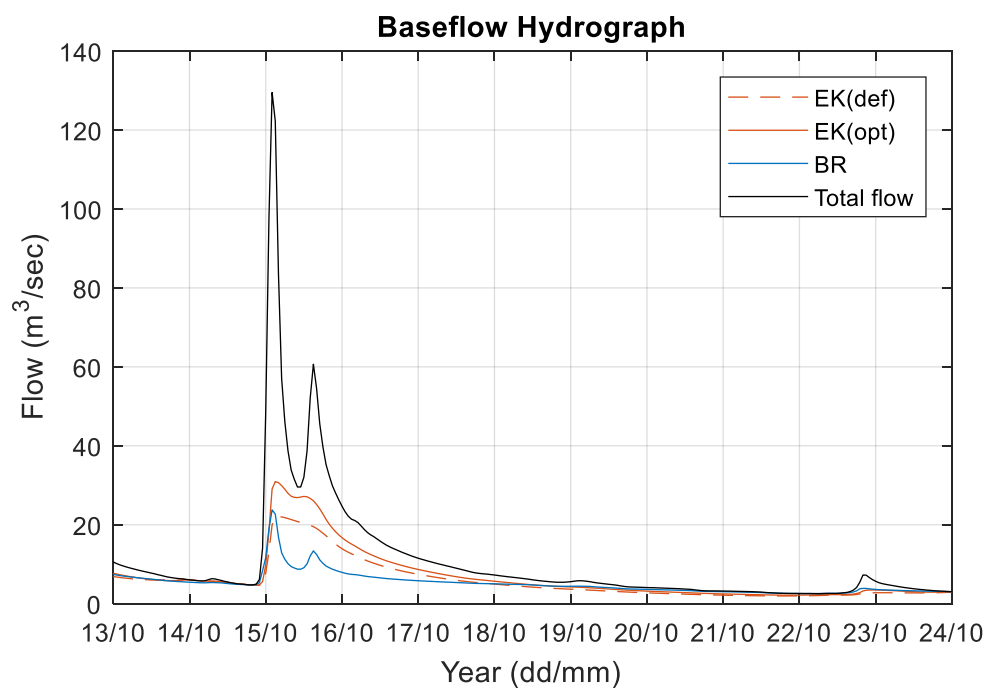

Figure B.13. Waipoua 2 total flow hydrograph with three estimates of baseflow: Eckhardt filter with a BFI of 0.8 (red dash), Eckhardt filter with a backward filter estimate of BFImax (red line) and Bump and Rise filter (blue line).