Santoft Groundwater Model and Rangitīkei Groundwater Management Zone Allocation Limit Technical Review

September 2024

horizons
REGIONAL COUNCIL

Prepared for:

Amy Lennard
Science Manager

August 2024
Report No. 2024/EXT/1874
ISBN 978-1-99-106183-6

Prepared by:

Dr Hisham Zarour
Stantec

CONTACT	24 hr freephone 0508 800 80		Help@horizons.govt.nz		www.horizons.govt.nz
SERVICE CENTRES	Kairanga Cnr Rongotea and Kairanga-Bunnythorpe Roads Palmerston North Marton 19 Hammon Street Taumarunui 34 Maata Street	REGIONAL HOUSES	Palmerston North 11-15 Victoria Avenue Whanganui 181 Guyton Street	DEPOTS	Taihape 243 Wairanu Road Taihape Woodville 116 Vogel Street
POSTAL ADDRESS	Horizons Regional Council, Private Bag 11025, Manawatū Mail Centre, Palmerston North 4442				F 06-952 2929

Santoft Groundwater Model and Rangitīkei Groundwater Management Zone Allocation Limit Technical Review

September 2024 **Ref: 123456789**

PREPARED FOR:

Dr Huma Saeed
Horizons Regional Council

PREPARED BY:

Dr Hisham Zarour

Stantec

Revision Schedule

Revision No.	Date	Description	Prepared by	Quality Reviewer	Independent Reviewer	Project Manager Final Approval
1	16/07/2024	Draft	H Zarour	V Dally	M Scaife	R Searle
2	02/09/2024	Final	H Zarour	V Dally	M Scaife	R Searle

Disclaimer

The conclusions in the report are Stantec's professional opinion, as of the time of the report, and concerning the scope described in the report. The opinions in the document are based on conditions and information existing at the time the document was published and do not take into account any subsequent changes. The report relates solely to the specific project for which Stantec was retained and the stated purpose for which the report was prepared. The report is not to be used or relied on for any variation or extension of the project, or for any other project or purpose, and any unauthorised use or reliance is at the recipient's own risk.

Stantec has assumed all information received from the client and third parties in the preparation of the report to be correct. While Stantec has exercised a customary level of judgment or due diligence in the use of such information, Stantec assumes no responsibility for the consequences of any error or omission contained therein.

This report is intended solely for use by the client in accordance with Stantec's contract with the client. While the report may be provided to applicable authorities having jurisdiction and others for whom the client is responsible, Stantec does not warrant the services to any third party. The report may not be relied upon by any other party without the express written consent of Stantec, which may be withheld at Stantec's discretion.

Summary

Horizons Regional Council commissioned Stantec to review the report titled *Rangitīkei Groundwater Management Zone Allocation Limit* (PDP, 2023) and the associated groundwater model. The report lacks detailed information on the model, which updates the model developed in 2017 and reported in the *Santoft Groundwater Model Report* (PDP, 2017). The 2023 model files were included in the review. This review focuses on the 2023 model but also utilises information from the 2017 report to understand the completed work. Future modelling reports must be standalone as recommended in best practice guidelines as this makes them more user-friendly.

Numerical groundwater models are computer simulations of hydrogeological systems designed to predict water levels and assess groundwater system behaviour under various conditions. They define hydrogeological systems through boundary and initial conditions and characterise them using appropriate parameters. They also incorporate stresses, including both natural and artificial inputs and outputs such as recharge and pumping. The primary output of these models is water levels (heads), from which flows can be calculated.

For a model to be considered fit for purpose, it must demonstrate good calibration and accurately reflect the groundwater system it represents. It is important to recognise that while multiple models can represent a system and replicate its past behaviour, they may produce different predictions due to inherent variability, known as non-uniqueness. Sensitivity analysis is essential in the calibration process, as it helps manage this variability. Since no single model can be deemed 'true,' it is vital that modelling results presented to decision-makers include estimates of uncertainty (Barnett *et al.*, 2012). While model uncertainties can be managed through careful data collection, sensitivity analysis, and uncertainty quantification, they cannot be completely eliminated. Regular updates are necessary to integrate new data, address emerging questions, and refine stress representations.

The reviewed models represent a significant advancement in understanding the hydrogeological system in the Rangitīkei Groundwater Management Zone (RGMZ) and its responses to stresses such as abstraction. While the 2023 model demonstrates good calibration by reasonably replicating historical groundwater levels, it also exhibits some unrealistic representations and discrepancies that degrade confidence in its predictions. Furthermore, the predictions are presented definitively without adequate uncertainty analysis. It is crucial to account for uncertainties in data, the conceptual model, historical simulations, and predictive modelling.

The reviewed reports and model hypothesise that rainfall recharge and bore abstraction are key controls on the groundwater system. PDP (2023) recommends revising the groundwater management zone (GMZ) boundaries and imposing a moratorium on abstraction and allocation until limits are reduced to sustainable levels. While evaluating the validity of these recommendations falls outside the scope of this review, it concludes that the conceptual and numerical models currently lack sufficient reliability to support such critical decisions.

The key review findings are:

- **Reporting**: The 2023 report does not address or comment on differences from the 2017 model, and some inaccuracies have been identified. For example, hydraulic conductivity values in the 2023 model deviate from the reported ranges, which undermines confidence in both the model and its predictions. Incorrect parameter values could lead to erroneous predictions, potentially resulting in inappropriate management decisions and actions with significant socio-economic and environmental consequences. Some boundary conditions are also inaccurately reported. Additionally, some key values are unclear; for instance, it is not specified whether the storativity values in the model refer to storage coefficient (S) or specific storage (S_s), and whether conductance (C) is measured in m/d as reported or in m²/d, as is commonly used in models. These parameters are crucial for historical simulations and future predictions. Furthermore, future modelling reports should present data in more informative formats, such as cross-sections, particularly for the conceptual model, hydrostratigraphy, and hydraulic properties. Future reports should present all data used in appendices for transparency.
- **Objectives**: The modelling objectives are well defined and appropriate. However, a target model confidence level classification class is not defined.
- **Model Planning:** The target quality of the model is not evident in the reports. This quality must be proportionate to the model objectives and available resources, including data. The limitations, planned exclusions, and the need to update and recalibrate the 2017 model are not reported.
- **Model Verification**: Model verification, or post-audit, involves comparing model results against a dataset that was not used during its calibration. The 2023 model update should have initially involved evaluating the 2017 model predictions against data accumulated from July 2014 to July 2022. This approach would have assessed the predictive accuracy of the 2017 model and provided valuable insights for the development and conceptualisation of the new model.

- **Data:** The models notably did not incorporate surface water monitoring data (both one-off measurements and continuous gauging), groundwater quality data, and groundwater level compliance monitoring data from consent holders. In some instances, the models relied on estimates and low-resolution data instead of higher-quality measurements and available best data (e.g., assumptions about surface water instead of survey data and elevation contours from topographic maps rather than digital elevation models). Additionally, groundwater level data from Bore 312006 were excluded from the 2023 model without explanation or justification in the report. Furthermore, relevant literature was not reviewed in a comprehensive manner.

- **Flow Direction and Hydraulic Gradients:** The general flow directions are reasonably characterised. However, the discussion of flow directions at shallow and deep depths, arbitrarily separated at 20 m depth, is potentially misleading. The vertical hydraulic gradient discussion is valid. Nevertheless, there is information that has not been utilised in conceptualisation and numerical modelling.

- **Interpolation and Conceptualisation:** Some of the interpolations presented are based on data exhibiting clear signs of instrumentation issues. These issues may affect the accuracy of the interpolated results and the conceptual model derived from them.

- **Geology**: The models are not informed by geology, including stratigraphy, lithology, and structure, leading to unrealistic model configuration and parameterisations. This oversight diminishes confidence in the model's predictive capability. The conceptual and numerical modelling should have utilised the national geological map (QMAP) and borelogs available to Horizons to ensure the model reasonably reflects geological conditions.

- **Surface Water-Groundwater Interaction**: The model is not informed by surface water data (flows, exchanges with groundwater, water and bed levels, bed properties). This affects the model's representativeness and calibration, increases its non-uniqueness, and reduces the reliability of the model’s predictions on surface water effects, especially since many features like coastal lakes are not represented, though ecologically and culturally significant. There is no evidence of checking the hydrological correctness of surface water elevation data, which can result in counterintuitive surface water levels and flow directions.

- **Conceptual Model:** The conceptual model is not adequately presented graphically and in analytical water budget format. In addition, there are arbitrary conceptual assumptions, e.g., the aquifer system is arbitrarily divided based on depth of 20 m to shallow and deep systems. The peripheral boundary conditions are not adequately justified. Future modelling must better characterise stress-response relationships, e.g., through the use of particle tracking and influence zone analyses.

- **Software:** The software used for modelling, pre- and post-processing, and calibration are industry standard and adequate. However, the rationale for switching from MODFLOW-NWT to MODFLOW 6 is not justified in the 2023 report, and the consequences of this change are unknown.

- **Model Domain**: The model domain does not align with the GMZ boundaries, which are delineated based on geological and hydrological conditions. This misalignment could affect the model's performance, potentially leading to inaccurate predictions despite acceptable historic data matching. Additionally, the title of the 2023 report emphasises allocation limits in the GMZ, necessitating alignment of the model domain with the GMZ boundaries. Otherwise, this objective would be impractical. The model’s peripheral boundary crosses topographical, hydrological, and geological boundaries, leading to potential inaccuracies.

- **Coastal Boundary:** The model’s western boundary is extended more than 2 km into the sea, which is unnecessary. Although it is appropriately represented using General Head Boundary (GHB) conditions, an inspection of the model files revealed that the layers to which these conditions apply were inaccurately reported. Additionally, the basis for the designated head and conductance values is not explained. The conductance values are generally high (approximately 28–472 m^2/d), causing the boundary to behave more like a Constant Head (CHD) boundary. This inappropriate representation of the coastal boundary likely resulted in a significant error in the calculated groundwater level at the coastline, which was found to be about 6 metres above sea level (masl) upon reviewing the model files.

- **Discretisation:** The transient model is discretised into 384 monthly stress periods, each comprising a single time step. However, no explanation is provided for the choice of the number of time steps per stress period. The model is laterally uniformly discretised into 500 m × 500 m cells. While this relatively large cell size is generally suitable for a regional-scale model, it does not adequately address all data, boundary conditions, calibration requirements, and, most importantly, the model objectives. It is unclear why the 2023 model did not utilise an unstructured grid, which is a key advantage of MODFLOW 6. This would have enabled grid refinement in areas of particular interest, such as monitoring and pumping bores and surface water features. Vertically, the model is arbitrarily divided into 11 layers, which is inadequate for realistically representing changes in hydraulic properties within the hydrogeological sequence.

- **Boundary Conditions:** There are oversights in the definition of boundary conditions in the conceptual model and their implementation in the numerical model. Notably, multiple boundary conditions, such as river (RIV), drain (DRN), and recharge (RCH), are simultaneously assigned to many cells across Layer 1. This is conceptually incorrect and could lead to double accounting and potential instability in the numerical model, resulting in unrealistic simulations and predictions. Future models should avoid conflicts among boundary conditions.

- **Numerical Model Implementation of Boundary Conditions**: The conceptualisation of surface water features reported is unrealistic and not correctly implemented in the numerical model. The elevation of water levels and beds is conceptualised based on depths below the land surface, as interpolated from topographical map contours. However, different values are used in the model, raising concerns about the calibration of the numerical model and casting doubt on the realism of its predictions.
- **Evapotranspiration:** Evapotranspiration from the water table is not considered in either the conceptual or numerical models. This differs from evapotranspiration from the soil, which is accounted for in soil moisture balance models used for estimating groundwater recharge. Future groundwater modelling should consider evapotranspiration from the water table during the conceptualisation phase and include it in numerical models if found appropriate, particularly in areas with a shallow water table.
- **Initial Conditions:** Transient models are sensitive to initial conditions. The first stress period in each of the 2017 and 2023 models is designated as steady state, providing initial conditions for the transient models. However, the details of these steady state models are not reported. Issues in the steady state models can carry over into the transient models, potentially affecting their accuracy and reliability of predictions.
- **Model Calibration:** The reported 2017 model calibration statistics are adequate. The 2023 model calibration statistics are not reported. Both the 2017 and 2023 models adequately simulate groundwater level long-term trends and seasonal changes. However, there are indications of errors in the model's structure, values, and results that diminish confidence in the model's realism and predictions. Notably, the calibrated hydraulic conductivity exhibits unjustifiable bullseyes and out-of-bounds values, which may indicate over-calibration using the Pilot Point method.
- **Recharge**: Recharge is conceptualised as a key stress in the models but was not adjusted during model calibration, reducing confidence in related parameters such as hydraulic conductivity and increasing model non-uniqueness and uncertainty. Recharge is not adequately reported. No details are provided on the soil moisture balance model used for its estimation. Additionally, groundwater recharge from irrigation returns is not considered.
- **Sensitivity Analysis:** The reported sensitivity analysis is minimal. The model's sensitivity to parameters not included in the calibration, as well as initial and boundary conditions and their locations, was not analysed.
- **Predictive Scenarios:** The predictive scenarios are well formulated in line with the modelling objectives and conceptualisation. However, they should be reconsidered if there are revisions to the conceptual and/or numerical models, which this review recommends.
- **Predictive Uncertainty:** Predictive uncertainty is not addressed in the current modelling. Future modelling should incorporate an assessment of predictive uncertainty in accordance with contemporary best practice guidelines.
- **Reporting:** The reports are generally prepared to a high professional standard. However, this review identifies several areas for improvement. Most importantly, future modelling reports should include a critique of the monitoring network and identify future data needs, such as the need for a geological model and dedicated monitoring bores.
- **Modelling Outcome:** The primary conclusion of the modelling is that groundwater levels will continue to decline if current allocation and abstraction levels remain unchanged. This review cannot confirm or refute this conclusion. However, it is the opinion of this review that the models do not provide a sufficiently robust basis for confidently recommending changes to current use and allocation limits, although such changes may still be warranted.
- **Peer Review:** A peer review undertaken in 2016 of a draft report on an early version of the 2017 model concluded that the model was not fit for purpose and identified several critical issues that needed to be addressed to render the model suitable for its intended use. This review concurs with the findings of the 2016 review and notes that its recommendations have not been implemented in the final version of the 2017 model, nor in the 2023 model. This review asserts that the Santoft/RGMZ groundwater model requires comprehensive revision and that all recommendations provided in the 2016 review, as well as those presented in this review, must be addressed to achieve a model that is trustworthy and fit for purpose.

In conclusion, while the reviewed models mark progress in understanding groundwater dynamics in the RGMZ, several critical areas require attention to enhance their reliability and usability. Key issues include discrepancies between reported and model parameters, inadequate consideration of boundary conditions, and a lack of sensitivity and uncertainty analysis. The models' ability to inform decision-making is compromised by these limitations, underscoring the need for a model revision. It is essential that future models adhere to best practice guidelines, including rigorous peer review, comprehensive data integration, and clear, transparent reporting. Addressing these recommendations will strengthen the foundation for making informed and effective groundwater management decisions.

Contents

List of tables

List of figures

List of Appendices

Abbreviations and acronyms

Abbreviation / acronym	Full Name
%	Percent
&	And
/	Per
2017 model	Santoft Groundwater Model reported in PDP (2017)
2023 model	Santoft Groundwater Model reported in PDP (2023)
3D	Three-dimensional
AEE	Assessment of environmental effects
AGMG	Australian Groundwater Modelling Guidelines (Barnett *et al.*, 2012)
b	Thickness of layer or aquifer [L]
C	Conductance of river, drain, or general head boundary [commonly L^2/T, L/T in the reviewed reports]
CHD	Constant Head Designation, also known as Constant Head Boundary
CLC	Model confidence level classification as defined in Barnett *et al.* (2012)
Cliflo	A climate and weather data platform developed by NIWA
cm	Centimetre
d	Day
DCCEEW	NSW Department of Climate Change, Energy, the Environment and Water, formerly DPE
DPE	NSW Department of Planning and Environment, now DCCEEW
DRN	MODFLOW Drain Package, boundary condition, or cell
e.g.,	*exempli gratia*, meaning 'for example'
EC	Electrical conductivity of water [µS/cm]
ed	Edition
et al.	*et alii*, meaning 'and others'
EVT	MODFLOW Evapotranspiration Package, boundary condition, or cell
FEFLOW	A finite element modelling software for simulating subsurface flow, heat, and mass transport in groundwater and porous media
FloPy	Python library for creating, running, and analysing groundwater models
GHB	General Head Boundary
GMZ	Groundwater Management Zone
GNS	Geological and Nuclear Sciences, Institute of. Also known as GNS Science.
GWMZ	Groundwater Management Zone (abbreviation used in the reviewed reports)
Ha	Hectar (10,000 m^2)
HK	Horizontal hydraulic conductivity [L/T]
Horizons	Manawatū-Whanganui region or Manawatū-Whanganui Regional Council, depending on context
HRC	Horizons Regional Council, the trading name for the Manawatū-Whanganui Regional Council
i	hydraulic gradient [–]
I	Model row index, $I = 1$ is the first (top) row
i.e.,	*id est*, meaning 'that is'
ID	Identification number or code
Inc	Incorporated
J	Model column index, $J = 1$ is the first (left) column
K	Depending on context, hydraulic conductivity [L/T] or model layer index, where $K = 1$ is the first (upper) layer
ka	kilo-annum, thousand years before the present
km	Kilometres
km^2	Square kilometres
LAWA	Land, Air, Water Aotearoa
LTD	Limited

Abbreviation / acronym	Full Name
m	Metres
m^2	Square metres
m^3	Cubic metres
Ma	Mega-annum, million years before the present
masl	Metres above sea level
mbgl	Metres below ground level
MGMZ	Manawatū Groundwater Management Zone
mm	Millimetre
MODFLOW	A widely used groundwater flow modelling software
MODFLOW 6	MODFLOW version 6
MODFLOW-NWT	Newton Formulation for MODFLOW
MSR	The mean sum of residuals
N/A	Not applicable, not available, depending on context.
NE	Northeast
NIWA	National Institute of Water and Atmospheric Research
NPS-FM	National Policy Statement for Freshwater Management (2020)
NRGMZ	Northern Rangitīkei Groundwater Management Zone
NSW	New South Wales
NW	Northwest
NZTM	New Zealand Transverse Mercator (mapping coordinate system)
PDP 2017	*Santoft Groundwater Model Report*
PDP 2023	*Rangitīkei Groundwater Management Zone Allocation Limit (report)*
PDP	Pattle Delamore Partners LTD
PEST	Parameter ESTimation, a software package designed to adjust model parameters to minimise discrepancies between observed data and model simulations (Doherty, 2023)
Plio	Pliocene
pp	Pages
QMAP	Quarter Million Map (New Zealand national geological map)
QoI	Quantity of interest
R^2	Correlation coefficient (between two datasets)
Rang_1	Upper Rangitīkei Water Management Zone
Rang_2	Middle Rangitīkei Water Management Zone
Rang_3	Lower Rangitīkei Water Management Zone
Rang_4	Coastal Rangitīkei Water Management Zone
RCH	MODFLOW Recharge Package, boundary condition, or cell
RGMZ	Rangitīkei Groundwater Management Zone
RIV	MODFLOW River Package, boundary condition, or cell
RMS	Root mean squared error
s	Second
S	Storage coefficient or storativity
SCHED4	Schedule 4 in the One Plan (Horizons, 2024)
SE	Southeast
SH	State Highway, e.g., SH3
SMSR	Scaled mean sum of residuals
SoE	State of Environment (monitoring and reporting)
SRMS	Scaled root mean squared error
S_s	Specific storage [L^{-1}]
SSD/SSI	State Significant Development and State Significant Infrastructure projects in NSW, collectively known as ‘Major Projects’
SW	Southwest
SWL	Static water level [L]

Abbreviation / acronym	Full Name
S_y	Specific yield [–]
T	Transmissivity [L^2/T]
VCSN	Virtual Climate Station Network developed by NIWA
VK	Vertical hydraulic conductivity [L/T]
WB	Whanganui (or Wanganui) Basin
WEL	MODFLOW Well Package, boundary condition, or cell
West_5	Southern Whanganui Lakes Water Management Zone
West_6	Northern Manawatū Lakes Water Management Zone
y	Year
µS	MicroSiemens

1. Introduction

1.1 Background

The Rangitīkei River catchment is located in the western part of the Manawatū-Whanganui region (Horizons), which straddles the northern part of the lower North Island of New Zealand (Figure 1-1). Geologically, the Rangitīkei River catchment is part of the Whanganui Basin (also spelled Wanganui Basin; WB for short). The WB is infilled with Pliocene (5.33–2.58 Ma[1]) and Pleistocene (2.58 Ma–11.7 ka[2]) aged alternating warm climate (global interglacial times) fine-grained marine strata and cold climate (global glacial times) fluvial deposits. Asymmetric growing folds generally traverse the WB from south-west (SW) to north-east (NE). Progressive growth of these folds elevated their crests above major river valley floors, isolating them from alluvial deposition over the last 400 ka, forming major river catchments including the Rangitīkei River catchment. Areas underlying active folds in the WB form interfluves, channelling the major river catchments down the fault zones and synclinal axes between anticlines. Cool climatic aggradational alluvial gravels were deposited in these loci, forming highly productive aquifer zones (Begg *et al.*, 2005). The hydraulic conductivity (permeability; K) of the marine strata is relatively low. They form aquitards. The alluvial deposits have relatively high hydraulic conductivity. They form aquifers. For hydrogeological analysis purposes aquifers and aquitards combinedly form hydrogeological systems, also known as aquifer systems and groundwater flow systems.

Ultimately, surface water and groundwater in the Rangitīkei River catchment drain into the Tasman Sea to the west. The catchment comprises 16 surface water sub-catchments that coincide with two groundwater management zones (GMZ)—the Northern Rangitīkei and the Rangitīkei GMZs delineated in Figure 1-1 by red and black lines, respectively. Due to the underlying geology, the groundwater resource potential in the Northern Rangitīkei Groundwater Management Zone (NRGMZ) is very small to non-existent. The Rangitīkei Groundwater Management Zone (RGMZ) hosts a significant groundwater system, especially downgradient of the surface expression of the contact between the Pliocene and the Pleistocene strata, i.e., from about Vinegar Hill to the sea. Groundwater in the RGMZ is an important resource for life sustenance, environmental, and socio-economic purposes.

The Manawatū-Whanganui Regional Council (Horizons Regional Council, or Horizons and HRC for short) is responsible for environmental management in the Horizons region, including groundwater resources. The One Plan is the Regional Policy Statement and Regional Plan within Horizons' jurisdiction. It sets the rules for management of groundwater resources in defined groundwater management zones (GMZ), including allocation limits. Horizons is also responsible for setting the limits for permitted and consented groundwater takes. The One Plan defines the extent of the RGMZ and Surface Water Management Zones within it as shown in Figure 1-1. Table 1-1 presents the allocation limits for the Rangitīkei water management zones.

Table 1-1. Rangitīkei water management units and allocation limits (from SCHED4 in Horizons, 2024).

Groundwater Management Zone	Surface Water Management Zones within the RGMZ	Groundwater Allocation Limit (m^3/y)
Northern Rangitīkei	Upper Rangitīkei (Rang_1)	Unspecified
	Middle Rangitīkei (Rang_2)	
Rangitīkei	Lower Rangitīkei (Rang_3)	75,000,000
	Coastal Rangitīkei (Rang_4)	
	Southern Whanganui Lakes (West_5)	
	Northern Manawatū Lakes (West_6)	

[1] Ma: mega-annum; million years before the present.
[2] ka: kilo-annum; thousand years before the present.

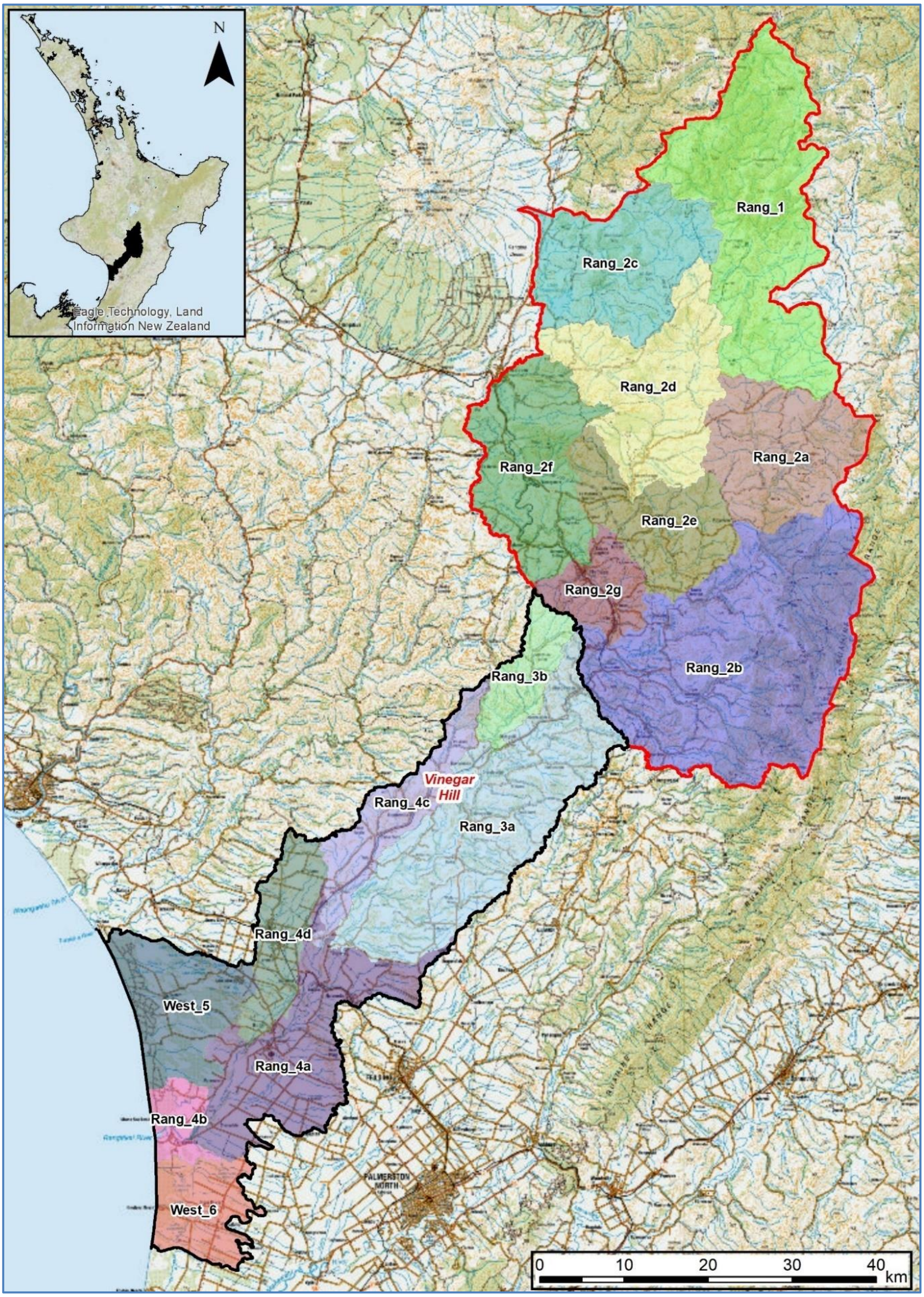


Figure 1-1. Map showing surface water and groundwater management zones in the Rangitīkei River catchment and its location in the North Island (in the inset).

Noticeable groundwater take consents and abstraction commenced in the RGMZ about 20 years ago (from about 2005). Groundwater abstraction in the RGMZ has been dominated by irrigation. From around 2015, there has been anecdotes on groundwater resource overallocation and associated groundwater level decline in parts of the RGMZ (e.g., Figure 1-2). Horizons and the RGMZ communities considered that the future trend of groundwater levels in the Santoft area is uncertain (PDP, 2017). Hence, Horizons embarked on groundwater resource assessments including the Santoft Groundwater Model Report (PDP, 2017). The key question the groundwater model reported in PDP (2017) intended to answer was '*whether a groundwater level decline is likely to continue under current* [then] *rates of abstraction or stabilise to a new steady state and, if the latter, how long that would take to occur.*' (PDP, 2017). The model reported on in PDP (2017) is referred to hereinafter as the 2017 model. Figure 1-3 shows the extents of the RGMZ and the model area ('domain').

The 2017 model represented the key features of the groundwater flow system. Based on it, PDP (2017) reported: '*A five-year forecast of groundwater levels based on continued abstraction at present* [then] *levels indicated that a continued decline in groundwater levels of around 0.4 m in deeper strata and up to 0.25 m in shallow strata was likely. The model did not indicate that groundwater levels would stabilise within the five-year forecast period.*' The 2017 model predicted subsequent small reduction in groundwater outflow to the coast and a small decline in groundwater discharge to rivers and streams. The sensitivity of groundwater receptors such as spring fed streams and the coastal lakes (dune lakes) to declines in shallow groundwater levels was not clear to the PDP (2017) investigators. PDP (2017) emphasised that other factors such as land use change and climate change may also impact these receptors. PDP (2017) recommended:

- *further assessment of the sensitivity of surface water bodies to reductions in discharge from groundwater, including assessment of the ecology and community values*
- *development of groundwater level thresholds to protect surface water bodies, if required*
- *use of the model to simulate application of the thresholds and evaluate implications for allocation limit setting.*

The report titled *Rangitīkei Groundwater Management Zone Allocation Limit* (PDP, 2023) was prepared to assist in setting an updated groundwater allocation limit for the RGMZ and to inform the assessment of groundwater take resource consent applications. The project included updating the 2017 model. The new groundwater model is referred to hereinafter as the 2023 model. The purpose of the 2023 model was '*to represent the groundwater system in the area which generally represents the patterns of groundwater levels observed in bores within the zone*' (PDP, 2023). This review is of the opinion that the 2023 model objectives should have been more relevant, specific, and practical.

According to PDP (2023), there are indications of a long-term declining trend in groundwater levels in the RGMZ in recent years, which appears to correlate with an increase in groundwater abstraction. PDP (2023) maintain that the 2023 model provides supporting evidence to this hypothesis. They note, however, that the current [2023] scale of decline does not appear to have resulted in substantial effects on surface water or the risk of saline intrusion. They ran the 2023 model to simulate how groundwater levels may change under different groundwater use options. They report that their simulations indicate that if groundwater use continues at current [2023] levels or increases by around a third, groundwater levels will continue to decline. Particularly under the increased abstraction scenario, the scale of the decline implied potentially adverse effects on surface water bodies and an increased risk of saline intrusion within a 10-year timeframe. PDP (2023) conclude that the current groundwater use and allocation limit are not at sustainable levels. Hence, they imply that groundwater use may need to reduce from the recent levels. PDP (2023) note that simulation of reducing groundwater use by around a third indicated that groundwater levels would likely stabilise at a similar level to current [2023] levels. Therefore, they recommend an interim allocation limit that is no more than the current level of consented takes, which is around 75% of the allocation limit set in the One Plan. They recommended a staged approach, which initially seeks voluntary reductions in water use from current consent holders. Thereafter, greater water efficiency requirements over a timeframe of around five years would be implemented. Finally, a consents review would be implemented based on further monitoring. PDP (2023) also recommend revision of the RGMZ boundary to better account for groundwater flow directions and alignment of the allocation zone with groundwater recharge zones.

1.2 Context, Concepts, and Scope

Horizons recently commissioned groundwater resource and allocation limit assessments for the RGMZ based on numerical groundwater modelling, a proven industry tool for over three decades. Groundwater models are computer simulations used to investigate the occurrence and behaviour of groundwater, aiding in resource assessment and management. They mathematically represent key features of a groundwater system, including its three-dimensional (3D) extent, hydraulic properties (such as hydraulic conductivity of the geological material), boundary conditions (like the interaction between groundwater and surface water), and stresses (such as recharge and abstraction). The outcomes of a groundwater model are calculated groundwater heads (levels and 3D flow directions) and flows (water budgets for the whole system and its various constituents). By definition, models are simplifications of reality, not exact replicas. They focus on important features rather than trying to represent everything, ensuring that the system is represented in a manner that suits their purpose. Therefore, the most important element in a model review is checking its fitness for purpose.

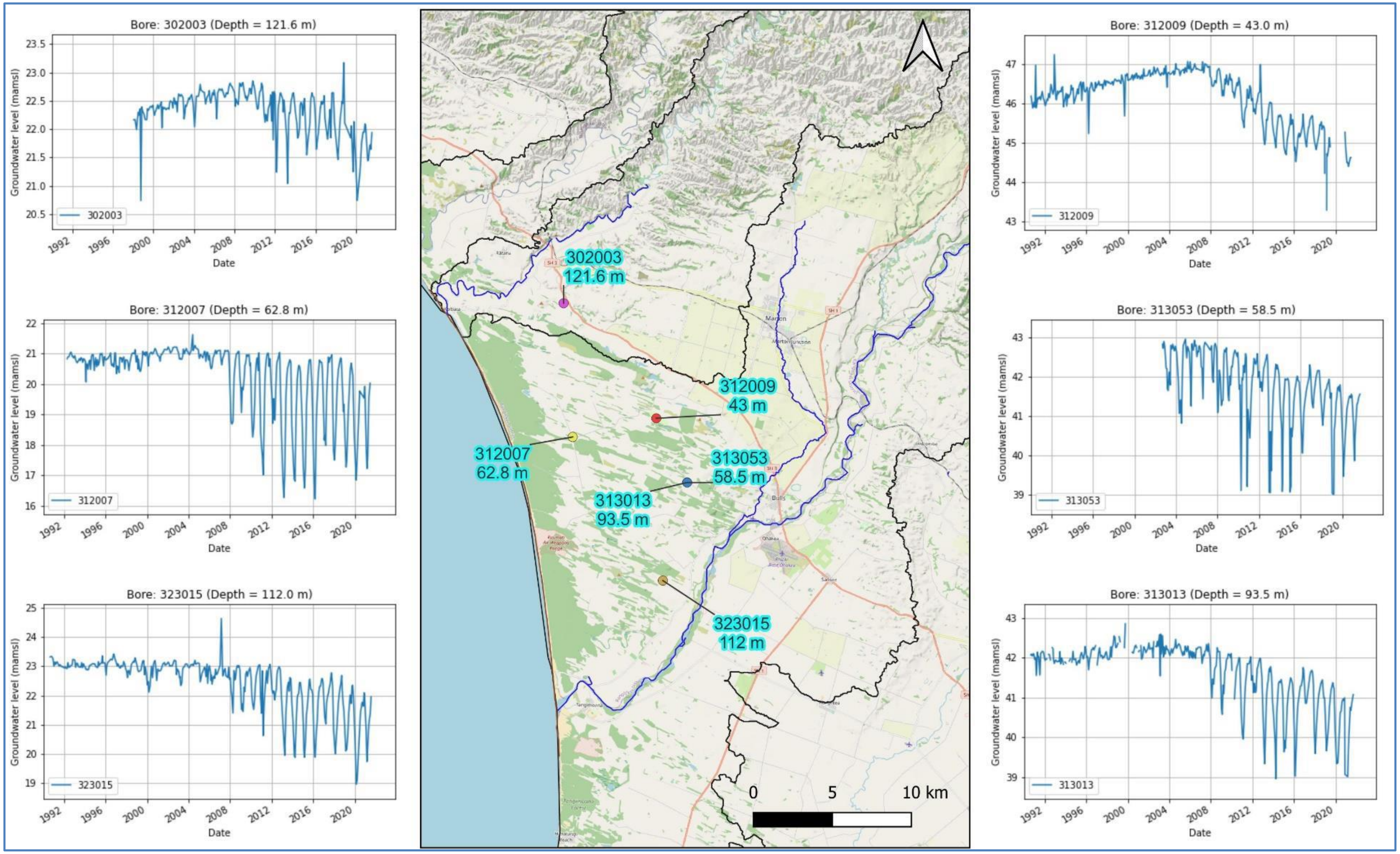


Figure 1-2. Groundwater level hydrographs and locations for selected monitoring bores in the Rangitīkei Groundwater Management Zone (RGMZ) models area (Figure 2 in PDP, 2023).

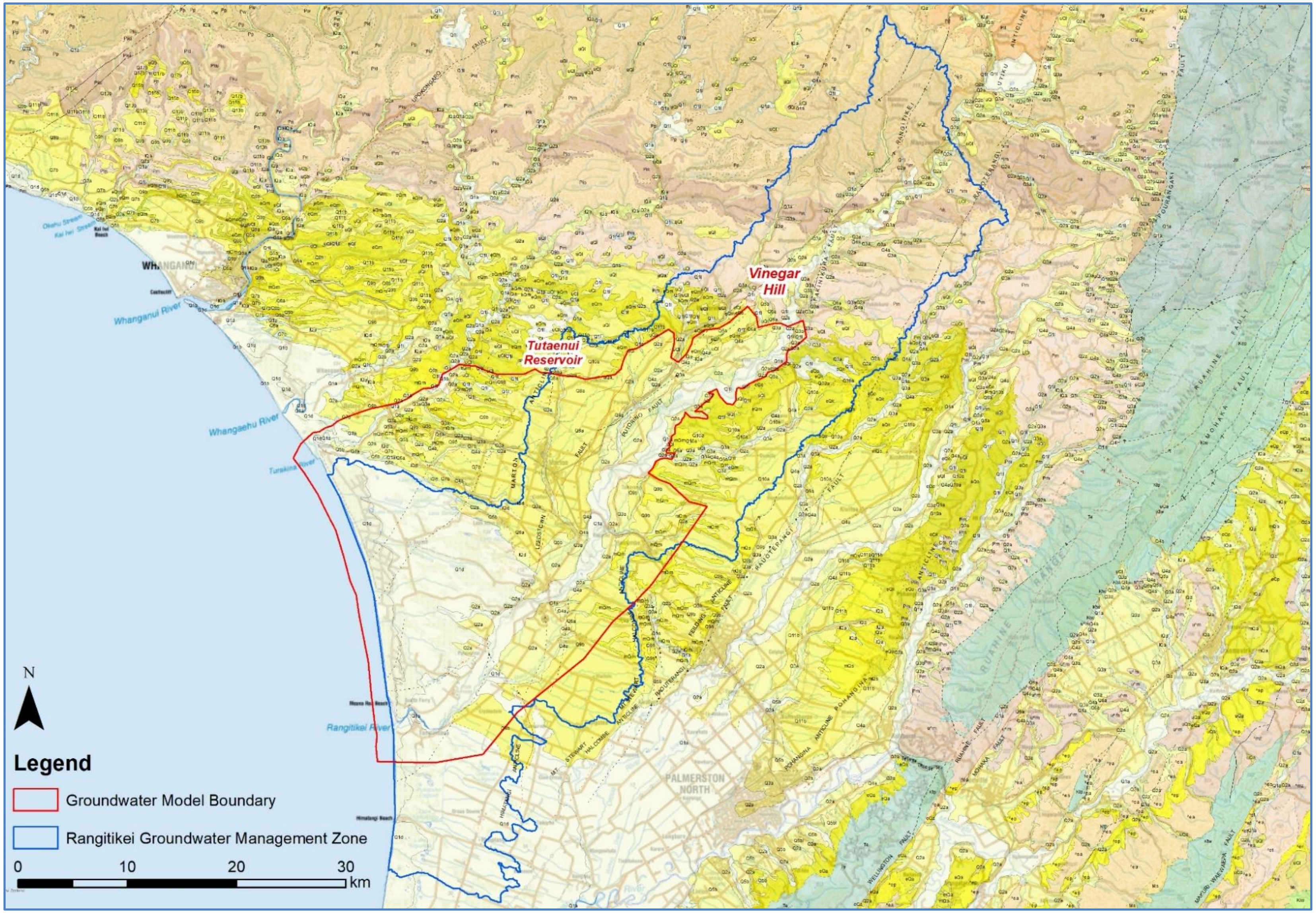


Figure 1-3. Location map showing the extents of the Rangitīkei Groundwater Management Zone (RGMZ) and the 2017 and 2023 groundwater models domain (see Appendix A for clearer presentation of the model domain).

Models are needed because groundwater resources are usually too large, too important, and too sensitive to directly experiment with. Additionally, the time scale involved is very long for practical experimentation. So, hydrogeologists build computer models to represent groundwater flow systems that can answer 'what-if' questions. For example, models can be built to assess the consequences of increased groundwater abstraction on the environment and resource users or evaluate how climate change might impact water supplies. The Australian Groundwater Modelling Guidelines (AGMG; Barnett *et al.*, 2012) argues that '*without modelling, it will never be possible to predict the future behaviour of groundwater systems.*'

It is implicitly assumed that if a model can replicate historical behaviour, it can make trustworthy future predictions. Adjusting model parameters to achieve a close match between model-calculated values and historical observations is termed model calibration. However, successful replication of past behaviour by a model does not guarantee the accuracy of its future predictions. While there is only one reality, there are countless possible models that could represent that reality—models are non-unique. Predictions made by different feasible models can and do vary. Therefore, model predictions are inherently uncertain. All modelling must include uncertainty analysis as part of its workflow (Figure 1-4).

Because models are built using historical data and knowledge current at the time of their development, they usually require updates. Models are commonly updated at specified intervals or when there is a need, such as incorporating newly collected data or addressing new questions. The 2023 update of the 2017 model for the RGMZ is a typical example.

Typically, a model update involves verification (post-audit) of the old model where suitable data are available. Barnett *et al.* (2012) explain that '*verification involves comparing the predictions of the calibrated model to a set of measurements that were not used to calibrate the model. The aim is to confirm that the model is suitable for use as a predictive tool.*' The 2017 model was calibrated against groundwater level data collected from July 1990 to July 2015. According to Section 7 in PDP (2017), it was used to make predictions for a five-year period (July 2015–July 2020). This review is of the opinion that the first step in updating the 2017 model in 2023 should had been to verify its predictions by comparing them to monitoring data that had become available by the time of the model update. PDP (2023) does not indicate that the 2017 had been verified as part of the 2023 model update. This would have provided useful information to the 2023 model update. Notwithstanding, it is noted that PDP (2023) states that the '*model prediction has been largely borne out through continued observation of groundwater levels since 2014, which have continued to decline.*' The report, however, does not compare the 2017 model predictions to monitoring data statistically or graphically.

It is essential to evaluate models through thorough peer review to confirm they are well-built, trustworthy, and fit for purpose, which commonly entails future predictions. Reviews are key components of any significant modelling exercise. Review is integral to all the steps in the typical modelling workflow presented in Figure 1-4. In the AGMG (Barnett *et al.*, 2012), Guiding Principle 9.2 states: '*Three levels of review should be undertaken: a model appraisal by a non-technical reviewer to evaluate model results; an in-depth peer review by experienced hydrogeologists and modellers; and a post-audit* [verification] *involving a critical re-examination of the model when new data is available or when the model objectives change.*' Barnett *et al.* (2012) clarify that '*the post-audit may happen long after the modelling project has taken place.*' Best practice guidelines recommend progressive reviewing of the models, i.e., reviewing models as they are developed. The reviews must be undertaken by suitably qualified and experienced persons. Comprehensive examination of models helps identify areas of strength and potential flaws or biases in them, thereby enhancing their overall quality and reliability. Accordingly, Horizons commissioned Stantec to peer review the RGMZ 2017 and 2023 groundwater models and reports. Verification of the models was not in scope for the review. Model verification entails comparing model calculations against measurements that have not been used in the model setup and calibration to assess the model's performance. The lessons learned from model verification processes are very useful for future modelling. Hence, verification of the 2017 model should have been the first step in the development of the 2023 model. Given that there is a newer version of the 2017 model (the 2023 model), there is no point now in verifying the 2017 model. Any future groundwater modelling in the RGMZ must start by verification of the 2023 model. Verification of the 2023 model was excluded from the scope of this groundwater model review due to work requirements and resource constraints.

The objective of this review is to evaluate the RGMZ 2017 and 2023 groundwater models to ensure they meet Horizons' requirements for groundwater resource assessment and allocation limit setting. This review focuses on assessing the models' effectiveness in representing hydrogeological conditions, their accuracy in reflecting historical groundwater behaviour, and their reliability for predicting future scenarios. Additionally, the review examines the models' assumptions (conceptualisation), methodologies, and performance to determine their suitability for current and future groundwater management needs. The review considers best practice guidelines, including the AGMG (Barnett *et al.*, 2012), NSW Minimum Groundwater Modelling Requirements for SSD/SSI Projects – Technical Guideline (DPE, 2022), and Information Guidelines Explanatory Note – Uncertainty Analysis for Groundwater Modelling (Peeters & Middlemis, 2023), to identify any necessary updates or improvements to enhance the models' effectiveness and fitness for purpose.

The review primarily targets the 2023 model's outcomes but also encompasses an evaluation of the 2017 model since it underpins the 2023 update. The 2023 model is not comprehensively detailed in the PDP (2023) documentation, necessitating a review of both the 2017 and 2023 reports. Additionally, the 2023 model files were reviewed to achieve a thorough understanding of the final model and confirm the accuracy of the RGMZ models documentation.

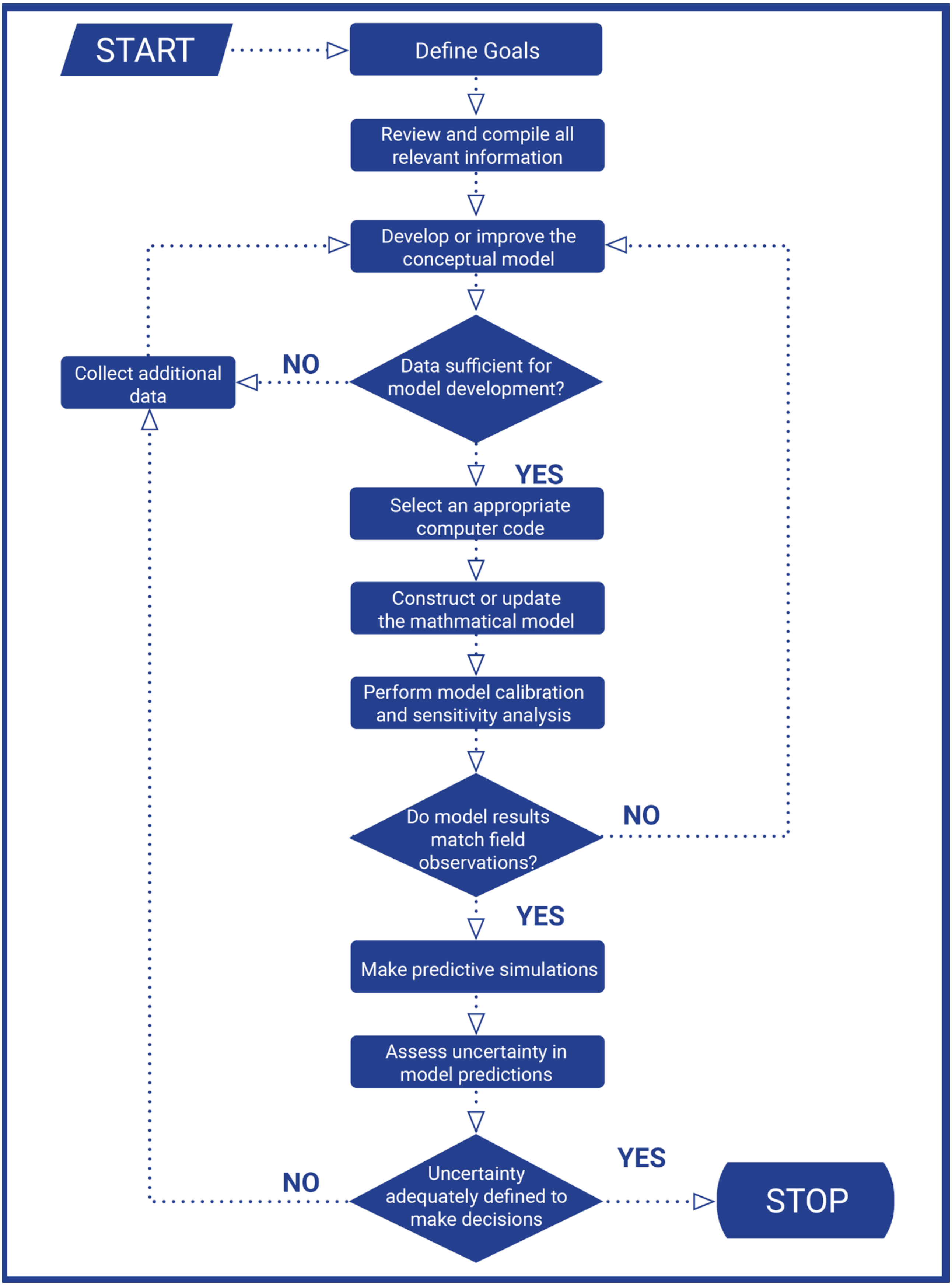


Figure 1-4. Typical groundwater modelling workflow.

2. General Hydrogeological Setting

The hydrogeology of the RGMZ is briefly described in Zarour (2008). It is part of the WB, which comprises up to about 5 km of Pliocene-Pleistocene (Plio-Pleistocene) sediments overlying greywacke basement rocks. The basement rocks crop out at the North Island Central Range and underlie the entire Zealandia continent. Crests of the SW-NE folds in the WB define topographic and surface water divides. It is assumed that the groundwater and surface water divides in the WB coincide (Zarour, 2008). The RGMZ is separated from the Turakina Catchment to the north by the Marton Anticline and from the Manawatū Catchment to the south by the Feilding, Mount Stewart-Halcombe, and Himatangi anticlines (Figure 1-3 and Figure 3-2). The RGMZ covers a land area of approximately 1,573 km² (157,316 ha)[3].

Groundwater in the WB variably occurs in the Plio-Pleistocene strata, depending on the strata's hydraulic properties and recharge potential. Typically, groundwater is deep at elevated areas (at and near topographic and water divides) where there are downward hydraulic gradients corresponding to groundwater recharge zone settings. In low-lying areas, groundwater is shallower, and there are upward hydraulic gradients corresponding to groundwater discharge zone settings. Rivers normally lose flow to underlying aquifers in relatively elevated topographies and gain flow from them at lower elevations. But overall, rivers in the WB constitute regional groundwater discharge features. In general, the regional groundwater flow direction in the WB is aligned with the main surface water flow direction and fold axes, from the NE to the SW. Intermediate and local flow directions vary depending on relevant scale hydrogeological conditions. Rainfall percolation and seepage from surface water where conditions allow are the main sources of groundwater recharge. Depending on hydrogeological conditions, the aquifers discharge into surface water and, ultimately, into the Tasman Sea.

An extensive groundwater source exists in the RGMZ within Last Glacial aggradational gravels that fill the valley and underlie the coastal plain. Groundwater in these deposits occurs at shallow depths and under unconfined conditions, although local confinement can be encountered in some bores. Late Quaternary deposits unconformably overlie Plio-Pleistocene material, which is less permeable to groundwater flow. The Late Quaternary strata are terraced by rivers. At the inland edges of the GMZs, groundwater can be very deep. For instance, in the nearly 300 m deep new Marton water supply bore (Horizons ID 303043) drilled in 2023 in the Tutaenui Reservoir area (see Figure 1-3 for location), the depth to artesian groundwater level exceeds 100 m (Scaife, 2024). Practically, the exploitable groundwater area within the Rangitīkei Catchment does not extend above Vinegar Hill (see Figure 1-3 for location), where the valley narrows, and the geology is predominantly older, more marine deposits dominated, and less permeable.

Rainfall across the catchment varies with geographic position and topographic elevation. Zarour (2008) estimated the average rainfall over the entire RGMZ at 956 mm/year, equivalent to 1.496 x 10^9 m³/year (about 1.5 billion cubic metres per year) over the entire GMZ. The One Plan groundwater allocation limit is 5% of the total rainfall on a GMZ. Hence, in the RGMZ it is 75 x 10^6 m³/year (75 million cubic metres per year) as indicated in Table 1-1.

The RGMZ groundwater system is noticeably connected with the Rangitīkei River system, and the surface and subsurface water systems interact almost continuously but variably in space and time (e.g., Figure 2-1). The flow of the Rangitīkei River varies depending on the location and time of year. On average, the mean annual flow at McKelvies near the river's mouth is approximately 36 m³/s (LAWA[4]), the equivalent to approximately 1.135 billion cubic metres per year, or 76% of the mean annual rainfall. The flow fluctuates based on weather conditions, seasonal changes, and water management practices.

Groundwater resource potential increases in the RGMZ down the valley, reaching its maximum in the coastal plains. The delta-shaped coastal area holds considerable groundwater potential, which was thought in 2008 to be above its use at that time. In this area, prevailing unconfined conditions result in large storativity, which, combined with high transmissivity (reflecting high hydraulic conductivity), provide high bore yields, as seen in the Santoft area, for example.

Bekesi (2001) notes that the deep groundwater in the RGMZ is generally hard and contains excessive quantities of iron and manganese. Data from the new Marton bore confirms this (Scaife, 2024). Bekesi (2001) also notes that the quality and quantity of groundwater available appear to be controlled by proximity to surface water.

---

[3] Calculated from the map in Figure 1-3, reported as 1,564 km² in Zarour (2008).
[4] https://www.lawa.org.nz/explore-data/Manawatū-whanganui-region/river-quality/Rangitīkei/Rangitīkei-at-mckelvies.

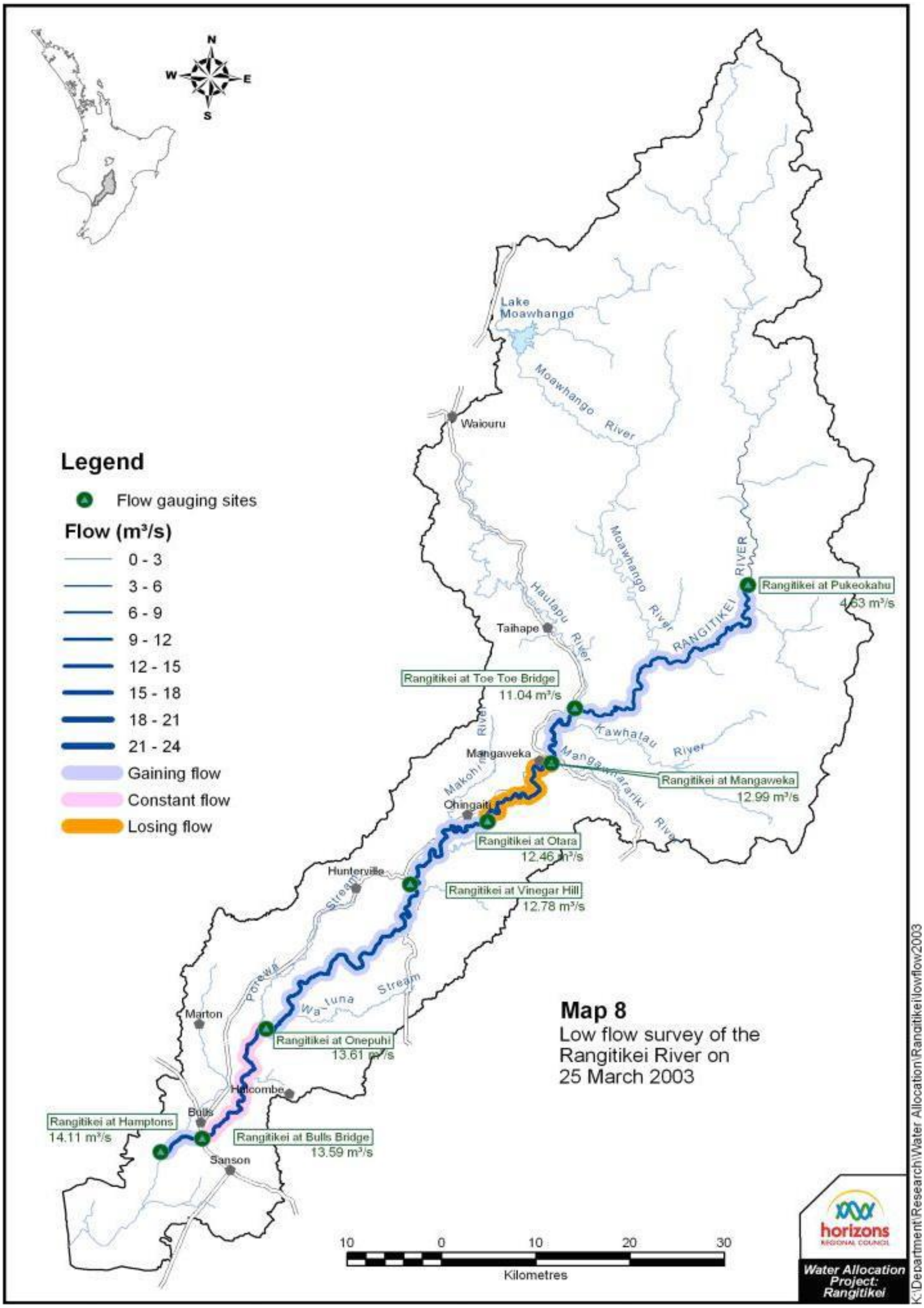


Figure 2-1. Low-flow survey of the Rangitīkei River on 25 March 2003, showing gaining and losing stretches of the river on the survey date (from Roygard and Carlyon, 2004).

Numerical data on groundwater consenting and abstraction are not provided in PDP (2017) or PDP (2023). Hence, this review cannot independently assess groundwater demand, supply, and consenting conditions in the RGMZ. PDP (2023) notes that many groundwater consents in the RGMZ are in the process of being renewed or will be renewed in the next two to three years (from 2023). Applications to increase the consented groundwater abstraction volume have also been submitted. PDP (2023) believes that, against a backdrop of declining groundwater levels and legislative direction to manage freshwater sustainably, any increase in the allocation of groundwater, or even the renewal of existing consents, may be unwise despite the available groundwater allocation within the One Plan limit. The report further infers that as a result, Horizons is investigating updating the groundwater allocation limit to ensure it is consistent with national statutory instruments, e.g., the National Policy Statement for Freshwater Management (NPS-FM, 2020).

The RGMZ has culturally and environmentally significant groundwater dependent wetlands and coastal dune lakes, which must be considered when assessing the environmental feasibility of groundwater development projects. Groundwater abstraction can result in depleting surface water, either by directly withdrawing from surface water features or by intercepting groundwater flow that feeds them. While there is substantial groundwater potential in the Santoft area, prudent development is essential to ensure resource sustainability. The connection with surface water, including rivers, streams, lakes, and wetlands, could necessitate restrictions on groundwater abstraction during low surface water levels and flows. Impacts on other users are also important when determining the feasibility of groundwater abstraction proposals. Additionally, attention must be given to the hydraulic connection between the ocean and groundwater in the coastal area, as seawater intrusion into the freshwater system would render it unusable indefinitely. The groundwater system's coastal boundary is discussed in Section 4.3.3 and illustrated in Figure 4-13. The coastal area in the RGMZ is a typical setting where seawater can encroach into the underlying aquifers if groundwater is not sustainably managed.

3. RGMZ Models

The RGMZ 2017 and 2023 models were developed using MODFLOW, the most widely used groundwater modelling software. MODFLOW's extensive validation, modular structure, and comprehensive support for various hydrological processes have established it as the industry standard for groundwater modelling. Both models were calibrated using PEST (**P**arameter **EST**imation), a software package designed to adjust model parameters to minimise discrepancies between observed data and model simulations. Effective calibration enhances the models' accuracy and predictive capabilities, ensuring they serve as reliable tools for groundwater analysis and management.

The RGMZ 2017 model was constructed using the MODFLOW-NWT code and the FloPy user interface. MODFLOW-NWT is an extension of the MODFLOW code that employs a finite-difference approach to solving groundwater flow equations, like previous MODFLOW versions, but it also utilises the Newton-Raphson method for handling nonlinearities. FloPy is a Python package designed to create, run, and post-process MODFLOW models efficiently. The 2017 model was developed and calibrated using a 25-year groundwater level data record from July 1990 to July 2015.

The RGMZ 2023 model is an update to the 2017 model. It was developed using the FloPy package and employs the MODFLOW 6 code. MODFLOW 6 is the latest version of the MODFLOW suite and utilises the finite volume method instead of the finite difference approach, offering advanced features for groundwater system simulations. Key advancements in MODFLOW 6 include the flexibility to use unstructured grids, where cell dimensions can vary and grids do not require consistent rows, columns, or layers, allowing for discontinuities or refinements in cell and layer sizes. This provides more sophisticated modelling options compared to the structured grid versions of MODFLOW. However, the RGMZ 2023 model uses the same grid structure as the 2017 model and does not incorporate unstructured gridding. The 2023 model was developed and calibrated using a 32-year groundwater level data record from July 1990 to July 2022, extending the data record by seven years compared to the 2017 model.

A common feature of both MODFLOW-NWT and MODFLOW 6 is their ability to handle thin cells and manage the drying and rewetting of cells through the Newton-Raphson method. However, there are no indications that modelling the RGMZ, as reported in PDP (2017) and PDP (2023), requires any of the special capabilities that MODFLOW-NWT and MODFLOW 6 offer beyond those provided by previous versions of MODFLOW, especially given the use of structured grid not unstructured grid, continuous layers across the entire model domain, and the fixed 500 m x 500 m cell size that is largely adequate for a regional scale model. The way cell drying and re-wetting was handled in the models is not reported.

PDP (2023) is shorter and provides less detail on the model compared to PDP (2017). As expected from its title, which emphasises groundwater allocation, the 2023 report primarily focuses on exploring various allocation options rather than offering a comprehensive description of the model development. In contrast, the 2017 report presented a more detailed account of the model's development.

According to PDP (2023), the 2017 model update involved changes to and an extension of recharge and abstraction, as well as model recalibration against an extended groundwater level timeseries.

1. <u>Updated recharge data</u>

   Like with the previous model, a Rushton method-based approach (Rushton & Redshaw, 1979; Rushton, 2003; Rushton, 2004) was adopted. However, data from Horizons climate station network and publicly available data from the NIWA Cliflo database were used instead of NIWA Virtual Climate Station Network (VCSN) data. The method was adapted to allow for a lagged recharge effect in some areas of the model where groundwater levels were more than 20 m deep, broadly the area inland of State Highway 3 (SH3). The lagged effect was achieved by applying a limit to the drainage rate from the soil store to the groundwater recharge store, which simulated the smoothing effect of seasonal recharge through a substantial unsaturated zone. PDP (2023) does not provide a comparison between the recharge datasets used in the 2017 and 2023 models. Hence, the impacts of the two changes made in recharge assessment is unknown, namely: (1) the use of a different data source, and (2) the adjustment of calculations to represent the perceived lag in recharge (travel time from the soil zone to the water table under unsaturated conditions).

   It seems that the adaptation to the recharge calculation method was made in response to the recharge assessment review comments by Barnett (2016). However, this is not stated in PDP (2017).

2. <u>Updated groundwater abstraction data</u>

   Given that the decline in groundwater levels in the Santoft area is perceived to be correlated with an increase in abstraction, PDP (2023) considered groundwater use data a critical input for the model. The model was updated with groundwater abstraction data mainly from records provided to Horizons by consent holders. Gaps in the record were infilled based on typical use patterns from nearby bores using water for the same purpose (i.e., irrigation, Stockwater, etc.), calculated as a proportion of the daily consented volume. While this methodology is not perfect, it is important to note that all modelling involves making simplifications, and modellers often have to use less than ideal datasets, including estimates. Figure 3-1 shows the proportion of actual use data to infilled data used in the model. PDP (2023) suggests that Figure 3-1 indicates that the consented rate of abstraction has remained relatively stable since around 2018, though there have been variations in actual use. Simply, this suggests that nearly no additional groundwater consents were granted, and no old consents expired or were surrendered.

3. Updated groundwater level observations and model recalibration

The 2023 model was calibrated to groundwater levels measured monthly at monitored operational bores within the model area (Figure 3-2). PDP (2023) notes that the model was calibrated to both annual maxima and monthly measured groundwater level data (Figure 3-3). They are of the opinion that calibrating the model against annual maxima allows long-term trends to be better matched during the calibration process. They assigned greater weight to annual maximum water levels in the calibration process to ensure that the model better matches long-term trends rather than short-term seasonal changes, as the aim of the model is to predict the effect of abstraction on long-term groundwater level trends.

It is noted that groundwater level data from Bore 312006 (Figure 3-4) were excluded from the 2023 model despite being used in the calibration of the 2017 model. This change is not noted in PDP (2023) and no reason for the bore exclusion from the new model calibration is provided.

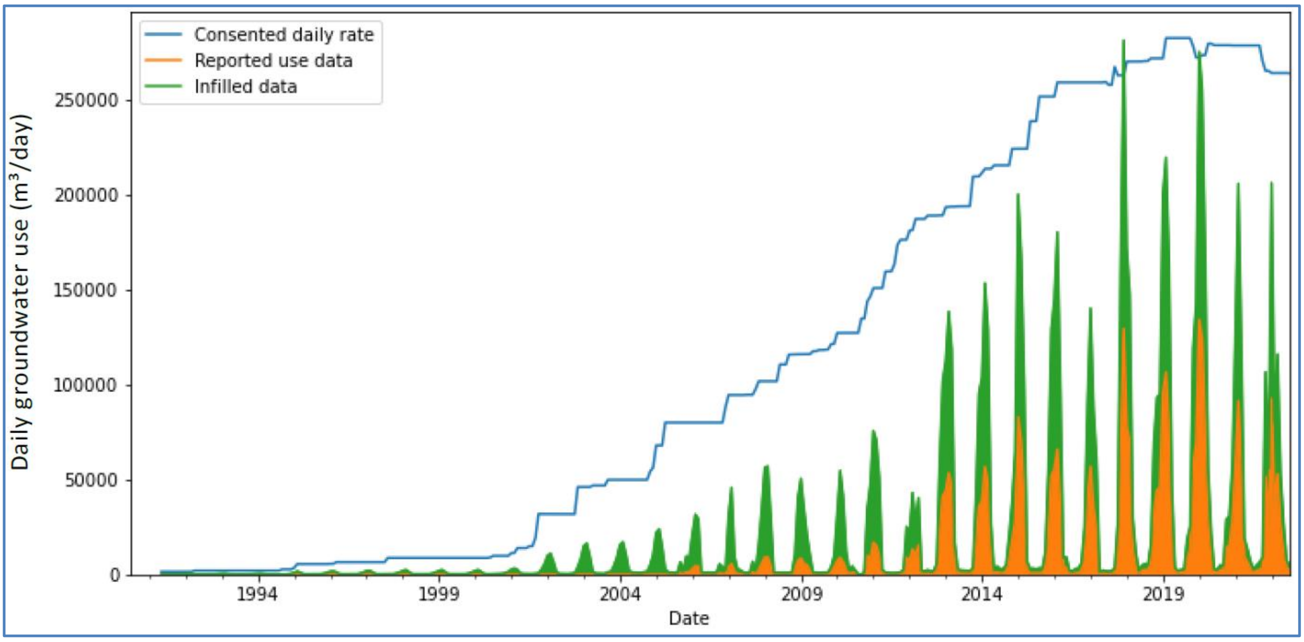


Figure 3-1. Actual use data and infilled data (from PDP, 2023). Consented daily rate in m³. Total daily use in m³ as average per month (Figure 6 in PDP, 2023).

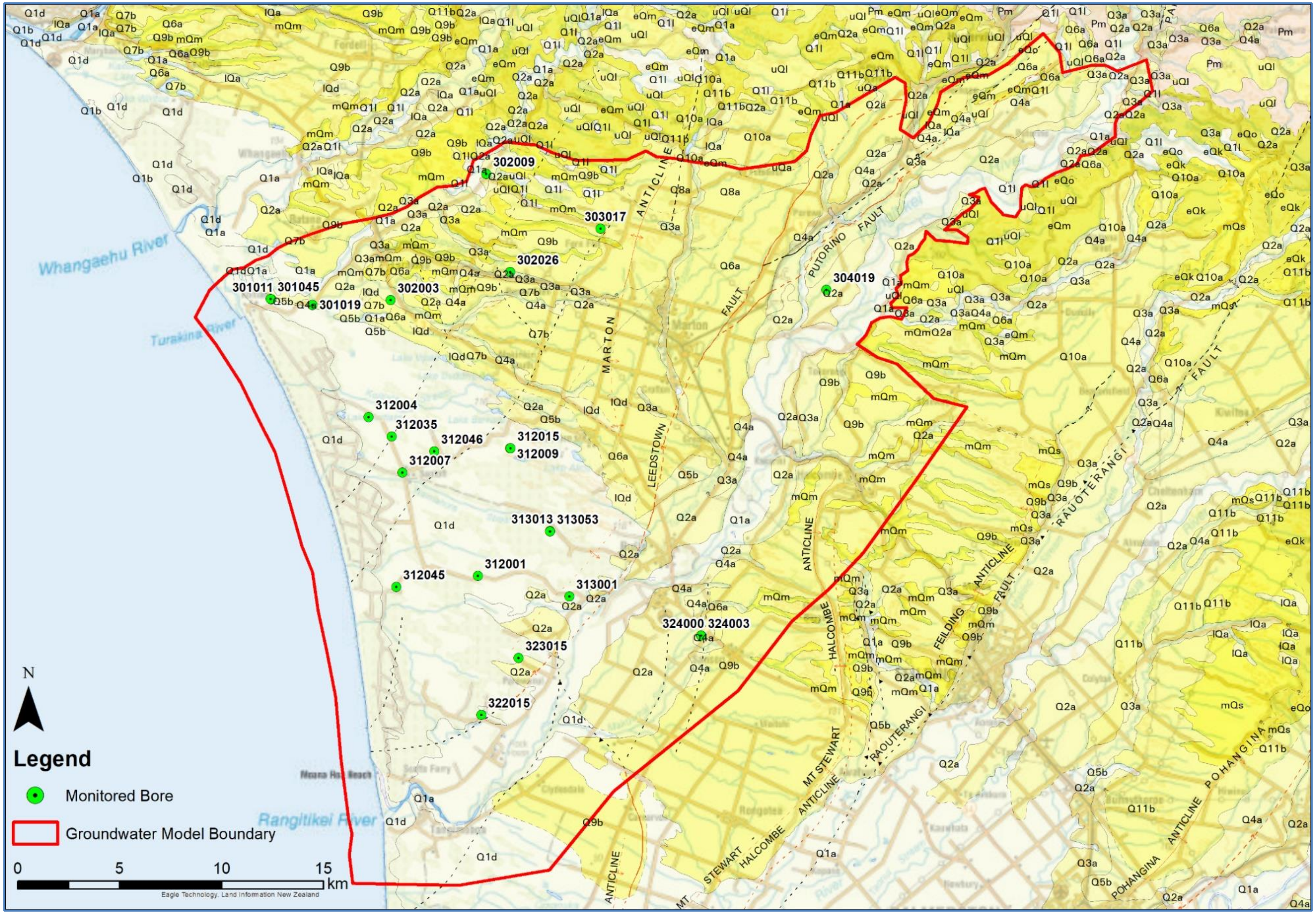


Figure 3-2. Location map showing groundwater monitoring bores used in the model calibration (some bores overlap on the map due to the small distance between them, e.g., Bores 324000 and 324003).

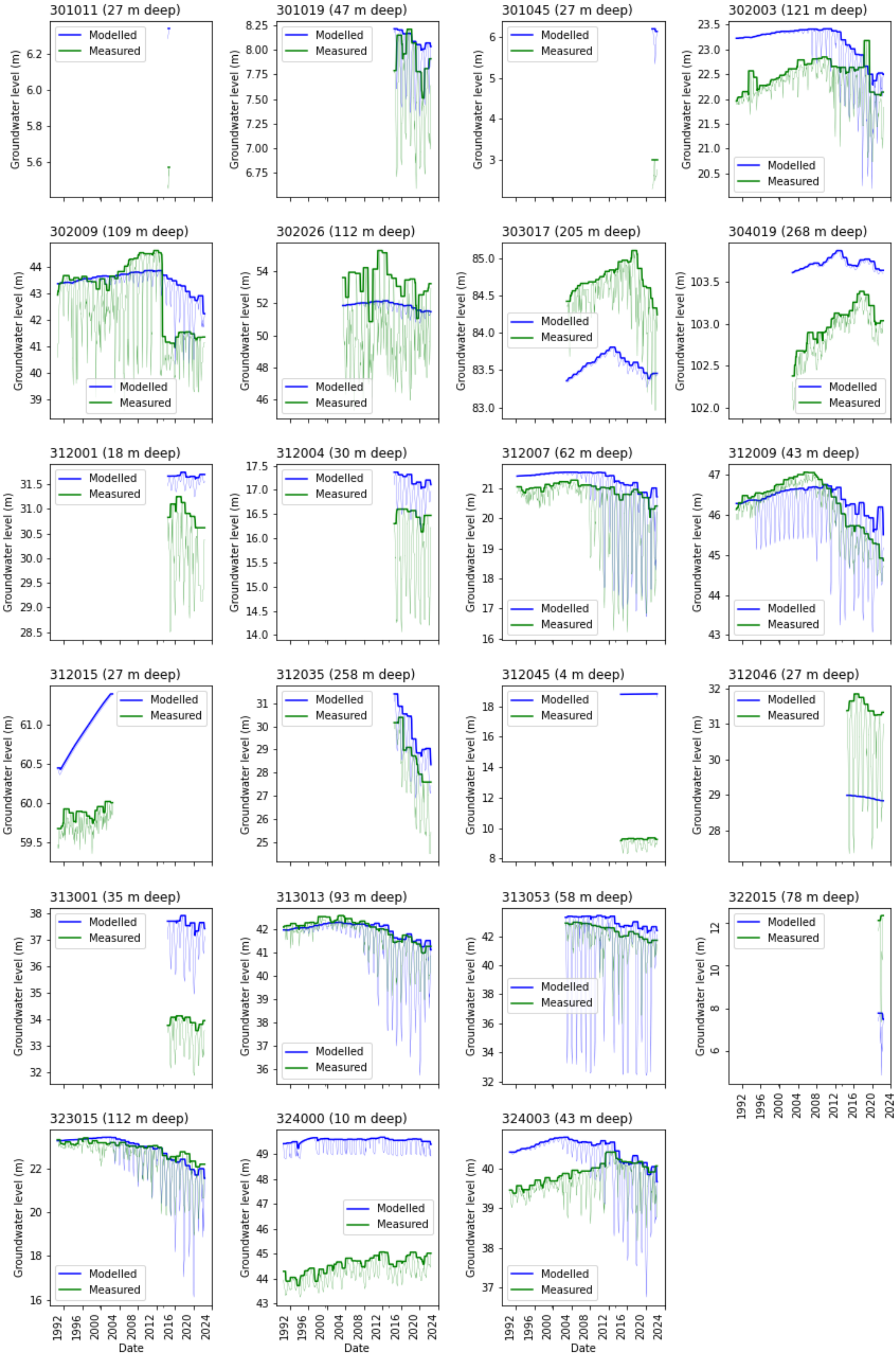

301011 (27 m deep)
301019 (47 m deep)
301045 (27 m deep)
302003 (121 m deep)
302009 (109 m deep)
302026 (112 m deep)
303017 (205 m deep)
304019 (268 m deep)
312001 (18 m deep)
312004 (30 m deep)
312007 (62 m deep)
312009 (43 m deep)
312015 (27 m deep)
312035 (258 m deep)
312045 (4 m deep)
312046 (27 m deep)
313001 (35 m deep)
313013 (93 m deep)
313053 (58 m deep)
322015 (78 m deep)
323015 (112 m deep)
324000 (10 m deep)
324003 (43 m deep)
Modelled
Measured
Groundwater level (m)
Date

Figure 3-3. Observed and modelled groundwater levels. The dark blue and dark green lines represent the modelled and measured annual maximum groundwater levels (Figure 9 in PDP, 2023).

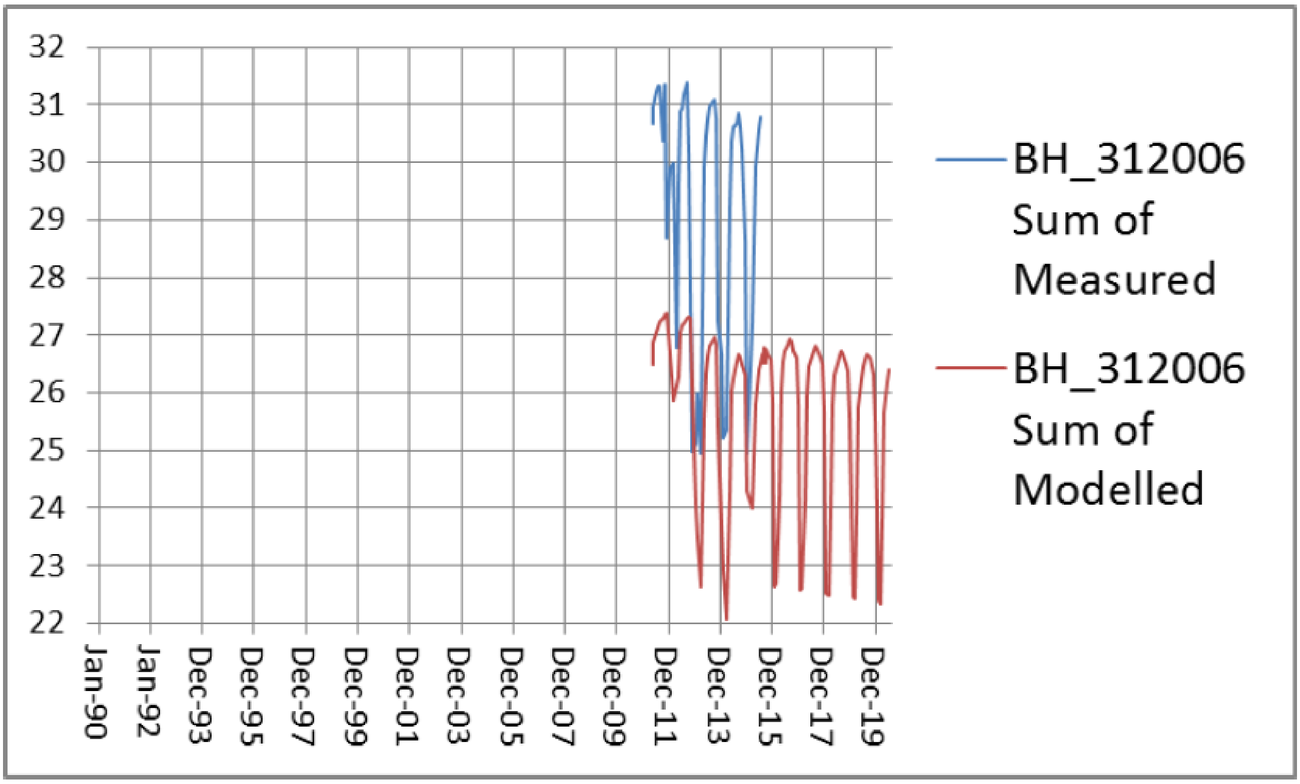


Figure 3-4. Observed and model-calculated groundwater levels at Bore 312006 (from PDP, 2017).

4. RGMZ Models Review

This section reviews and comments on the RGMZ modelling reported in PDP (2017) and PDP (2023). It is based on the reported information and examination of the provided 2023 model files, with additional information from publicly accessible sources, Horizons, and PDP. The review also included personal communication with the PDP modeller. The review is guided by the workflow presented in Figure 1-4 and addresses the model evaluation questions listed in DPE (2022).

4.1 Planning

Successful groundwater modelling starts with careful planning. Understanding the genesis of the problem and the existing knowledge base is essential for setting practical objectives and identifying limitations, including technical factors (e.g., data and software availability and adequacy) and other considerations (e.g., cost). A crucial aspect of the planning process is determining how the model will address relevant questions and contribute to solving the identified problems. This understanding helps set appropriate targets for model complexity and confidence levels. When updating previous models, it is important to clearly demonstrate the added value of the new modelling efforts.

4.1.1 Modelling Incentive

The incentive for the modelling is clear in the reports, being '*to assist with groundwater allocation in the Rangitīkei Groundwater Management Zone*' (PDP, 2023), given that a perceived '*declining trend in groundwater levels has been noted in some monitoring bores in the Rangitīkei*[5] *and Turakina GWMZs*[6]*, likely related to increasing abstraction.*' PDP (2017). PDP (2017) notes that '*the affected bores are generally located in the vicinity of Santoft township and that area has experienced notable groundwater development over the past 10 to 15 years.*' Accordingly, PDP (2017) states that the model objective was '*to provide a tool for investigating groundwater abstraction impacts on groundwater levels and groundwater discharge in the Santoft area.*' The key questions for the 2017 model were:

- '*Is groundwater level decline likely to continue under current rates of abstraction?*
- *Alternatively, are groundwater levels likely to stabilise to a new steady state? If so, how long will that take to occur?*'

The 2017 model exclusions are clear—'*the model is not intended to simulate local scale drawdown effects between abstraction bores.*'

PDP (2023) highlights that '*Horizons is investigating updating the groundwater allocation limit so that it is consistent with national, statutory instruments (e.g., the National Policy Statement for Freshwater Management (NPS-FM), 2020).*' However, it does not clarify why an update to the 2017 model is required or how the updated model will help meeting NPS-FM (2020) requirements.

The objectives of the model should include exploring and evaluating all possible reasons for the anecdotal groundwater level decline.

The modelling incentives and objectives in the reports should have provided basis for:

1. the conceptual model including defining the model spatial and temporal domains—inadequate
2. a suitable target model confidence level class, e.g., as described in Barnett *et al.* (2012)—not considered
3. modelling data requirements including, for example, the need for a geological model—overlooked
4. the selection of the numerical modelling approach including code, parameterisation approach, and space and time discretisations—not addressed
5. model calibration criteria including conceptual and numerical modelling sensitivity analyses, steady state and transient modes, manual and/or automated methods, the objective function and model performance metrics—incomplete
6. the definition of the predictive scenarios—adequate
7. predictive uncertainty assessment—overlooked
8. a framework for recommendations including further work needs (investigations and modelling) —overlooked.

[5] Mainly spelled as Rangitikei in PDP (2017) and PDP (2023).
[6] GWMZ in the reviewed reports. GMZ herein.

4.1.2 Numerical Modelling

Although not explicitly discussed or stated in the reports, numerical modelling is the best option to address Horizons' objectives as it offers the best possible representation of the system and superior flexibility compared to other methods. Overall, planning for numerical modelling of the RGMZ is acceptable. The modelling was undertaken after compiling reasonable knowledge of the system, albeit not entirely adequate, particularly concerning hydrostratigraphy. The 2017 and 2023 models' lack of a proper geological framework and unrealistic representation of surface water features are significant issues that compromise their reliability. These deficiencies could lead to inaccurate predictions and undermine the models' overall validity. Addressing these critical shortcomings, along with the other issues identified in this review and the previous review (Barnett, 2016), is essential for ensuring the robustness and reliability of any new modelling.

The reported numerical groundwater modelling was embarked upon in 2017 after undertaking groundwater level monitoring for 25 years as part of Horizons' State of Environment (SoE) programme, limited hydrogeological investigations like the October 2014 piezometric survey in the Rangitīkei and Turakina GMZs, basic private assessments mainly as part of assessments of environmental effects (AEE) for resource consent purposes, and eigen modelling—a linear algebra-based modelling technique.

The numerical model objectives are clearly stated and appropriate. It is clear how the model will contribute to meeting the assessment objectives. However, commensurate target model complexity and model confidence level classification (CLC; see Section 4.9) were not determined prior to modelling. The reports do not present any planned limitations and exclusions for the numerical model. This is different from the general modelling exclusions mentioned in the previous section. It is about the model build and performance assessment metrics. For example, there is no discussion of the adequacy of monitoring data and their representativeness of the system in space and time. It is also noted that the planning overlooked the potential use of groundwater level data collected and provided to Horizons by consent holders in compliance with their consent conditions.

The 2023 model is an update to the 2017 model. The update included a change of the modelling code from MODFLOW-NWT to MODFLOW 6. These are important model planning decisions. The reports do not discuss the reasons for using MODFLOW-NWT in the 2017 model and the rationale for the subsequent change to MODFLOW 6 in the 2023 model.

The overall design of the 2017 model was not changed in the 2023 model. The new model adopts the same geometry, planar cell spacing, layering, and boundaries as the old model. The reports do not discuss the adequacy of the initial design and why it was maintained in the new model.

The 2017 model was recalibrated as part of the 2023 update. The reasons for the recalibration are understandable, but PDP (2023) does not state them. The report does not provide adequate critique of the performance of the previous model and the lessons learned from the previous modelling work. As noted above, the 2023 model update should have started with verification of the 2017 model.

4.2 Conceptualisation

A conceptual model is a mental image of the system being modelled. Like all models, conceptual models are simplifications of reality and are subject to uncertainty and non-uniqueness.

A groundwater conceptual model is a simplified framework that describes the key components and processes of a hydrogeological system. It helps to understand and communicate the system's structure, function, and interactions and forms the basis for more detailed analytical or numerical models. It typically includes descriptions of geology, hydrostratigraphy, hydrogeology, flow systems, boundary conditions, hydrological interactions, and spatial and temporal variations. It aids communication with stakeholders (scientists, engineers, decision-makers, the public), guides the development of mathematical models, and provides a basis for hypothesis testing. According to DPE (2022), it may be developed as maps, cross-sections, block models, and tables of aquifer characteristics, and may also be maintained in a digital format. The conceptual model must be checked using simple analytical models and water balance calculations.

Conceptual modelling is a prerequisite for developing more detailed and quantitative models.

Section 2 in PDP (2017) and Section 2.1 in PDP (2023) present the groundwater system conceptualisation for the 2017 and 2023 models, respectively. PDP (2023) incorrectly references the previous report as '*PDP (2015)*'.

The conceptual model description in PDP (2023) is shorter than in PDP (2017). Neither report presents a conceptual model graphically (e.g., cross-sections or block diagrams), nor do they provide an analytical water balance of the modelled system, which is an important element in hydrogeological conceptualisation.

4.2.1 Transfer of Project Objectives to Conceptual Model

The modelling reports offer conceptual notes rather than a conventional conceptual model. The conceptualisation attempts to explore the reasons for groundwater level changes observed in space and time and to identify potential links to various conditions and stresses, such as hydraulic properties, connections to surface water, rainfall recharge, and groundwater abstraction. In this sense, the research and modelling objectives have been successfully transferred to the conceptual model and, subsequently, to the numerical simulation and predictive modelling. Future work must present the conceptual model more clearly, preferably graphically too.

Determining the analysis unit is a fundamental first step in any analysis. The title of the PDP (2017) report is '*Santoft Groundwater Model Report*', which implies that the model's focus is on the Santoft area. However, the term '*Santoft area*', as repeated in the report, is not defined in suitable geographical or hydrogeological terms. The title of the PDP (2023) report is '*Rangitīkei Groundwater Management Zone Allocation Limit*', implying that the focus is on the RGMZ. However, the 2023 report and model focus on the Santoft area, which remains undefined in clear hydrogeological terms. Future modelling must consider the feasibility of defining groundwater management subzones and corresponding allocation limits within the RGMZ. The current modelling does not provide a suitable basis for achieving this.

Hence, it is not clear how the research and modelling objectives were transferred to the 2023 conceptual model. The modelling domain defined for the models is not suitable for assessing the suitability of the allocation limits of the RGMZ as defined in the One Plan because the two areas differ (Figure 1-3). The models could be used to assess whether there is over-abstraction in their domains, but not to calculate the allocation limit for the RGMZ. This is considered an inconsistency in transferring the modelling objectives to the conceptual and numerical models.

Notwithstanding the above comments on the unexplained incongruity between the model domains and the extent of the RGMZ, the conceptual model description is technically consistent with the stated modelling objectives. However, the conceptual model description in the reports does not outline a suitable level of model complexity and confidence level. While the reports identify the key hydrogeological processes, they do not provide an adequate conceptual discussion of the model's spatial coverage and peripheral boundary conditions.

4.2.2 Literature Review and Existing Data

No numerical models have been developed for the RGMZ prior to the 2017 model and there are very limited hydrogeological studies of the system. The 2017 and 2023 models utilise available data from different sources. PDP (2017) used NIWA Virtual Climate Station Network (VCSN) data in groundwater recharge calculations. These data have not been available to the 2023 model, which therefore had to use Horizons climate station network and publicly available data from the NIWA Cliflo database. PDP (2023) does not present a comparison between the two datasets or the recharge estimates used in the old and new model.

The two reports do not include literature review sections. The literature reviews undertaken are not thorough. For example, PDP (2023) overlooks recent relevant work like Rees *et al.* (2019). The reports and models do not adequately utilise groundwater quality data from Horizons databases. Those data would have helped in the characterisation of the groundwater system and the interaction with surface water. The modelling also does not utilise compliance data like groundwater level data that are required to be collected and provided to Horizons by consent holders as part of their consent conditions.

Pumping tests are useful for aquifer typification and are commonly the main source of aquifer properties for modelling. Section 2.6 in PDP (2017) state that '*Pumping tests have been performed in a variety of locations in the study area.*' Figure 4-1 shows the locations of the pumping test bores, bore depths, and estimated transmissivity (T) values. The pumping test data are not presented numerically or graphically. PDP (2017) methodology for converting pumping test T to K for modelling is appropriate. However, K values derived from pumping tests are not reported and apparently there are inconsistencies between the T values presented in Figure 4-1 and the K values in the model presented in Figures 18 a through 18 k in PDP (2017). K values used in the 2023 model are not presented in map format. PDP (2023) does not discuss or present differences in K values adopted in the 2017 and 2023 models as a result of model recalibration. For this review, K values used in the 2023 model were extracted from the provided model files and are presented in Appendix B together with the reported 2017 model K values. There are noticeable differences between the two models (also see Section 4.3.3).

It is noted here that the map presentation of the K data in PDP (2017) is not adequately clear. It is also noted that few of the K value maps presented in Appendix B show bullseyes that could indicate unrealistic parameterisation and/or model calibration issues.

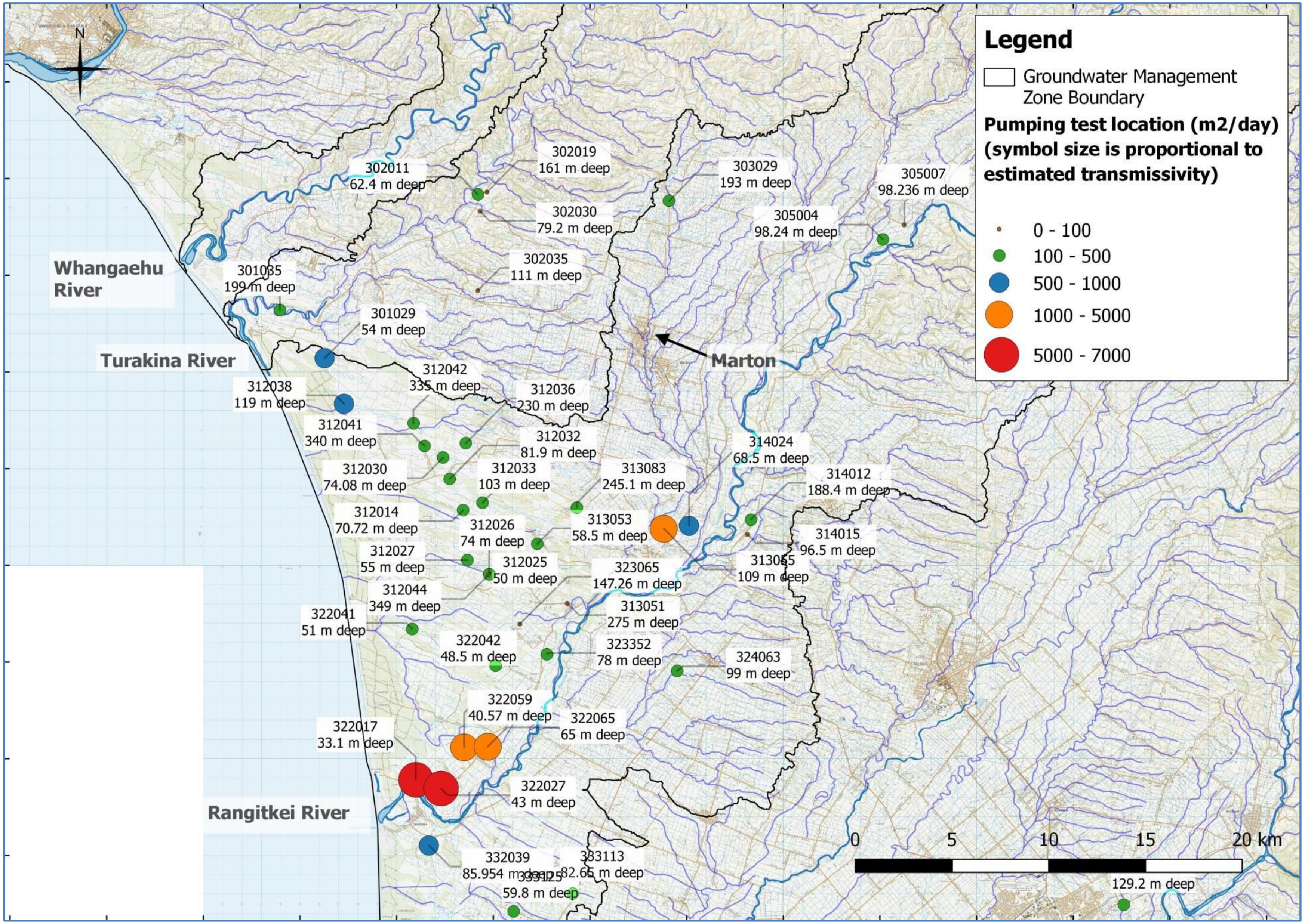


Figure 4-1.Transmissivity values from bores in the model (Figure 10 in PDP, 2017).

4.2.3 Collation of Observed Heads and Flow Data

The modelling projects collated groundwater level (head) data from Horizons SoE monthly groundwater level monitoring network and used them in hydrogeological analysis and model conceptualisation and calibration. The analysis did not incorporate consents compliance monitoring groundwater level data.

It is noticed that groundwater level data from Bore 312006 were used in the 2017 model (Figure 3-4). However, these data were excluded from the 2023 model. PDP (2023) does not note or justify this omission.

SoE monthly groundwater level monitoring data used in the modelling are presented graphically as hydrographs (timeseries plots) in both reports, e.g., Figure 8 in PDP (2017) and Figure 2 in PDP (2023), which is reproduced in this review as Figure 1-2. These data are also included in PDP (2023) as part of model calibration plots (observed vs simulated groundwater level hydrographs) and in predictive scenario simulation plots. SoE monthly groundwater level monitoring data are not presented in map format.

Section 2.4 in PDP (2017) describes the groundwater flow system, mainly based on a piezometric survey of bores that Horizons undertook in the Rangitīkei and Turakina GMZs in October 2014. However, it does not note the value of these data to numerical model calibration. The report presents the results of the 2014 piezometric survey in the form of two piezometric maps for bores shallower and deeper than 20 metre below ground level (mbgl). The reports do not present the rational for this arbitrary division of the aquifer system. The lack of a data-driven basis for this division could lead to misinterpretations of groundwater behaviour and conceptual and numerical modelling errors, including groundwater-surface water interaction and allocation limit estimates.

Based on the above-mentioned piezometric maps, PDP (2017) suggests that groundwater flow generally follows the topography, with piezometric mapping evidence to groundwater discharge into the Rangitīkei River upstream of Bulls. The report should have used concurrent river gauging data to check on this hypothesis (e.g., Figure 2-1). PDP (2017) notes that the maps depict vertical head gradients. It argues that the 'shallow aquifer' piezometric map indicates a distinct flow direction towards the Rangitīkei River, particularly upstream of Bulls, which the report suggests is consistent with information regarding gaining and losing reaches along the Rangitīkei River. This is inconsistent with Figure 2-1, which indicates that the Rangitīkei River flow upstream of Bulls is constant, at least at the time of that gauging (25 March 2003). The figure suggests that the river stops gaining flow from about Onepuhi, about 15 km upstream of Bulls. PDP (2017) maintains that the 2014 piezometric survey also provides evidence of flow gain in the Rangitīkei River from the 'deep aquifer'. These conclusions are dependent on the arbitrary division of the aquifer system, the contouring method, and do not utilise surface water data. Hence, they should be interpreted with caution, as they may not be representative of the system.

PDP (2017) notes that upward hydraulic gradients are evident in some locations, indicating groundwater discharge settings. It states it bases this opinion on groundwater level observations in some deeper bores like the 335 m deep Bore 312042 which has a static water level of 2.4 m above ground level. Bore 312042 is not part of the SoE groundwater monitoring network and it is unknown whether it has analysable groundwater level timeseries record. The report also notes that to the south-west of Bulls, closer to the coast and around Santoft Forest, there is upwards groundwater gradient indicating groundwater discharge settings.

PDP (2017) notes that the greatest depths to water are observed in bores to the north of Bulls, over 100 mbgl. However, it does not clarify whether that is depth to the water table or artesian head. The report notes that generally there is a strong downward hydraulic groundwater at elevated parts of the RGMZ. However, it does not explain whether that indicates a groundwater recharge area setting or the presence of perched groundwater. Recent work by Scaife (2024) suggests that in the Tutaenui Reservoir area there is a shallow perched aquifer separated by a considerably thick unsaturated sequence from a deeper semi-confined/confined aquifer.

The SoE groundwater level monitoring includes bores that could be considered as pairs/clusters for vertical hydraulic gradient analysis purposes including model conceptualisation and calibration (Table 4-1). A good example includes Bore 324000 (10 m deep) and Bore 324003 (43.3 m deep) located just to the south-east of the Ohakea Airbase. This bores pair shows downward hydraulic gradient, which could be indicative of groundwater recharge area setting (Figure 4-2). However, the difference in groundwater level in these two bores could also be due to perching or the fault that passes through their area (Figure 3-1). This situation requires more investigation, e.g., through pumping tests. Note that the reports do not consider geological structure and, subsequently, the models do not represent them.

Table 4-1. Monitoring bores clusters (monitoring bores located close together at various depths).

Cluster	Bore ID	Monitoring Record			Model layer	Depth (mbgl)	Depth difference (m)
		Measurement	From	To			
1	301011	5	1/07/2014	6/11/2014	3	27.0	0.9
	301045	13	5/05/2021	23/08/2022	3	27.9	
2	312015	132	17/10/1990	4/09/2002	1	27.0	16
	312009	341	11/10/1990	23/08/2022	4	43.0	
3	313053	223	17/09/2002	18/08/2022	4	58.5	35
	313013	346	4/09/1990	18/08/2022	5	93.5	
4	324000	362	4/09/1990	25/07/2022	1	10.0	33.3
	324003	335	4/09/1990	19/08/2022	5	43.3	

It is noted that the bores in Cluster 1 in Table 4-1 could be troublesome for model calibration as they occur in the same model cell (same model layer, column, and row). This is not discussed in the model conceptualisation or calibration sections in both PDP reports. Some of the bore clusters in Table 4-1 may not be suitable for vertical hydraulic gradient analysis due to no or limited record overlapping. To further clarify, using data from cluster 1 in Table 4-1 for hydrogeological analyses and model calibration could cause serious mistakes. The average head in Bore 301011 over the period 1/07/2014 to 6/11/2014 is 5.49 metres above sea level (masl). The average head in Bore 301045 over the period 5/05/2021 to 23/08/2022 is 2.56 masl. Ignoring that these are two different bores, one could conclude that there has been a drop of about 3 m in the groundwater level at that point (defined by model layer, row, and column). Using these data in the groundwater model could result in seriously misleading conclusions that are supported by modelling. Nonetheless, the 2023 model indicates stable high groundwater levels at the location and screen depth of the two bores (Figure D-1 and Figure D-3).

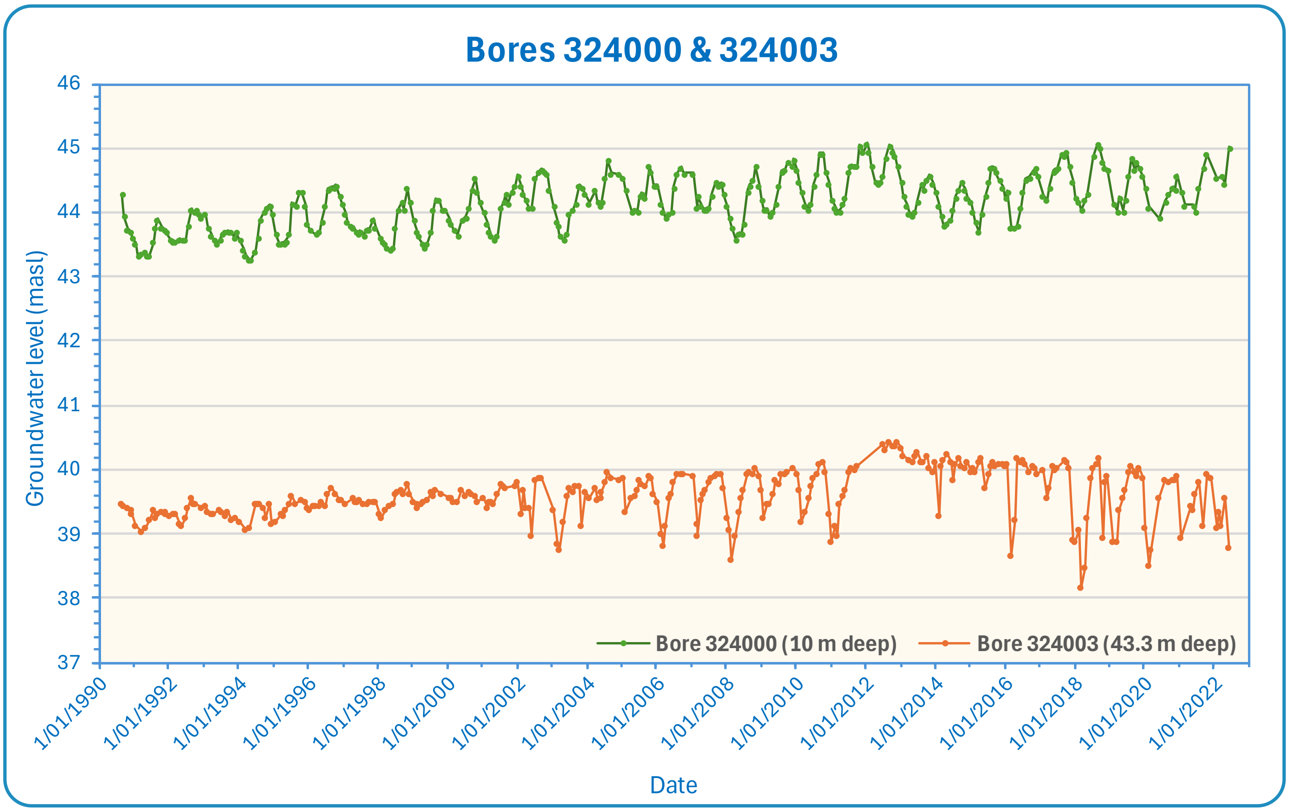


Figure 4-2. Observed groundwater level hydrographs for nearby Bores 324000 (10.0 m deep) and 324003 (43.3 m deep), showing downward hydraulic gradient (i.e., groundwater recharge area setting).

Figure 4-3 presents observed groundwater levels in Bores 313053 (58.5 m deep) and 313013 (93.5 m deep). The data indicate a general downward hydraulic gradient, i.e., recharge area settings. The bores are about 3 km to the south of Lake Alice and 5 km to the west of Bulls, near the beginning of the Waimahora Stream. This information should have been used in the conceptual model to characterise groundwater-surface water interaction and help delineate groundwater recharge areas. The existence of lakes in a groundwater recharge setting area could mean that the lake is perched above the regional groundwater table and/or it is associated with a perched groundwater system. Either way, proper conceptualisation of the situation is required for realistic numerical modelling, which could need to incorporate unsaturated flow simulation.

The models have not utilised river stage (level) elevation or flow data (surface water losses and gains). Horizons concurrent river gauging surveys provide useful data on surface water flow losses and gains. These data have not been used in the models. Flow data would have helped in the conceptualisation and the calibration of the numerical models.

Horizons hold hydrological data for some waterways that could have helped with the groundwater conceptual and numerical modelling, including:

- stage (water level)
- flow
- cross sections (waterway width, bottom elevation, water level)
- concurrent gauging (most of the above and estimates of losses and gains).

The models have not utilised the above data that come from one-off surveys and continuous monitoring. Subsequently, the surface water representation in the models is inadequate. In addition, failure to calibrate the model against surface water levels and flow exchanges with groundwater increases the models' non-uniqueness and, subsequently, uncertainty.

The locations of natural groundwater discharge are discussed in general terms and not quantified. Future work must attempt to gather, analyse, present this data for more robust conceptual and numerical modelling.

From Section 3.5 in PDP (2017) and Section 2.2 in PDP (2023), it is understood that groundwater recharge from rainfall was calculated for both models using a soil moisture balance modelling approach (see Section 3 and Section 4.2.6). The used model is not described and no information on it was made available for this review. However, it is believed to be adequate based on general knowledge of PDP rainfall groundwater recharge calculations approach and because the estimates in the reports are consistent with previous assessments (e.g., Zarour, 2008).

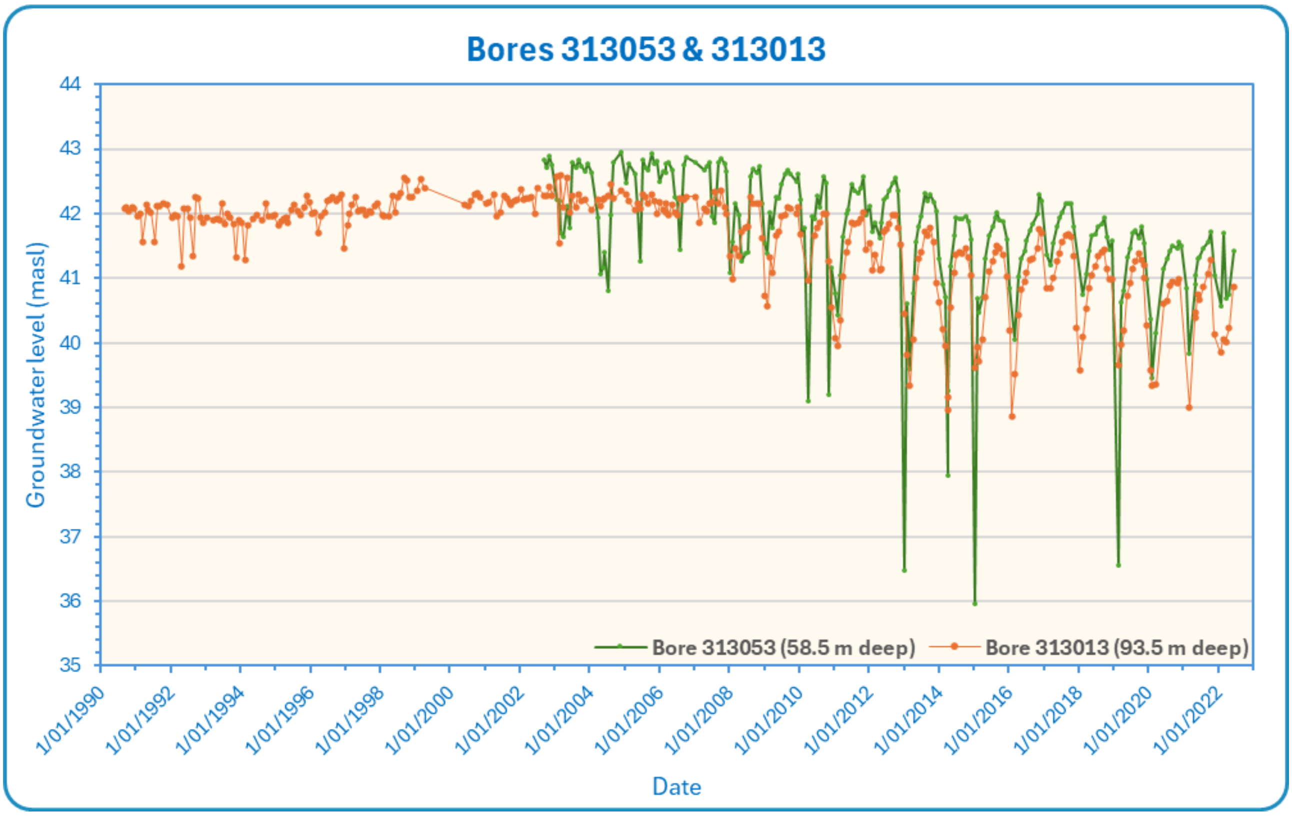


Figure 4-3. Observed groundwater level hydrographs for nearby Bores 313053 (58.5 m deep) and 313013 (93.5 m deep), showing downward hydraulic gradient (i.e., groundwater recharge area setting).

4.2.4 Geology: Has the Geological Framework Been Understood?

Section 2.1 in PDP (2017) describes the topography. Section 3.1 notes that the surface elevation was based on the 1:50,000 scale topographic contours and averaged across each 500 m x 500 m grid cell. PDP (2023) does not discuss the topography. This review is of the opinion that a digital elevation model like that available from Land Information New Zealand (LINZ) would have been a better source of topographic data than the adopted source. Of course, LiDAR data would be a better source of elevation data for the land and surface waterways. The availability of such data across the model domain extent is unknown to the reviewer. However, if LiDAR data are available, a finer model grid would be required to make good use of them.

Section 2.3 in PDP (2017) notes that the surface water drainage network is relatively dense to the north-east of SH3, whereas it is comparatively sparse to the south-west of the highway. It suggests that the surface geology to the north-east of the highway is less permeable, resulting in greater surface water flow and less rainfall infiltration to groundwater (recharge). It also notes that to the north-east of the highway, greater rainfall, higher elevations, and steeper land slopes would encourage greater rainfall runoff and, subsequently, relatively less soil drainage (groundwater recharge).

The landforms analysis provided indicates that where shallow groundwater levels are close to the surface (e.g., within more recent deposits near the coast), the low permeability material implies that some surface drainage must occur. Although there are few mapped surface water courses in the coastal area, it is likely that there are a number of field drains that may not be mapped. Additionally, there are several swamps and small-scale lakes in the area that act as shallow groundwater drainage (discharge) points. No maps of field drains and wetlands are presented, which limits the ability to establish drain boundary conditions (DRN) for the numerical model and assess wetlands' dependence on groundwater, which is expected to be assessed by the numerical models.

The geology discussion in Section 2.2 in PDP (2017) is based on Begg *et al.* (2005) (Figure 4-4). It suggests that the location of registered bores in the area is largely consistent with the occurrence of Units 3 and 4 defined in Begg *et al.* (2005). However, the geology description does not include structural details, such as faults and their role as conduits or barriers to groundwater flow, or any potential influences of folding. The dip of the strata, which could benefit the model's hydrostratigraphic conceptualisation, is also not mentioned.

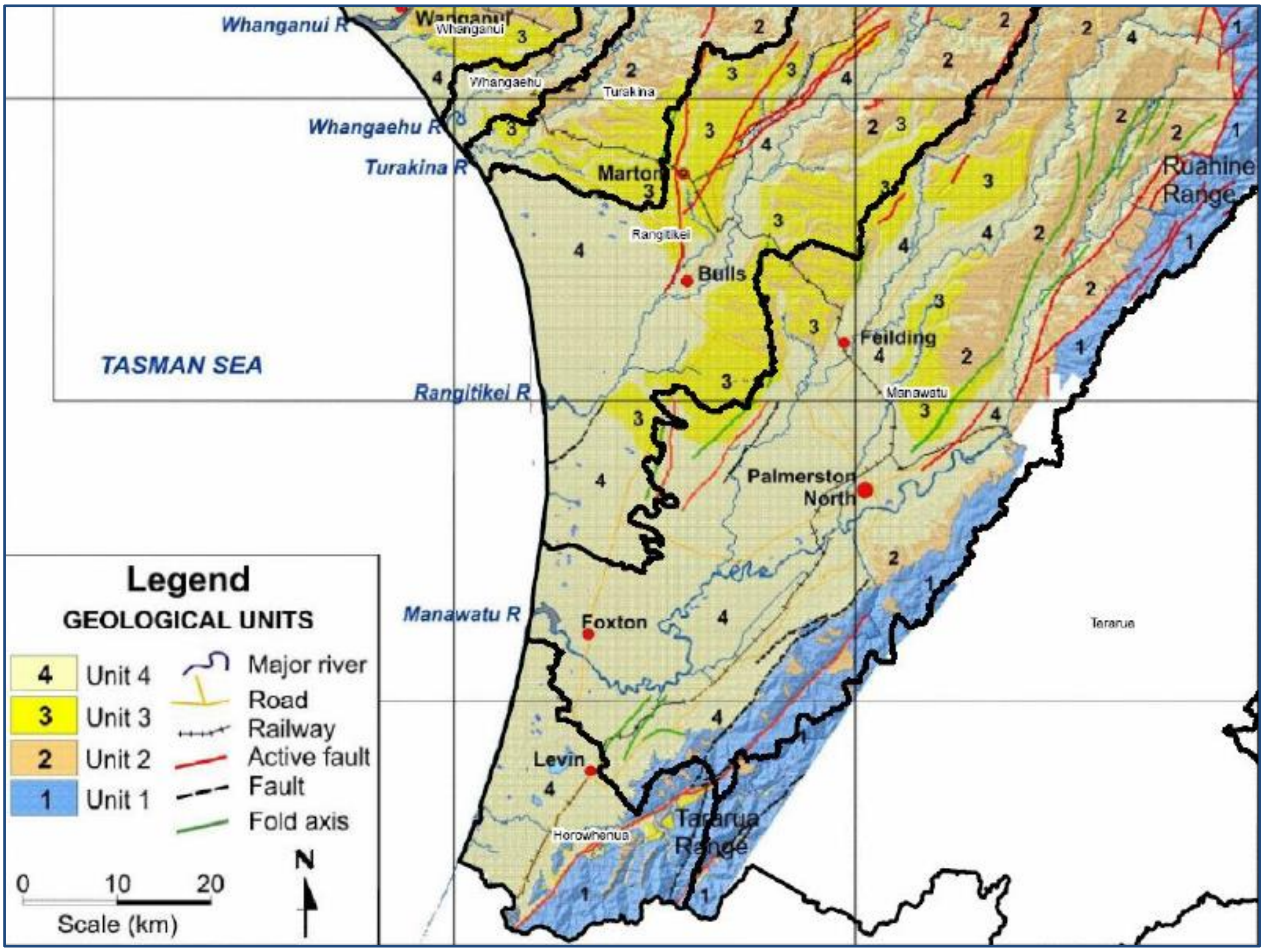


Figure 4-4. Simplified hydrogeology (Begg *et al.,* 2005).

PDP (2017) argues that drillers' borelogs do not allow for the identification of boundaries between different subsurface geological units. It notes that the deepest bore in the area (Horizons ID: 312040) is 573 m deep and located in the coastal plain, implying that the deposits extend to at least that depth. Previous geological work, including oil drilling and seismic surveys, indicates that the Plio-Pleistocene sediments are much thicker. The reports seem to be unaware of the Santoft 1A Well drilled in 1964/1965 near the coast and other oil exploration wells in the area (Zarour, 2018). Fig. 2 in Melhuish *et al.* (1996) reproduced below as Figure 4-5 implies that the Santoft 1A Well encountered Torlesse greywacke basement at a depth of more than 2,500 m. Examination of the 2023 model files reveals that the elevation of the bottom of the deepest model layer (Layer 11) along the coastline is 305–320 m below sea level, i.e., somewhere in the Castlecliffian sequence. The reports do not explain the rationale for the assigned model bottom elevation.

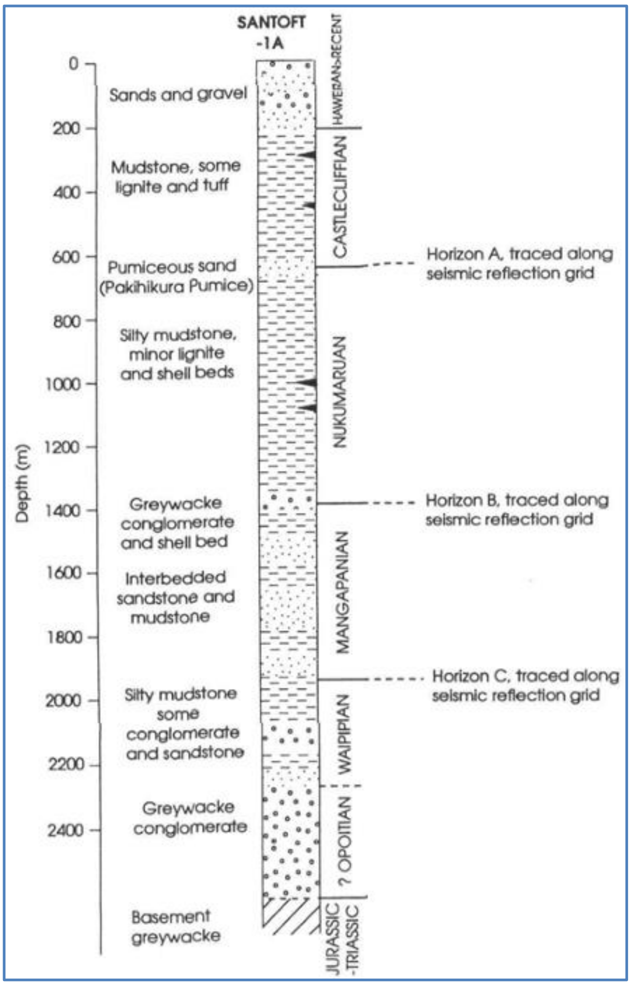


Figure 4-5. Summarised well log for Santoft-1A (extracted from Fig. 2 in Melhuish *et al.*, 1996).

Stratigraphy is the study of rock and sediment layers. Hydrostratigraphy specifically focuses on the water-bearing properties of these strata, in terms of their ability to store and transmit groundwater. Groundwater models are expected to reasonably represent hydrostratigraphy. Layers in a numerical groundwater model do not have to completely align with hydrostratigraphic units. A model layer can encompass more than one hydrostratigraphic unit, and a hydrostratigraphic unit can be divided over multiple model layers, depending on the modelling objectives and constraints.

Notwithstanding, the RGMZ conceptual and numerical models are not informed by geology. The RGMZ numerical models adopt arbitrary layering that does not reflect the hydrostratigraphy of the system. This is considered a notable shortcoming of these models. This was a major deficiency highlighted in the 2017 model review (Barnett, 2016), which was not addressed in the 2023 model update. PDP (2023) does not explain why that review point was not addressed.

It is understandable that the data, especially drillers' borelogs, may not be ideal for constructing a 3D geological model for the area. However, a conceptual geological model could have been included to provide a framework for groundwater modelling. An example of such a model is provided in Figure 4-6.

Overall, it is not clear how the conceptual and numerical models benefited from the provided geology discussion in the reports.

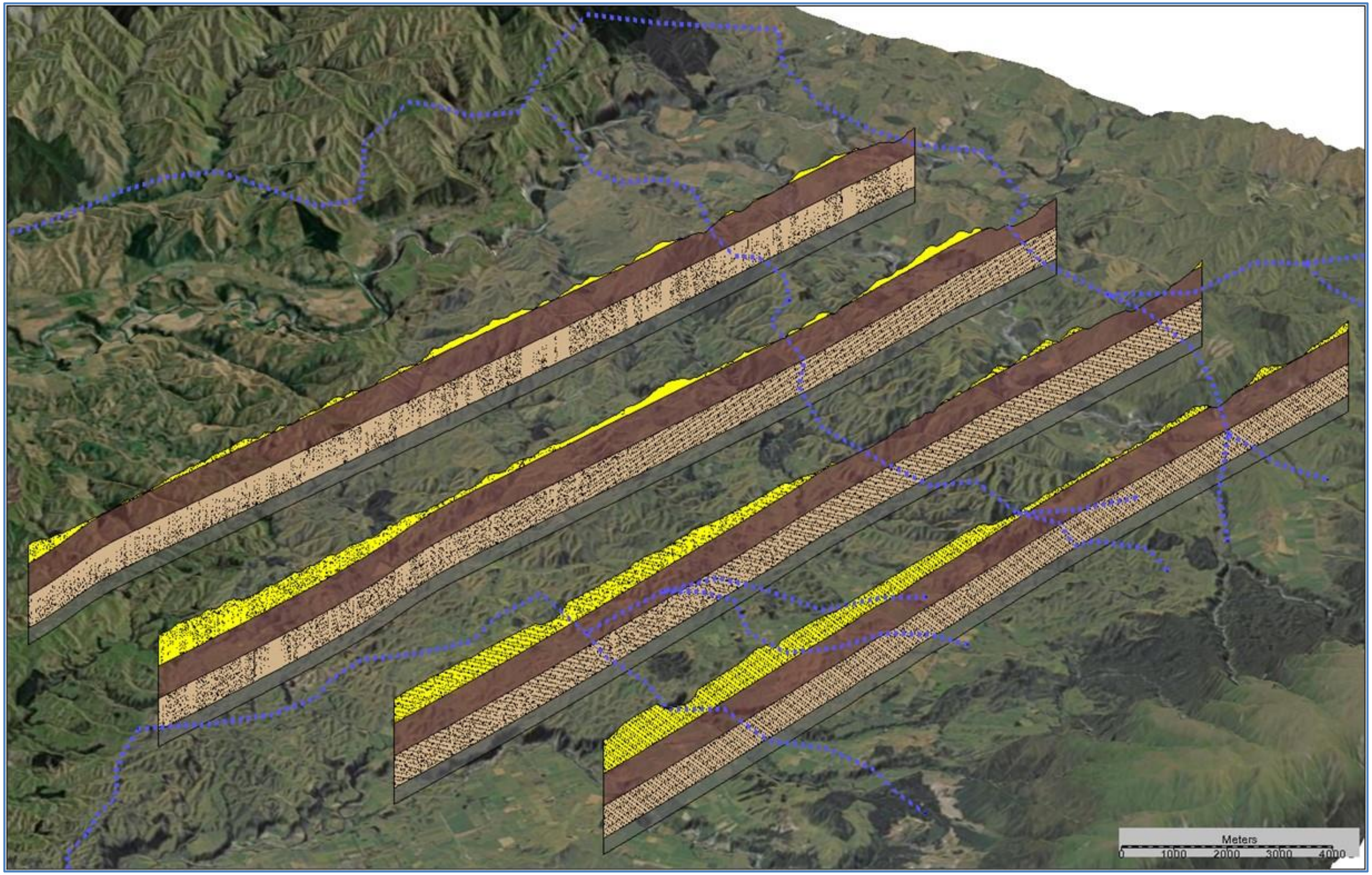


Figure 4-6. Cross sections east of the Rauoterangi Fault cut through a 3D geological block model – 2x vertically exaggerated (from Scaife & Zarour, 2023).

4.2.5 Hydrogeology: Has the Aquifer System Been Described Adequately?

As discussed above, the hydrostratigraphy, including aquifer geometries (extensions and thicknesses) and types (perched, unconfined, semi-confined/leaky, and confined), is not adequately presented in the conceptual sections of the reports. There are also no comments on preferential flow paths, or the lack thereof. However, the reports provide a reasonable understanding of the hydrogeology. The main shortcoming in the hydrogeological characterisation in the reviewed work is the lack of a hydrostratigraphic framework for the analyses and modelling.

4.2.6 Have Data on Groundwater Stresses Been Collected and Analysed?

PDP (2017) and PDP (2023) hypothesise that the main stresses that affect the RGMZ hydrogeological system are groundwater recharge from rainfall (inflow) and abstraction (outflow). Both reports attempt to establish cause-effect relationships between stresses and impacts, mainly in terms of observed groundwater level changes as flow data have not been incorporated in the analysis.

Section 2.5 in PDP (2017) considers temporal groundwater level variations. It suggests that there are two notable temporal signatures in the groundwater level records:

1. a slow, long-term rising trend in some bores (304019, 303017, 312009 and 302003), followed by a declining trend in those bores located towards the coast (312009 and 302003). The rising trend continues in the two bores located further inland (304019 and 303017).
2. relatively small seasonal variations in groundwater levels in most of the bores, particularly bores 312009 and 302003.

The analysis seems to have been undertaken qualitatively (visually). It should have been undertaken qualitatively, for example using trend analysis statistical methods and groundwater level hydrograph decomposition methodology to separate the different components of groundwater level timeseries. This would have provided better insights into the underlying processes and their influences on groundwater levels.

PDP (2017) suggests that the rising trend in groundwater levels observed in some bores around Santoft prior to 2005 is most likely due to increased recharge compared to a previous steady state. It attributes the increased recharge to climatic variations (increases in rainfall and/or decreases in evapotranspiration) and/or changes in overlying landuse resulting in enhanced rainfall infiltration. This review is of the opinion that the landuse change is an unlikely reason for change in groundwater recharge as there has not been a large-scale disposal of water to land in the area prior to 2005. Other landuse changes will not result in a large-scale long-term groundwater level rise.

PDP (2017) use cumulative deviation of annual rainfall from the long-term mean (CDFM) to assess rainfall trends (Figure 4-7). The report suggests that the CDFM analysis indicates a 'stable pattern' from around 1935 to 1980, a 'declining pattern' between 1980 and 1989, and 'rising pattern' from 1990 to around 2011. The report argues that it is difficult to confirm the above hypothesis because the groundwater level records start from 1990. It also notes that the CDFM declining trend from around 2011 was not apparent in groundwater level data at the time of the report preparation. The report highlights that the noticed declining groundwater levels trend from 2005 cannot be related to the decline in rainfall from 2011.

PDP (2017) notes that Feilding, in the Manawatū Groundwater Management Zone (MGMZ) to the south of the RGMZ, is outside the model area, but its rainfall data plotted in Figure 4-7 from 1970 are similar to those observed at Marton, implying that the older data (from 1920 to 1970) are likely to represent the long-term rainfall pattern in the model area. The wet/dry period analysis is generally consistent with the analysis by Zarour (2018) of MGMZ rainfall data (Figure 4-8). Zarour (2018) notes that 2014 was the driest year on the record he analysed (1985–2015, inclusive). It followed three years of less than average rainfall (2010, 2012, and 2013). The multi-year low rainfall resulted in a drop in the groundwater levels in most monitored bores in the MGMZ, some reaching their lowest levels ever. Still, other groundwater levels in some other monitored bores in the MGMZ were at lower levels in the 1990s. It is noted that the impacts of the historic February 2005 flood in the Manawatū on groundwater levels in the MGMZ are not clear, or very short-lived to be noticed through monthly monitoring. This suggests that long-term climatic trends and cumulative effects could be more influential on the hydrogeological system than short-term events, even if those events were exceptional.

Zarour (2018) calculated rainfall deviation against the cumulative mean, while PDP (2017) apparently calculated it against the simple arithmetic mean of the data from the start to the end of the record. Although both methods led to similar conclusions, the use of the cumulative mean is preferred. Using the cumulative mean ensures that data collected at a certain time are not compared with data from the future. For instance, rainfall in 2010 should be compared with the mean rainfall up to 2010, not against a long-term mean extending into the future, say to 2022.

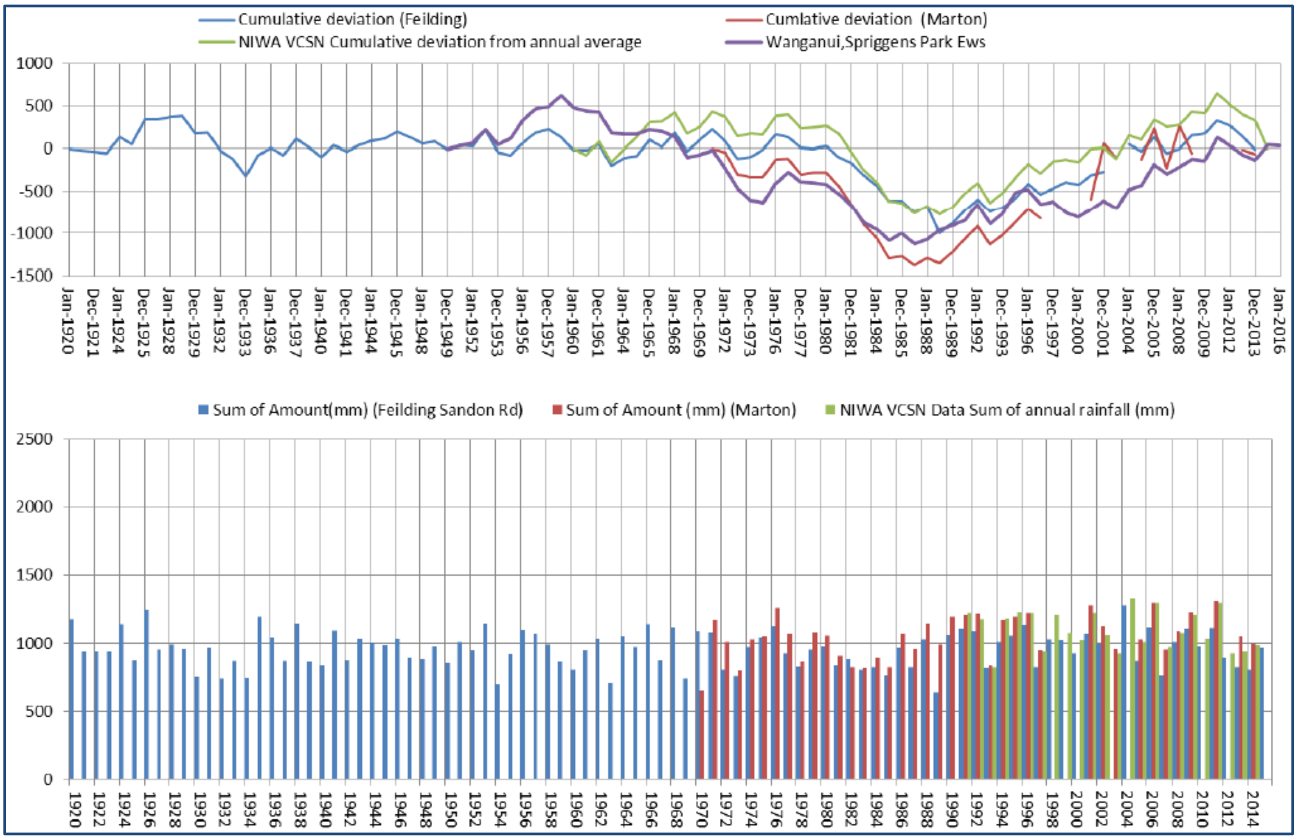


Figure 4-7. Long-term plots showing the cumulative deviation from the long-term mean annual rainfall at selected sites (Figure 9 in PDP, 2017).

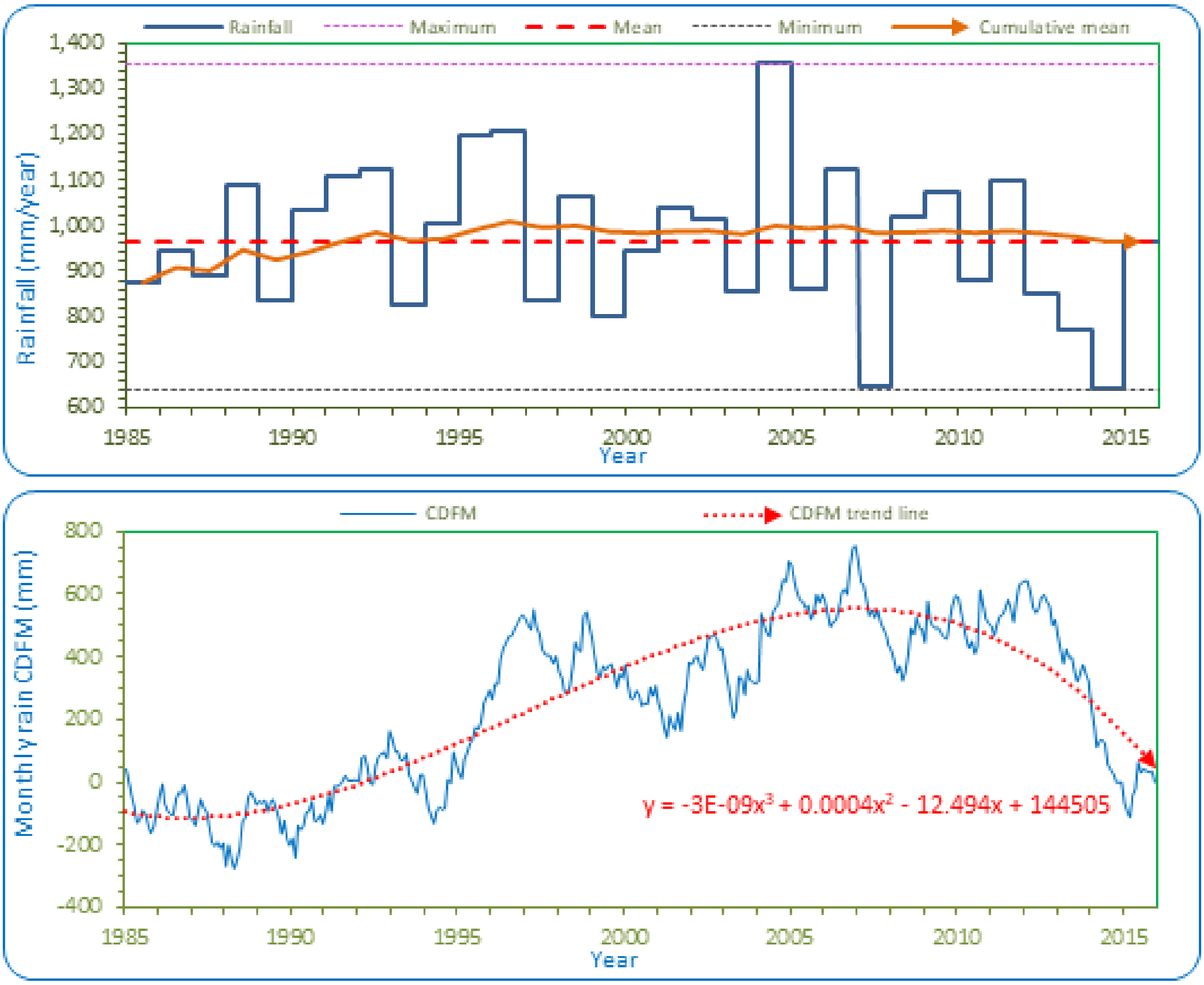


Figure 4-8. Monthly rain cumulative deviation from mean (CDFM) at Palmerston North for the period 1985–2015 (from Zarour, 2018).

Section 2.5.2 in PDP (2017) discusses seasonal trends, noting the greatest seasonal variations occur in Bore 312007 (Figure D-11) and Bore 323015 (Figure D-21) located closest to the coast, with changes of nearly 5 m and 6 m, respectively. The report notes that prior to the onset of seasonal pumping in the area, seasonal changes in these bores were up to 0.5 m. The reports should have recommended further investigations on these two bores, including pumping tests with monitoring bores to check on the storativity around these bores. Barnett (2016) indicated that storativity could be an important factor in shaping the groundwater level response to stresses in the model area. Similar tests should also be undertaken on bores that show the smallest fluctuations.

Conversely, the least seasonal variations occur in Bore 312009 (Figure D-12). This review notices a small increase in seasonal variability in Bore 312009 from around 2009. PDP (2017) hypothesises that the subdued seasonal groundwater level fluctuation in Bore 312009 could be due to large storage capacity in the shallow strata and/or very slow vertical infiltration, consistent with the silt, mud, and clay lithologies observed in drillers' logs. However, this review finds that this hypothesis contradicts the report's other hypothesis that seasonal changes are related to pumping. Additionally, the presence of low hydraulic conductivity material would result in confinement or semi-confinement of the deeper strata and/or perching of the shallow groundwater. The reviewed conceptual and numerical models do not reflect such conditions.

Section 2.5.2 in PDP (2017) concludes that '*…groundwater level responses in the Santoft area are related to rainfall recharge, albeit with a subdued response to seasonal rainfall variations. Larger seasonal variations in the latter part of the record are related to seasonal variation in groundwater abstraction. Therefore, it is considered appropriate to simulate the groundwater system based on a recharge model derived from observed rainfall and evapotranspiration.*' The reviewer understands this to mean that the model must include rainfall recharge and groundwater abstraction. However, it is noted that the modelling reports do not present the used groundwater recharge and abstraction data in their main text or appendices.

Figure 4-9 presents the steady state rainfall recharge data extracted from the 2023 model files. The steady state period in the model corresponds to 1 July 1989 to 1 July 1990. This dataset was taken as a sample of recharge data in the model. Recharge data for the transient 384 stress periods were not checked due to time restrictions.

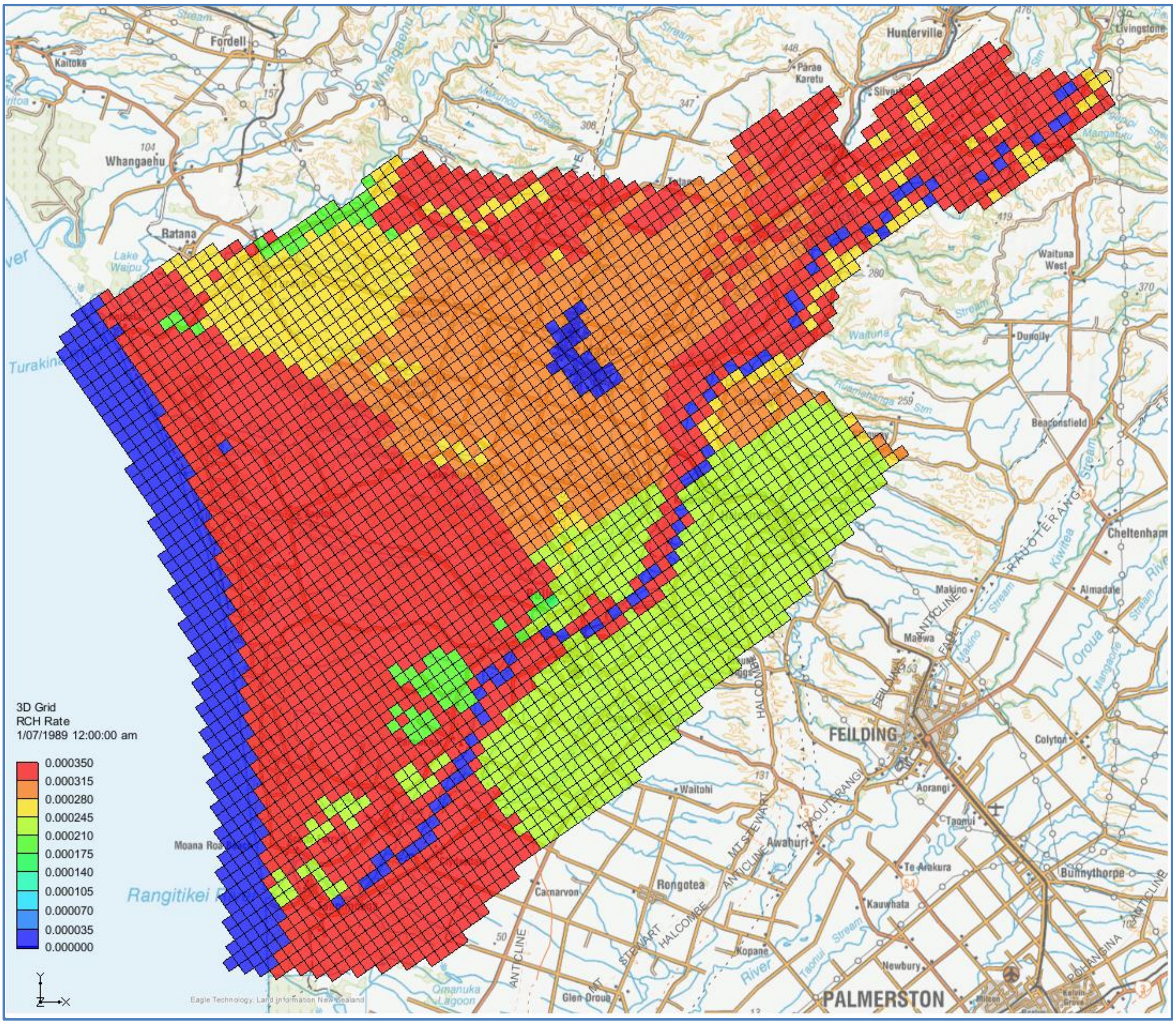


Figure 4-9. Steady state recharge as assigned to the steady state stress period (1 July 1989–1 July 1990) in the 2023 model.

Note that recharge from irrigation return flow is not considered in the conceptual and numerical models. Some irrigation water comes from surface water. Irrigation returns recharge must be considered in future models and, if needed, included in them.

The steady state rainfall recharge assigned to the model varies across its domain from 0.00017411 m/d to 0.00034895 m/d, equivalent to 64 mm/y and 127 mm/y, respectively. This range is plausible. However, the map in Figure 4-9 shows relatively high recharge in cells also designated as drain (DRN) cells. According to Section 3.1 in PDP (2017), these cells are primarily located in the coastal area, where low vertical permeabilities and neutral or upward vertical hydraulic gradients imply limited infiltration to deeper strata, i.e., no recharge. No recharge is assigned in Marton Township cells, Lake Koitiata, and some cells along the Rangitīkei River. Otherwise, relatively high recharge is assigned across the model domain. This contradicts the conceptualisation presented in PDP (2017) and PDP (2023) and is contrary to hydrogeological principles, as no rainfall groundwater recharge occurs on surface water features. Misimplementation of recharge in numerical groundwater models affects the reliability of their calibration since recharge is strongly correlated with hydraulic conductivity. Overestimation of recharge is balanced by an increase in the model hydraulic conductivity by the model during calibration to be able to match the observed heads. Hence, misimplementation of recharge eventually impacts the models' water budgets and the reliability of predictions.

Section 2.1 in PDP (2023) attempts to relate changes in groundwater levels and electrical conductivity (EC)[7] to abstraction. It further tries to explore inter-aquifer relationships (vertical flows through the system in the coastal area). However, some of the opinions offered in the report are not supported by the data presented. For example, the report claims that Figure 4-10 shows an inverse relationship between abstraction and EC. While the assumption is generally reasonable, it is not evident in the figure. There seems to be a problem with the monitoring instrumentation in the bore. Figure 4-10 suggests that in autumn and winter 2018, there is no pumping, the groundwater level is recovered, but there are EC spikes. The figure further shows that in the first part of 2019, groundwater level assumed to be drawn down by abstraction coincides with low EC. These observations contradict with the report's hypothesis that lower groundwater levels are associated with higher EC. There are additional issues with the monitoring record, including missing groundwater level data in 2021. Hence, no conclusions should be drawn from Figure 4-10 datasets until they are validated.

This review agrees with the reports that groundwater recharge from rainfall and abstraction from bores are major stresses on the RGMZ hydrogeological system. However, it recommends consideration of additional stresses including surface water-groundwater interaction (natural waterways like rivers, streams, wetlands, as well as artificial features like drains). Pre-numerical modelling assessment should incorporate further analytical assessments, including zone-wide and local-scale water budget analyses. For example, analytical calculations should be performed on groundwater discharge into the Tasman Sea and compared to the numerical model assessment. Baseflow assessment for selected surface waterways should also be undertaken and compared to the numerical model results.

The reports do not consider evapotranspiration from the water table to be a stress on the groundwater system, which could be the case. However, future assessments must investigate this potential stress, especially in shallow groundwater parts in the GMZ. Stresses that must also be investigated include baseflow, river leakage, and irrigation returns.

[7] Electrical conductivity (EC) is a proxy measure of salinity. The greater the water salinity the greater its EC.

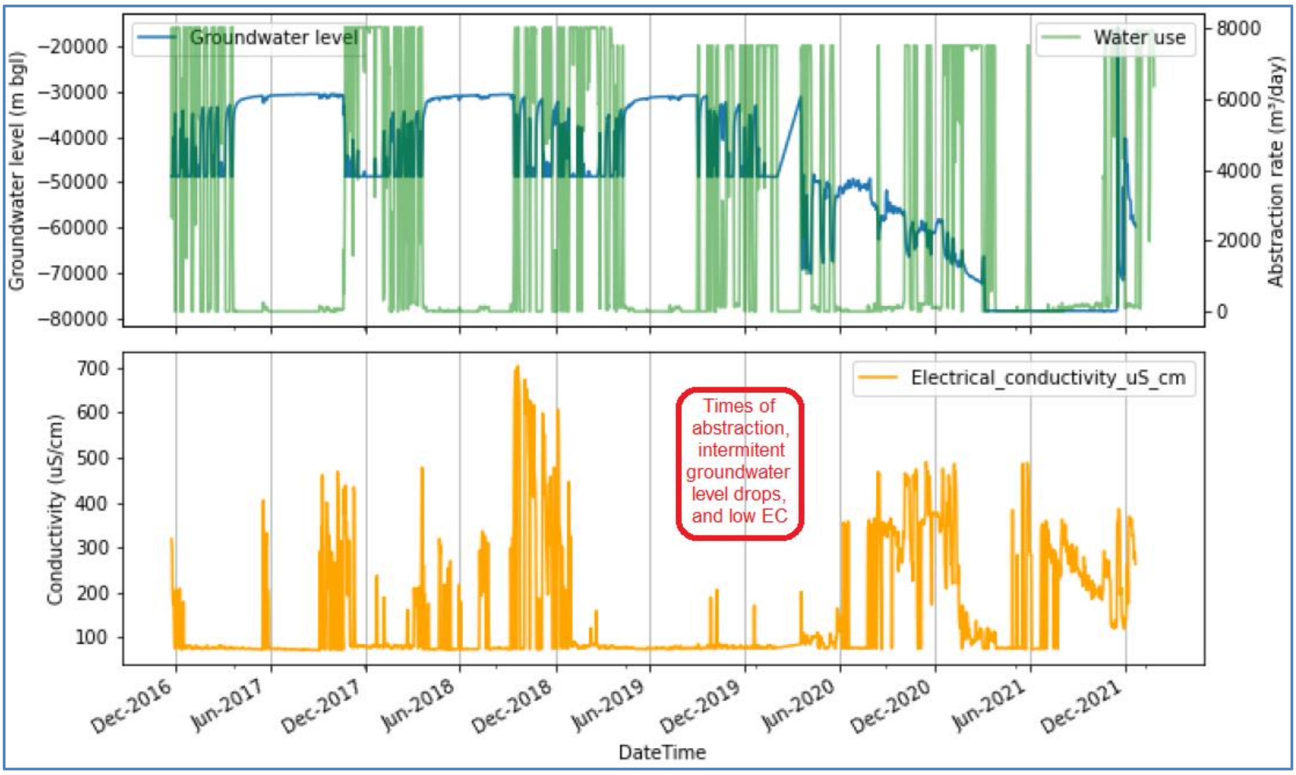


Figure 4-10. Groundwater levels, electrical conductivity (EC), and water use data for the Hyde Park Farms consent (ATH-2006011251). Figure 3 in PDP (2023). Comment in red added as part of the review.

4.2.7 Have the Responses to Stresses Been Understood?

The responses to stresses have not been fully understood, including the horizontal and vertical flow directions and gradients. Further investigations could benefit the system's understanding and future modelling. Revised modelling should target the understanding of stresses-responses relationships. For example, particle tracking and capture zone analysis using particle tracking must be included in future modelling to provide basis for cause-effect analyses.

4.2.8 Data Uncertainty

Sections 2.3 and 3.1 in PDP (2023) notes that the numerical model was not calibrated to surface water flows (drains, streams, lakes, wetlands, swamps, and rivers) due to the absence of data for the small features and difficulties in defining baseflow in large rivers. Therefore, it expects model predictions of the effects of abstraction on surface waterways to be subject to a wider range of uncertainty compared to the predictions of groundwater levels to which the model is calibrated. However, this implies that the relative effect of abstraction on surface waterways in shallow groundwater levels are within the predictive capacity of the model. This type of data-related uncertainty should have been reduced through the validation (comparison) of the numerical model against analytical estimates of flow exchange and water balances, which this review recommends for future model development. Using available and additional concurrent flow gauging will help reducing groundwater-surface water interaction uncertainties.

Section 3.1 in PDP (2017) notes that river stage elevations were unknown, contributing to model uncertainty. This review assumes that generally there were no data on surface water stage or bottom elevations. Missing data for the model also include bed conductance (C) parameters for surface water features and General Head Boundary (GHB) conditions, which are discussed below in Section 4.3.3.

Section 3.1 in PDP (2017) notes that hydraulic conductivity values are subject to uncertainty due to the lack of knowledge on aquifer thickness. Further discussion on hydraulic conductivity is presented in Section 4.4.4.

The reports do not discuss uncertainty in groundwater recharge estimates which are related to data types (landuse, climate, soil, etc.), coverage (number and location of observation sites), quality, measurement accuracy, and record length.

Uncertainty related to subsurface geology is discussed. However, there are geological data (e.g., structure) that have not been considered in the hydrogeological analyses and the groundwater models.

Data measurement uncertainty is not explicitly noted in the reports. Some elements of this uncertainty are inherent and cannot be resolved. They could even be irrelevant to regional scale models. Some other elements relate to unavailability of data, which the reports should have made recommendations to collect through surveys, monitoring, and assessments.

It must be noted that there are no dedicated groundwater monitoring bores in the RGMZ. The monitored bores are operational bores, pumped for various reasons. Hence, there is inherent uncertainty in groundwater level observations mainly due to (a) groundwater levels may not have been fully recovered to static water level (SWL) conditions at the time of measurement, (b) the water level at the measured bore may be impacted by pumping from neighbouring bores, and (c) bore efficiency may impact on the measurements.

Future modelling reports must comprehensively cover data uncertainty. For each data type, future reports must describe the sources and importance of associated uncertainties, how the uncertainties have been dealt with, and recommend actions to reduce them if necessary.

4.2.9 Conceptual Model Certainty

Hydrogeological conceptualisation uncertainties can arise from any of the above discussed conceptual model matters. Uncertainty cannot be eliminated altogether; it can be reduced and managed through well-planned data collection, sensitivity analysis, and uncertainty quantification. It must be considered in the numerical model implementation and subsequent decision-making.

PDP (2017) and PDP (2023) do not discuss the conceptual model uncertainty. There is no evidence in PDP (2017) and PDP (2023) on testing their common conceptual model through, for example, simple water balance or exploratory modelling. There is also no indication to investigating alternative conceptual models and there is no discussion of uncertainty in the conceptual model.

The main conceptual uncertainties identified in this review are related to the lack of a hydrostratigraphic framework and the definition of the lateral model extent, which does not coincide with the RGMZ boundary.

4.2.10 Reporting and Presentation

Noticeably, the conceptual model used in the 2017 and 2023 models is not presented in graphical format (e.g., spatial maps, cross-sections, schematic geological and hydrogeological diagrams). The conceptual model sections in the reviewed modelling reports do not present water budgets to explain, support, and validate the numerical model. While the provided conceptual model description is useful, it could have been better communicated. Future modelling reports should represent the conceptual model in line with best practice recommendations.

4.3 Design and Construction

4.3.1 Initial Modelling Considerations

The numerical model designs align with the conceptual model described in the reports. However, the reports do not discuss the choice of modelling method, seemingly assuming that a transient, 3D distributed numerical model is the obvious choice for representing the groundwater system and addressing the research questions.

The rationale for selecting numerical methods (finite difference in the 2017 model and finite volume in the 2023 model) is not provided, though both methods are well-suited to the groundwater system and modelling objectives. Similarly, the selection of modelling and calibration software is not justified. However, the software packages used (MODFLOW-NWT in the 2017 model, MODFLOW 6 in the 2023 model, and PEST in both models) are appropriate for the system and objectives. The transition to MODFLOW 6 allows for the use of an unstructured grid in the new model, though this option was not utilised. This decision, while not explained, facilitates direct comparison between the results of both models, though the recalibration—resulting in different parameters such as hydraulic properties and recharge—complicates this comparison. This review assumes that the recalibration is simply due to the availability of additional data for the new model, which is a valid reason.

4.3.2 Are the Spatial and Temporal Domain and Discretisation Appropriate?

Groundwater modelling requires defining the modelling volume and time, which represent the spatial and temporal domains, respectively. As part of the simplification inherent in all models, the model domains are discretised—the spatial domain is divided into units like cells, while the temporal domain is divided into 'stress periods' (intervals) that collectively represent the total modelled period.

Conventionally, the spatial domain of a groundwater model is discretised into layers, rows, and columns using the finite difference approach (e.g., MODFLOW-NWT). Alternatives such as the finite element approach (e.g., FEFLOW) and the finite volume method (e.g., MODFLOW 6) offer more flexibility in model discretisation as they do not require structured grids.

With regards to time, groundwater models are categorised as either ‘steady state’ or ‘transient’:

- **Steady State Modelling**: Assumes that the groundwater system is in equilibrium, meaning conditions such as hydraulic head and flow do not change over time. These models are used for long-term average conditions and do not account for temporal variations.
- **Transient Modelling**: Accounts for changes in the groundwater system over time, capturing the dynamic nature of variables like hydraulic head and flow. Transient models use time-stepping (through stress periods) to simulate how the system responds to varying stresses, such as seasonal changes in recharge or pumping.

In a transient model, external stresses (e.g., pumping rates, recharge rates, boundary conditions) are assumed to be constant within each stress period. Thus, a transient model can effectively be considered a series of consecutive steady state model runs. Commonly, transient models start with a steady state model to provide initial conditions and other useful information.

Stress periods are divided into intervals called time steps. Time steps improve computational accuracy and allow the model to capture the transient (time-varying) behaviour of the groundwater system within a stress period. Finer time steps result in more detailed simulations but increase computational demands (e.g., model run time). Selecting appropriate time steps is crucial for balancing accuracy and computational efficiency.

PDP (2023) does not elaborate on the spatial and temporal model setups. It notes that the 2023 model design is the same as the 2017 model. It is clear the 2023 model extends seven years longer.

Sections 3.2 through 3.4 in PDP (2017) describe the spatial and temporal discretisations for the 2017 model.

According to Section 3.2, the model domain is divided into 11 layers to represent vertical hydraulic gradients (Figure 4-11). The top layer (Layer 1) has a variable thickness with a minimum of 15 m, while Layers 2–4 each have a thickness of 10 m. Layers 5 through 11 each have a thickness of 40 m. Hence, the total model thickness exceeds 300 m, which PDP (2017) argues is consistent with the depth of bores drilled for abstraction in the region. This review is of the opinion that the model total thickness and layer structure should have been based on hydrostratigraphy, aiming at representing the full thickness of all the units exploited in the RGMZ (aquifer units) and the layers between them (aquitard units).

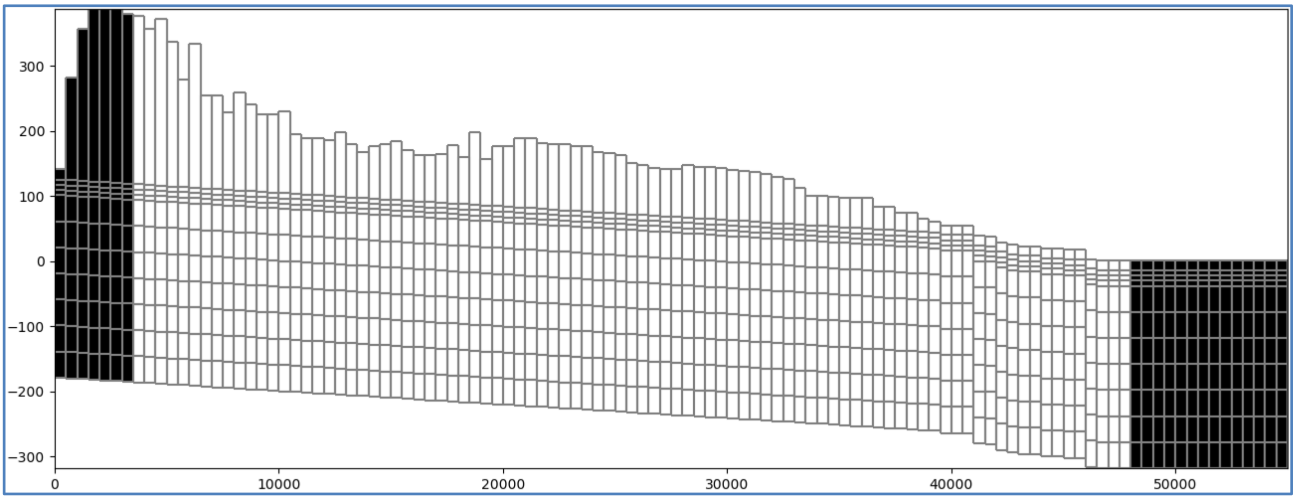


Figure 4-11. Model layers through cross section along column 30 (Figure 12 in PDP, 2017).

PDP (2017) indicates that the model layering scheme was designed to reflect the confining effects of low permeability strata in the model area. Layer 1 was variable thickness was intended to prevent it from drying out during simulations and to represent the significant vertical gradients observed inland. By increasing the thickness of Layer 1 inland, the model provided unconfined storage to buffer the effects of seasonal rainfall recharge on deeper layers. This approach ensured that recharge was continuously applied to Layer 1, while the thinner Layers 2–4 delayed recharge to deeper layers. It is worth noting that both MODFLOW-NWT and MODFLOW 6 offer features for handling cell drying, which the RGMZ models apparently did not utilise.

Although PDP (2017) offers a reasonable hydrogeological rationale for the model’s layering, this configuration is somewhat arbitrary and does not fully represent the hydrostratigraphy or the distribution of hydraulic characteristics in the system. This affects both the model’s representativeness and its predictive capabilities.

According to PDP (2017), the model domain is discretised into 500 m x 500 m cells arranged in rows and columns. The model grid is rotated 55° from north to align with the general direction of groundwater flow. Grid cells outside the model domain are inactivated, resulting in an active model area of approximately 957 km². PDP (2017) notes that while finer grid sizes can improve model accuracy, a 500 m grid is deemed appropriate for the model's aims due to considerations of model run times and data requirements. This review has not tested a finer model grid, but it is noted that MODFLOW-NWT and MODFLOW 6 offer options for grid refinement and unstructured grids, which could enhance the model's representativeness and outputs.

Both the 2017 and 2023 models are transient simulations. Examination of the 2023 model files reveals that the 2017 and 2023 transient simulation comprised 301 and 385 stress periods, respectively. Each model starts with one year of steady state simulations (1 July 1989–1 July 1990), which is not mentioned or comment on the reports. The subsequent stress periods in both models are one month long each. The 2017 model simulation end on 1 July 2015 and the 2023 model simulation ends on 1 July 2023. Each transient simulation stress period in both models comprises only one time step. The models could have benefited from additional time steps, but at a cost mainly in terms of run time. In this review, the effects of smaller time steps on the 2023 model run time was tested. It was found that a model with one time step per stress period takes about 3 minutes and 15 seconds to run. Running the model with daily time steps takes about 1 hour, 17 minutes and 20 seconds, i.e., it is about 25 times longer.

Section 3.4 (Temporal settings) in PDP (2017) notes that the use of the MODFLOW-NWT code helped limit issues around drying and rewetting cells in the model area, particularly towards the north-east of the model where the Rangitīkei River enters the model area and topographical gradients are steep. This note should have been included in the discussion on code selection rather than in Section 3.4.

Overall, the model spatial coverage and temporal extent and discretisations are reasonable and clearly defined. The effects of the model extents and discretisations on model run times is not reported. However, more detailed spatial and/or temporal discretisation would proportionally increase the model run time. This review is of the opinion that grid refinement would have enhanced the model representativeness of hydraulic conditions and stresses as well as helping with boundary conditions assignment. The model could have been run with one time step per stress period in the model calibration phase, but with smaller time steps for reporting.

4.3.3 Are the Chosen Boundary Conditions Appropriate?

As discussed above, the peripheral model boundaries do not coincide with the RGMZ boundary as defined in the One Plan. This review considers this to be a conceptual problem in the 2017 and 2023 models. Firstly, it does not enable answering questions on the target GMZ as it only covers parts of it and extends into other GMZs. Secondly, the peripheral model boundaries breach hydrogeological principles. For example, they cross surface and subsurface catchments. Where possible, peripheral model boundaries must be aligned to natural boundaries, e.g., contact between water-bearing geological material and basement rock, rivers, the coastline, etc. Artificial boundaries can be used if necessary and there are ways to represent them, e.g., GHB conditions.

The model domain is set out about 2 km into the Tasman Sea (Figure 1-3). According to Section 3.1 in PDP (2017), this is to allows for modelled subsurface discharge to occur offshore. This review is of the opinion that this is unnecessary. The boundary should have been set at the coastline at least in Layer 1. There are various options for representing the coastal end of the other layers.

The sea is suitably represented by GHB boundary conditions (Figure 4-12). Section 3.1 in PDP (2017) notes that the coastal GHB boundary conditions were set only in Layers 1–3 and no boundary conditions were set in Layers 4–11 at the coast. However, examination of the 2023 model files reveal that GHB conditions are applied to Layers 1–6 and no boundary conditions are designated in Layers 7–11. Where no boundary conditions are assigned, the peripheral cells default to no-flow boundary conditions. The report notes that in effect, this modelling decision implies that deeper strata are modelled as blind, and any discharge must occur via upwards leakage into shallower layers. It argues that is in keeping with the above surface artesian groundwater pressures in deeper bores at the coast. This review notes that seawater is denser than groundwater. As a result, there is an interface between groundwater and seawater, which pushes flow upwards in the coastal area. This is known as the Ghyben-Herzberg relationship (Figure 4-13). The coastal model boundary should have been set to represent the Ghyben-Herzberg relationship rather than being represented by six layers with GHB cells underlain by five layers with no-flow boundaries.

Examination of the 2023 model files reveal that there are 232 GHB cells in each of Layers 1–6. The report notes that the head assigned to them is 0.1 masl, which is confirmed from the examination of the 2023 model files. The report explains this as to account for the different density of sea water. This review notes that at the coast, the head increases with depth, which the model does not represent.

PDP (2017) notes that the GHB conductance (C) was varied during the calibration process, but it does not report the values. Examination of the 2023 model files reveals that GHB c values are constant in each layer but vary between the layers. It is 28.08 m²/d for Layers 1–3, 72.62 m²/d for Layer 4, 472.18 m²/d for Layer 5, and 359.20 m²/d for Layer 6. The report does not provide this information. The values seem to be arbitrary and not representative of hydraulic properties. The impacts of the designation of the model coastal boundary cells to GHB and no-flow conditions and the assigned GHB C values on the model performance have not been tested. The coastal boundary representation contributes to the model uncertainty.

PDP (2017) describes the locations of the remaining model peripheral boundaries but does not clarify their conditions or which model layers they apply to. The model boundaries are reported to have been delineated as follows:

- The northern boundary largely follows the contact between mudstone/sandstone and the alluvial terrace
- The north-eastern boundary follows the contact between mudstone/sandstone and the Rangitīkei River alluvial gravels
- The south-eastern boundary extends beyond the line of the Rangitīkei River to the south and follows a groundwater flow path defined from groundwater contours in Figure 6 in the report (Figure 4-14), around 5 km away from the Rangitīkei River. The report notes that it is not clear whether deeper groundwater movement follows the orientation of the Rangitīkei River hence the model boundary was extended in this area.

Examination of the numerical model data indicates that the peripheral boundaries to the north, north-east, and south-east were not designated in all layers, i.e., they default to no-flow conditions. However, some of the cells in Layer 1 are designated boundary conditions to represent specific features, like major rivers assigned RIV boundary conditions (Figure 4-15) and Drain (DRN) boundary conditions (Figure 4-16) to represent surface water discharge in some parts of the model close to the coastline (Section 3.1 in PDP, 2017). Designation of peripheral cells as RIV or DRN cells override the default no-flow boundary conditions. Note that the offshore assignment of DRN cells shown in Figure 4-16 is inappropriate.

Figure 4-12. General Head Boundary (GHB) cells in Layers 1–6 in the model (brown triangles).

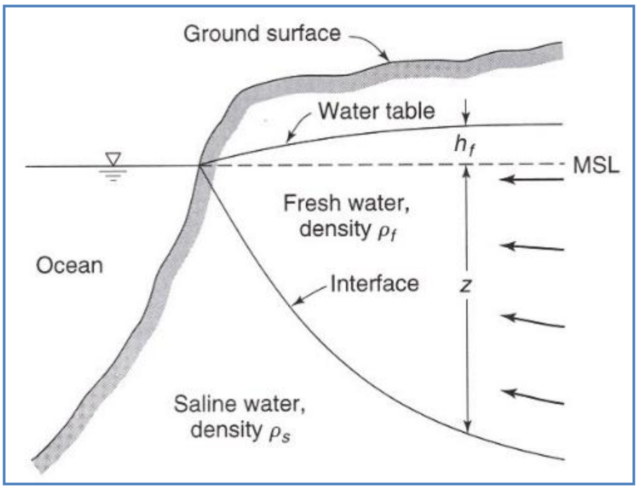


Figure 4-13. Idealised sketch of occurrence of fresh and saline groundwater in an unconfined coastal aquifer (after Todd and Mays, 2005).

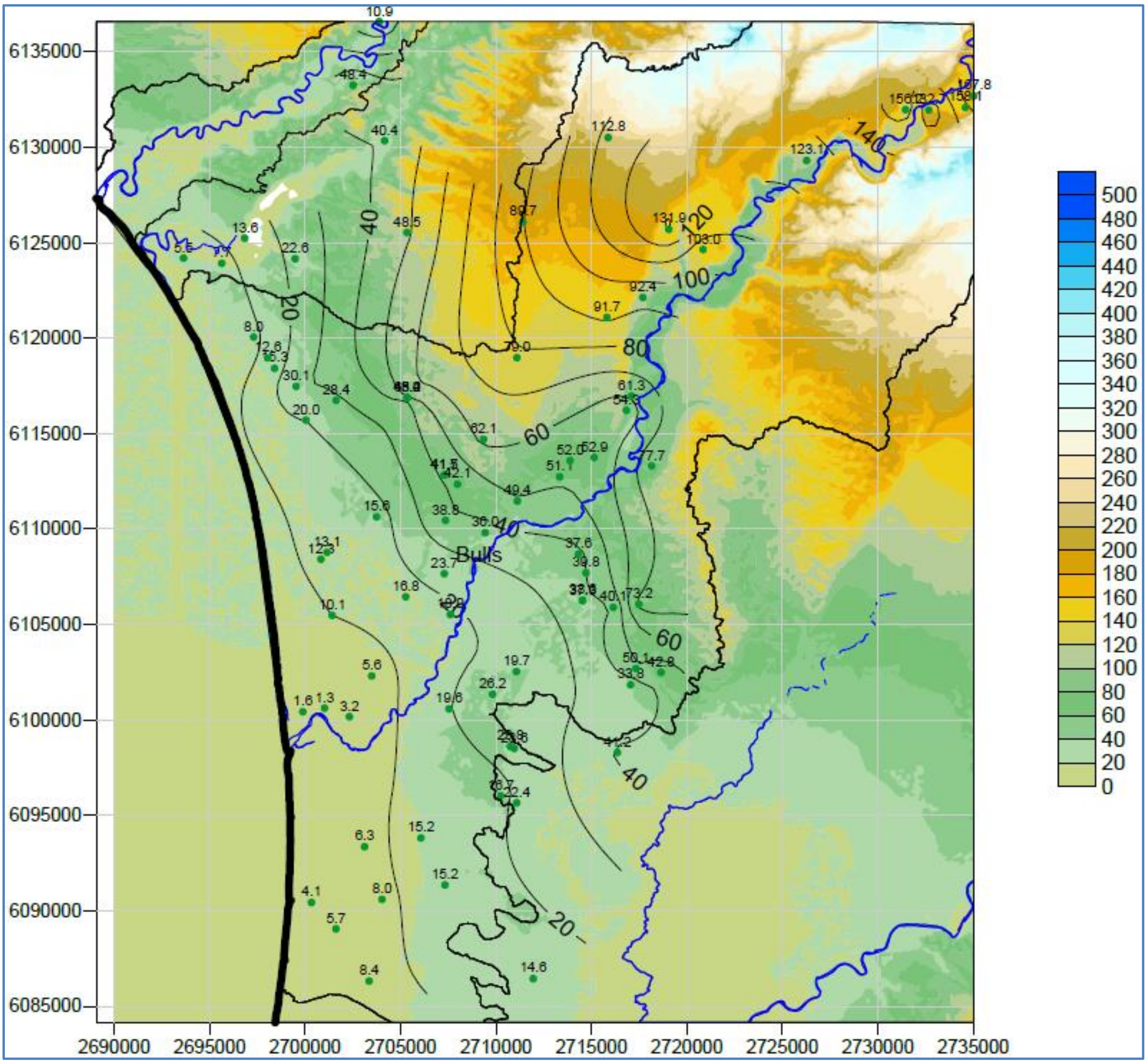


Figure 4-14. Groundwater level contours from October 2014 piezometric survey data from bores more than 20 m deep (Figure 6 in PDP, 2017).

Figure 4-15. Rivers and streams in the area (blue lines) and river (RIV) cells in Layer 1 in the model (blue squares).

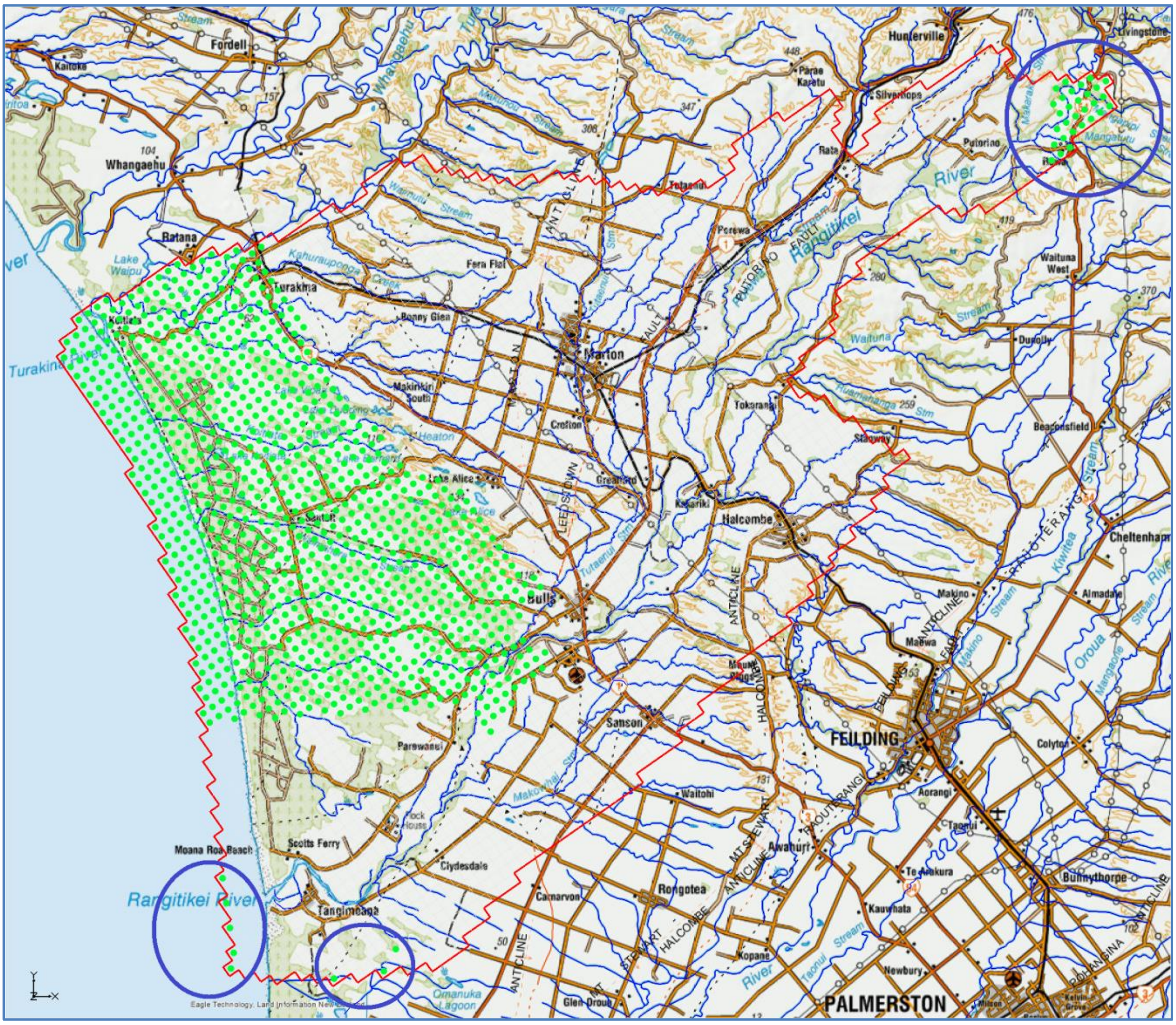


Figure 4-16. Rivers and streams in the area (blue lines) and drain (DRN) cells in Layer 1 in the model (green dots). The blue circles mark likely inappropriate designation of DRN cells. Not the inappropriate offshore assignment of drain cells.

Section 3.1 in PDP (2017) clarifies that RIV boundary conditions were specified in Layer 1 only. Checking the 2023 model files confirms this. The report clarifies that RIV boundary conditions require data on river bed conductance (c), river bed elevation, and river stage elevation. These parameters are not known with certainty. PDP (2017) notes that the river bed conductance was varied during the calibration process within predetermined bounds, the elevation of the river bed was set to 2 m below the surface elevation, and the river stage was set as 1 m above the river bed, i.e., the water depth in all RIV cells is set at 1 m. This review notes that approximations and uncertainty in surface elevation data transpire into the river bed and stage estimates. Examination of the 2023 model files revealed that the water depth in all RIV cells is not 1 m as reported. It is 0.5 m. The RIV C is set at 220.85 m²/d for ‘major rivers’ and 8.28 m²/d for ‘minor rivers’. It was also found that the bottom of the river is set at 1 m below the surface (not 2 m as reported) and the stage at 0.5 m below the surface (not 1 m as reported). The impacts of the reporting discrepancies on the model's performance and uncertainty were not evaluated in this review, as such an assessment would have required a substantial modelling effort. However, the inconsistency between the inspected model and the reviewed report diminishes confidence in both. This review also highlights that the assignment of uniform RIV parameters across all river systems and reaches is unrealistic and must be addressed in future modelling efforts. The identified omissions in the rivers' conceptualisation and numerical modelling are significant enough to question the model’s realism and its ability to accurately represent the modelled hydrogeological system.

PDP (2017) notes that in the coastal area, the low vertical permeabilities observed from pumping tests and neutral or upwards vertical hydraulic gradients imply that there is relatively limited infiltration to deeper strata. Swamps and small lakes occur in a number of areas around the coastal plains in the model area. Those surface water bodies provide discharge points for shallow groundwater in the area and are represented in the model as DRN cells. These cells only allow groundwater discharge. The required parameters include a drain bottom elevation and conductance. The Adrain cell elevations in the models were set 0.5 m below the surface elevation and conductance was varied during the model calibration.

Figure 4-16 shows the cells designated DRN boundary conditions in Layer 1 extracted from the model files. The blue circles in the figure indicate likely erroneous designation of DRN cells, as they occur in areas that are not conceptualised to contain such boundary conditions.

Examination of the 2023 model files reveal that the DRN C is constantly set at 50 m²/d for all designated cells. The drain cell elevations in the models were set 1 m below the surface elevation (not 0.5 m as reported). Like with the RIV boundary condition, the impacts of these reporting discrepancies on the model performance and uncertainty were not checked in this review. Furthermore, this review is of the opinion that the assignment of constant DRN parameters for all river systems and reaches is unrealistic. This must be revised in future modelling.

The reports do not state that the rivers and drains represented in the model were checked for hydraulic correctness, i.e., their elevations were not checked to see if they allow surface water flow.

It is noted that some RIV and DRN cells overlap. This could be a problem with the numerical model implementation. 'MOD' in MODFLOW refers to the software's modularity, i.e., it comprises modules that are applied independently but iteratively. MODFLOW processes each boundary condition module (known as package) separately and sequentially within a single time step. The interaction and priority between these boundary conditions depend on how the packages are implemented in the model and the sequence in which they are called. In the 2017 and 2023 models, both packages will be active, and the flows will be calculated independently, creating a potential conflict. If the groundwater head in the cell is higher than both the river stage and the drain elevation, both the RIV and DRN packages will simulate outflow from the groundwater system. This might result in higher than expected outflow from the cell. If the groundwater head is between the river stage and the drain elevation, the river might contribute water to the groundwater system (if the river stage is higher), while the drain might remove water from the groundwater system (if the groundwater head is higher than the drain elevation). Assigning both RIV and DRN boundary conditions to the same cell could have implications for model stability and accuracy. It could lead to unrealistic or unstable simulations, depending on the specific conditions and parameters set for each package. Careful calibration and validation are necessary to ensure that the model accurately represents the real-world system and that the combined effects of these boundary conditions are reasonable. In practice, it is generally advisable to avoid assigning conflicting boundary conditions to the same cell unless there is a specific and justified hydrological reason to do so. If necessary, it may be better to use adjacent cells to represent different boundary conditions or to refine the model grid to better capture the interactions between different hydrological features. Accordingly, this review identifies this issue as a problem area that must be resolved in future models.

PDP (2017) notes that coastal lakes in the RGMZ were not simulated in the model, partly due to their size relative to model grid cells, and because very little information regarding their hydraulic characteristics and connection with groundwater is available. This review is of the opinion that the importance of these features must be analysed. If they are found hydrogeologically significant, they must be included in future models.

Conductance (C) is an important parameter for few of the boundary conditions in the model (GHB, RIV, DRN). It indicates the ease with which water can move through a boundary. It is commonly expressed in units of m²/d but can also be expressed in units of m/d, depending on how it is defined. In PDP (2017) it is expressed in units of m/d, implying it is defined per unit width or length of the boundary. This should be confirmed.

Conductance values were extracted from the 2023 model and compared to the values reported in Table 3 in PDP (2017). The outcome is summarised in Table 4-2. Conductance values in the 2023 model are different from the reported values.

Table 4-2. Conducatance (C) values assigned to various features in the 2017 and 2023 models.

Feature	PDP (2017) Conductance (C)	2023 model Conductance (C)*
Major rivers (river bed conductance)	500 m/d	220.85
Minor rivers (river bed conductance)	16.5 m/d	8.28
Drain conductance	7.68 m/d	50.00
General head boundary conductance	250 m/d	28.08 m²/d in Layers 1, 2, 3 72.62 m²/d in Layer 4 472.18 m²/d in Layer 5 359.20 m²/d in Layer 6
* Units could not be confirmed.		

Conductance is dependent on hydraulic conductivity (i.e., $C \propto K$). It should have been varied to reflect the specific conditions of the relevant boundary conditions. However, examination of the 2023 model does not reflect this relationship in relevant boundary conditions (GHB, DRN, RIV).

This review is of the opinion that the peripheral boundaries are not adequately assigned in terms of their types and their parameters.

Groundwater abstraction was simulated in the model using MODFLOW WEL package (Figure 4-17). There are no accurate data on groundwater abstraction within the model domain. Section 3.2 in PDP (2017) explains the methodology followed for estimating abstraction. According to the report, there were 80 consented abstractions in the 2017 model period, of which only 20 had metered abstraction data. The abstraction records did not extend back before 2005. The 2023 model files were not thoroughly inspected for abstraction data because of the difficulty associated with this type of transient data. However, the abstraction estimation methodology reported in PDP (2017) seems to be reasonable. In future model reports, these data are required to be provided in an electronic appendix. Best practice guidelines like the AGMG (Barnett *et al.*, 2012) recommend that the modelling report should describe all data collected and information created through the modelling process. They further recommend that the report should be accompanied by an archive of all the model files and all supporting data so the results presented in the report can, if necessary, be reproduced and the model used in future studies.

4.3.4 Initial Conditions

Initial groundwater model conditions refer to the starting hydraulic heads and flow conditions assigned to the model at the beginning of a simulation. In steady state models, these initial conditions represent an equilibrium state where inflows and outflows balance, providing a stable baseline from which to evaluate long-term average conditions. Steady state models are tolerant to inaccuracies in initial conditions. They will mainly impact the model run time. The closer the initial conditions estimations to the steady-state model solution, the less model iteration required and the faster the solution is reached. On the other hand, accurate initial conditions are crucial in transient models as they set the starting point for simulating temporal variations in groundwater flow and storage over time. Properly defined initial conditions ensure that the model accurately reflects the real-world groundwater system from the outset, enhancing the reliability and validity of subsequent predictions and analyses.

PDP (2017) and PDP (2023) do not report the initial conditions for the models. Examination of the 2017 model files indicate that the initial head was set at 200 masl at Stress Period 1, the steady state stress period. As explained above, this value is not important. However, the first step steady state simulation outcome is crucial for the transient model.

Examination of the 2023 model files suggest reasonable initial conditions for the transient simulation in all layers (e.g., Figure 4-18). The data for all other layers were inspected but are not presented in this review. It is nearly impossible in this review to quantitatively validate that the initial conditions are appropriate. Hence, no further work has been done on this matter.

Figure 4-17. Wells (groundwater abstraction bores) and well (WEL) cells in Layers 1–8, 10 and 11 in the model (red squares). There are no wells in Layer 9. Wells in different layers are presented in Appendix C.

4.4 Calibration and Sensitivity Analysis

4.4.1 Have All Available Types of Observations Been Used in the Calibration?

The 2017 and 2023 models were only calibrated against groundwater level data. The calibration did not include surface water levels, groundwater and surface water flows, and groundwater-surface water exchanges albeit availability of surface water level data from various sources and concurrent surface water gauging that would have provided useful data for model conceptualisation and calibration.

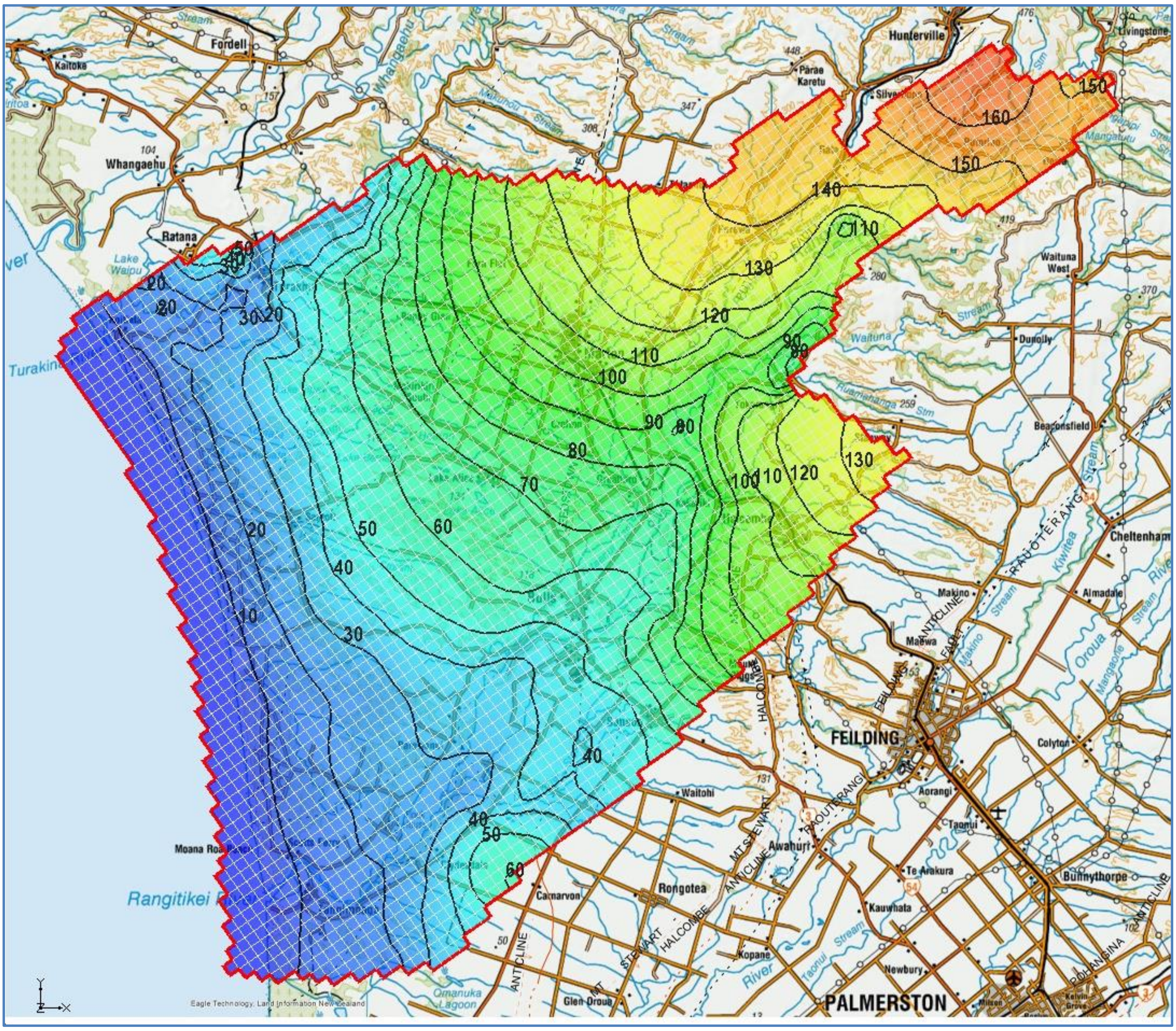


Figure 4-18. Piezometric map showing the elevation of the modelled water table (Layer 1 groundwater head) in masl at the initial steady state time step, which provides initial conditions to the transient model.

4.4.2 Does the Calibration Methodology Conform to Best Practice?

Section 4.1 in PDP (2017) indicates that the model has been parametrised using Pilot Points (e.g., HK). The location of Pilot Points is shown in Figure 4-19. Examination of the 2023 model indicates that zone parameterisation has also been used (e.g., for GHB, RIV, and DRN C). The report also explains that zoned Pilot Points were used in the model calibration. The calibration was undertaken automatically using PEST (Doherty, 2016)[8].

PDP (2017) presents most of the required information on the objective function, e.g., Table 2 in the report. Section 6.1 in the report discusses parameter sensitivities based on composite sensitivities. This is adequate for understanding parameter identifiability and uncertainty in the model. The outcome of the analysis is presented in bar-graph format for Pilot Points. It is difficult to relate this information to the model layers and areas. Maps would have been helpful. Also, given that the model layers do not represent stratigraphic strata, the parameter identifiability are hydrogeologically uninformative.

[8] The most recent edition is Doherty J (2023). PEST: Model-Independent Parameter Estimation User Manual (8th ed.). Watermark Numerical Computing.

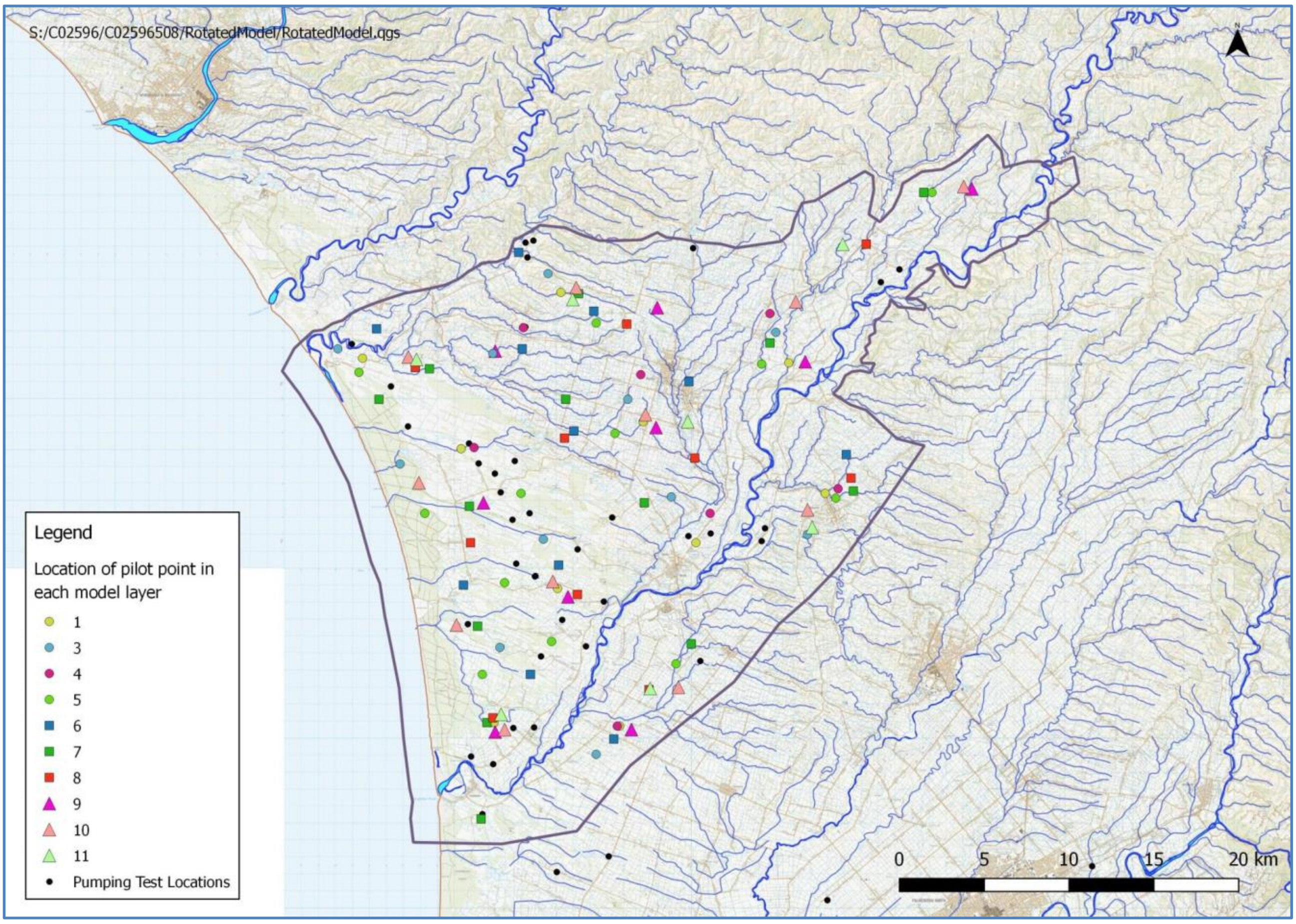


Figure 4-19. Pilot Point locations (Figure 16 in PDP, 2017).

Section 6.1 clarifies that the analysis indicates that the most sensitive parameters are the inland vertical hydraulic conductivity values in Layers 2, 3, 4 and 5, together with the value of hydraulic conductivity estimated at Pilot Points in Layer 5. The model is stated to be relatively insensitive to river bed conductance and general head boundary conductance.

Without making a judgement on the outcome, the methodology used is contemporary, robust, and conforms to best practice guidelines. The parametric sensitivity and uncertainty analyses reported would be useful to future modelling. However, they must be performed comprehensively in future models as they are model-specific.

It must be noted that recharge was not included in the model calibration and sensitivity analysis. Recharge and hydraulic conductivity are strongly correlated. Excluding it from the model calibration increases the level of the model's non-uniqueness. Section 6.3 does not specifically discuss recharge. However, it notes that there is a degree of non-uniqueness present in the model based on parameter correlation analysis, but without specifying the correlated parameters inspected.

4.4.3 Has the Sensitivity of the Calibration Been Assessed Against Variability in Conditions?

The model sensitivity to changes in boundary conditions was not assessed, especially in terms of the peripheral boundary types and location. It was not assessed against the initial conditions and stresses. It was assessed against parameters that did not include recharge. Future models must include assessing the model for the location and type of peripheral boundaries, and stresses like pumping and recharge.

4.4.4 Calibration Outcome

As noted above, the model was parameterised using Pilot Points, zones, and zoned Pilot Points. PDP (2017) notes that Layer 1 was split into zones to account for the different geological strata and Pilot Point interpolation was restricted to those zones (Figure 4-20). The effect of this on the 2023 model calibrated horizontal hydraulic conductivity can be seen in Figure B-2 but is less clear in Figure B-1. The report clarifies that other layers were not zoned due to lack of knowledge of the distribution of their properties. The report notes that zones were used for the definition of vertical hydraulic conductivity across the model area to allow for differences in vertical hydraulic conductivity particularly between the coastal area and inland and where apparently higher values of leakage occur close to the true right bank of the Rangitīkei River. The locations of those zones are shown in Figure 4-21. It is not known with certainty from PDP (2017) or the 2023 model data whether these zones were applied to all layers or to Layer 1 only. The report states that the zones were not set to different starting values in the PEST calibration process, and the vertical hydraulic conductivity values for all layers and zones were varied during the calibration process within a range of 1 x 10^{-3} and 1 x 10^{-6} m/d. The reason for setting zones that are all similar is not clarified.

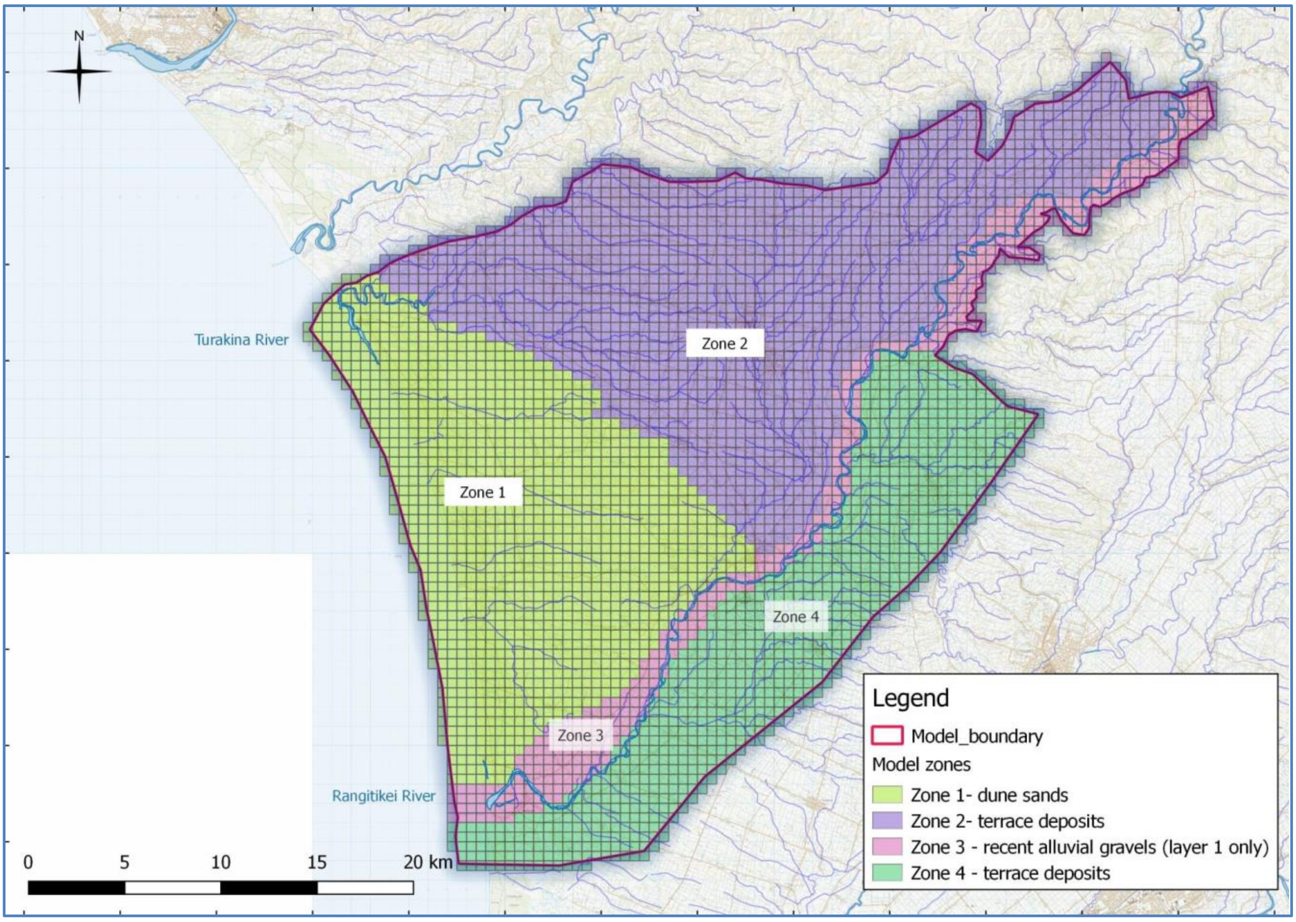


Figure 4-20. Model Zones for horizontal hydraulic conductivity in Layer 1 (Figure 17a in PDP, 2017).

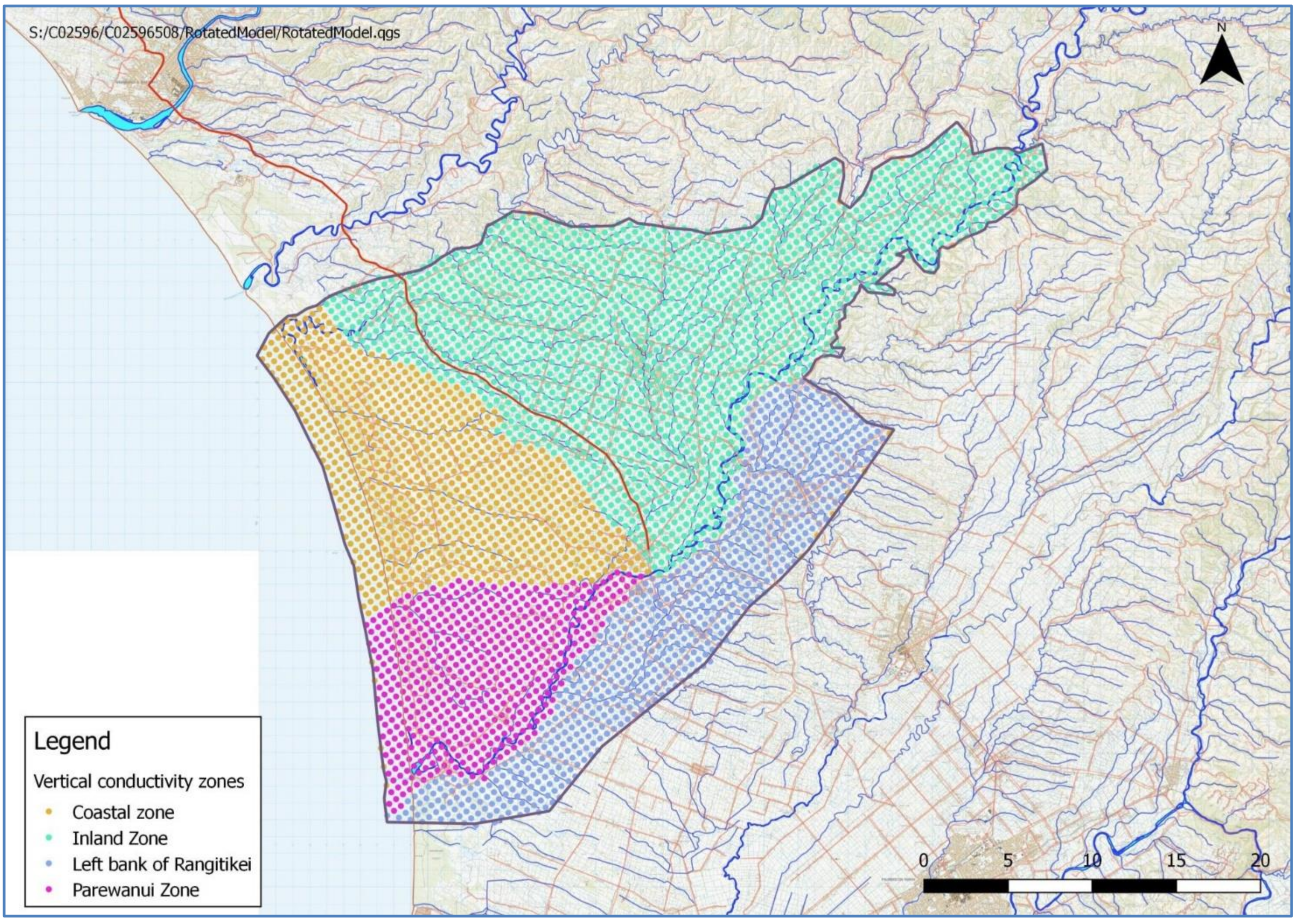


Figure 4-21. Model vertical conductivity zones (Figure 17b in PDP, 2017).

The report indicates that in all model layers, 'storativity' was varied in the range between 1 x 10^{-3} and 1 x 10^{-6} and Specific yield was varied between 0.1 and 0.3 during the calibration process. Commonly, the model uses specific storage (S_s) not storativity (S). It is noted that S is dimensionless, whereas S_s is in m^{-1} units as it is calculated as S/b, with b being the aquifer thickness in metres. This could be a reporting omission. According to Rau *et al.* (2018), S_s has a physical upper limit of about 1.3 x 10^{-5} m^{-1}. Hence, it is important to confirm that the storage term in the model is realistic.

The report presents maps showing the calibrated model hydraulic conductivity in each layer. The hydraulic conductivity data were extracted from the provided 2023 model files. Table 4-3 compares the extracted values against the values reported in PDP (2017). There is noticeable difference between the two datasets albeit the two models producing similar simulation results and reasonably good calibration measures. This is a good example on the possibility of producing different plausible models to represent the same system, i.e., model non-uniqueness.

The data in Table 4-3 suggest that the horizontal hydraulic conductivity in the 2023 model was not constrained within the value ranges presented in PDP (2017). In addition, some values seem to be unrealistic. For example, horizontal hydraulic conductivity values of about 334 m/d and 105 m/d for Layers 5 and 6, respectively, are above expectations. Layer 5 is at least 45 m deep, and Layer 6 is at least 85 m deep. High horizontal hydraulic conductivity values as found in the 2023 model for these layers are not expected at their respective depths. In addition, the hydraulic conductivity values for these layers in the 2023 model imply that the transmissivity in Layers 5 and 6 are as high as about 13,350 m^2/d and 4,200 m^2/d, respectively. Such high transmissivity values have not been reported at these depths in the area.

Table 2 in PDP (2017) present the model calibration metrics. They are within best practice guideline ranges. The 2023 model calibration metrics are not reported. Observed and model-calculated heads in both models are presented in hydrograph and scatter plot formats in their respective reports. The data were checked in the 2023 model files.

Table 4-3. Horizontal hydraulic conductivity (HK) values in m/d in the calibrated model as reported in PDP (2017) compared to the values extracted from the reviewed 2023 model files.

Model layer	HK in PDP (2017)		HK extracted from the 2023 model files		Figures in this review
	Minimum	Maximum	Minimum	Maximum	
1	1.00	2.57	1.00	85.30	Figure B-1 & Figure B-2
2	0.69	1.04	1.00	1.47	Figure B-3 & Figure B-4
3	0.63	1.11	1.00	1.71	Figure B-5 & Figure B-6
4	0.83	2.10	1.00	351.75	Figure B-7 & Figure B-8
5	0.10	100.00	1.00	333.91	Figure B-9 & Figure B-10
6	0.10	100.00	1.00	104.96	Figure B-11 & Figure B-12
7	0.10	100.00	1.00	78.72	Figure B-13 & Figure B-14
8	0.10	0.84	1.00	14.47	Figure B-15 & Figure B-16
9	0.29	0.83	1.00	36.67	Figure B-17 & Figure B-18
10	0.59	1.07	1.00	4.47	Figure B-19 & Figure B-20
11	1.00	1.00	1.00	1.73	Figure B-21 & Figure B-22

It was noticed that the groundwater level calibration plots compare groundwater level measurements in monitoring bores to model calculated groundwater level in the cells they are associated with. The cells are 500 m x 500 m. Hence, the bore can be up to about 350 m off the centre of the cell. Therefore, the model estimated groundwater level at the monitoring bores was calculated. The hydrographs in Appendix D compare the measured groundwater levels in the bores to corresponding model calculated groundwater levels both at the centres of the cells and at the bore location. The difference could be negligible for some bores (e.g., Bore 301045, Figure D-3). However, it is over 5 m in the case of Bore 312009 (Figure D-12). This variability can affect the assessment of the model calibration. Model calculated values at the bore site should be used in the assessment of its calibration.

Similarly, it was noticed that in scatter plots the measured groundwater levels were compared to the model calculated groundwater levels at the centre of the cell and the end of the stress period within which the measurement date falls (Figure 4-22). This means that the time difference between the measured and model calculated groundwater levels can be up to 30 days. Therefore, the measured groundwater levels were plotted against model groundwater levels calculated at the bore site and the date of the measurement (Figure 4-23), which is the recommended approach. The difference between Figure 4-22 and Figure 4-23 cannot be seen easily, but it is clear in the data for Bore 313013 (Figure E-3). The scatter plot is broken into segments to enhance its scale in Appendix E.

All the plots show reasonable match. Notwithstanding, the hydraulic conductivity values presented in Table 4-3 suggest imperfect model calibration outcome with data breaching their conceptual determined ranges.

PDP (2017) present the calibrated horizontal hydraulic conductivity values in maps. It does not present these data in cross sections. It presents model calculated groundwater head in a contour map representing Layer 4 on 1 October 2014. It also presents all observed and model calculated groundwater level data as hydrographs. The report does not present cross sections of hydraulic conductivity or groundwater heads. PDP (2023) only presents observed and model calculated groundwater level data as hydrographs.

Horizontal hydraulic conductivity and groundwater head data were extracted from the 2023 model files to compare them to the measured and test data reported in PDP (2017).

Section 2.6 in PDP (2017) discusses hydraulic properties based on pumping tests of bores, the locations of which are shown in Figure 4-1. The report notes that the highest transmissivity is in an area close to the coast on the north bank of the Rangitīkei River, where it exceeds 5,000 m²/day. It highlights that there are several large-scale groundwater abstraction consents in that area, which exploit the higher aquifer transmissivity. The report suggests that the high transmissivity values in that area could reflect a highly permeable river palaeochannel.

PDP (2017) indicates that lower transmissivity values (generally less than 500 m²/day) have been derived from pumping tests in bores located on the coastal plains between the Rangitīkei River and the Turakina River, suggesting relatively lower permeability. It suggests that these lower transmissivity values occur across a range of depths, and there does not appear to be a clear link between permeability and bore depth. The report maintains that lower permeability in this area supports its hypothesis that groundwater levels in bores from this region (e.g., Bores 312007 and 312009) suggest slow percolation of rainfall recharge through the unsaturated strata.

Apparently, the report considers transmissivity values up to 500 m²/day to be low. This review considers this classification inappropriate. According to Cherry & Davis (1991) and Mather & Chapman (2007), low transmissivity is defined as being less than 50 m²/day. Fetter (2001), Sophocleous (2002), and Todd & Mays (2005) define low transmissivity values as up to 100 m²/day. Assuming an average screen length of 5 m and using the same method for deriving hydraulic conductivity from transmissivity as in PDP (2017), it appears that the report considers the upper bound of low hydraulic conductivity to be 100 m/d. This contrasts with the definition in Domenico & Schwartz (1998), Fetter (2001), Freeze & Cherry (1979), and Hiscock (2005), which set the upper bound of low hydraulic conductivity at 1 m/day.

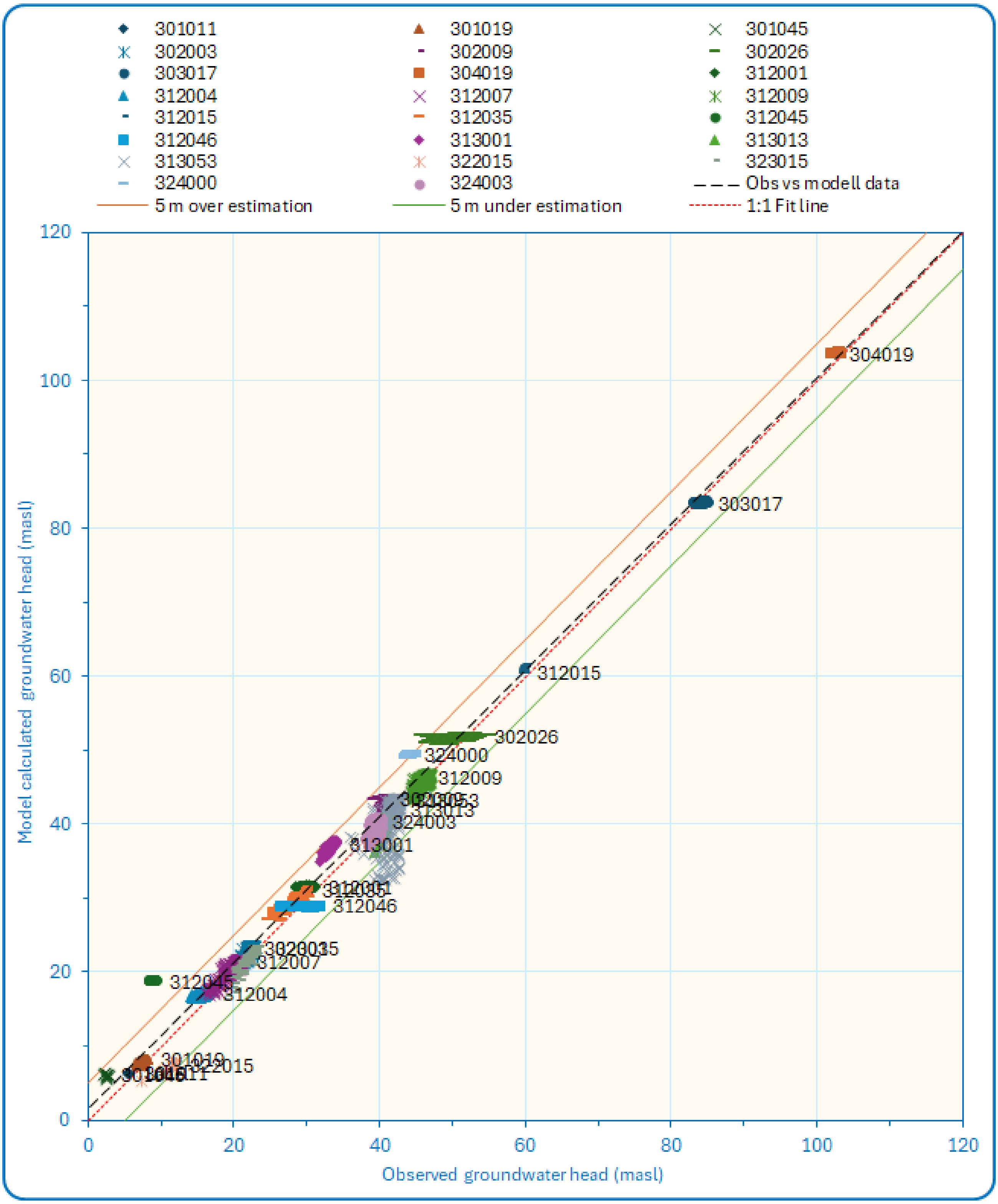


Figure 4-22. Scatter plot of measured groundwater levels and the 2023 model-calculated groundwater levels at cell centres at the end of the stress periods in which the dates of measurement fall.

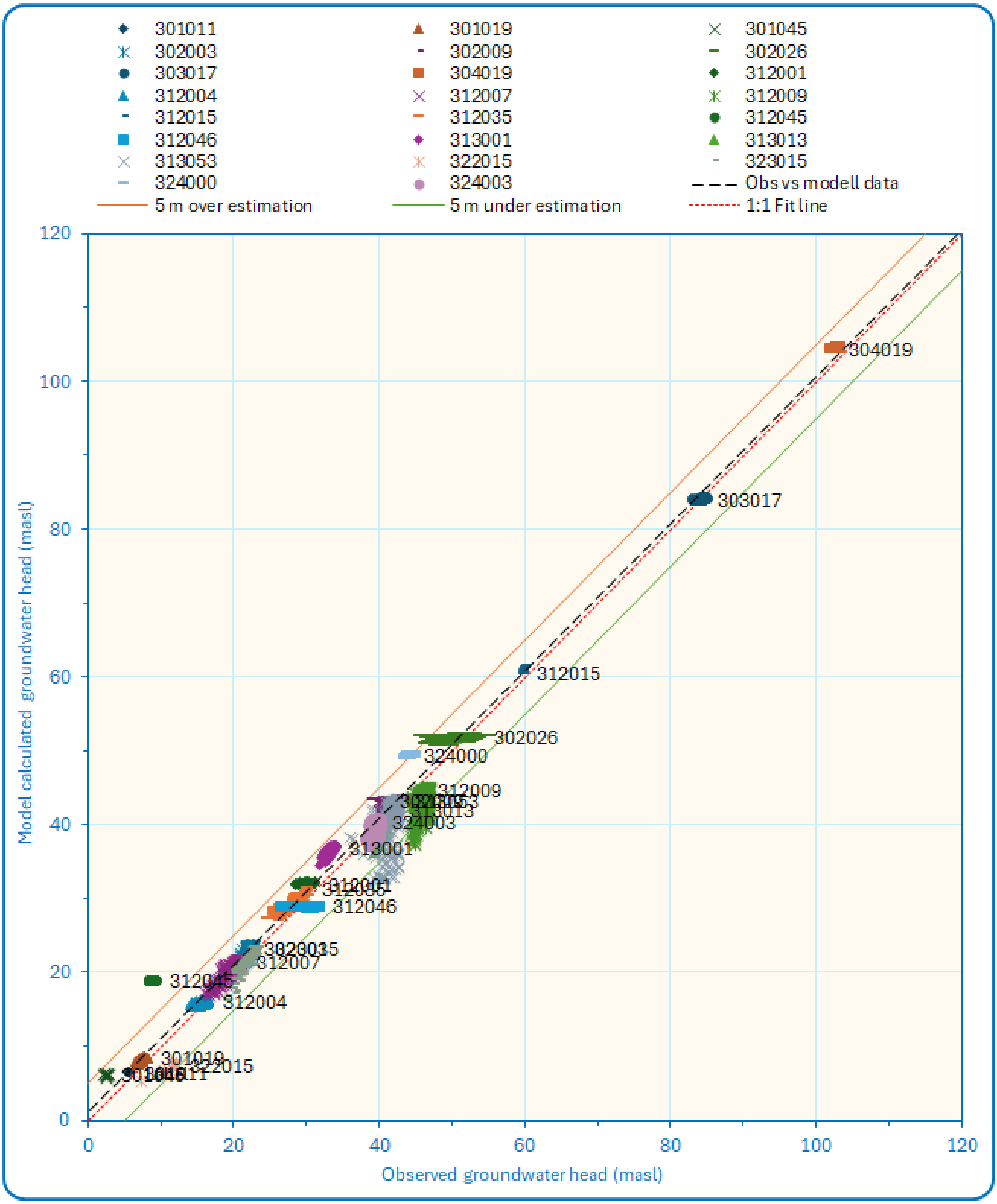


Figure 4-23. Scatter plot of measured groundwater levels and the 2023 model-calculated groundwater levels at the locations of bores on the dates of measurements.

Inspecting the provided 2023 model files reveals that most cells were assigned horizontal hydraulic conductivity values of less than 5 m/day (e.g., Figure 4-24). Noticeably, high horizontal hydraulic conductivity values have been assigned to the Turakina River area, which this review believes to be unrealistic and inconsistent with the reports' discussions. This review finds the 2023 model assigned hydraulic conductivity unrealistic and inconsistent with the conceptualisation presented in the reports.

Hydraulic conductivity data presented in the maps in the figures listed in Table 4-3 indicate bullseyes as noted in Section 4.2.2 above (see Appendix B). Bullseyes can be seen also in the maps in the hydraulic conductivity maps presented in PDP (2017), but they are not very clear due to the used contour interval and colour scheme. Bullseyes in calibrated hydraulic conductivity maps often indicate unrealistic distribution of the parameter.

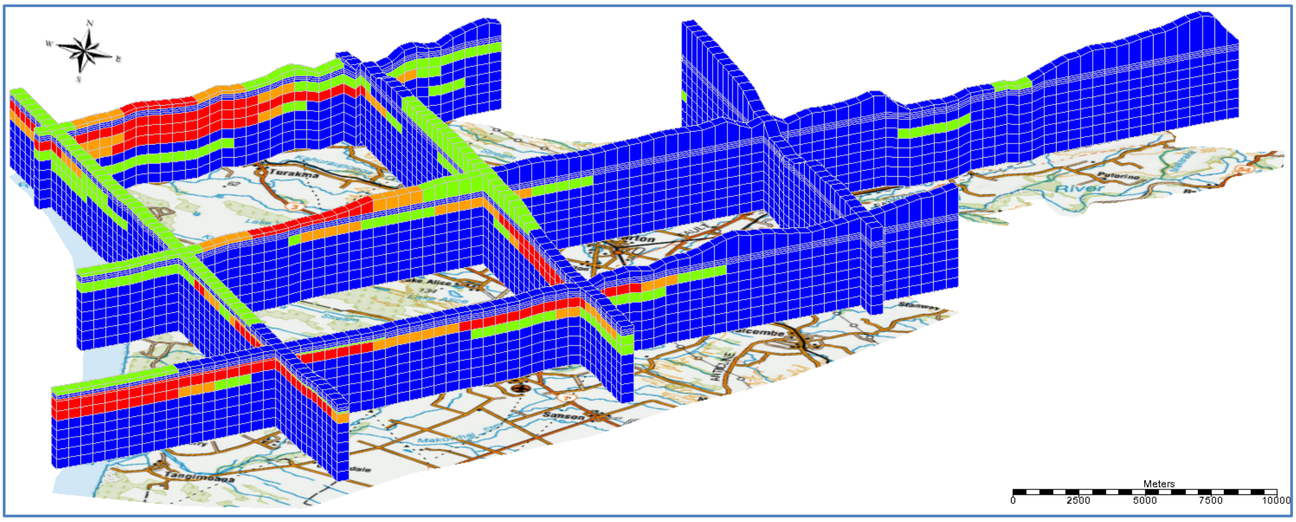


Figure 4-24. Fence diagram showing assigned horizontal hydraulic conductivity values along selected transect, colour-coded to highlight prevalance of low values. The value range interval is selected to highlight contrast. Blue 1–5 m/d, green 5–10 m/d, amber 10–15 m/d, red 15–260 m/d. For location of transcets, see Figure 4-25. Fixed intervall diagrams in Appendix F.

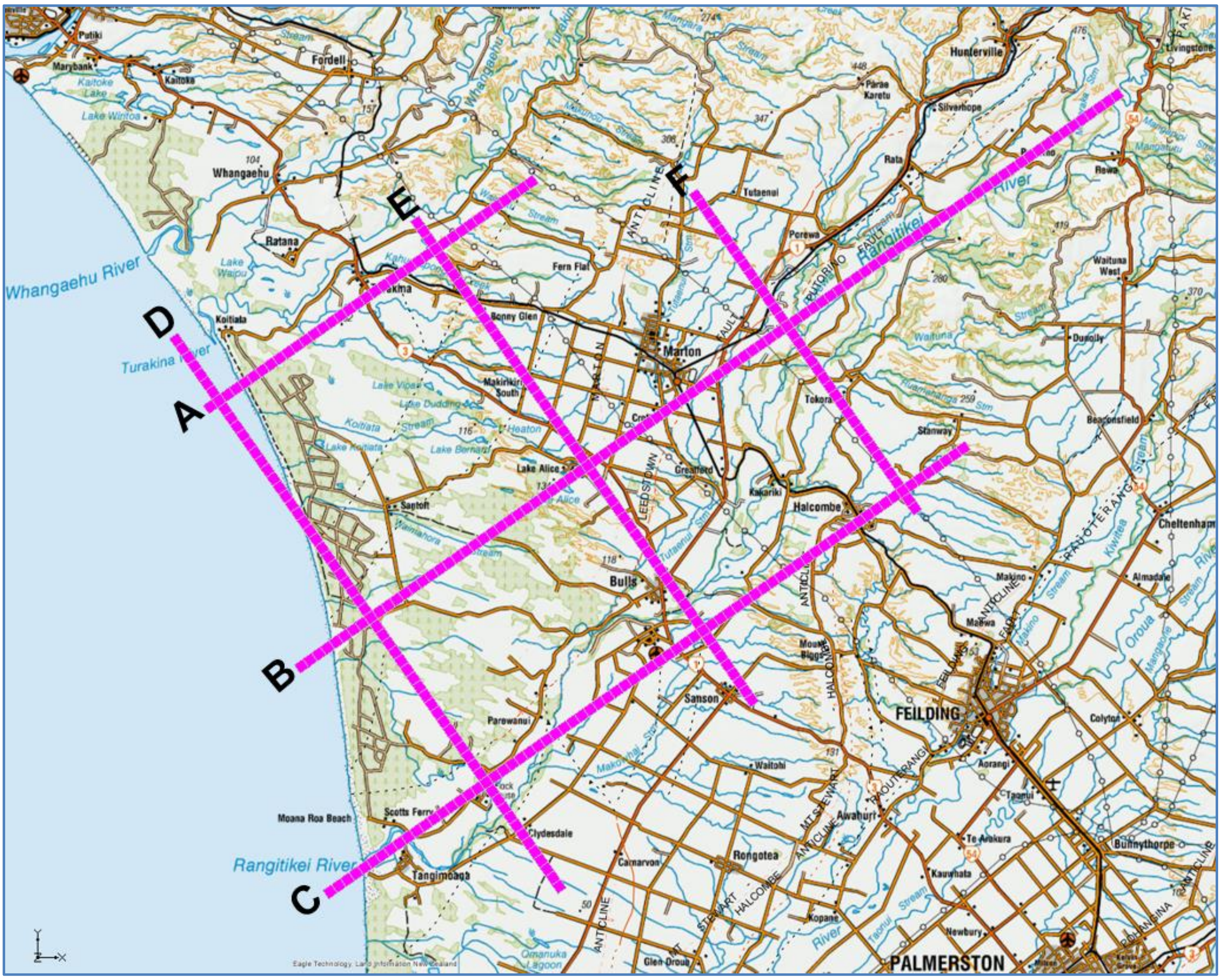


Figure 4-25. Location lines for horizontal hydraulic conductivity (HK) cross sections and calibrated model head sectional views.

PDP (2017) states that storativity values from pumping tests are typically around 1 x 10^{-4} or 1 x 10^{-5}, suggesting that the strata are confined or semi-confined. However, it notes that pumping tests from a few bores up to 75 m deep located around 6 km inland from the coast have a reported storage of around 0.1 (which is well within the unconfined aquifer storativity values range). PDP (2017) clarify they did not review these estimates. This review notes that most pumping tests in the Horizons region do not involve observation bores. Hence, they cannot be used to estimate storativity (a methodology limitation). In addition, the shape of the time-drawdown curves for many tests indicates semi-confined (leaky) or unconfined aquifer conditions. Hence, this review considers the storage parameters (specific yield and storage coefficient) to be largely unknown in the area. Modelling is recommended to utilise literature values suitable for the lithology. The sensitivity of the model to storage parameters must be assessed and, if appropriate, storage parameters must be included in the model calibration. Steady state calibration prior to transient calibration of the models is advisable as this could help constrain the other parameters (e.g., hydraulic conductivity and recharge), which could help with finding appropriate values for the storage parameters.

Groundwater data extracted from the calibrated 2023 model are plotted as sectional views (Appendix G), with the locations of the sectional lines shown in Figure 4-25. These sectional views reveal that groundwater generally flows along the topographic slope from northeast to southwest. They indicate a downward gradient inland, horizontal flow in an approximate strip 7–9 km from the coastline, and an upward gradient in the coastal area. The influence of the assigned horizontal and vertical hydraulic conductivity values on the model layers is evident, with flow refraction through low hydraulic conductivity layers (aquitards). The figures also indicate groundwater-surface water interaction, although not very clearly. However, the extent, configuration, and hydraulic properties of the modelled layers may differ from actual conditions. The fact that the model shows signs of good calibration does not necessarily imply it is realistic or that it has strong predictive capability.

The models review highlights the need for more data for hydrogeological analysis purposes, particularly geological information. This review recommends further hydrogeological assessments to enhance the understanding of the RGMZ hydrogeological system for modelling and other purposes. A competent geological model can greatly help.

4.4.5 Calibrated Groundwater Head

The key output of a groundwater model historic and predictive simulations is groundwater level (head). Groundwater heads from historic simulations are compared to observed data as a method to assess the model performance. The closer the fit between model calculated heads and observed measurements the better the model calibration. For models to be reliable, they must show good calibration results and be realistic, i.e., representative of hydrogeological settings and conditions. Figure 4-26 presents observed and model calculated head scatter plots from the reviewed reports, generally indicating calibration within guideline ranges. Figure 4-22, Figure 4-23, Appendix D, and Appendix E enable clearer comparison of the observed and calculated heads extracted from the 2023 model files.

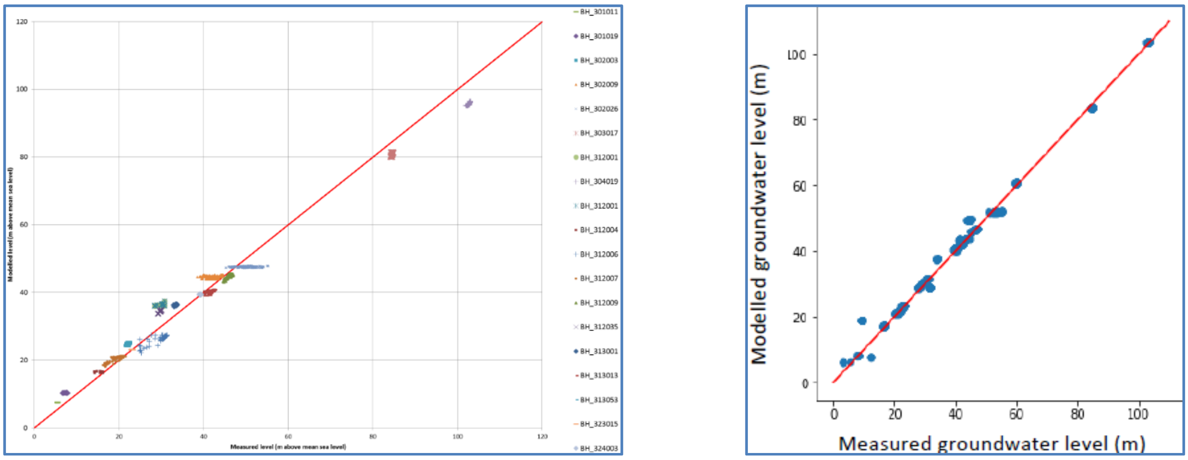


Figure 4-26. Scatter plots model calculated groundwater heads to observed heads from Figure 22 in PDP (2017) to the left and Figure 8 in PDP (2023) to the right.

PDP (2023) does not present contour maps of modelled groundwater levels. Instead, it provides maps of predicted changes in groundwater levels for different scenarios. This makes it difficult to assess the quality of the model calibration. The report should have included groundwater level contour maps for all model layers at key times and compared them to the outputs of the 2017 model to indicate differences in performance between the two models. The 2017 model files were not available for this review. However, a basic comparison between the heads simulated by the two models is presented in Figure 4-27. The 2017 calibrated model heads, representing groundwater levels in Layer 4 as of October 2014, are taken from Figure 19 in PDP (2017). Corresponding data were extracted from the 2023 model files. Significant differences between the two models become evident around the 40–50 masl groundwater levels in the 2017 model map. The difference continues and increases upgradient. PDP (2023) does not note the differences between the heads calculated in the 2017 and 2023 models and does not comment on them. The difference between the two maps reflects differences in the models' calibration (used parameters).

The resolution of the modelled groundwater levels map presented in Figure 19 in PDP (2017) is inadequate to show conditions near the coastline. Careful examination of the 2023 model files reveals that the model calculated head at the coastline is excessively exaggerated. Figure 4-28 presents the model-calculated groundwater head in Layer 1 around the coastline from the start to the end of the simulation period at eight-year intervals. The contour interval in the figure varies to clarify the situation. The head values are around 6 masl at the coastline, indicating a serious error in the implementation of the GHB boundary conditions at the western end of the model domain. This error is due to the offshore location selected for the western peripheral model boundary and the parameterisation of the GHB conditions designated there. These modelling errors introduce a high level of uncertainty in the model's historical simulations and future scenario predictions, necessitating correction in future model revisions. Note that PDP (2017) reports that GHB boundary conditions were set at the coast in Layers 1–3 but examination of the 2023 model files reveal that they are applied to Layers 1–6. Such reporting omission degrades confidence in the model even further.

The calibrated 2023 model groundwater head timeseries were examined through animation. The animation videos provided as Appendix H show that groundwater changes mainly occur near the lower reaches of the main rivers within the model domain (the Rangitīkei River and the Turakina River). Potential reasons are not explored in the reviewed reports.

It is noticed that the shallow groundwater level contours in PDP (2017) differ from the groundwater level contours extracted from the 2023 model files in terms of values and the general flow pattern (Figure 4-29). This is an essential check of the model calibration that indicates that the model is not representative of real-life conditions.

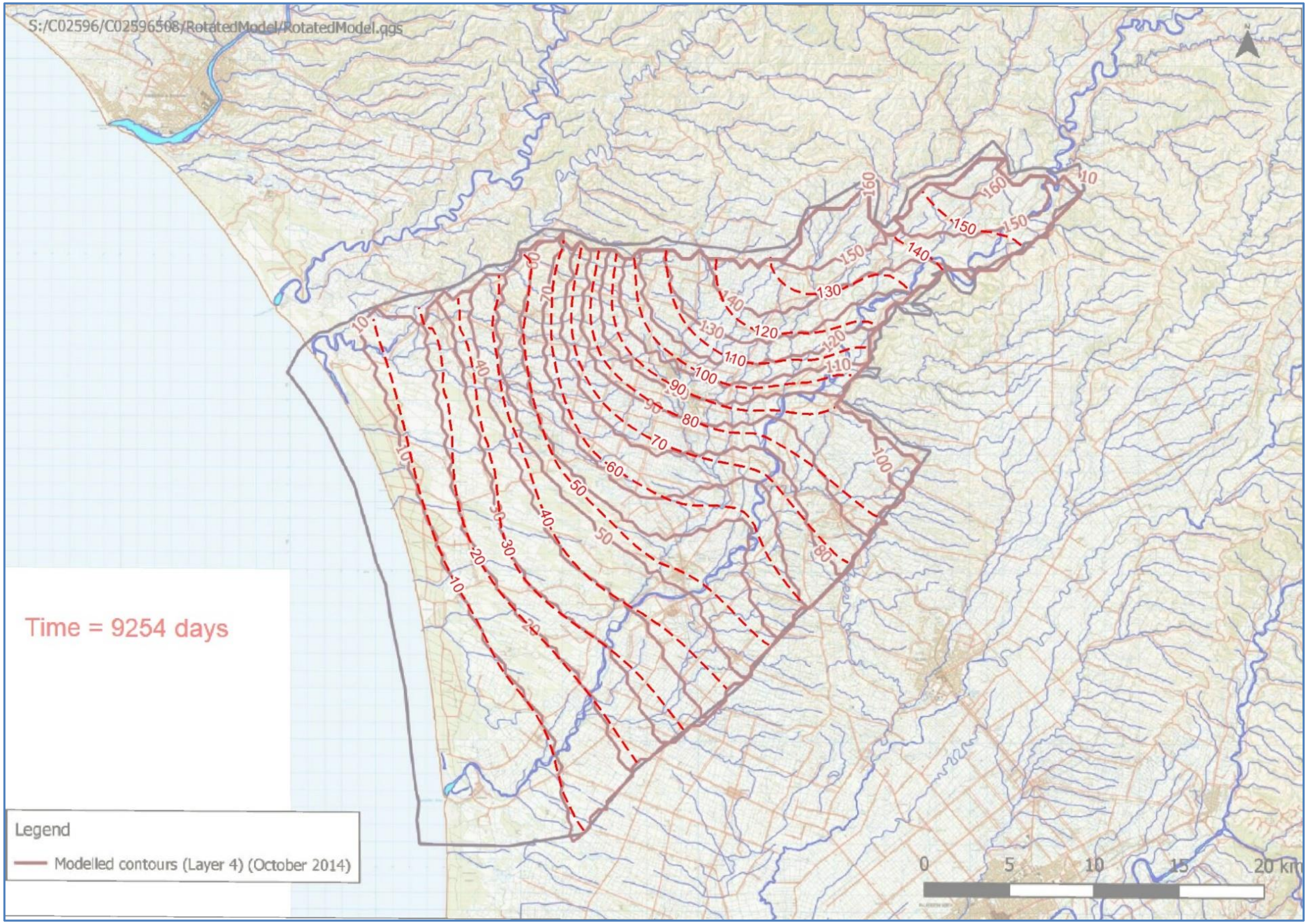
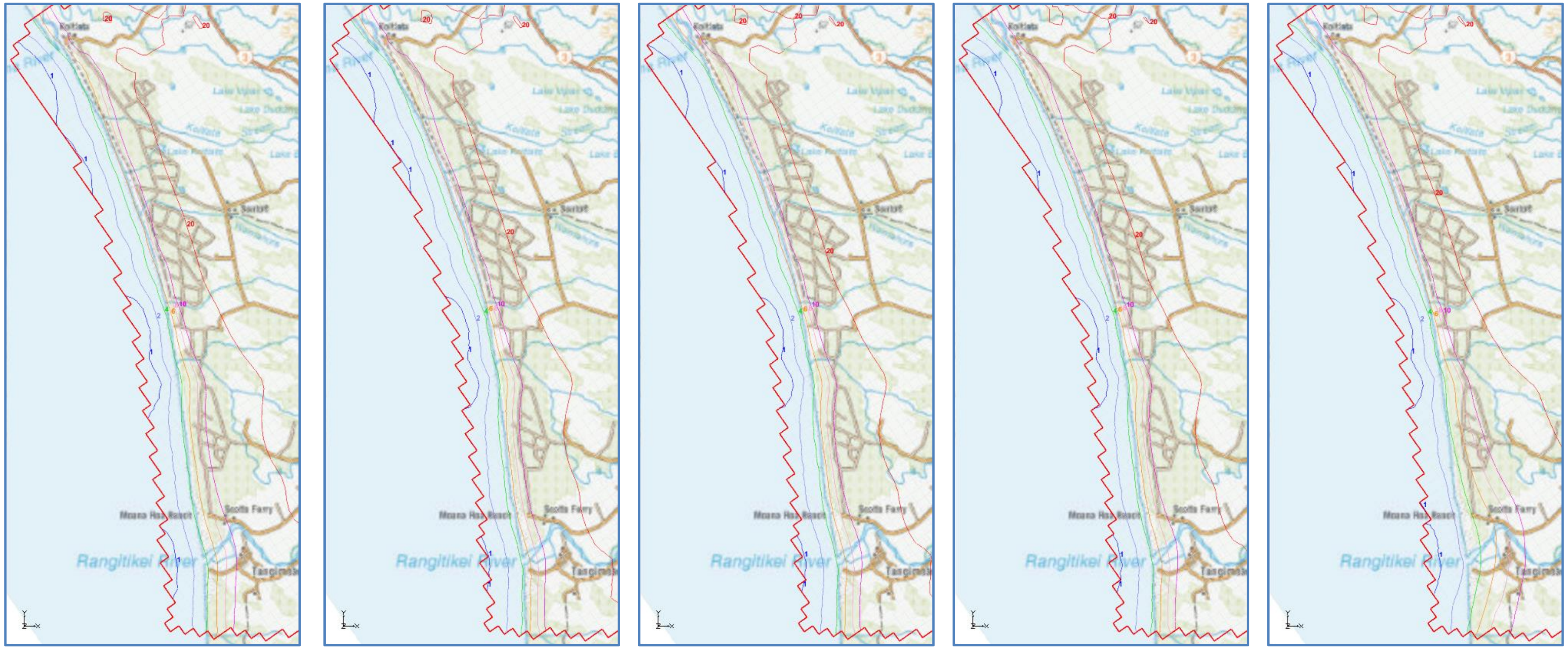


Figure 4-27. Groundwater level contours for Layer 4 in October 2014 as presented in in Figure 19 in PDP (2017) compared to the same data extracted from the 2023 model files.



1/07/1990 1/07/1998 1/07/2006 1/07/2014 1/07/2022

Figure 4-28. Groundwater level (head) contours up to 20 masl in Layer 1 at different times extracted from the 2023 model (variable contour interval to enhance visibility).

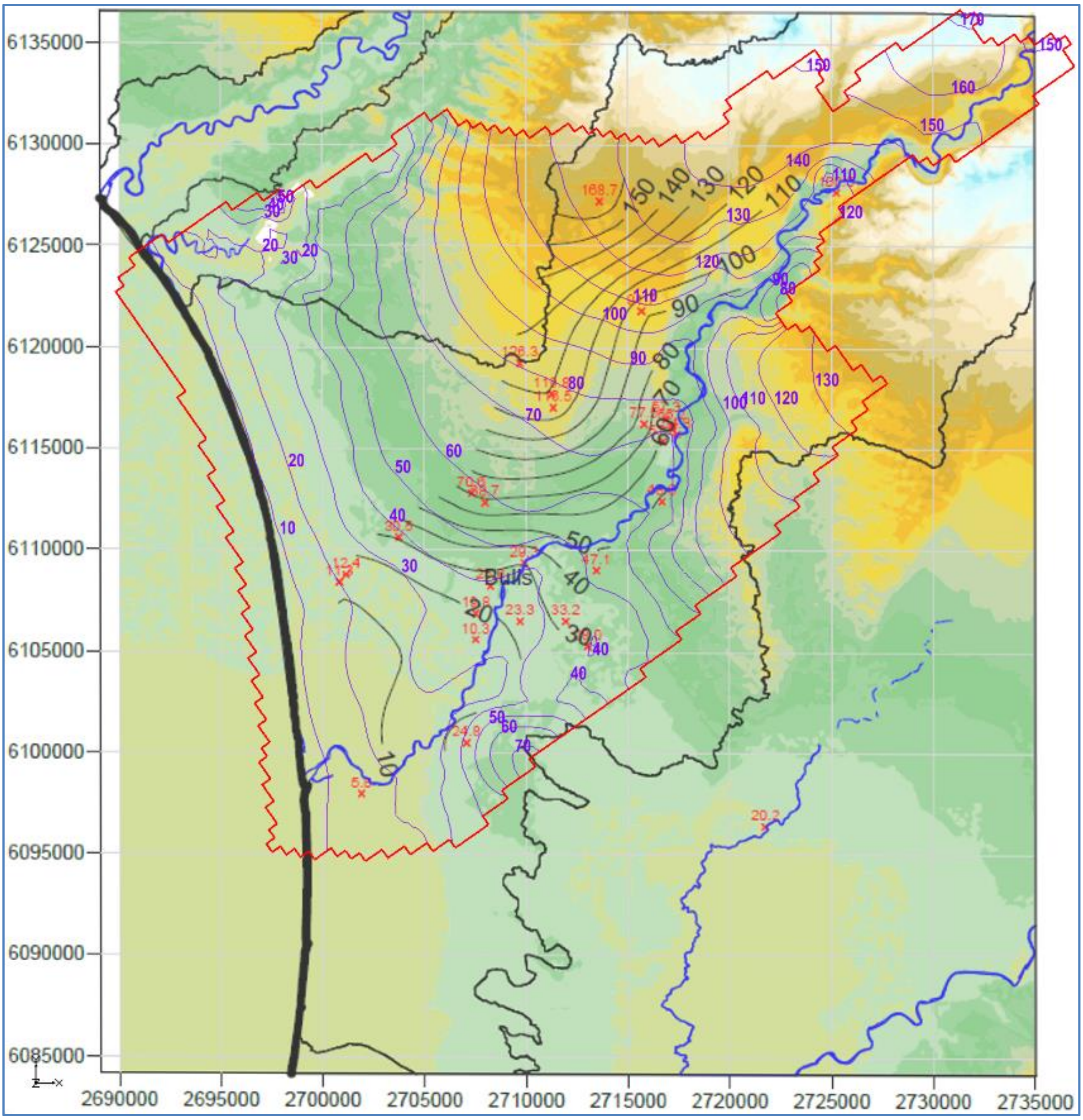


Figure 4-29. Comparison between shallow groundwater level contour lines (black) presented in Figure 5 in PDP (2017) and groundwater level contour lines for Layer 1 in the 2023 model (purple). Both maps are for 1 October 2014 conditions.

4.4.6 Calibration Statistic

Table 2 in PDP (2017) reports calibration statistics for the first model. The correlation coefficient (R^2) is 99.65%, which is very high, suggesting excellent model calibration (also see Figure 4-22 and Figure 4-23). Calibration metrics are not presented in PDP (2023). According to Barnett *et al.* (2012), the goodness of fit between model calculations and observations is often measured using a simple statistic, including:

- **RMS**: The root mean squared error.
- **SRMS**: The scaled root mean squared error.
- **MSR**: The mean sum of residuals.
- **SMSR**: The scaled mean sum of residuals.

The above metrics are not reported in PDP (2017) and PDP (2023). They have not been calculated in this review. However, it is expected that the models meet the guideline criteria for the above measures. This, however, does not mean that the model predictions are reliable because of the type and magnitude of the discovered errors found in the conceptual and numerical models.

4.5 Prediction

4.5.1 Are the Predictive Scenarios Designed Adequately?

PDP (2017) reports only one model predictive scenario aimed to explore the effects of continued abstraction at current rates. While this is a valid scenario and in line with the modelling objectives, that prediction is now obsolete. Its real value is in its usability for verification of the 2017 model to understand necessary model enhancements. Hence, the 2017 model predictions will not be discussed further.

Section 3.1 in PDP (2023) argues that its model '*reasonably represents the long term groundwater level trends observed in bores in the area and therefore it can be used to assess a variety of 'what if' scenarios regarding future groundwater management in the area. The 'what if' scenarios aim to provide information on what may happen to groundwater levels in the area based on different management options for groundwater allocation.*' It notes that because the model is not calibrated to river flows or surface water fluxes, direct predictions of the effects of abstraction on receptors are expected to be less certain than groundwater level predictions.

The report lists three predictive scenarios intended to assist with the development of an updated groundwater allocation limit:

- **Scenario 1**: A scenario where groundwater allocation is reduced to a point where the long term decline in groundwater levels ceases over the next 10 years. Abstraction rates are reduced for 10 years into the future.
- **Scenario 2**: A scenario simulating a cap on allocation at existing levels of abstraction (i.e., no new consents are granted but existing consents are renewed). Abstraction continues at the same rate as the current year for 10 years into the future.
- **Scenario 3**: A scenario considering a proportional increase in existing levels of abstraction up to the existing allocation limit. Abstractions are increased to the limit for 10 years into the future.

The above scenarios are appropriate for the stated modelling objectives.

According to the report, rainfall groundwater recharge for the 10-year prediction period (2022–2032) is assumed to be similar to the preceding 10 years (2012–2022). Abstraction rates are varied to suit the above scenarios.

While the above climate and abstraction related assumptions are reasonable, attention is drawn to the above discussed issues with the implementation of recharge in the numerical model and other reasons that could limit the model's predictive reliability. Climate change is not expected to be a crucial consideration over the defined predictions timeframe. Overall, the assumed stresses and timescale are appropriate for the stated objectives, albite that the stresses quantities and distributions require revisions (mainly recharge).

The report does not explicitly define a baseline (null) scenario. However, the predictive modelling results include 'naturalised groundwater levels'. They are defined as simulated groundwater levels assuming no abstraction from the start of the transient modelling to the end of the prediction period, i.e., from 1990 to 2032.

4.5.2 Presentation

The predictive scenarios modelling results are principally presented as groundwater level hydrographs (e.g., Figure 4-30) showing:

- calibrated model groundwater levels (1990–2022)
- predicted groundwater levels for as defined in the scenario (2022–2032)
- naturalised groundwater levels (1990–2032).

For each predictive scenario, the report discusses the modelling outcomes and presents the differences in shallow groundwater levels between the predictive scenario and the null scenario as a map, e.g., Figure 4-31. Additionally, the report includes a bar graph comparing the average model water budget for the last year of predictions (2031–2032), e.g., Figure 4-32.

The predictive modelling presentation is adequate and commensurate with the objectives of the modelling project.

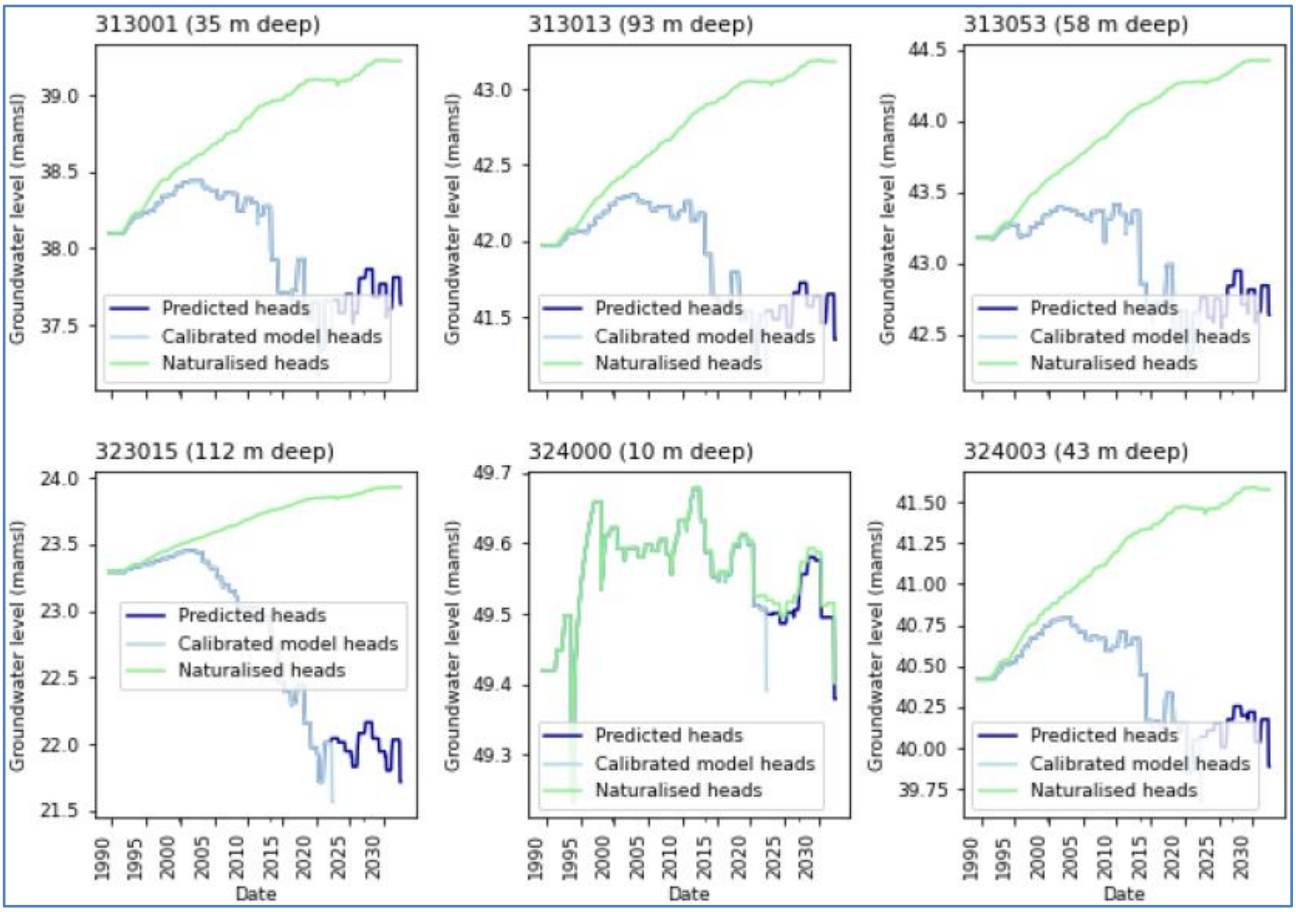


Figure 4-30. Example on groundwater level hydrographs for a predictive scenario (extracted from Figure 15 in PDP, 2023).

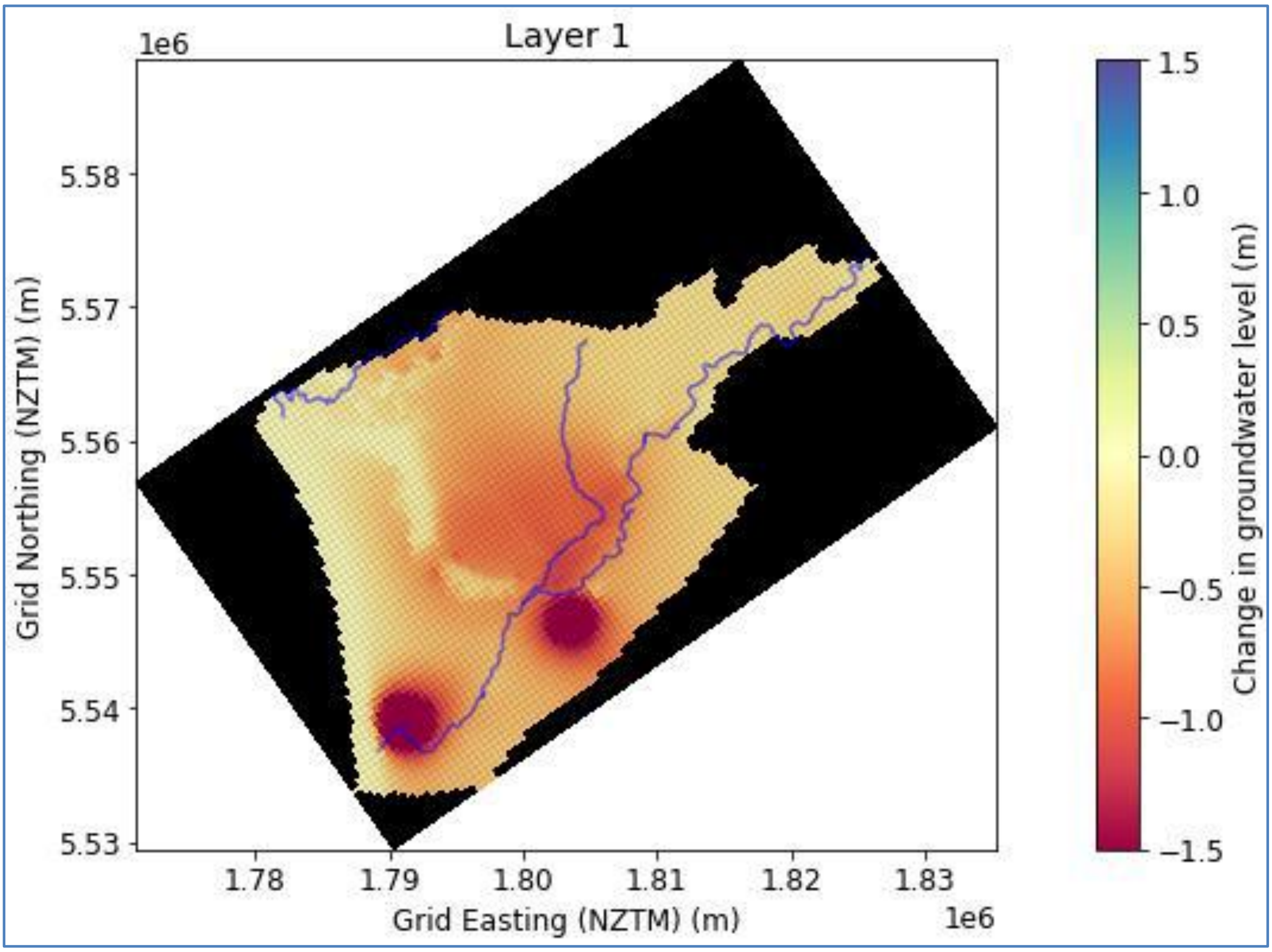


Figure 4-31. Example on map presenting the difference in the shallow groundwater levels between the null and predictive scenarios (Figure 16 in PDP, 2023).

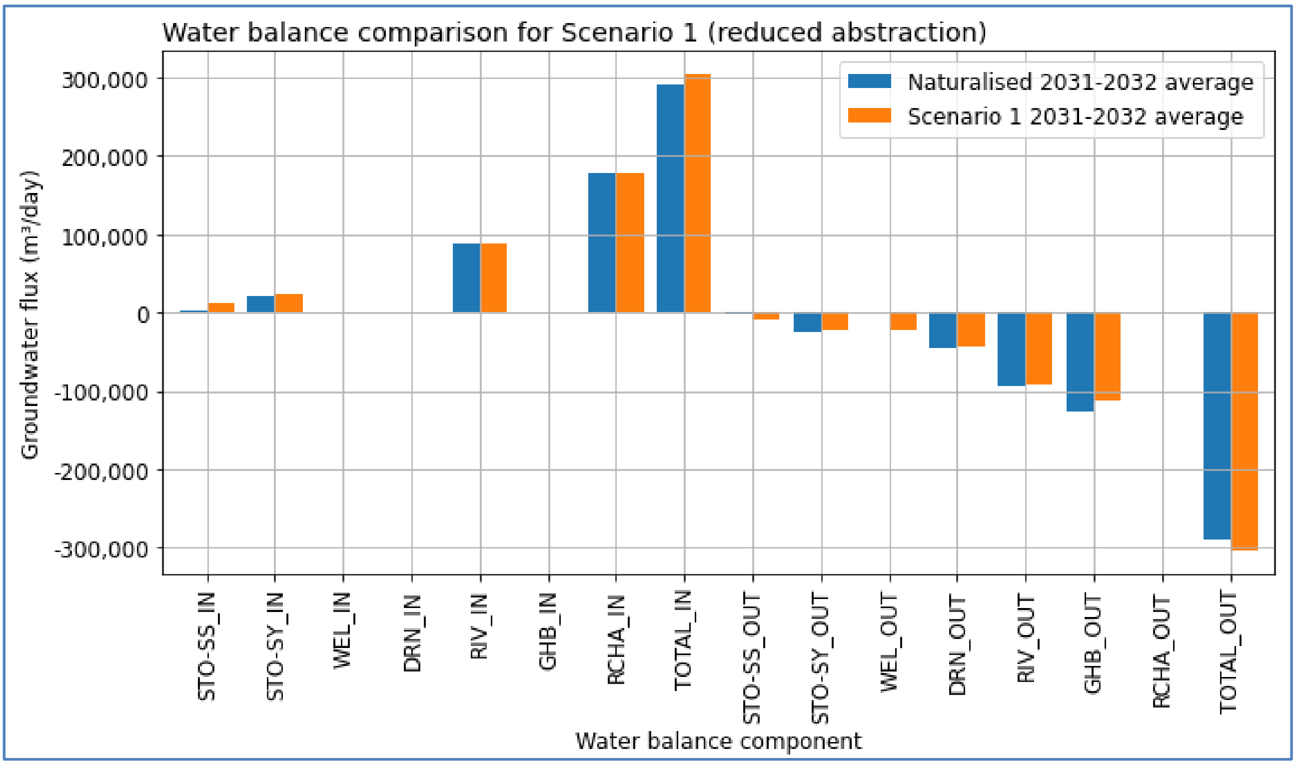


Figure 4-32. Example on the change in water balance between the null and predictive scenarios (Figure 17 in PDP, 2023).

4.5.3 Have the Results Been Checked for Anomalous Results?

The predictive modelling results do not appear to include counterintuitive water balance estimates. Depending on the scenario, there could be significant differences from the baseline. However, the objective of the modelling is to explore these differences, and large deviations from the baseline are anticipated for some scenarios.

4.5.4 Scenarios and Impacts Assessment, Including Cumulative Impacts

The model reports and an examination of the model files show that historic and predictive modelling encompass all identified stresses in the model domain. Thus, cumulative impacts are included in the modelling.

4.5.5 Predictive Scenarios Uncertainty

PDP (2023) does not address the uncertainty in predictive scenarios. Section 4.5.1 above explains that the scenarios assume climate conditions from the past 10 years will repeat over the next 10 years, an assumption that is inherently uncertain. Predictions are based on assumed abstraction levels and consented limits, which adds further uncertainty, especially since irrigation demand, the dominant groundwater use, is climate dependent. Future modelling should consider additional deterministic scenarios and incorporate stochastic modelling. Scenario uncertainty, along with other discussed modelling uncertainties, results in a complex and challenging-to-assess uncertainty landscape.

4.6 Predictive Uncertainty

According to Barnett *et al.* (2012), because all models have uncertainty, it follows that no model output should be reported as a single model result unless that single result is accompanied by a due-diligence effort at estimating the associated expected uncertainty.

DPE (2022) explains that there are three basic types of uncertainty analysis (Figure 4-33):

1. Scenario analysis with deterministic modelling
2. Linear uncertainty analysis with deterministic modelling
3. Bayesian analysis with stochastic modelling.

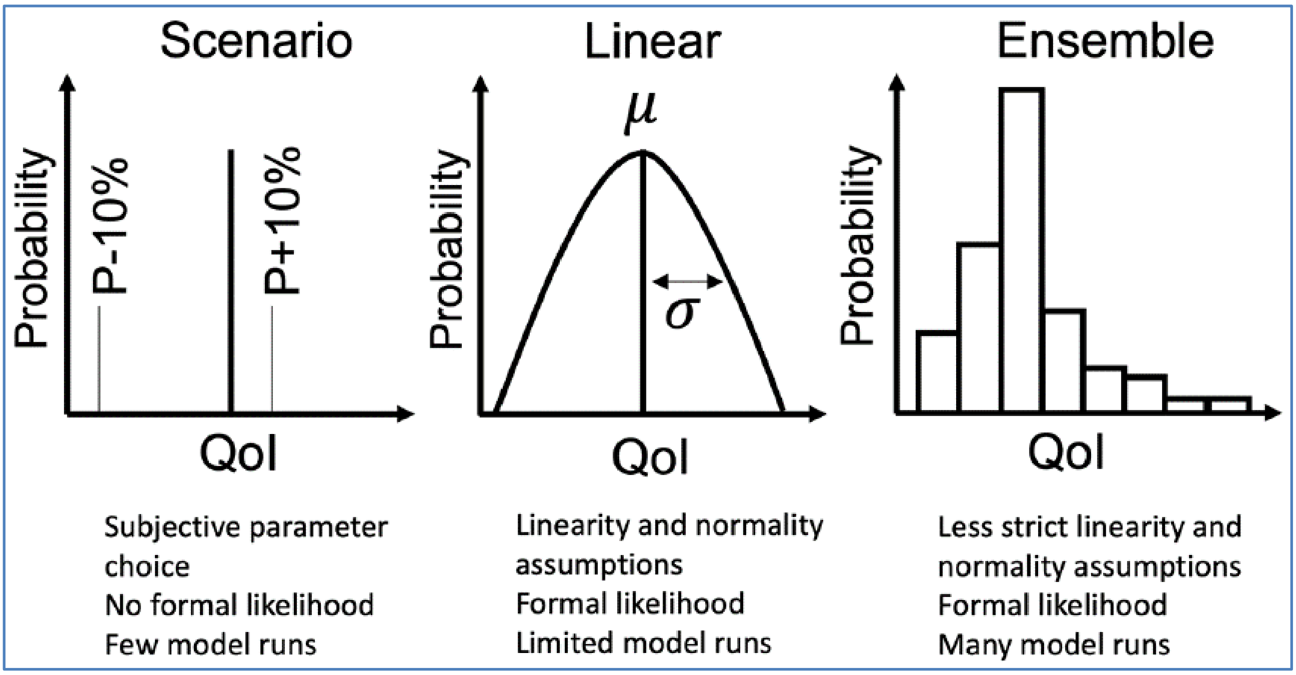


Figure 4-33. Types of uncertainty analysis in groundwater modelling (Peeters & Middlemis, 2023).

4.6.1 Are the Sources of Predictive Uncertainty Discussed?

PDP (2023) discussion of predictive uncertainty is limited to expecting predictions of groundwater abstraction effects on surface water to be less certain than predictions of effects on groundwater levels.

Section 7.2 in PDP (2017) is on prediction uncertainty. It relates uncertainty in the model predicted groundwater levels to the uncertainty in the calibrated model. It analysed uncertainty for the forecast period 2015 to 2020 in the same way as it did for the calibrated model. It concludes that '*the error variance of the forecast decline is relatively small, and imply that a declining trend is within the range of uncertainty.*' The report also acknowledges other sources of uncertainty including uncertainty of observations and parameters (e.g., hydraulic conductivity from pumping test estimates of transmissivity), and missing or inaccessible data that were estimated in the modelling (e.g., surface water data). The report does not discuss model conceptual and structural uncertainty (e.g., the model domain extent).

The discussion of the potential sources of predictive uncertainty is not adequate.

4.6.2 Assessment

Section 7.2 in PDP (2017) assesses the model uncertainty through testing the sensitivity of the winter groundwater levels gradient.

This review recommends the use of one of the uncertainty analysis methods listed in the previous section in future modelling.

4.6.3 Communicating Uncertainty

The approach to estimation of uncertainty in PDP (2017) is clearly described. Communicating predictive uncertainty analysis in future work will greatly be controlled by the selected methodology.

4.7 Surface Water-Groundwater Interaction

Surface water is very important to groundwater modelling in the RGMZ. Appendix B in PDP (2023) states: '*The model is most sensitive to variations in the river bed conductance for small streams away from the Rangitīkei River because these influence the volume of recharge to the model.*' This highlights the importance of surface water and parameters associated with it in the 2023 model. In contrast, Section 6.2 in PDP (2017) states: '*… conductance of the drains, rivers and general head boundary are not key model parameters in terms of matching the observed groundwater levels.*' The new report does not highlight or justify this difference between the two models. It is noted that despite such a critical difference, the two

models produce very similar groundwater head estimates, highlighting the concept of non-uniqueness (different models can produce similar results). However, the predictions of different models most likely differ.
Importantly, the models' boundaries do not respect hydrological principles; they cross surface water divides (catchment boundaries) and cut through waterways (Figure 4-34).

The conceptual and numerical models lack vital information on surface water. For example, flow rates, base flow (groundwater losses to surface water), groundwater gains from surface water (river recharge), stage (water level), bed elevation (to provide with stage depth of water), and bed conductance (c) are missing. Hence, the models are not calibrated to surface water stage or flows. The sensitivity of the models to surface water stage and bed elevation was not tested.

PDP (2017) describes the general hydrological environment in the RGMZ area but provides no detailed information on surface waterways. The reports mainly discuss the groundwater-surface water interaction based on groundwater level data (shape of piezometric contours and the presence of free artesian flow). Piezometric contouring did not incorporate surface water one-off measurements, continuous monitoring, or concurrent gauging. The hydrological interpretation offered makes only superficial use of hydrological studies. For example, PDP (2017) articulates '*gauging surveys along the Rangitīkei River do not obviously imply that there are large losses from the major rivers (i.e., the Rangitīkei and the Turakina Rivers) that would have large effects on groundwater levels.*' This opinion is reiterated in different sections in PDP (2023). However, it is not supported by evidence. Figure 2-1 suggests that the Rangitīkei River's losses and gains are measurable, i.e., not small.

The water budgets presented in PDP (2023) for the calibrated model and predictive scenarios (e.g., Figure 4-32) show that the features represented as rivers and drains (DRN_OUT, RIV_OUT, RIV_IN) have a greater impact than groundwater abstraction (WEL_OUT), which is assumed to be a major stress on the groundwater system by the conceptual model. This indicates that future models should focus on improving the representation of these features.

Dune lakes in the Santoft area are ecologically and culturally significant features that could also be critical components of the hydrogeological system. However, they are not represented in the groundwater model, and the modelling report fails to discuss their relevance and importance within the model. Without explicit representation, the interactions between dune lakes and the groundwater system, as well as the effects of groundwater abstraction on these lakes, remain unknown.

As mentioned above, there is an overlap of boundary conditions representing surface waterways in some cells in the model that can be problematic to the numerical models, making the models unrealistic and/or creating numerical problems. This is not noted in the reports or shown in Figure 14 in PDP (2017) reproduced here as Figure 4-35 but can be spotted from comparing Figure 4-15 and Figure 4-16.

Overall, the surface water-groundwater interaction is not adequately conceptualised in relation to the area's hydrology and model objectives, nor is it effectively characterised and quantified in space and time for satisfactory incorporation into the numerical model.

The above discussion highlights a high level of uncertainty in the models' representation and predictions regarding surface water. This is important to note because PDP (2023) indicates that model-predicted '*declines in shallow groundwater levels… will affect surface water bodies such as the coastal lakes*'. This could be an important water management consideration for Horizons and the community.

While PDP (2017) recommended the assessment of the sensitivity of surface water bodies to reductions in discharge from groundwater, including the assessment of the ecology and community values, it did not recommend data gathering and investigations to enable better surface water representation in groundwater models to develop groundwater level thresholds to protect surface water bodies.

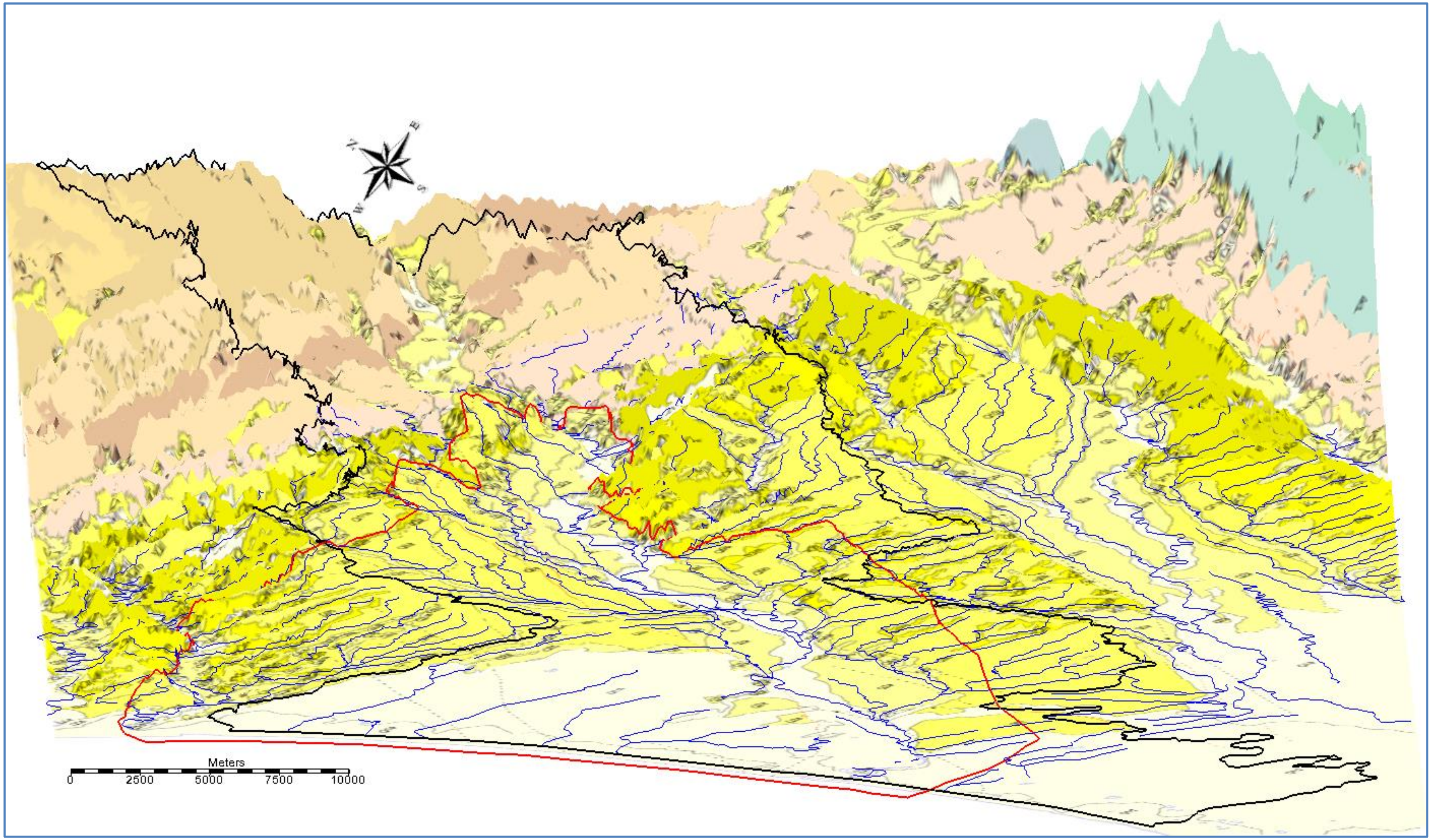


Figure 4-34. Three-dimensional representation of the geology, groundwater management zone boundary (black line), model domain (red line), and rivers and streams (blue lines).

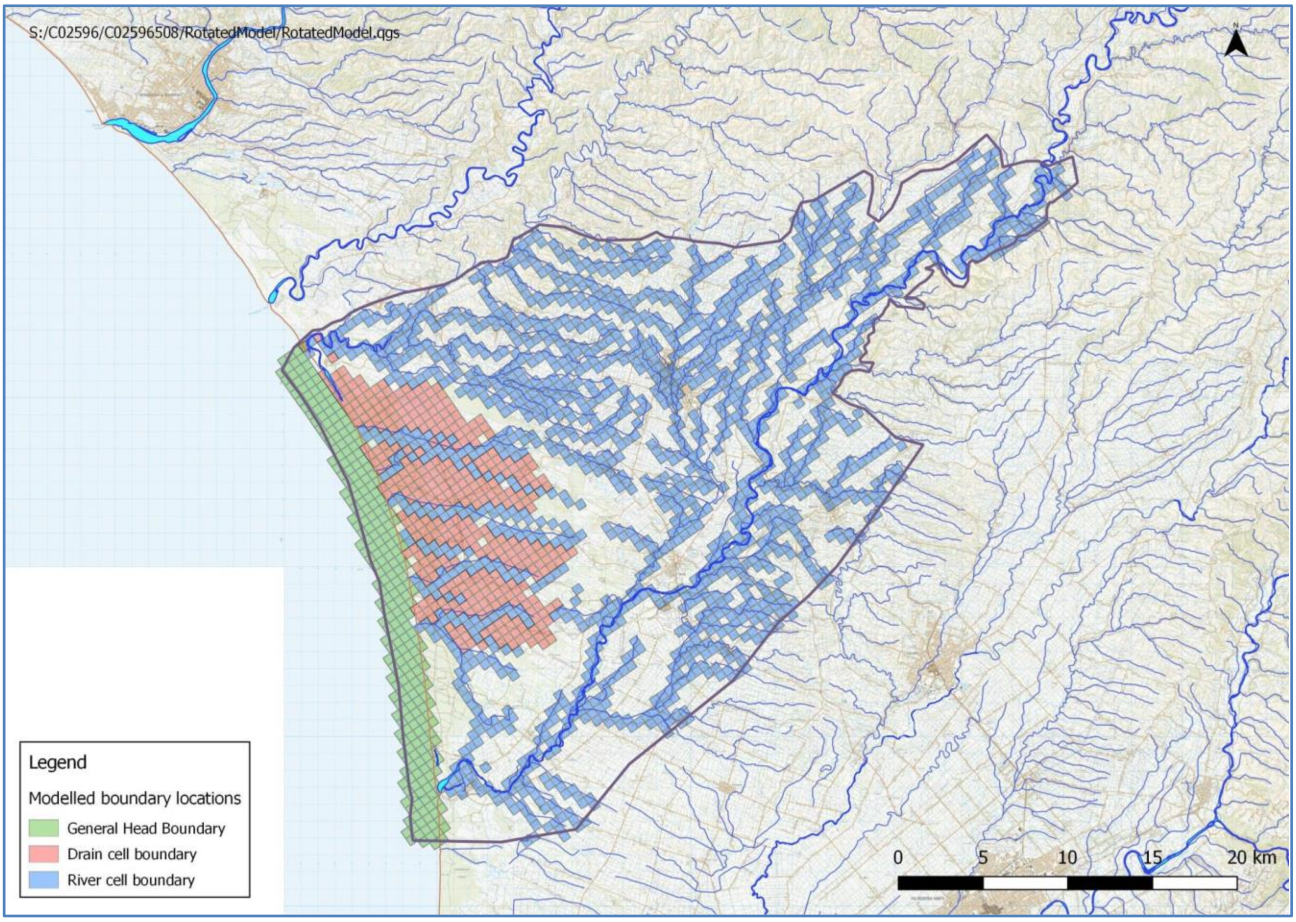


Figure 4-35. Location of modelled river, drain and general head boundaries (Figure 14 in PDP, 2017).

4.8 Reports Adequacy

As indicated earlier, PDP (2017) contains more detailed information compared to PDP (2023). To gain a full understanding of the 2023 model, reviewing the 2017 model report is essential. Best practice guidelines suggest that groundwater modelling reports should be both standalone and comprehensive, offering all the necessary information in a single document.

PDP (2017) and PDP (2023) are well-structured according to their respective objectives as reflected in their titles. The first report focuses on modelling, whereas the second focuses on allocation. As a result, the modelling aspects are not adequately described in PDP (2023). Both reports are well written and prepared to professional standards.

PDP (2017) and PDP (2023) present groundwater level and hydraulic properties data in map and time series plot formats. Future reports must also consider presenting vertical changes in the data using cross sections, slices, sectional views, and fence and block diagrams. This would help in understanding the system and the model better.

In PDP (2017), some grid-based maps suffer from clarity issues due to the chosen colour coding schemes (e.g., hydraulic conductivity and groundwater difference maps).

Some figures in the modelling reports must be re-designed. For example, the map displaying transmissivity values from pumping tests, presented as Figure 10 in PDP (2017), uses circles of proportional sizes to represent different value ranges, with divisions set at 0–100, 100–500, 500–1000, etc. However, this creates ambiguity for the reader, as it is unclear whether a value of 100 is represented by the 0–100 symbol or the 100–500 symbol. This example underscores the importance of thorough proofreading of the modelling reports, in addition to conducting technical model reviews, to ensure clarity and accuracy in data presentation.

Modelling reports need to be transparent, accurate, and include all necessary components at an appropriate level of detail. While it is understandable and acceptable to rely on PDP (2017) for information on the 2023 model, any deviations from the descriptions presented in the earlier report should be clearly captured in the new report. In previous sections, some discrepancies between the model descriptions and the results of model inspections have been noted, which could affect confidence in both the report and the model.

It is noticed that PDP (2017) does not present a list of figures in the contents section. This is a common component of reports and helps the readers. Furthermore, it noted that PDP (2017) and PDP (2023) do not include a list of cited references. This is a critical component of technical reports.

The reports do not present all used data in a manner that would enable reviewers and analysts to fully understand the model design and outcomes. Future versions of the RGMZ modelling reports must include all data used in hydrogeological analyses and modelling in paper and/or electronic appendices.

4.9 Model Confidence Level Classification

In the AGMG, Barnett *et al.* (2012) proposed a confidence level classification (CLC) scheme for groundwater models, categorising them into Class 1, Class 2, or Class 3 based on factors such as data quality, calibration rigour, prediction reliability, and key indicators. This framework is designed to guide the appropriate use of models, helping stakeholders understand their reliability and limitations, and establishing clear expectations for model development.

Class 1 models, typically developed with limited data, are suitable for lower-risk scenarios or preliminary assessments. In contrast, Class 2 and Class 3 models, which require more detailed data and rigorous calibration, are used for higher-risk assessments and critical decision-making. This classification system influences the model development process and assists reviewers in evaluating whether the model meets the agreed-upon criteria. It is essential to reassess the confidence level classification throughout the model's development to ensure it remains fit for purpose.

Guiding Principle 2.3 in the AGMG (Barnett *et al.*, 2012) recommends that a target model CLC be agreed upon and documented early in the project to clarify expectations. However, this practice is often neglected, with CLC assessments frequently conducted only after a model's completion, which can give the impression that it is a secondary consideration. This issue is evident in the 2017 and 2023 models, where not only was a target CLC not defined, but CLC assessments or similar evaluations were entirely overlooked.

In this section, the CLC for the 2023 model is assessed (Table 4-4). The aim of this assessment is to evaluate the model's potential applications and identify its limitations. The criteria are traffic-light coloured—green represents Class 3, amber Class 2, and red Class 1.

The CLC is intended as a semi-quantitative or qualitative rather than a quantitative assessment. The analysis in Table 4-4 indicates that the 2023 model meets several Class 2 criteria, incorporates some Class 3 features, but also aligns with important Class 1 criteria. Therefore, the model can be described as a Class 1–2 model with some Class 3 attributes.

According to Table 2-1 in the AGMG (Barnett *et al.*, 2012), a Class 3 model is required for providing reliable information for sustainable yield assessments in high-value regional aquifer systems, such as the groundwater system in the RGMZ. A Class 3 model is also necessary for accurately simulating groundwater-surface water interactions to the extent required for dynamic linkage to surface water models. While Class 2 models are appropriate for predicting impacts in medium-value aquifers, and Class 1 models for long-term impact predictions in low-value aquifers, it is evident that a Class 3 model is essential to meet Horizons' and the community's groundwater management objectives in the RGMZ in general and the Santoft area in particular.

Table 4-4. Table 4-4. Model Confidence Level Classification (CLC) evaluation—adapted from AGMG (Barnett *et al.*, 2012).[9],[10],[11]

Characteristic / indicator	Class	Comment
1. Data		
• Groundwater **head** observations and **bore logs** are available but may not provide adequate coverage throughout the model domain.	2	**Head**: Groundwater head data come from operational bores, which could be impacted by pumping effects and measurements may not reflect fully recovered heads (SWL). In addition, there are not enough bore pairs or clusters to help understand the 3D groundwater head distribution. **Bore logs**: Class 1 – See below.
• Metered groundwater-extraction data may be available but spatial and temporal coverage may not be extensive.	2	There are metered extraction data, the quality of which is unknown. They do not provide full coverage over the model calibration period. There is a need to fill data gaps and estimate unavailable data.
• Few or poorly distributed existing wells from which to obtain reliable **groundwater** and **geological** information.	1	**Head**: Class 2 – See above. **Bore logs**: Well logs are not adequate in terms of areal distribution, depth coverage, and quality of description to enable clear definition of the aquifer geometry or characterisation of the groundwater system. The model is not informed by and does not represent the geology.
• Climate data only available from relatively remote locations.	1	Measured and modelled climatic data used to estimate recharge but values seem to be unrealistic and cannot be verified. Evapotranspiration (from shallow aquifers) is not represented in the model. Irrigation data are not available, not estimated from climate and land use data, and not incorporated in the model.
• Little or no useful data on land-use, soils or river flows and stage elevations.	1	**Land-use and soils:** No information provided in the model report or files on land-use and soil. These data are assumably used in rainfall recharge calculations. However, the rainfall recharge data inspected in the 2023 model files are not considered realistic. Irrigation is not represented in the model. **River data:** The model does not represent river flows and stage elevations.
2. Calibration		
• Long-term trends are adequately replicated where these are important.	3	Largely, long-term trends are replicated, notwithstanding estimation residual.
• Seasonal fluctuations are adequately replicated where these are important.	3	Largely, seasonal fluctuations are replicated, notwithstanding the magnitude of fluctuation is not adequately reproduced.
• Transient calibration is current, i.e., uses recent data.	3	Model calibrated against groundwater heads up to July 2022.

[9] Table 2-1 in the AGMG (Barnett *et al.*, 2012) confuses the terms 'validation' and 'verification'. Herein, the term validation in the first column means verification, i.e., post-audit.
[10] This table lists only the characteristics and indicators applicable to the 2023 model to determine the corresponding Confidence Level Class (CLC).
[11] The criteria are colour-coded: green indicates Class 3, amber indicates Class 2, and red indicates Class 1.

Characteristic / indicator	Class	Comment
• Validation is either not undertaken or is not demonstrated for the full model domain.	2	Post-audit not undertaken on the 2017 model as part of the development of the 2023 model, despite adequate groundwater head data availability for this purpose (July 2014–July 2022; eight years' worth of data). In addition, all the surface water data records can be used for the 2017 model post-audit because they were not used in its calibration.
• Calibration statistics are generally reasonable but may suggest significant errors in parts of the model domain(s).	2	Calibration statistics not presented for the 2023 model. The groundwater head calibration hydrographs and scatter plots indicate reasonable calibration. However, the model is considered to be a poor, unrealistic representation of the system.
• Long-term trends not replicated in all parts of the model domain.	2	See Figure 3-3.
• Transient calibration to historic data but not extending to the present day.	2	Calibration against groundwater head extends to July 2022. This was reasonably current at the time of the development of the 2023 model.
• Seasonal fluctuations not adequately replicated in all parts of the model domain.	2	See Figure 3-3.
• Observations of the key modelling outcome data set are not used in calibration.	2	This includes groundwater-surface water exchanges and surface water levels (stage).
• Calibration illustrates unacceptable levels of error especially in key areas.	1	Error in surface water-groundwater exchanges and coastal groundwater discharge are unknown/unreported.
• Calibration is based on an inadequate distribution of data.	1	There are not enough monitoring locations in the area around Marton and generally in the area covered by sediment older than the Last Glacial strata (Figure 3-2). In addition, the monitoring depth coverage is poor.
• Calibration only to datasets other than that required for prediction.	1	Surface water-groundwater interaction is a key groundwater management consideration. However, surface water data are not appropriately used in the model setup or included in the model calibration and sensitivity analysis.
3. Prediction		
• Length of predictive model is not excessive compared to length of calibration period.	3	Calibration period: 1990–2022, i.e., 32 years Prediction period: 2022–2032, i.e., 10 years
• Temporal discretisation used in the predictive model is consistent with the transient calibration.	3	Monthly stress periods in both. Notwithstanding, the data quality is questionable because all the monitoring network bores are operational bores (mainly irrigation bores).
• Level and type of stresses included in the predictive model are within the range of those used in the transient calibration.	3	The main modelled stress types are rainfall recharge and groundwater abstraction. Rainfall recharge in the calibrated model is replicated in the predictive models. Abstraction in the predictive base line scenario is the same as the calibrated model. Abstraction is decreased and increased to acceptable levels in other scenarios.
• Steady-state predictions used when the model is calibrated in steady-state only.	3	The model is calibrated in transient mode and transient predictions are presented. No steady state predictions are made.
• Validation suggests relatively poor match to observations when calibration data is extended in time and/or space.	2	The 2017 model was not verified. However, it is expected that post-audit of both the 2017 and 2023 models would show relatively poor match to observations when calibration data are extended in time and/or space.
• Model validation suggests unacceptable errors when calibration dataset is extended in time and/or space.	1	Post-audit not undertaken.

Characteristic / indicator	Class	Comment
4. Key indicators		
• Model predictive time frame is less than 3 times the duration of transient calibration.	3	Yes
• Stresses are not more than 2 times greater than those included in calibration.	3	Yes
• Temporal discretisation in predictive model is the same as that used in calibration.	3	Yes
• Mass balance closure error is less than 0.5% of total.	3	Yes. However, this is not a suitable measure of the performance of modern models. Modern software nearly always closes the mass balance at 0.5% of total or less.
• Not all model parameters consistent with conceptualisation.	2	There are values in the model files (e.g., HK) that are many times greater than the maximum values in presented in the report. These values are unrealistic.
• Spatial refinement too coarse in key parts of the model domain.	2	The model grid is not refined anywhere within the domain. The thicknesses of the layer are not the same, but the thickness of each layer is constant. The constant cell size of 500 m x 500 m is not suitable for representing surface water features. This has been highlighted in the review of the 2017 model but was not addressed in the 2023 model.
• Model is uncalibrated or key calibration statistics do not meet agreed targets.	1	Unrealistic parameters values and unrealistic distribution are noticed in the model files, but not reported.
• Model parameters outside the range expected by the conceptualisation with no further justification.	1	The model files include parameter values that are outside of the reported ranges. They are unrealistic.
• Unsuitable spatial or temporal discretisation.	1	The layering and zoning of Pilot Points is completely uninformed by the geology.
• The model has not been reviewed.	1	The 2017 model was reviewed as reported. Important review comments were ignored during the update (the development of the 2023 model). The 2023 model was not reviewed progressively during development.

4.10 Peer Review

In September 2016, Brian Barnett of Jacobs Group (Australia) Pty Limited (Jacobs) conducted a review of a draft report on an early version of the 2017 model. Barnett (2016) review was limited to the information provided in the report and did not include inspection of the model files. As a result, Barnett (2016) noted that it was sometimes difficult to determine whether an identified issue stemmed from a deficiency in the modelling report or a shortcoming in the model structure or approach. In contrast, this review of the 2023 model involved an examination of both the report and the model files.

Barnett (2016) review highlighted several significant concerns and provided recommendations for addressing the identified issues. A summary of Barnett (2016) review is presented below.

4.10.1 Hydrogeological Conceptualisation

Barnett (2016) expressed disappointment that the conceptualisation did not expand on the geology or utilise bore log information and geological interpretation to postulate underground geological structures relevant to the model design. Consequently, the early version of the 2017 model he reviewed was constructed with a series of constant thickness layers that bear little or no resemblance to the subsurface geology. This issue has not been addressed in either the 2017 or 2023 models. This review considers this to be a major shortcoming in the model, which must be rectified in future versions.

Barnett (2016) also noted that groundwater level trends were not systematically analysed and recommended using the cumulative deviation from mean (CDFM) monthly rainfall method to better understand the underlying causes of observed trends. This review acknowledges that the technique was adopted but it appears to have been applied incorrectly, as discussed in Section 4.2.6. The review concludes that the conceptualisation fails to explain the observed groundwater level trends adequately. This review emphasises that the modelling of rainfall recharge does not correspond well to the reported CDFM analysis.

Additionally, Barnett (2016) pointed out that the reviewed report correctly identified recharge as a key factor influencing the sustainable yield of the aquifer. He recommended incorporating the spatial variability in the depth to the water table and storage capacity in the unsaturated zone into the recharge assessment. Although the recharge calculations method for the 2023 model were adapted in response to this recommendation, the details provided on the calculation methodology are insufficient for a thorough assessment in this review. Barnett (2016) also recommended using recharge values calculated from the soil moisture model as initial rates for the groundwater model, with further adjustments considered during the calibration process, i.e., not consider them definite values. This review notes that this recommendation was ignored in both the 2017 and 2023 models—recharge was neither a calibration parameter nor included as an uncertain parameter in an uncertainty analysis. This shortcoming must be addressed in future models.

4.10.2 Model Design

Barnett (2016) recommended adjusting the model's layering configuration to better represent the geology. This recommendation has not been adhered to in either the 2017 or 2023 models. This review strongly supports Barnett (2016) recommendation, asserting that future models must reflect the surface and subsurface geology, even if only at a conceptual level.

Barnett (2016) also noted that the coarse grid used in the model (500 m x 500 m cell size) was inappropriate for modelling rivers and streams. He highlighted that the combined area covered by the river cells in the model far exceeds the actual surface area of the rivers and streams they aim to simulate. Barnett recommended re-evaluating the river boundary conditions and possibly re-conceptualising them to eliminate the clear discrepancies between the heads assigned to tributaries and those assigned to the main rivers. However, this recommendation was ignored in both the final version of the 2017 model and the 2023 model. This review is of the opinion that a finer grid is necessary to enable a realistic representation of surface water features, which could easily be achieved using an unstructured grid.

Since Barnett (2016) did not have access to the model files, his review did not detect the erroneous assignment of boundary conditions, including overlapping boundary conditions, and discrepancies between the reported and actual model designs—issues that have been identified in this review.

Last but not least, based on work by GNS in 2014, which was not reviewed as part of this evaluation and is not referenced in PDP (2017) and PDP (2023), Barnett (2016) considered groundwater evapotranspiration to be an important component of the catchment water balance, particularly where the water table is shallow. He noted that models ignoring this process often compensate in other mass balance components, leading to errors and unrealistic behaviour in the groundwater system. Barnett recommended incorporating the MODFLOW EVT package into the model, but this suggestion was ignored in both the 2017 and 2023 models. This review asserts that future models must include EVT to ensure accurate representation of groundwater processes.

4.10.3 Calibration

Barnett (2016) notes that the calibration targets are limited to time series records of groundwater head measured in an array of monitoring bores and that the parameter refinement was limited to the hydraulic conductivity distribution using PEST Pilot Points. This approach assumes estimated recharge and storage parameters are correct, which Barnett (2016) considers to be not warranted. Barnett (2016) also highlights that no justification is provided for the assumed storage values. He believes that given that the lower values may be more appropriate to enable better model calibration. These issues identified in Barnett (2016) were not addressed in the 2017 and 2023 models.

In line with Barnett (2016) recommendations, PDP (2017) included model calibration statistics in table format, but not the usual parameters required in popular guidelines like the AGMG (Barnett *et al.*, 2012). PDP (2023) does not present any model calibration metrics.

Barnett (2016) notes that the model calibration would be improved if it were expanded to include baseflow flux targets obtained from the measured streamflow in the major rivers including the Rangitīkei River.

4.10.4 Predictive Scenarios

Barnett (2016) did not comment on the appropriateness of the 2017 model's predictive scenarios. However, he noted that their timeframe (2015–2020) is relatively short compared to the calibration period (1990–2015), and the rates of groundwater abstraction were of similar magnitude to those used in the calibration. Consequently, he hypothesised that the level of predictive uncertainty would likely be similar to the level of error observed in the calibration model's match to observations.

This review emphasises that any future modelling predictive scenarios must be developed in collaboration with, and with the agreement of, Horizons and relevant stakeholders.

4.10.5 Uncertainty Analysis

Barnett (2016) emphasised that current industry best practices place significant importance on acknowledging and quantifying uncertainties in groundwater models. He pointed out that a groundwater model will never provide entirely accurate predictions of future groundwater heads and fluxes. Instead, the value of a groundwater model lies in its ability to offer useful insights into future groundwater behaviour, including estimates of relative differences under varying future conditions.

Barnett (2016) concluded that the reviewed work was deficient in effectively ignoring predictive uncertainties. He recommended that modelling uncertainties be thoroughly addressed. This review notes that these recommendations were not implemented in either the 2017 or 2023 models. Adequate uncertainty analysis must be included in any future modelling.

4.10.6 Reporting

The report edition reviewed by Barnett (2016) lacked a clear statement of the modelling objectives, a shortcoming that was corrected in the 2017 and 2023 reports.

Barnett (2016) commented on specific discrepancies in the report he reviewed, including technical issues such as the occurrence of dry cells. His discussion suggests uncertainty about whether these issues stemmed from the reporting or the model setup, a point he noted at the outset of his review.

It appears that the authors of the 2017 report attempted to address some of Barnett (2016) review comments, such as the need for hydraulic conductivity maps for all layers. However, this review identified residual errors and previously undetected errors in the 2017 and 2023 reports, which must be avoided in future modelling reports.

Barnett (2016) did not notice that the report version he reviewed lacked a 'references' section. The final 2017 report and the reviewed 2023 report do not have a references section, which this review considers to be a serious reporting shortcoming.

4.10.7 Conclusions

Barnett (2016) summarised issues of concern he identified in table format. His assessment is guided by the AGMG (Barnett *et al.*, 2012). He mainly requires:

1. reconsideration of recharge and including it in the uncertainty analysis
2. re-evaluated and reconceptualization of the rivers and streams
3. use of the MODFLOW EVT Package
4. calculation and reporting of calibration statistics
5. recalibration of the model with possible redefinition of targets and the inclusion of additional variables like recharge and storage parameters
6. analysis, quantification, and reporting of predictive uncertainty
7. definition of the modelling objectives.

Ultimately, Barnett (2016) concludes that the model he reviewed is not fit for purpose.

4.11 Model Fitness for Purpose

The main objective of a models' review is to assess its fitness for purpose. The objective of the reviewed models is predicting the impacts of groundwater abstraction on groundwater and surface water. The modelling objectives should have included exploring and evaluating all possible reasons for the anecdotal groundwater level decline.

In addition to the discussion presented in various sections of this review report, the 2023 model confidence level classification was assessed (Section 4.9) and the model attributes were evaluated in checklist format (Appendix I), both adapted from the AGMG (Barnett *et al.*, 2012).

While the reviewed model exhibits excellent calibration signs, it is believed to suffer from symptoms that render it unfit for its stated purpose. Nonetheless, it provides a solid basis for further modelling that would meet the identified objectives. The model confidence level is estimated as Class 1–2, with some Class 3 attributes (Section 4.9). This indicates the need for improvement for the model to become fit for the purpose of predicting the impacts of groundwater development in the RGMZ and identifying the key factors that determine these impacts.

5. Conclusion

The One Plan is the Regional Policy Statement and Regional Plan for the Manawatū-Whanganui region. It defines groundwater management zones and sets the rules for their management through various tools, including allocation limits. The groundwater allocation limit in the Rangitīkei Groundwater Management Zone (RGMZ) is 75 million cubic metres per year.

Noticeable groundwater take consents and abstraction commenced in the RGMZ about 20 years ago (around 2005). Groundwater abstraction in the RGMZ has been dominated by irrigation. Since around 2015, there have been anecdotes about groundwater resource overallocation and associated groundwater level decline in parts of the RGMZ. Horizons and the RGMZ communities considered that the future trend of groundwater levels in the Santoft area was uncertain. Hence, Horizons embarked on groundwater resource assessments, including commissioning the Santoft Groundwater Model in 2015, which is presented in PDP (2017). The key questions for the model were whether a groundwater level decline was likely to continue under existing rates of abstraction or stabilise to a new steady state and, if the latter, how long that would take to happen. The 2017 model forecasted groundwater levels to further decline by about 0.4 m in the deeper strata and up to 0.25 m in the shallow strata over the forecast period of five years. It also predicted a small reduction in groundwater outflow to the coast and a small decline in groundwater discharge to rivers and streams.

In 2023, Horizons commissioned an extension to the 2017 model to bring it up to date and scrutinise the RGMZ allocation limit. The extension involved updating the model with groundwater recharge and abstraction data collected from July 2014 to July 2022. Noticeably, recharge in the new model used climate data from a different source and an adapted calculations method to account for perceived lagged recharge in deep groundwater areas. The model parameters were also changed as the model was re-calibrated (PDP, 2023). Hence, the 2023 model update is more than a straightforward extension of the 2017 model. It is rather a different model especially with the noticed discrepancies between the model description in PDP (2017) and the 2023 model files (e.g., assignment of boundary conditions to layers). The 2023 assessment indicated a long-term declining trend in groundwater levels in the RGMZ, which appears to correlate with an increase in groundwater abstraction. However, it suggested the decline does not appear to have resulted in substantial effects on surface water or the risk of saline intrusion. It predicted changes in groundwater levels under different groundwater abstraction scenarios. It concluded that if groundwater use continued at current levels or increased, groundwater levels would continue to decline. Particularly under the increased abstraction scenario, the scale of the decline implied potentially adverse effects on surface water bodies and an increased risk of saline intrusion within a 10-year timeframe. The study concluded that groundwater use and allocation are not at sustainable levels. It suggested that groundwater use may need to reduce by around a third to stabilise at a similar level to 2023 levels. The study recommended an interim allocation limit that is no more than the current level of allocation, which is around 75% of the allocation limit. This would ultimately be followed by a consents review based on further monitoring. The study also recommended a revision of the RGMZ boundary. The 2023 model and report do not provide basis for the definition of subzones within the RGMZ. Future modelling objectives can include definition of a groundwater allocation subzone for the Santoft area with an appropriate groundwater allocation limit.

Following best practice guidelines, Horizons commissioned this peer review of the models and reports. Most often, models are reviewed as reported. This review followed best practice guidelines and also inspected the model data. The objectives of the review are to confirm that the models are well-built, trustworthy, and fit for purpose, which is predicting the impacts of various groundwater abstraction scenarios on groundwater and surface water. Verification of the 2017 model was not in the scope of the review. Verification entails comparing the model predictions to subsequent observations. The 2023 model development should have started with verification of the 2017 model.

The review primarily focuses on the 2023 model's outcomes but also encompasses an evaluation of the 2017 model since it underpins the 2023 update. It assesses the models' effectiveness in representing hydrogeological conditions, their accuracy in reflecting historical groundwater behaviour, and their reliability for predicting future scenarios. Additionally, the review examines the models' assumptions (conceptualisation), methodologies, and performance to determine their suitability for current and future groundwater management needs.

The review outcome can be summarised as follows:

Reports and Data Reviewed

1. Overall, the reports are prepared to a professional standard.
2. The newer report is not standalone; the old report is required to understand the model.
3. The 2023 model files were provided in native format.
4. There are noticeable discrepancies between the modelling reports and the examined model files, e.g., boundary conditions assignments to layers and model input and output values.
5. All the data used in the modelling should have been presented in electronic appendices.

Modelling Planning and Approach

6. Numerical modelling is the best option to address Horizons' objectives.
7. The modelling incentives and objectives are clear and appropriate.
8. The objective of the model should include exploring and evaluating all possible reasons for the anecdotal groundwater level decline.
9. The software packages used for modelling and calibration are adequate and proven (MODFLOW and PEST).
10. The predictive scenarios are well defined.

Model Preparation

11. Horizons was generally well prepared for the modelling through groundwater monitoring, initial assessments, and eigen modelling.
12. The literature review and data compilation are incomplete.

Conceptual Model

13. Some elements of the conceptual model have not been transferred accurately into the numerical model, e.g., rainfall recharge in areas covered with water.
14. The conceptual model should have been presented graphically and in the form of water budgets in addition to the text description.
15. Conceptualisation should have included analytical water budgets, with a discussion of uncertainties in them.
16. There are hydrological and groundwater quality data that have not been compiled and used in conceptualisation and numerical modelling. Compliance groundwater level data were not used.
17. Conceptual model uncertainty has not been considered.
18. Uncertainty in measurements should have been discussed in the reports and considered in the models. For example, the groundwater level data are measured in operational bores, not dedicated monitoring bores. This affects the quality of the data due to potential incomplete groundwater level recovery at the time of measurement.

Numerical Model Implementation

19. The models cover a relatively large area, i.e., they are of regional scale, which required relatively coarse grid cells. This can be addressed using unstructured griding. The last model was built using the MODFLOW 6 code. However, it did not make full use of the software's capabilities, mainly the possibility of using unstructured grid. This would have enabled effective refinement of the grid at areas of interest.
20. The model spatial and temporal discretisations do not accommodate the characteristics of the data available and the requirements of the features of interest.
21. The model peripheral boundaries do not coincide with those of the RGMZ. They are not technically defensible, creating a resource management conflict.
22. The last model did not make full use of the software capabilities, mainly unstructured grid. This would have enabled effective refinement of the grid in areas of interest.
23. The model needs a geological framework.
24. There are some errors in the setup of boundary conditions. For example, there are cells where recharge, drain, and river conditions overlap, rendering the numerical model unrealistic and potentially creating numerical instabilities and errors. In addition, there are boundary conditions assigned totally in implausible locations like DRN cells in the ocean
25. Some discrepancies are noticed between the reported and model boundary conditions and parameters. For example, river depth and stage and drain depth in the report are different from those in the model.
26. The model layer structure is not defensible.
27. The location of the coastal boundary in the sea and its parameterisation are inadequate. They do not properly consider seawater density.
28. There seems to be prejudice that affects the modelling approach with regard to the main controls on groundwater levels. This must be resolved through systematic analysis and cross-examination of all relevant data.
29. Incorporation of evapotranspiration from the water table in the model must be considered. This is different from evapotranspiration from the soil and evaporation from open water bodies.

Model Calibration and Sensitivity Analysis

30. The model update included recalibration.

31. The numerical model sensitivity analysis is incomplete. It does not include many important parameters and boundary conditions' locations and properties.
32. The model calibration statistics are within guideline bounds, and the match between model calculations and observations in both models is good. Generally, the models simulate well the long-term and seasonal groundwater trends. This highlights the general non-uniqueness of models and must not be taken as the key and only measure of the models' predictive capabilities.
33. Some parameters should have been used in the model calibration but were excluded.
34. The model calibrated hydraulic conductivity displays bullseyes, which is often a symptom of unrealistic automated calibration.
35. Some calibrated data in the new model are out of the parameter ranges presented in the old report.
36. Recharge in the models is calculated using two different data sources. The reports do not compare the recharge values used in the models.
37. Recharge should have been included in the calibration.
38. Some interpretations are not supported by data, e.g., the relationship between groundwater abstraction, level, and electrical conductivity, albeit being within what is expected of a groundwater system.
39. Steady-state modelling results are not assessed or reported. This is important as the transient models that provide the predictions depend on them for their initial conditions. Transient models are greatly sensitive to initial conditions.
40. The models are not calibrated to surface water stage and flows.

Predictive Uncertainty

41. The modelling lacks systematic and qualitative predictive uncertainty analysis.

Additional Observations

42. The reports' recommendations should have included enhancements to the monitoring network and further hydrogeological investigations to assist with future modelling.
43. There is unclarity in some of the model parameters. For example, it is unclear whether the confined storage parameter reported in the old report is the storage coefficient (S) or specific storage (S_s). There is also unclarity in the definition of the conductance parameter, reported in m/d whereas it is commonly expressed in m²/d.
44. Some data presented in the model are not fully utilised. For example, there are bore pairs that could be used to characterise vertical hydraulic gradients and, subsequently, recharge and discharge settings in certain areas.

6. Recommendations

This review recommends the following:

- Revise the RGMZ conceptual and numerical models to address identified discrepancies and inaccuracies.
- Adopt a systematic approach to modelling in line with contemporary guidelines to ensure thoroughness and accuracy.
- Achieve Confidence Level Class 2 in future models, incorporating as many attributes as possible from Class 3, according to the AGMG (Barnett *et al.*, 2012).
- Ensure future modelling reports are standalone, comprehensive, and accurate. All data used in the modelling should be presented, and the conceptual model should be depicted graphically and include analytical water budgets.
- Explore cause-effect relationships using techniques such as particle tracking and influence zone delineation.
- Report and review models progressively, including all model files in the review process to ensure transparency.
- Enhance the groundwater and surface water information base in preparation for model development. This should include developing a geological model, upgrading monitoring with dedicated bores and multi-level piezometers, and undertaking further investigations.
- Incorporate specialised surface water and geological expertise in groundwater modelling.
- Integrate all relevant data (e.g., surface water, groundwater quality) and conduct a thorough literature review in the conceptualisation and numerical modelling phases.
- Include recommendations on further data needs and improvements to monitoring networks as part of future modelling objectives.
- Start all model updates with verification of the predecessor model to ensure continuity and accuracy.
- Align the model domain with the GMZ boundaries. Any changes to the model domain should be well-justified if necessary.
- Tailor spatial and temporal discretisations to suit the model objectives, geological and hydrological settings, data, and calibration requirements. Consider using unstructured gridding (e.g., Voronoi, quadtree, layer pinch out) if appropriate.
- Investigate various options for representing the coastal boundary, including the location and seawater density.
- Avoid conflicts and double counting in boundary conditions to prevent inaccuracies and numerical instabilities.
- Include sensitivity analysis in historical modelling to adjust parameters and determine appropriate calibration targets and tolerances.
- Incorporate predictive uncertainty analysis to quantify and address uncertainties in predictive modelling.
- Assess the importance of processes like evapotranspiration from the water table during the conceptualisation and numerical modelling phases.

These recommendations aim to improve the representativeness, robustness, reliability, and informativeness of the groundwater modelling process and its outcomes.

7. References

Barnett B, 2016: *Santoft Groundwater Model Modelling Review*. Jacobs Group (Australia) Pty Limited. 14 September 2016.

Barnett B, Townley LR, Post V, Evans RE, Hunt RJ, Peeters L, Richardson S, Werner AD, Knapton A & Boronkay A, 2012: *Australian Groundwater Modelling Guidelines*. Waterlines Report No. 82, National Water Commission, Canberra.

Begg JG, Palmer A & Gyopari M, 2005: *Geological synopsis of the Manawatu-Horowhenua area for a review of the region's hydrogeology*. Report prepared for Horizons Manawatu Regional Council. GNS Client report 2005/172. Project Number: 440W1159.

Bekesi G, 2001: *Manawatu-Wanganui*. Chapter 19 in Rosen MR & White PA (Editors). *Groundwaters of New Zealand*. New Zealand Hydrological Society Inc & the Caxton Press. Wellington, New Zealand, pp. 387-396.

Cherry JA & Davis RK, 1991: *Groundwater Hydrology*. Wiley.

Doherty J, 2023. *PEST: Model-Independent Parameter Estimation, User Manual*. 8th ed. Watermark Numerical Computing.

Domenico PA & Schwartz FW, 1998: *Physical and Chemical Hydrogeology*. Wiley.

DPE, 2022: *Minimum Groundwater Modelling Requirements for SSD/SSI Projects*. Technical guideline prepared for the Water Division, NSW Department of Planning and Environment as part of the Groundwater Modelling Toolbox Project. January 2022.

Fetter CW, 2001: *Applied Hydrogeology*. 4th ed. Springer.

Freeze RA & Cherry JA, 1979: *Groundwater*. Prentice-Hall.

Hiscock KM, 2005: *Hydrogeology: Principles and Practice*. Wiley.

Horizons Regional Council (Horizons), 2024: *One Plan – The Consolidated Regional Policy Statement, Regional Plan and Regional Coastal Plan for the Manawatū-Whanganui Region*. Report number: 2014/EXT/338. Horizons Regional Council. March 2024.

Mather AS & Chapman A, 2007: *Hydrogeology of the Land*. Routledge.

Melhuish A, van Dissen R & Berryman K, 1996: *Mount Stewart-Halcombe Anticline: A Look Inside a Growing Fold in the Manawatu Region, New Zealand*. New Zealand Journal of Geology and Geophysics, 39, 123-133.

New Zealand Government, 2020: *National Policy Statement for Freshwater Management 2020* (NPS-FM 2020). Ministry for the Environment.

Pattle Delamore Partners LTD (PDP), 2017: *Santoft Groundwater Model Report*. Horizons Report 2017/EXT/1556. Horizons Regional Council. October 2017.

Pattle Delamore Partners LTD (PDP), 2023: *Rangitīkei Groundwater Management Zone Allocation Limit*. Report No. 2023/EXT/1823. Horizons Regional Council. September 2023.

Peeters LJM & Middlemis H, 2023. *Information Guidelines Explanatory Note – Uncertainty Analysis for Groundwater Modelling*. A report prepared for the Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development through the Department of Climate Change, Energy, the Environment and Water, Commonwealth of Australia.

Rau GC, Acworth RI, Halloran LJS, Timms WA, Cuthbert MO, 2018: *Quantifying compressible groundwater storage by combining cross‐hole seismic surveys and head response to atmospheric tides*. Journal of Geophysical Research: Earth Surface 123(8), 1910-1930.

Rees C, Palmer J, Palmer A, Singh R, 2019: Landscape evolution and hydrogeochemical characteristics of the Pourewa Stream catchment, lower North Island, New Zealand. New Zealand Journal of Geology and Geophysics, 62(2), 147-170.

Roygard J & Carlyon G, 2004: *Rangitīkei River Water Resource Assessment – Allocation Limits and Minimum Flows*. Technical Report to Support Policy Development. Water Allocation Project. Report No: 2004/EXT/606. Horizons Regional Council.

Rushton KR & Redshaw MR, 1979: *Water Movement in Soils and Groundwater Recharge: A Theoretical and Practical Approach*. Journal of Hydrology, 41(1-2), 11-30.

Rushton KR, 2003: *The Application of Groundwater Recharge Models in Water Resource Management*. Hydrogeology Journal, 11(1), 27-37.

Rushton KR, 2004: *Groundwater Recharge Estimation Using Soil Moisture Models*. Proceedings of the International Conference on Groundwater Recharge Management, 1-12.

Scaife M & Zarour H, 2023: *Vinegar Hill Stockwater Optioneering – Groundwater*. Prepared for: Farmers Liaison Group / Manawatū District Council. Project Number: 310003323. Stantec. July 2023.

Scaife M, 2024: *Tutaenui Reservoir Bore Assessment of Environmental Effects*. Technical Report to Support Marton Water Strategy. Report Reference: 310003260. Report prepared by Stantec for the Rangitīkei District Council.

Sophocleous M, 2002: *Introduction to Hydrogeology*. Academic Press.

Todd DK & Mays LW, 2005: *Handbook of Groundwater Engineering*. CRC Press.

Zarour H, 2008: *Groundwater Resources in the Manawatu–Wanganui Region – Technical Report to Support Policy Development*. Report No: 2008/EXT/948. Horizons Regional Council. December 2008.

Zarour H, 2018: *Characterisation and Numerical Simulation of the Lower Manawatu Catchment Hydrogeological System*. A thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Earth Sciences at Massey University, Manawatu, New Zealand.

Appendices

Appendix A Model Domain Maps

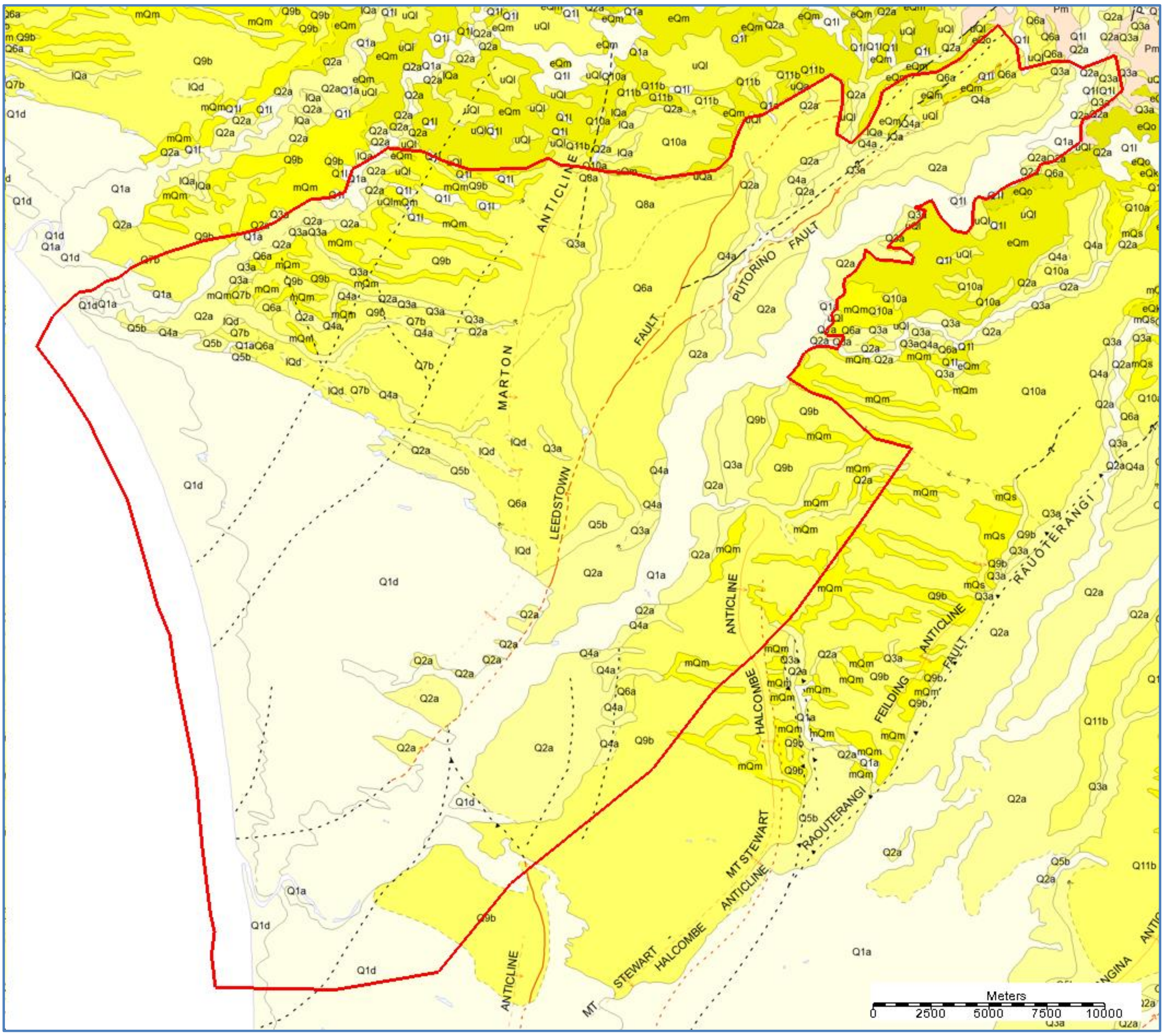


Figure A-1. Model domain extent over geology basemap.

Figure A-2. Model domain extent over topographic basemap.

Appendix B Calibrated Model Horizontal Hydraulic Conductivity Data and Maps

Table B-1. Horizontal hydraulic conductivity (KH) values in m/d in the calibrated model as reported in PDP (2017) compared to the values extracted from the reviewed 2033 model files.

Model layer	HK in PDP (2017)		HK extracted from the 2023 model files		Figures in this review
	Minimum	Maximum	Minimum	Maximum	
1	1.00	2.57	1.00	85.30	Figure B-1 & Figure B-2
2	0.69	1.04	1.00	1.47	Figure B-3 & Figure B-4
3	0.63	1.11	1.00	1.71	Figure B-5 & Figure B-6
4	0.83	2.10	1.00	351.75	Figure B-7 & Figure B-8
5	0.10	100.00	1.00	333.91	Figure B-9 & Figure B-10
6	0.10	100.00	1.00	104.96	Figure B-11 & Figure B-12
7	0.10	100.00	1.00	78.72	Figure B-13 & Figure B-14
8	0.10	0.84	1.00	14.47	Figure B-15 & Figure B-16
9	0.29	0.83	1.00	36.67	Figure B-17 & Figure B-18
10	0.59	1.07	1.00	4.47	Figure B-19 & Figure B-20
11	1.00	1.00	1.00	1.73	Figure B-21 & Figure B-22

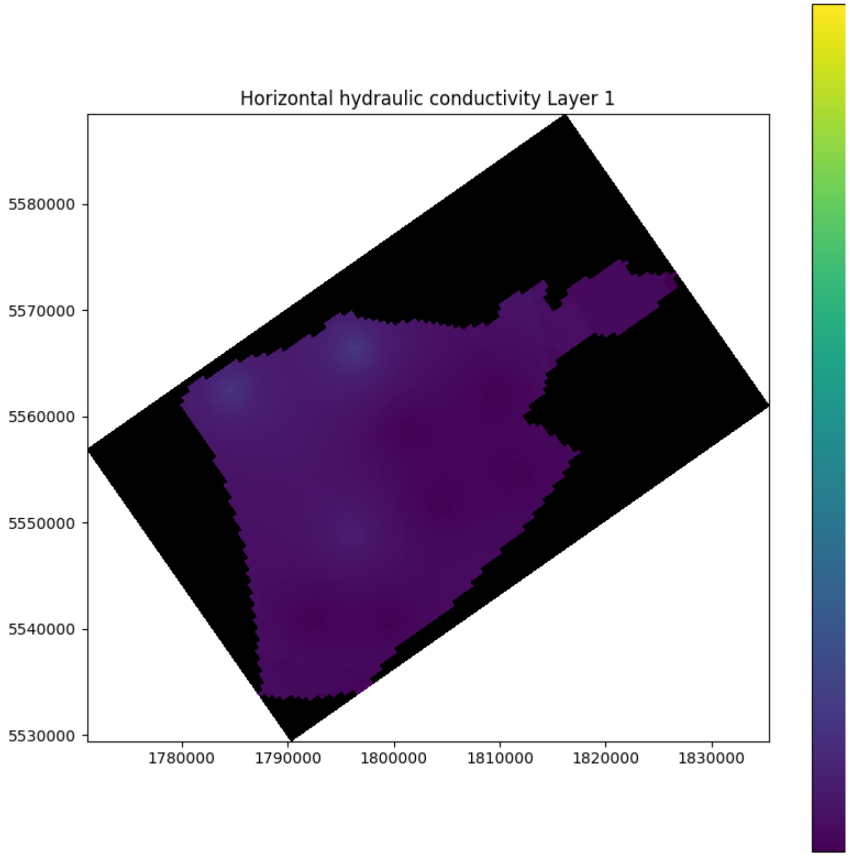


Figure B-1. Layer 1 calibrated model horizontal hydraulic conductivity (KH) as reported in PDP (2017).

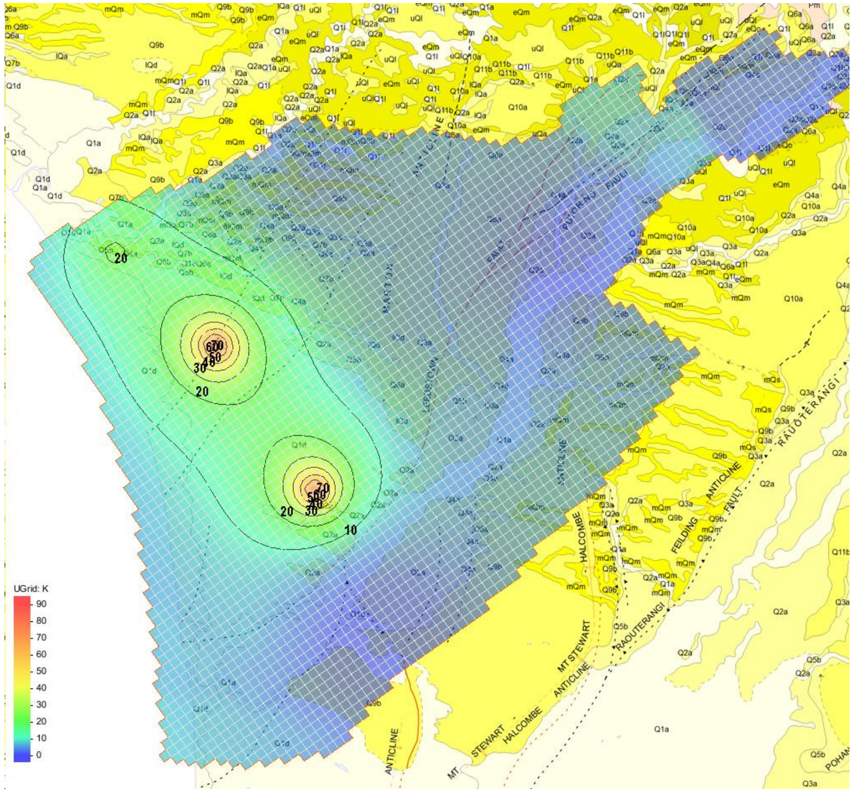


Figure B-2. Layer 1 calibrated model horizontal hydraulic conductivity (KH) extracted from MODFLOW 6 files.

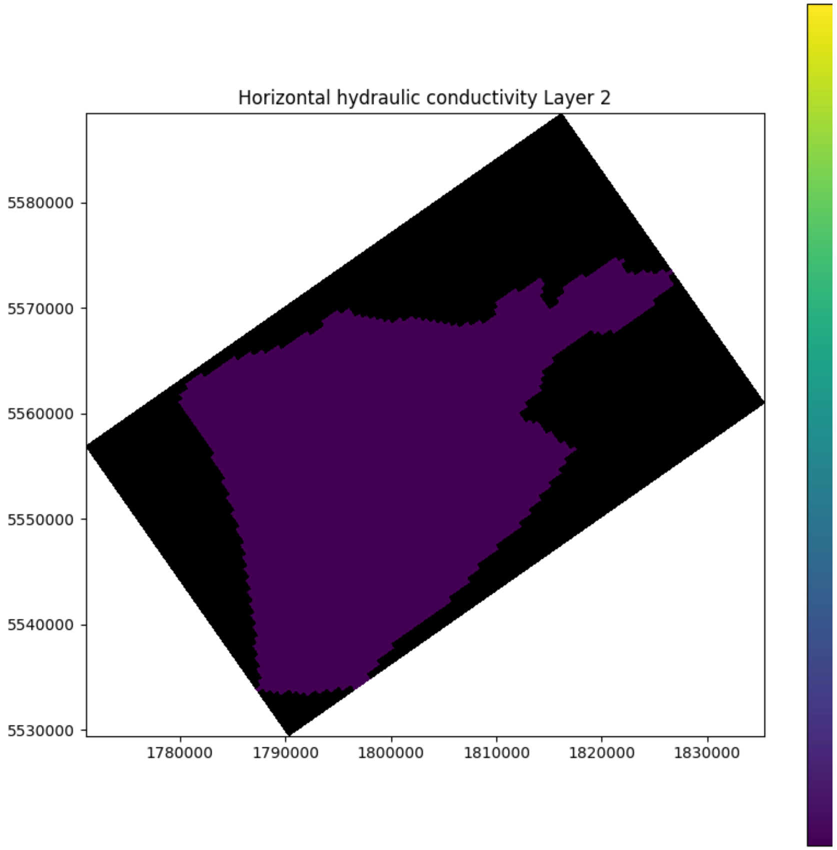


Figure B-3. Layer 2 calibrated model horizontal hydraulic conductivity (KH) as reported in PDP (2017).

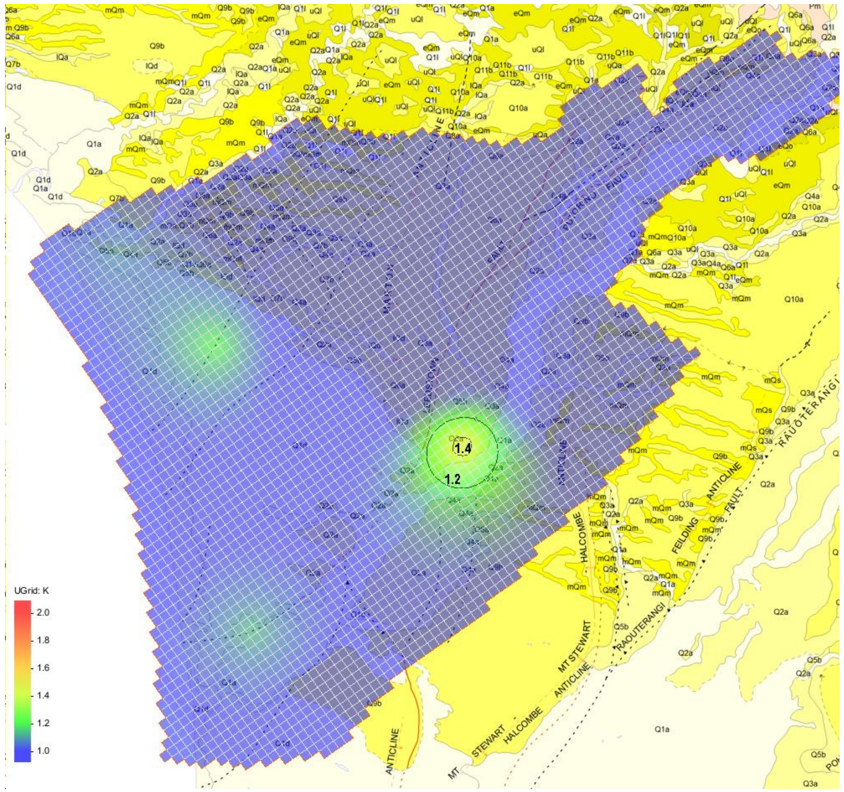


Figure B-4. Layer 2 calibrated model horizontal hydraulic conductivity (KH) extracted from MODFLOW 6 files.

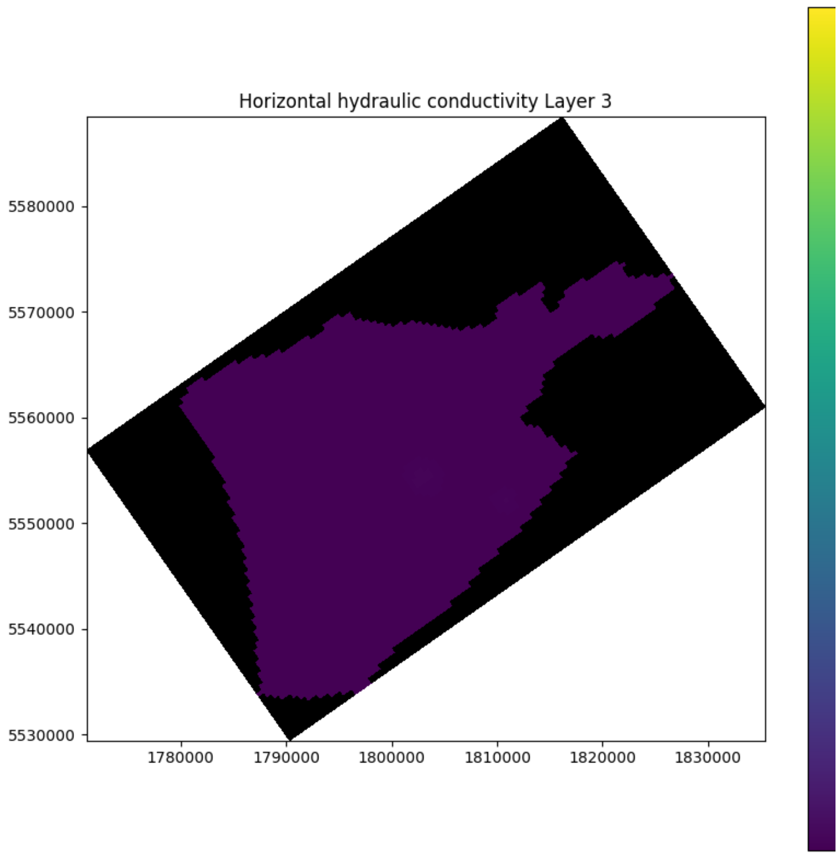


Figure B-5. Layer 3 calibrated model horizontal hydraulic conductivity (KH) as reported in PDP (2017).

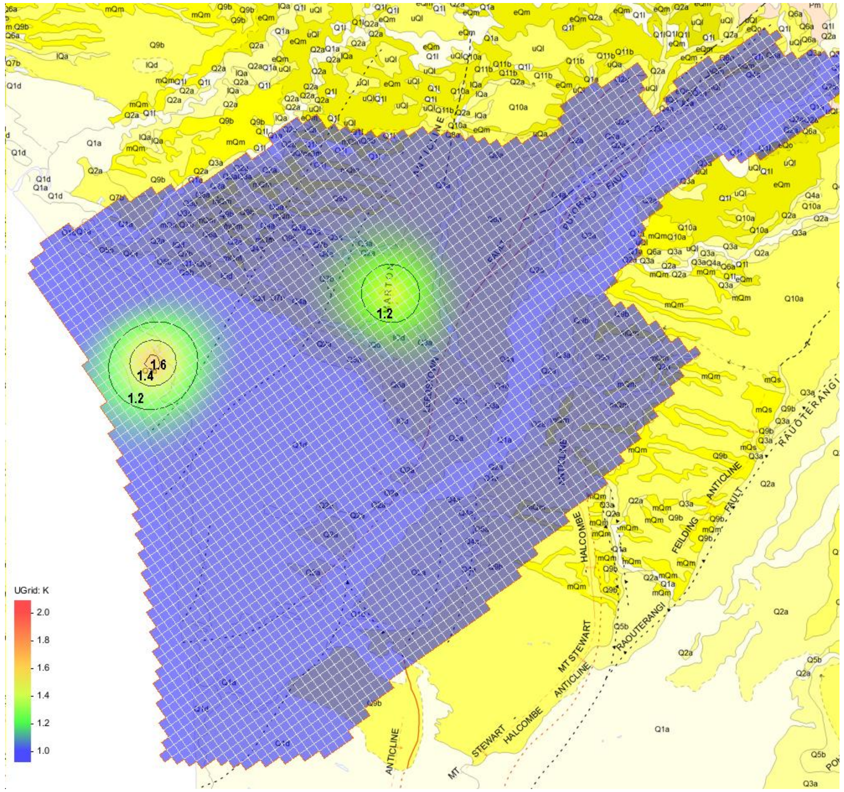


Figure B-6. Layer 3 calibrated model horizontal hydraulic conductivity (KH) extracted from MODFLOW 6 files.

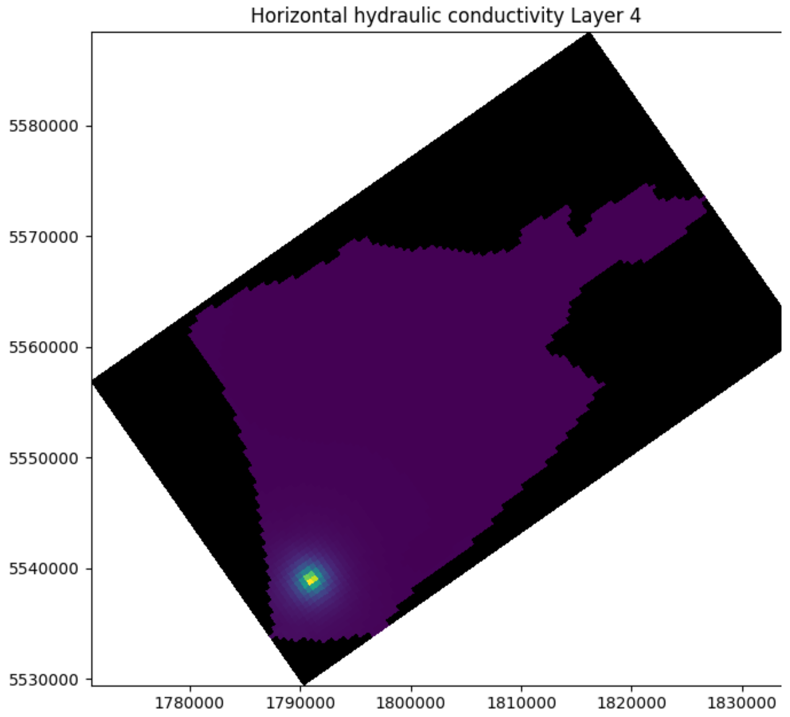


Figure B-7. Layer 4 calibrated model horizontal hydraulic conductivity (KH) as reported in PDP (2017).

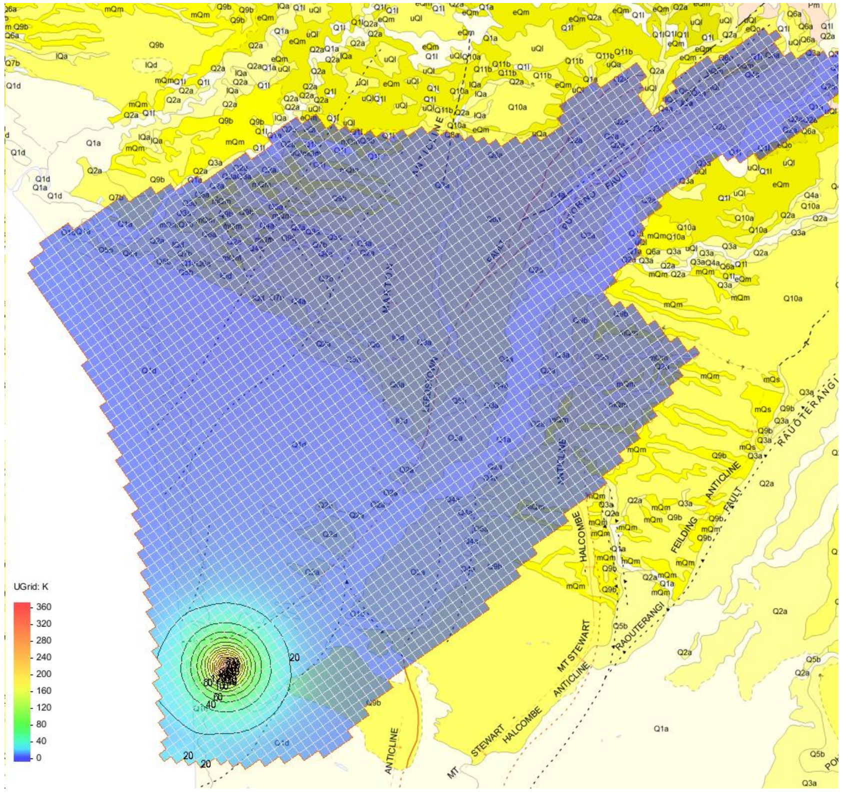


Figure B-8. Layer 4 calibrated model horizontal hydraulic conductivity (KH) extracted from MODFLOW 6 files.

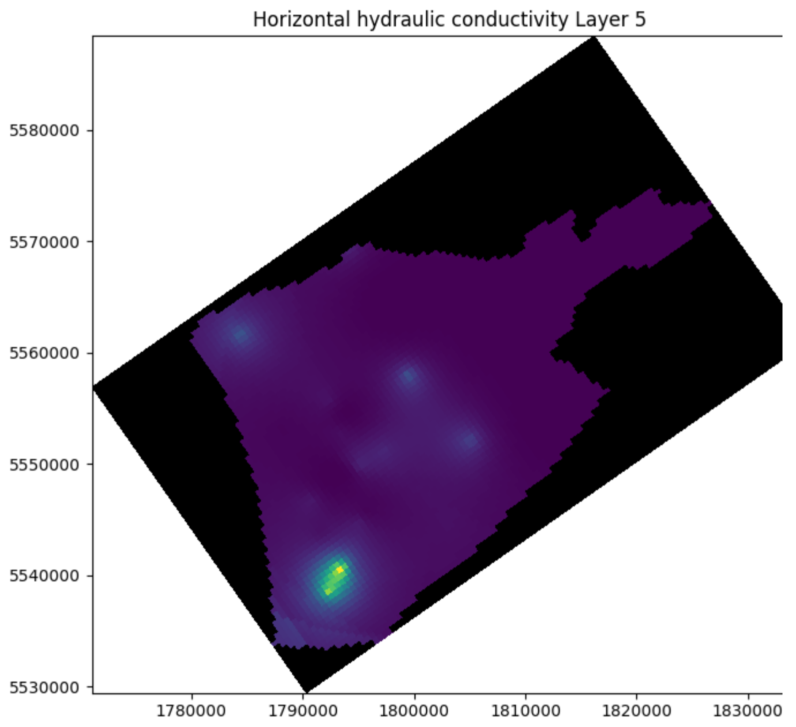


Figure B-9. Layer 5 calibrated model horizontal hydraulic conductivity (KH) as reported in PDP (2017).

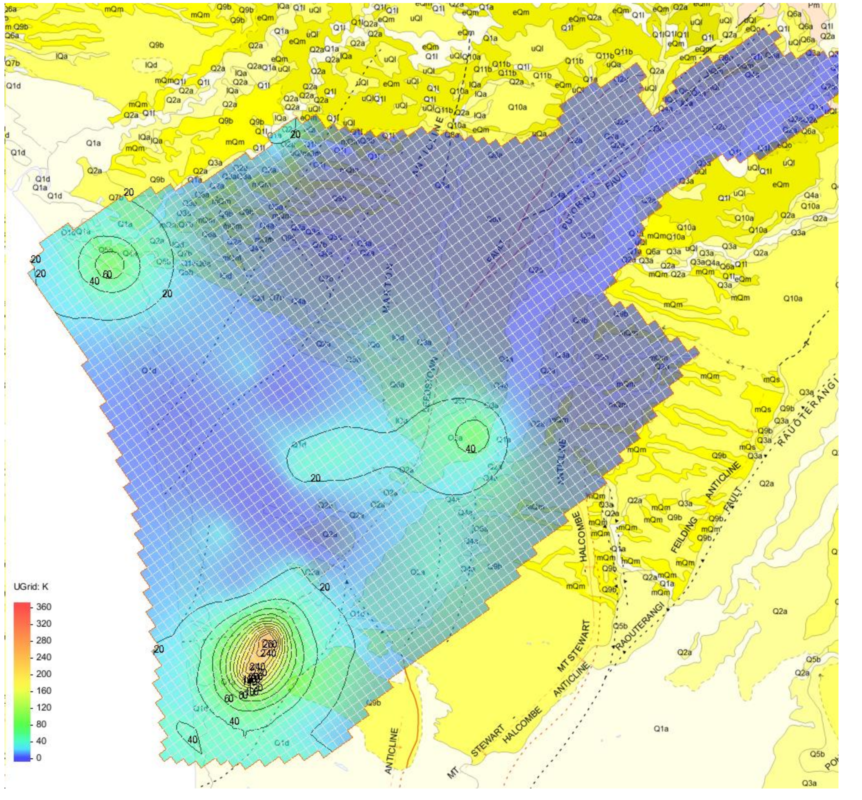


Figure B-10. Layer 5 calibrated model horizontal hydraulic conductivity (KH) extracted from MODFLOW 6 files.

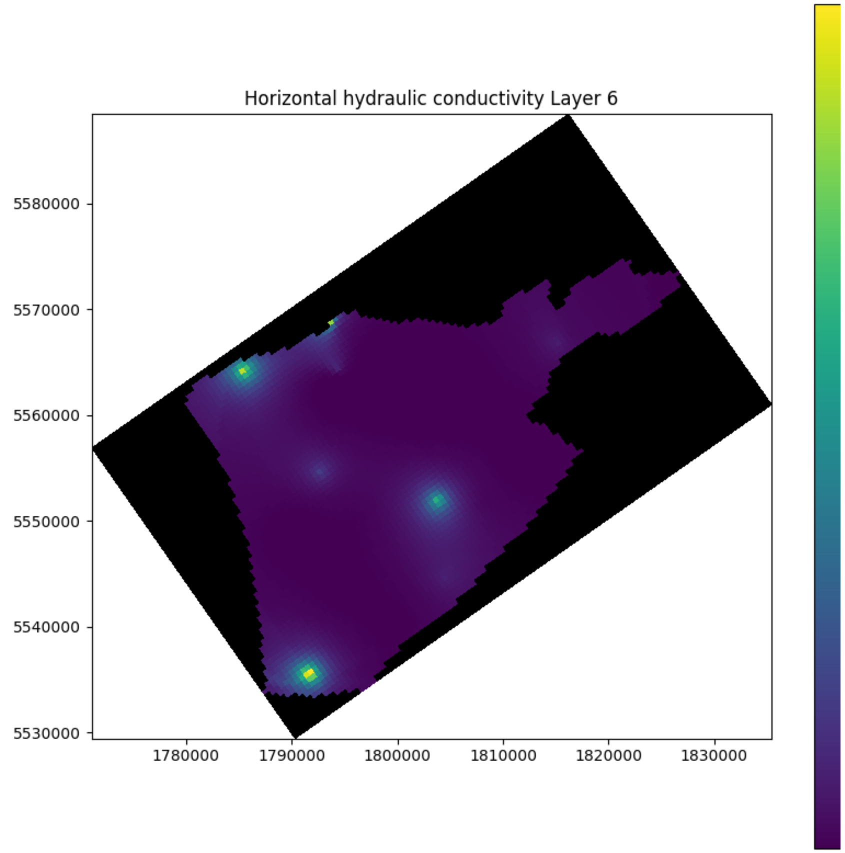


Figure B-11. Layer 6 calibrated model horizontal hydraulic conductivity (KH) as reported in PDP (2017).

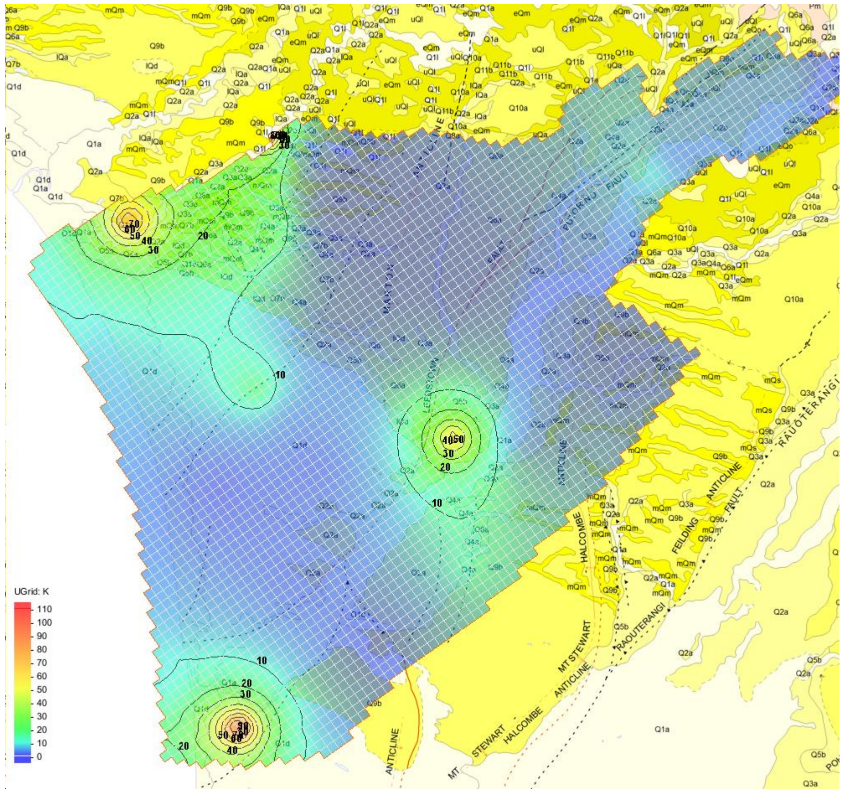


Figure B-12. Layer 6 calibrated model horizontal hydraulic conductivity (KH) extracted from MODFLOW 6 files.

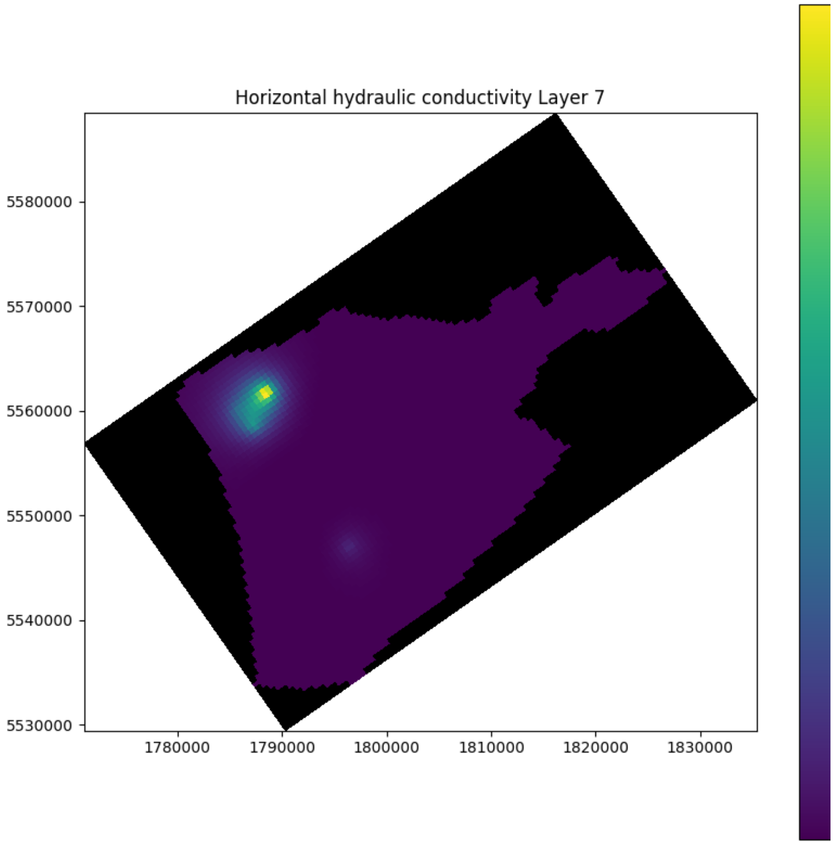


Figure B-13. Layer 7 calibrated model horizontal hydraulic conductivity (KH) as reported in PDP (2017).

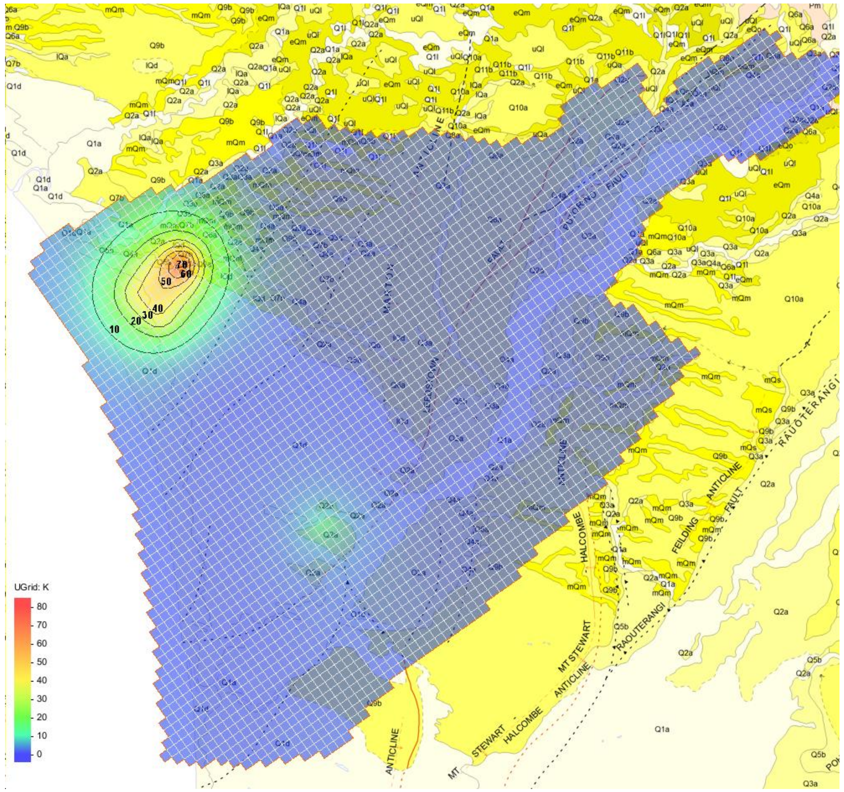


Figure B-14. Layer 7 calibrated model horizontal hydraulic conductivity (KH) extracted from MODFLOW 6 files.

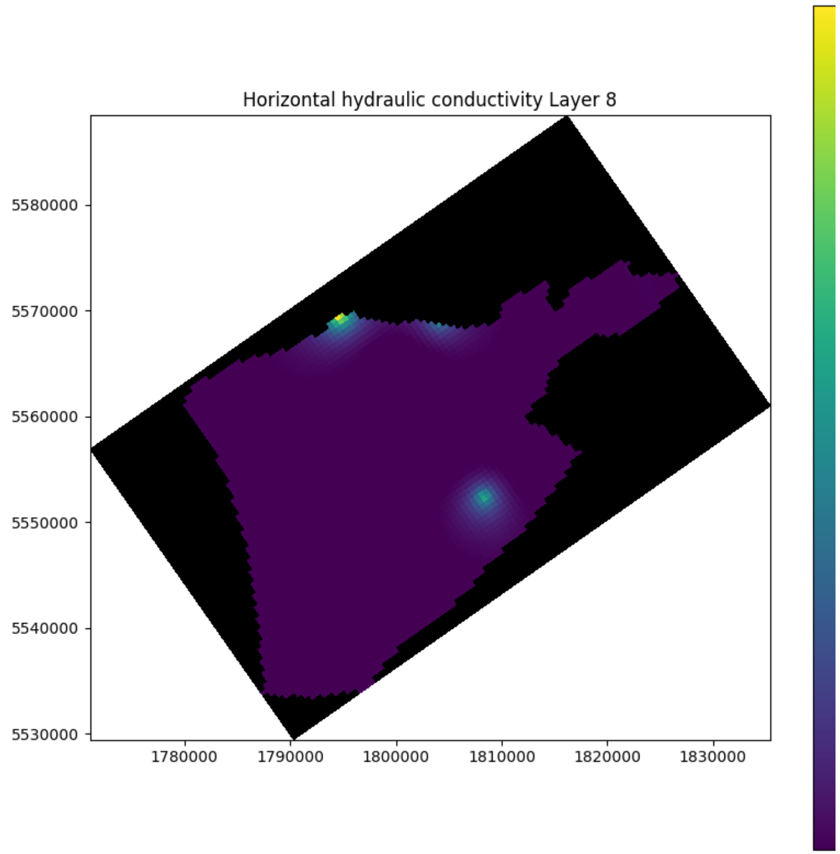


Figure B-15. Layer 8 calibrated model horizontal hydraulic conductivity (KH) as reported in PDP (2017).

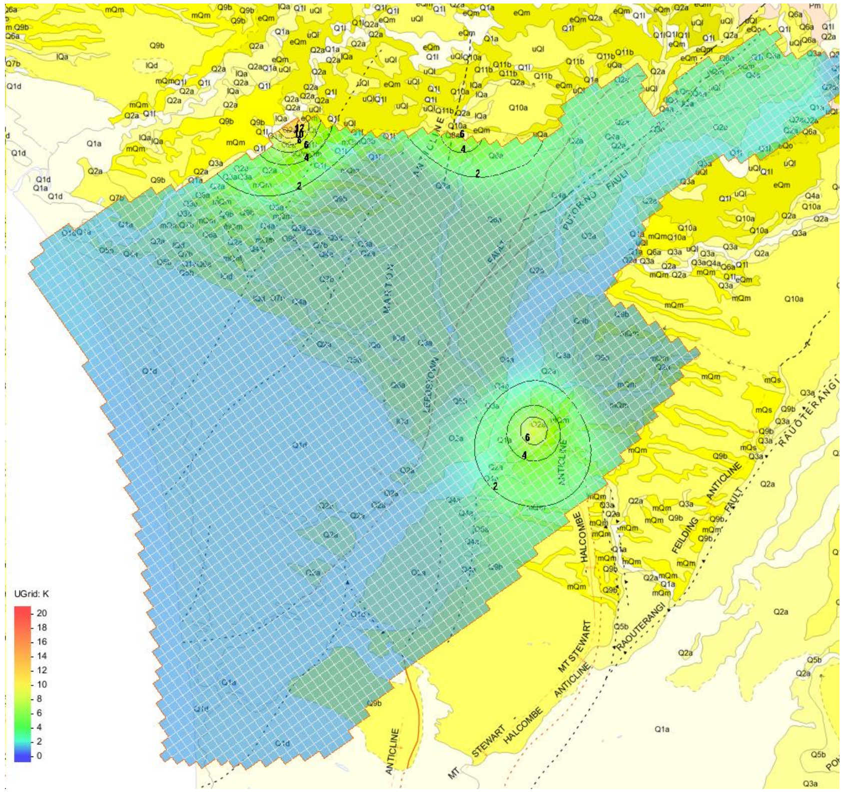


Figure B-16. Layer 8 calibrated model horizontal hydraulic conductivity (KH) extracted from MODFLOW 6 files.

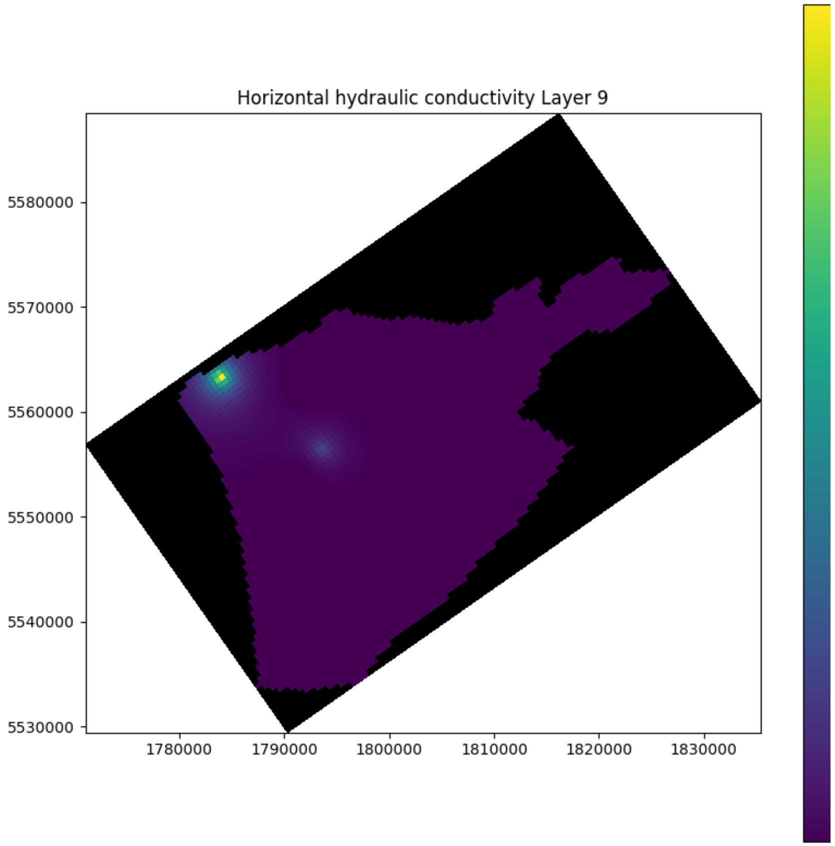


Figure B-17. Layer 9 calibrated model horizontal hydraulic conductivity (KH) as reported in PDP (2017).

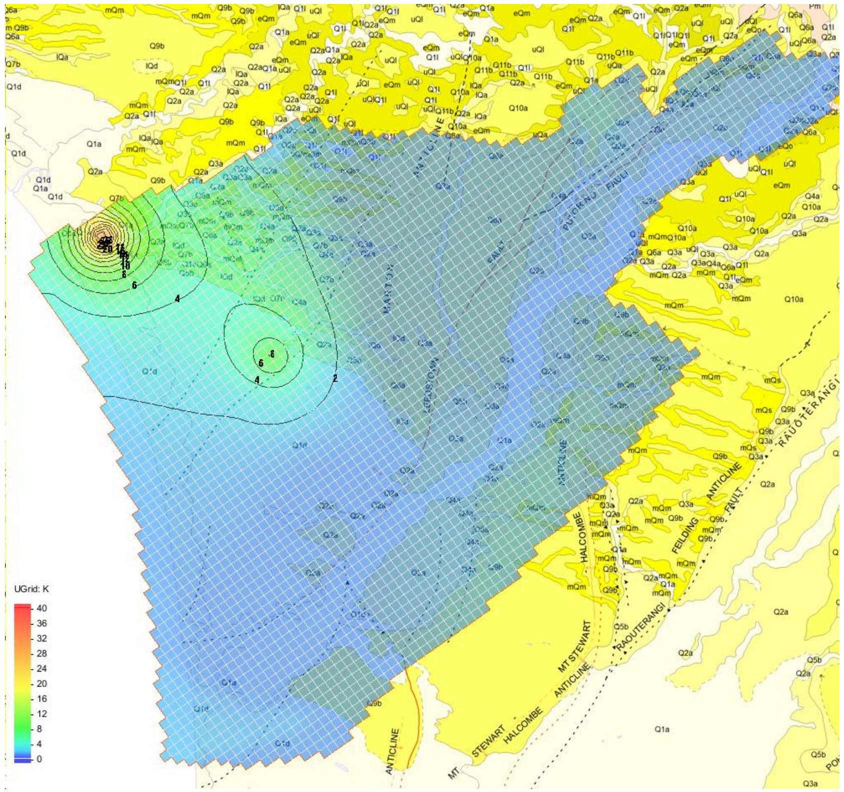


Figure B-18. Layer 9 calibrated model horizontal hydraulic conductivity (KH) extracted from MODFLOW 6 files.

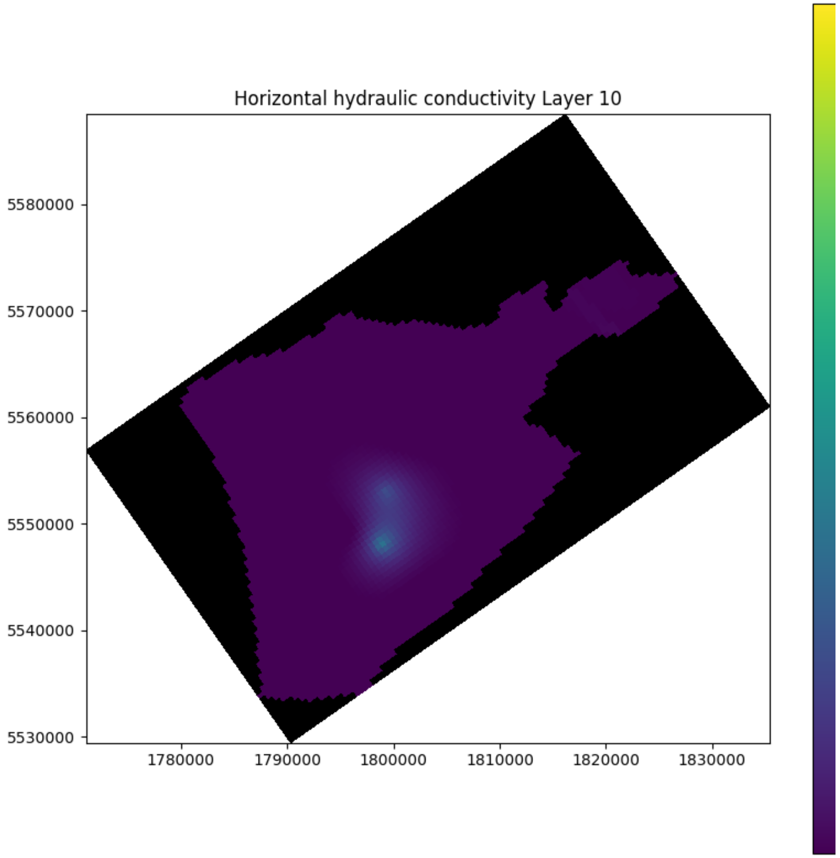


Figure B-19. Layer 10 calibrated model horizontal hydraulic conductivity (KH) as reported in PDP (2017).

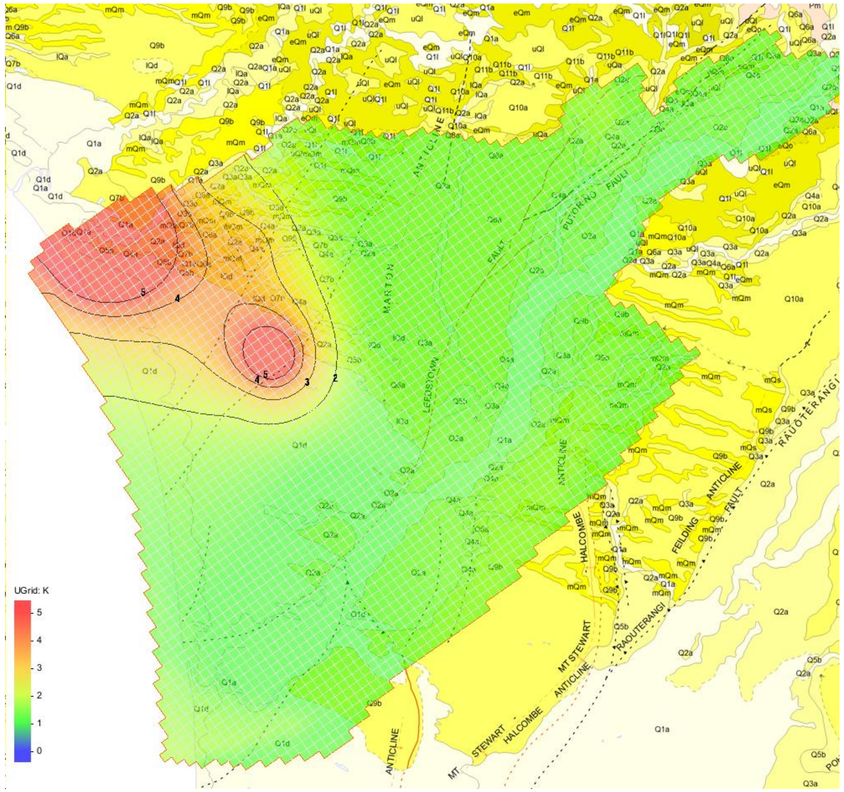


Figure B-20. Layer 10 calibrated model horizontal hydraulic conductivity (KH) extracted from MODFLOW 6 files.

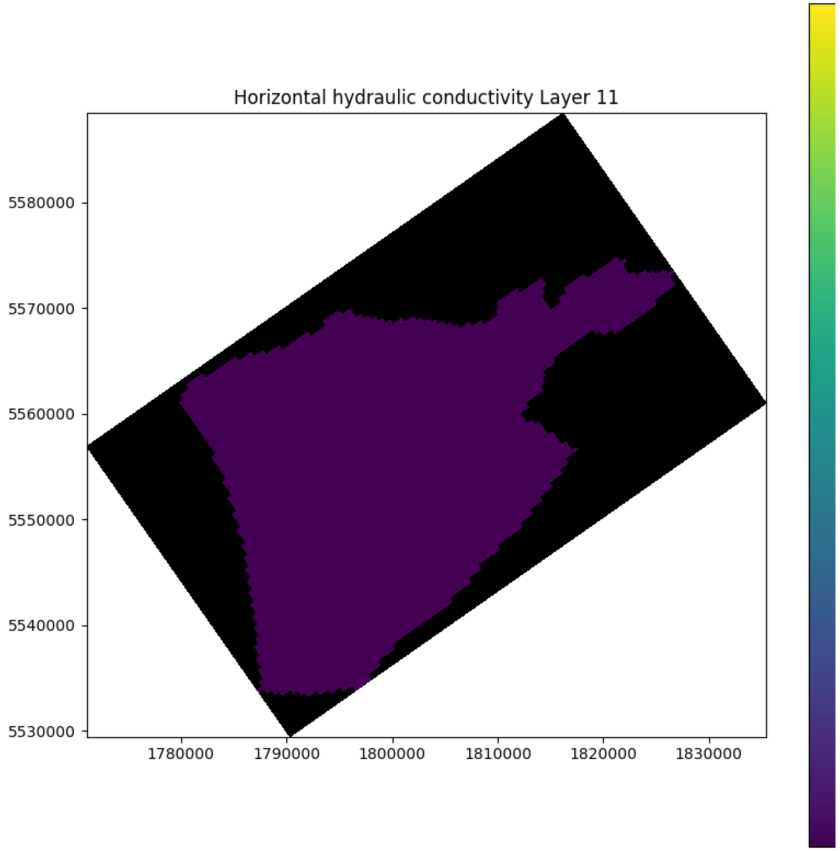


Figure B-21. Layer 11 calibrated model horizontal hydraulic conductivity (KH) as reported in PDP (2017).

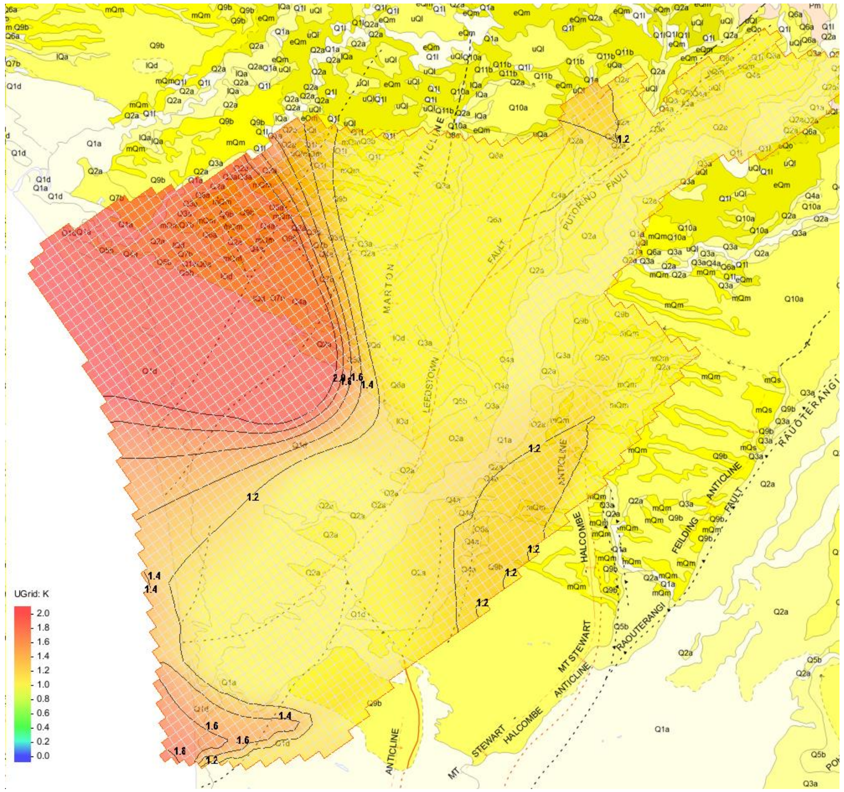


Figure B-22. Layer 11 calibrated model horizontal hydraulic conductivity (KH) extracted from MODFLOW 6 files.

Appendix C Well (Wel) Cells in the Model Layers


Figure C-1. Wells (groundwater abstraction bores) and well (WEL) cells in Layer 1 in the model (red squares).

Figure C-2. Wells (groundwater abstraction bores) and well (WEL) cells in Layer 2 in the model (red squares).

Figure C-3. Wells (groundwater abstraction bores) and well (WEL) cells in Layer 3 in the model (red squares).

Figure C-4. Wells (groundwater abstraction bores) and well (WEL) cells in Layer 4 in the model (red squares).

Figure C-5. Wells (groundwater abstraction bores) and well (WEL) cells in Layer 5 in the model (red squares).

Figure C-6. Wells (groundwater abstraction bores) and well (WEL) cells in Layer 6 in the model (red squares).

Figure C-7. Wells (groundwater abstraction bores) and well (WEL) cells in Layer 7 in the model (red squares).

Figure C-8. Wells (groundwater abstraction bores) and well (WEL) cells in Layer 8 in the model (red squares).

Figure C-9. Wells (groundwater abstraction bores) and well (WEL) cells in Layer 10 in the model (red squares).

Figure C-10. Wells (groundwater abstraction bores) and well (WEL) cells in Layer 11 in the model (red squares).

Appendix D Groundwater Level Hydrographs – Model-calculated and Observed

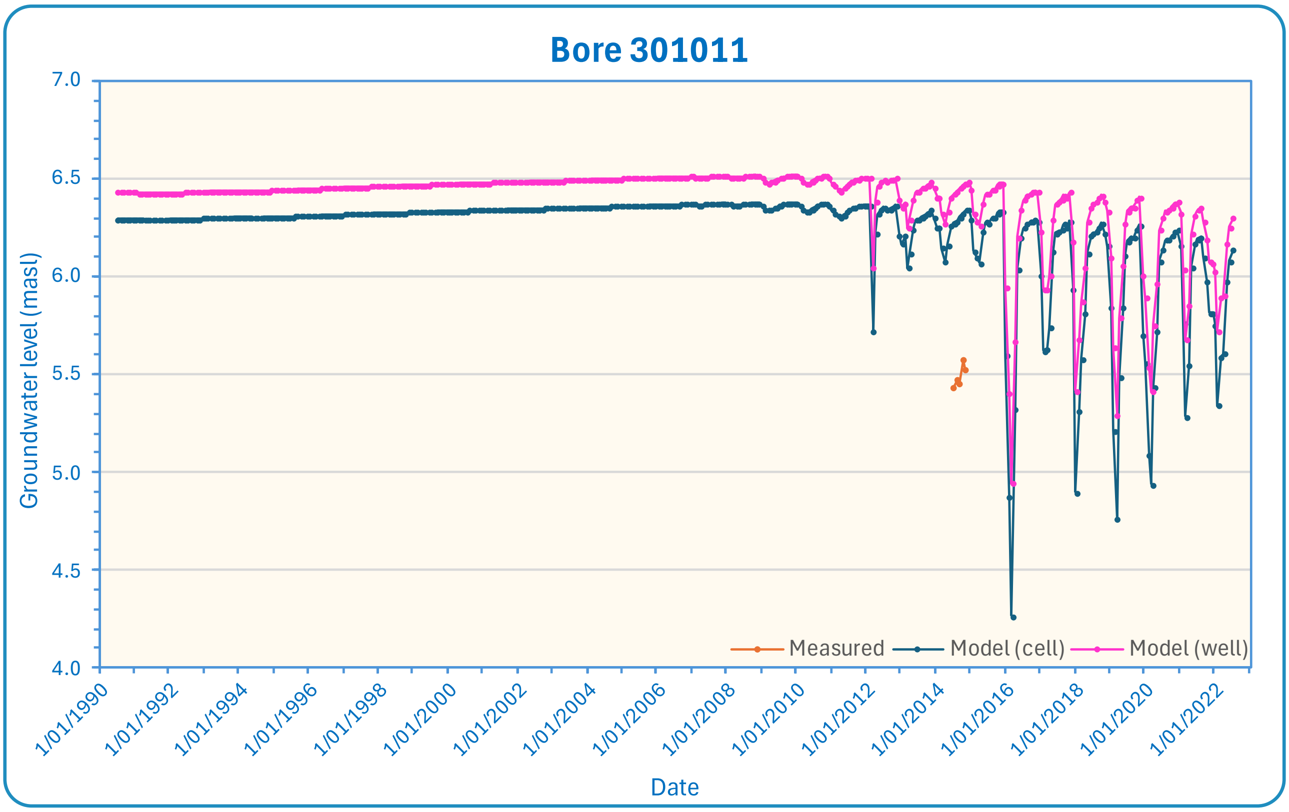


Figure D-1. Observed and model calculated groundwater level for Bore 301011.

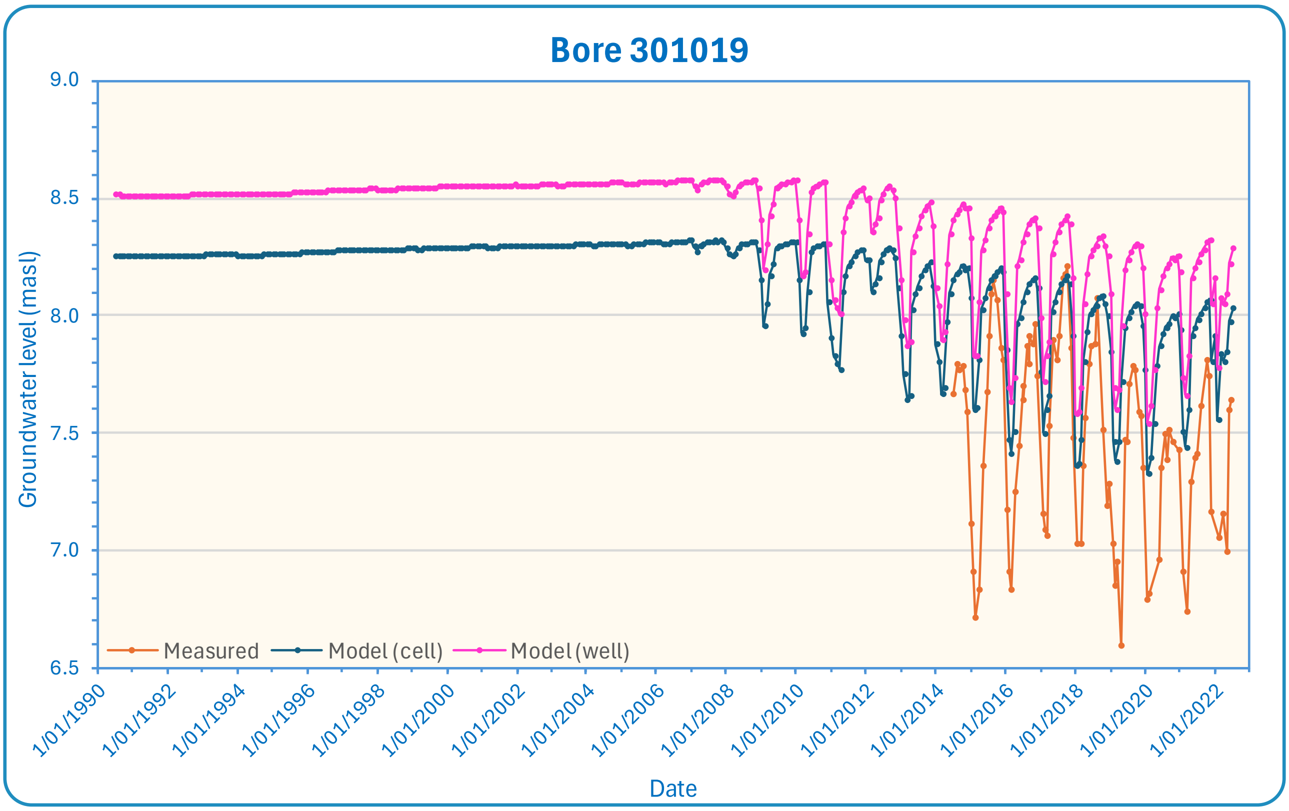


Figure D-2. Observed and model calculated groundwater level for Bore 301019.

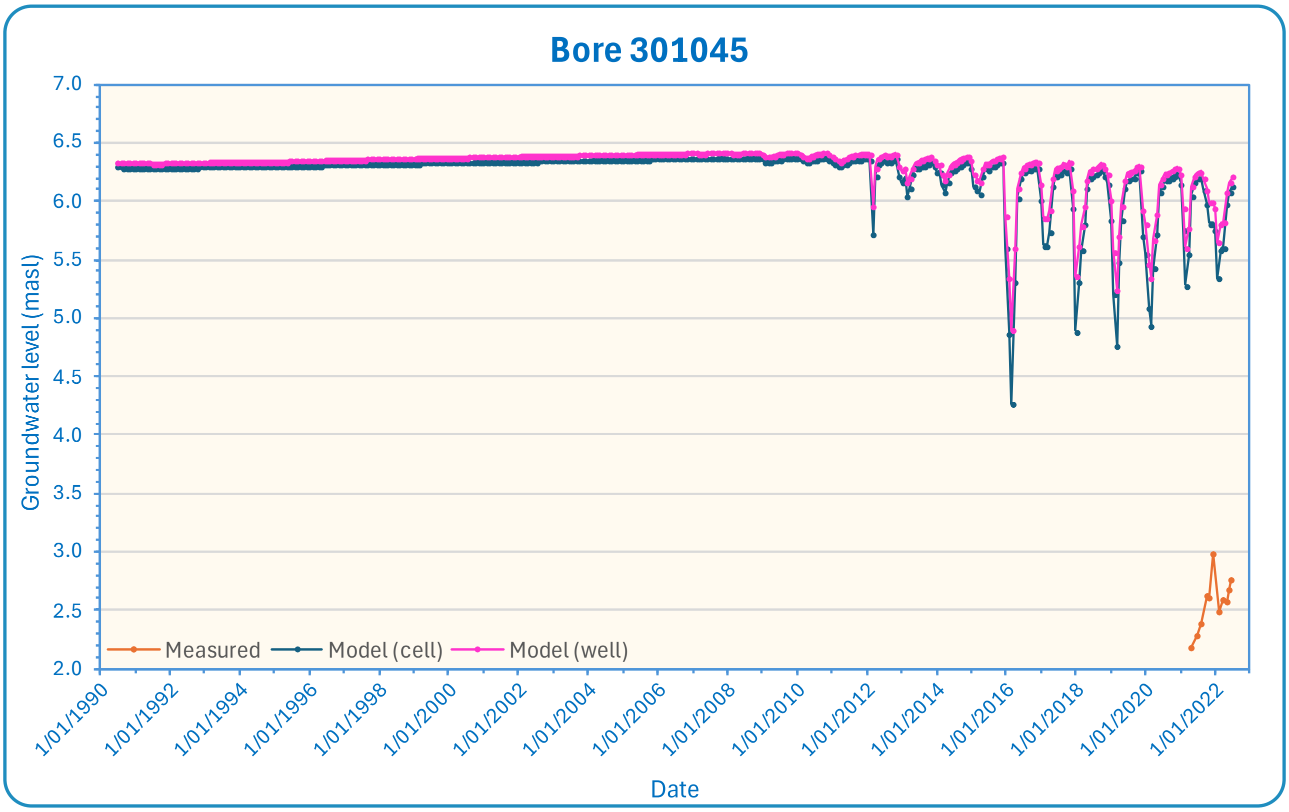


Figure D-3. Observed and model calculated groundwater level for Bore 301045.

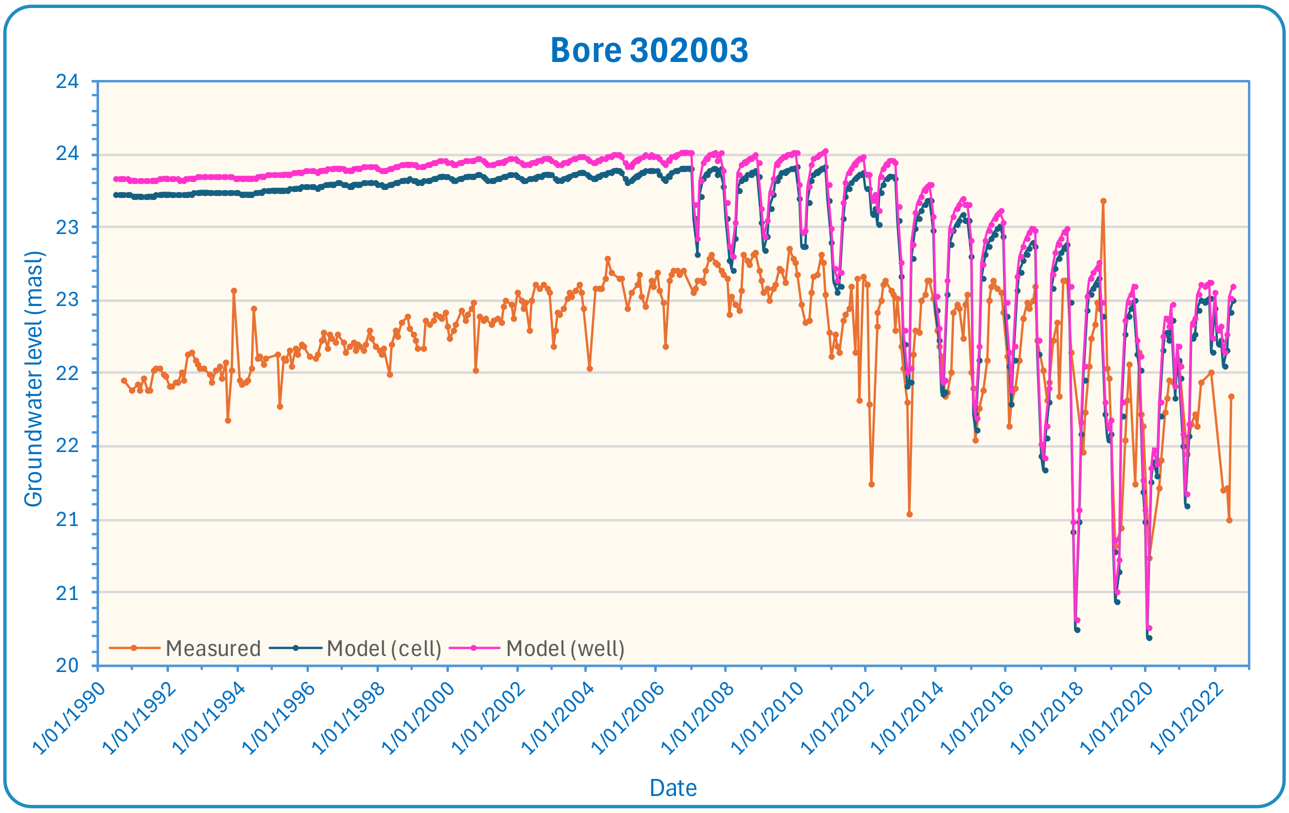


Figure D-4. Observed and model calculated groundwater level for Bore 302003.

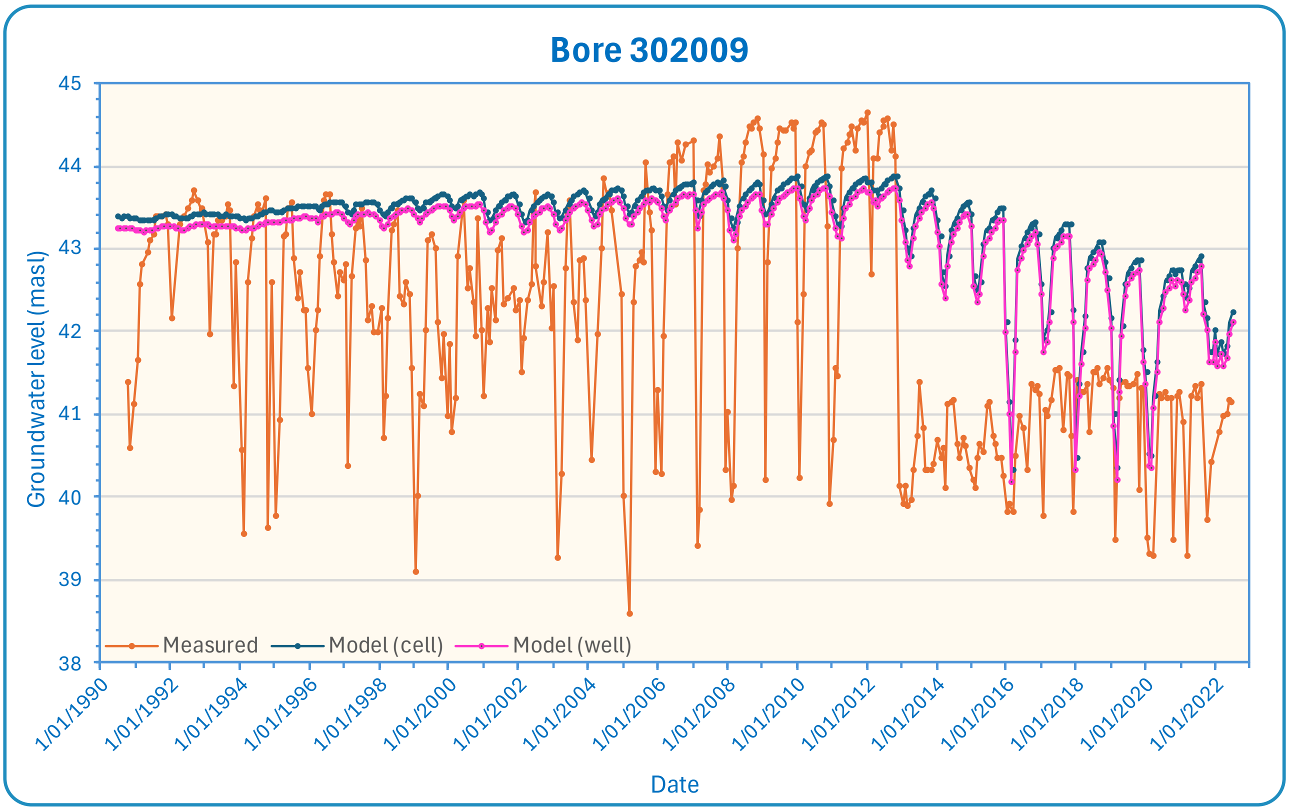


Figure D-5. Observed and model calculated groundwater level for Bore 302009.

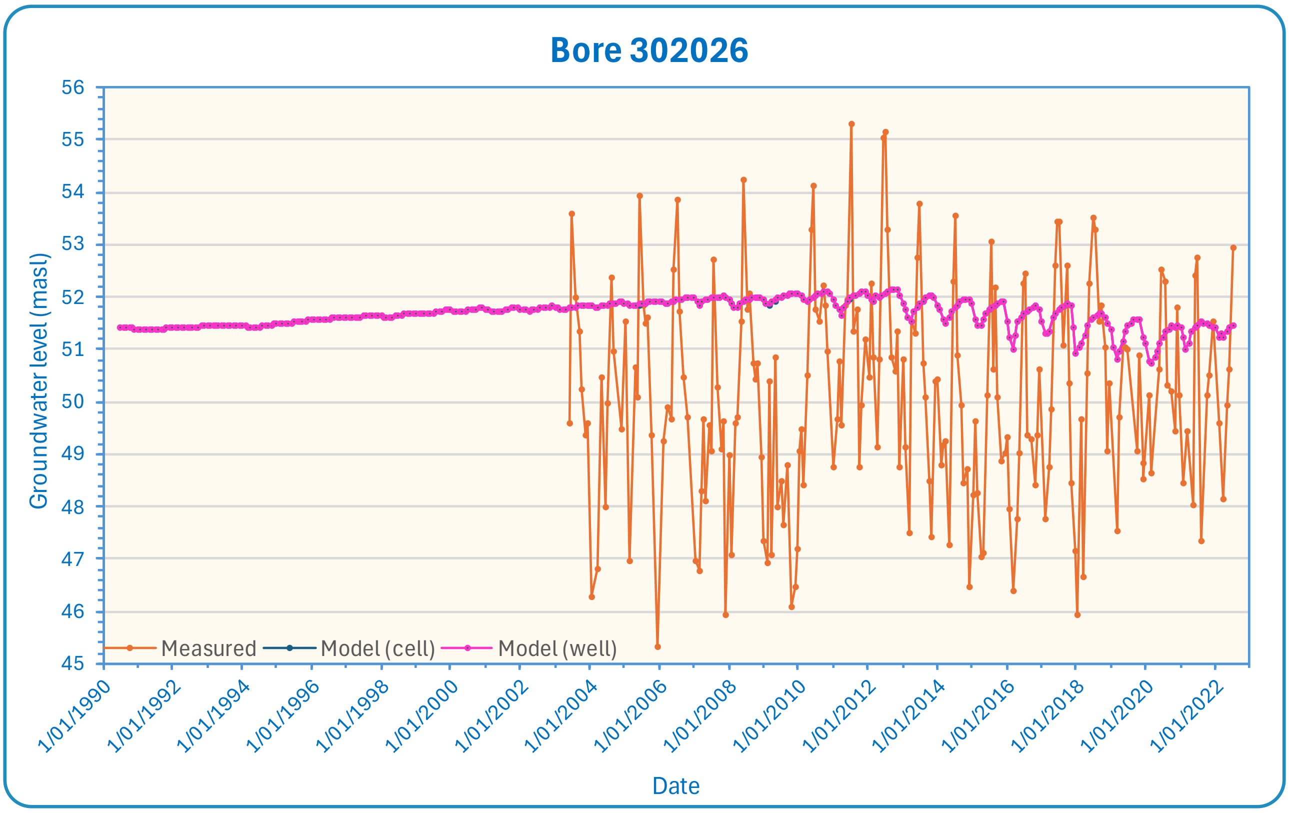


Figure D-6. Observed and model calculated groundwater level for Bore 302026.

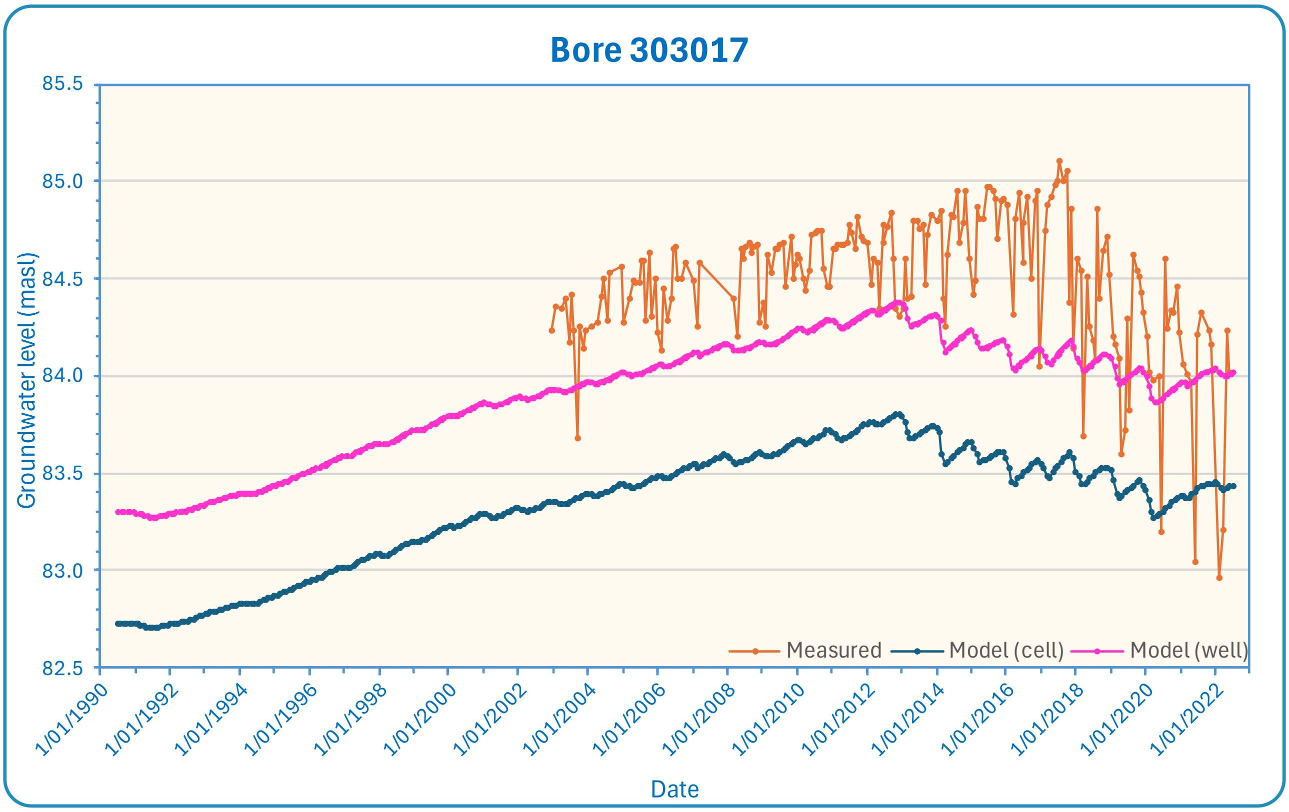


Figure D-7. Observed and model calculated groundwater level for Bore 303017.

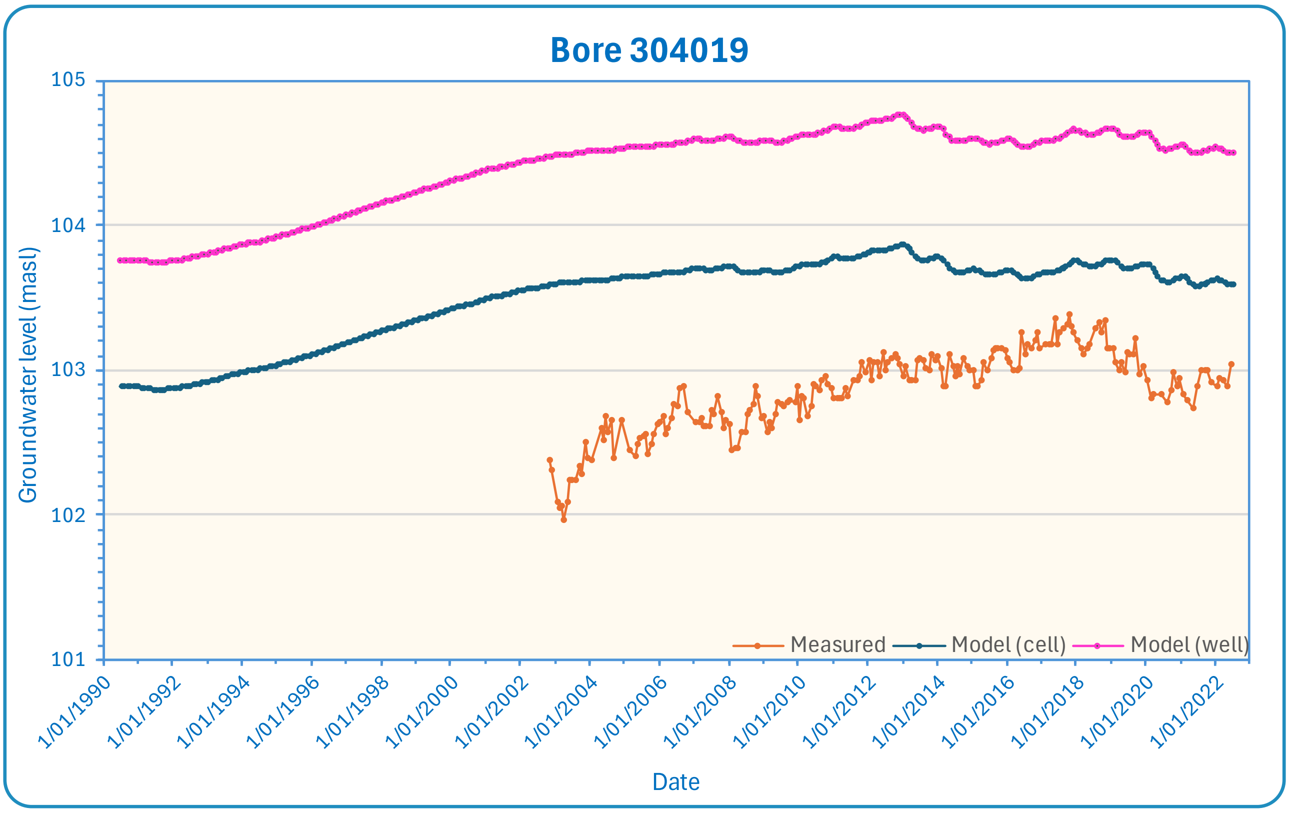


Figure D-8. Observed and model calculated groundwater level for Bore 304019.

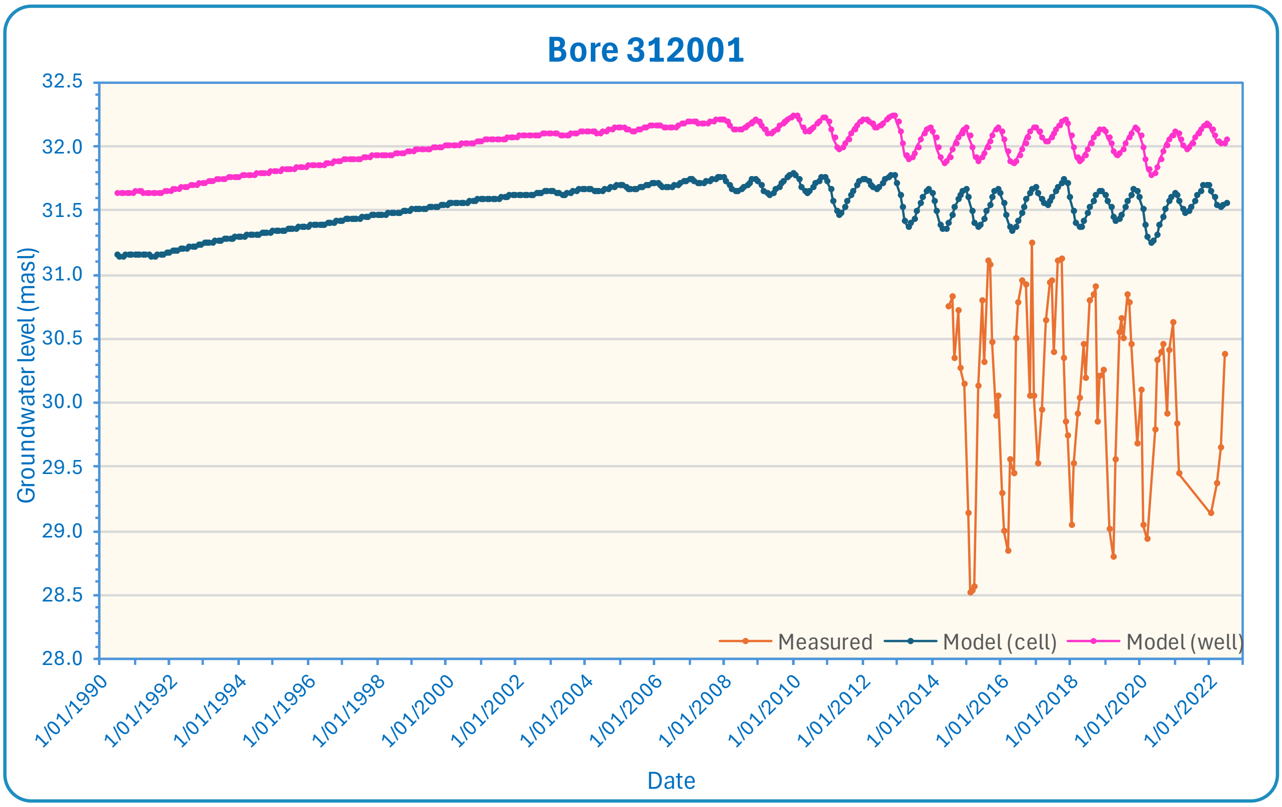


Figure D-9. Observed and model calculated groundwater level for Bore 312001.

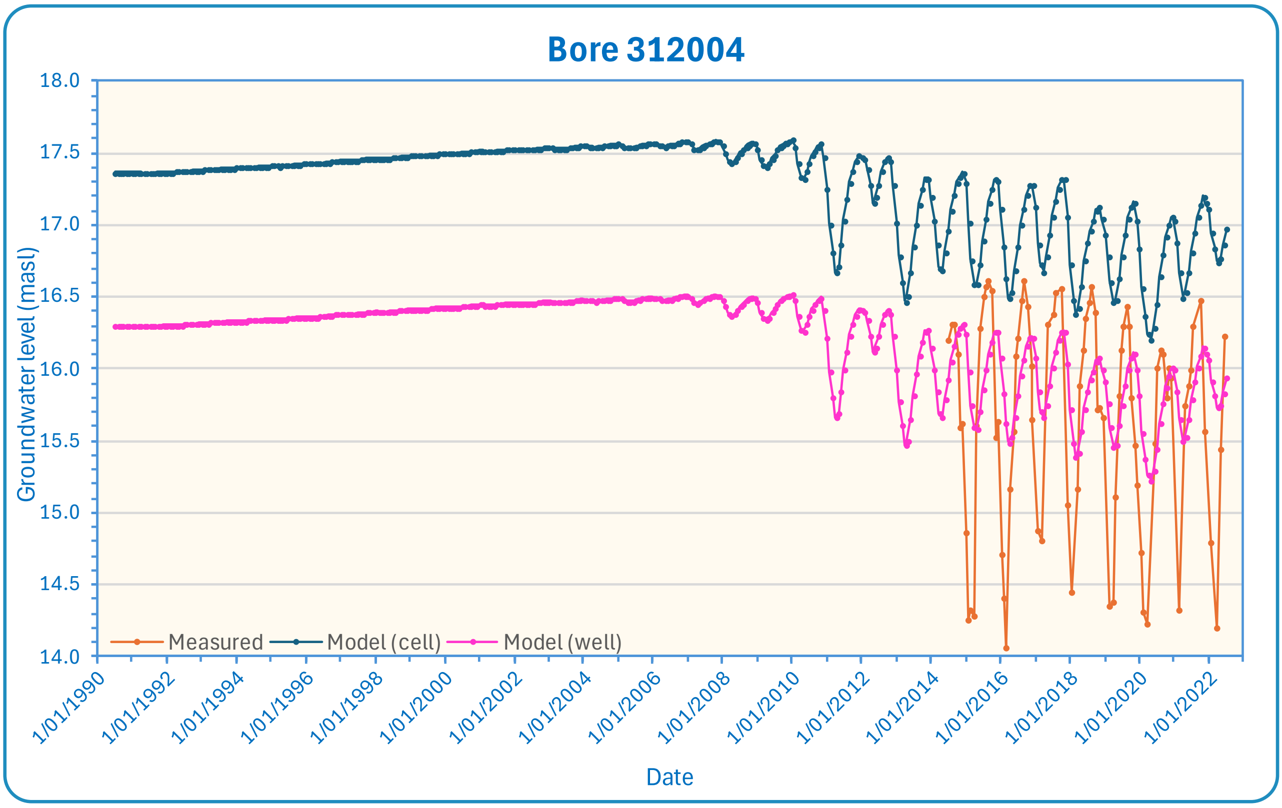


Figure D-10. Observed and model calculated groundwater level for Bore 312004.

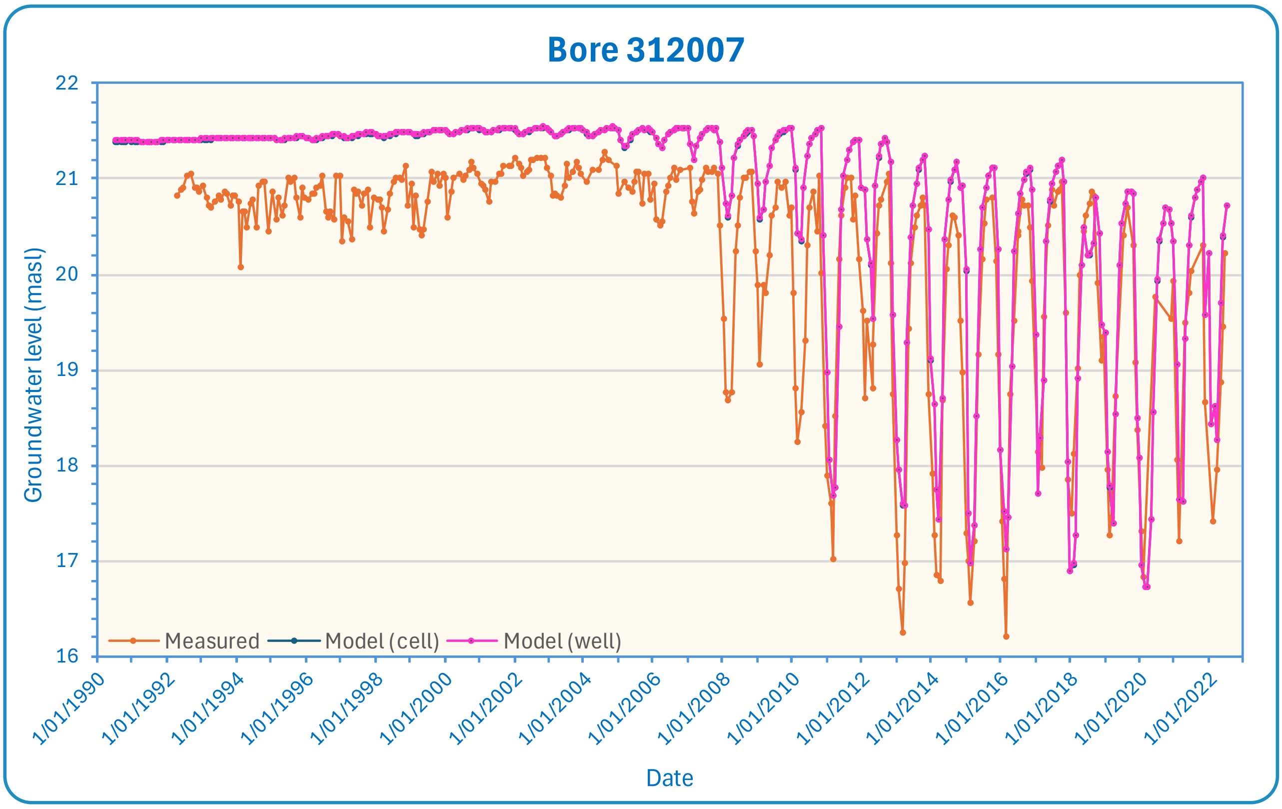


Figure D-11. Observed and model calculated groundwater level for Bore 312007.

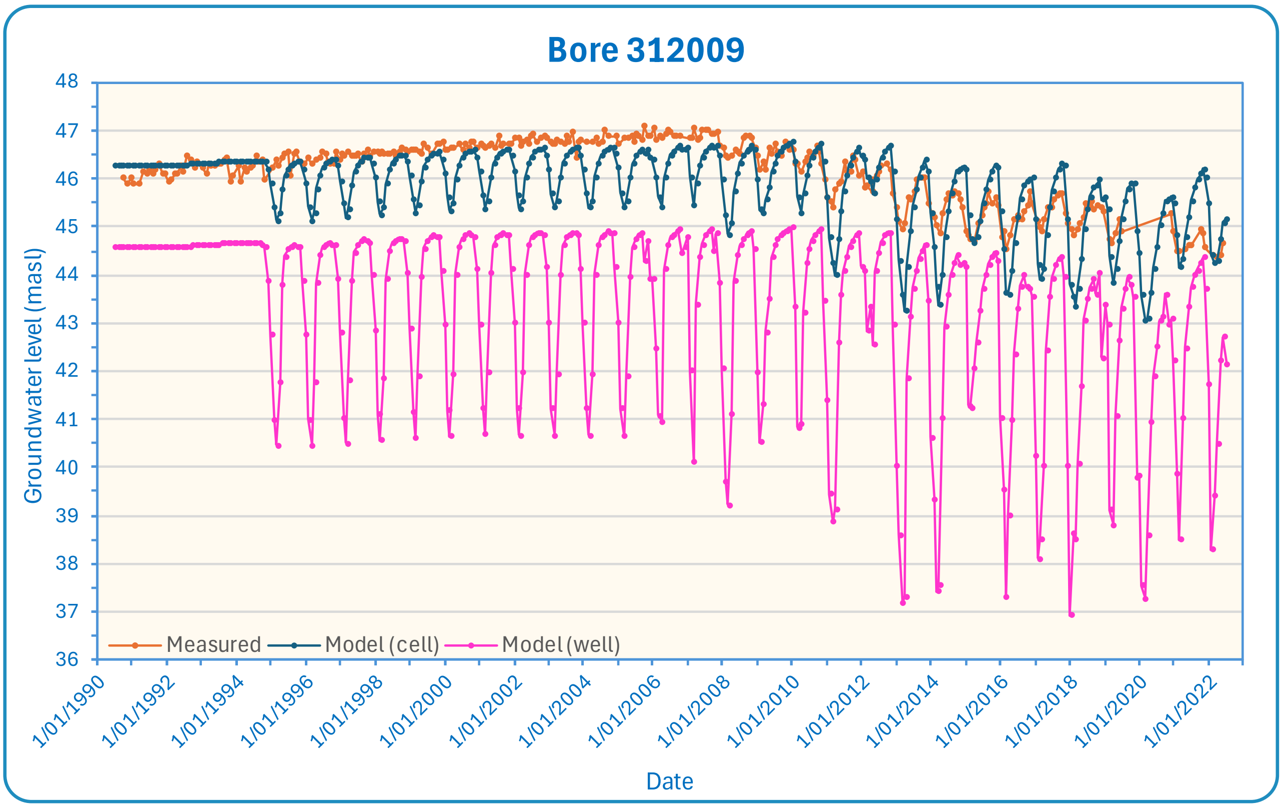


Figure D-12. Observed and model calculated groundwater level for Bore 312009.

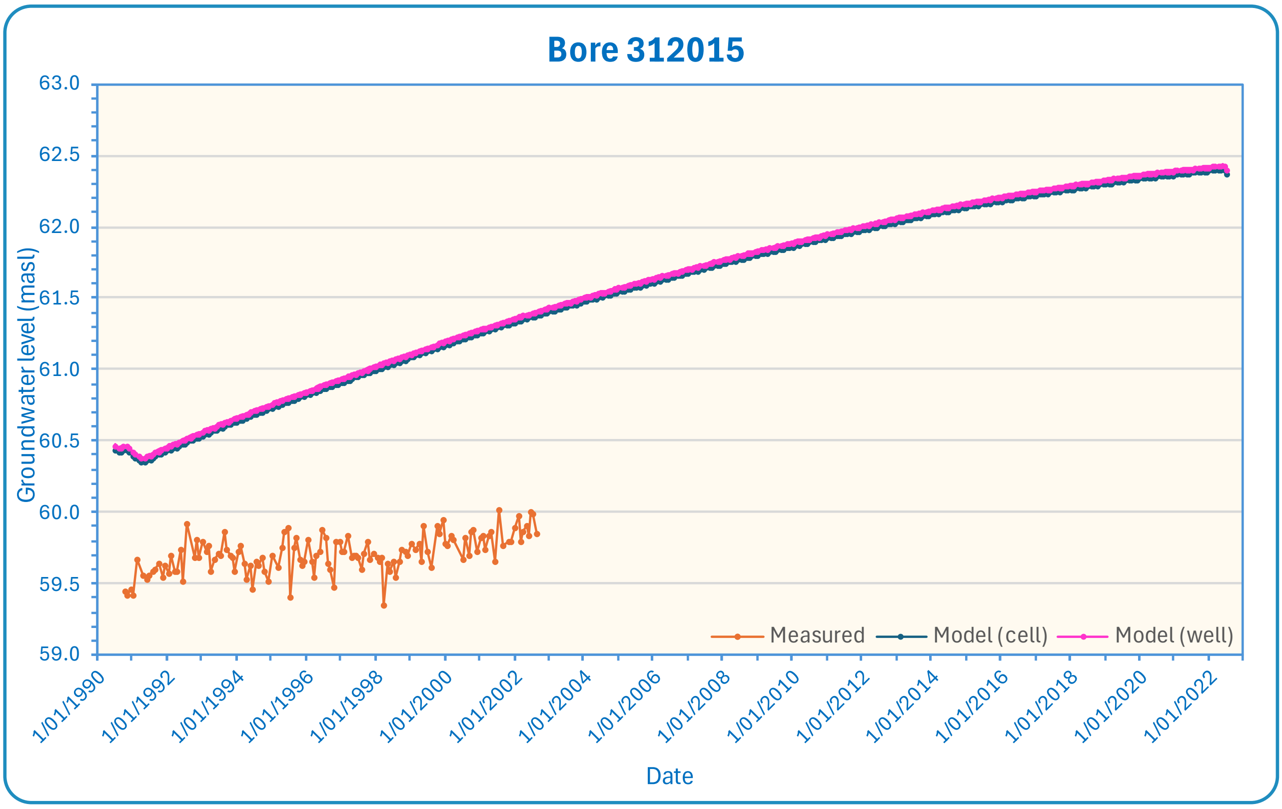


Figure D-13. Observed and model calculated groundwater level for Bore 312015.

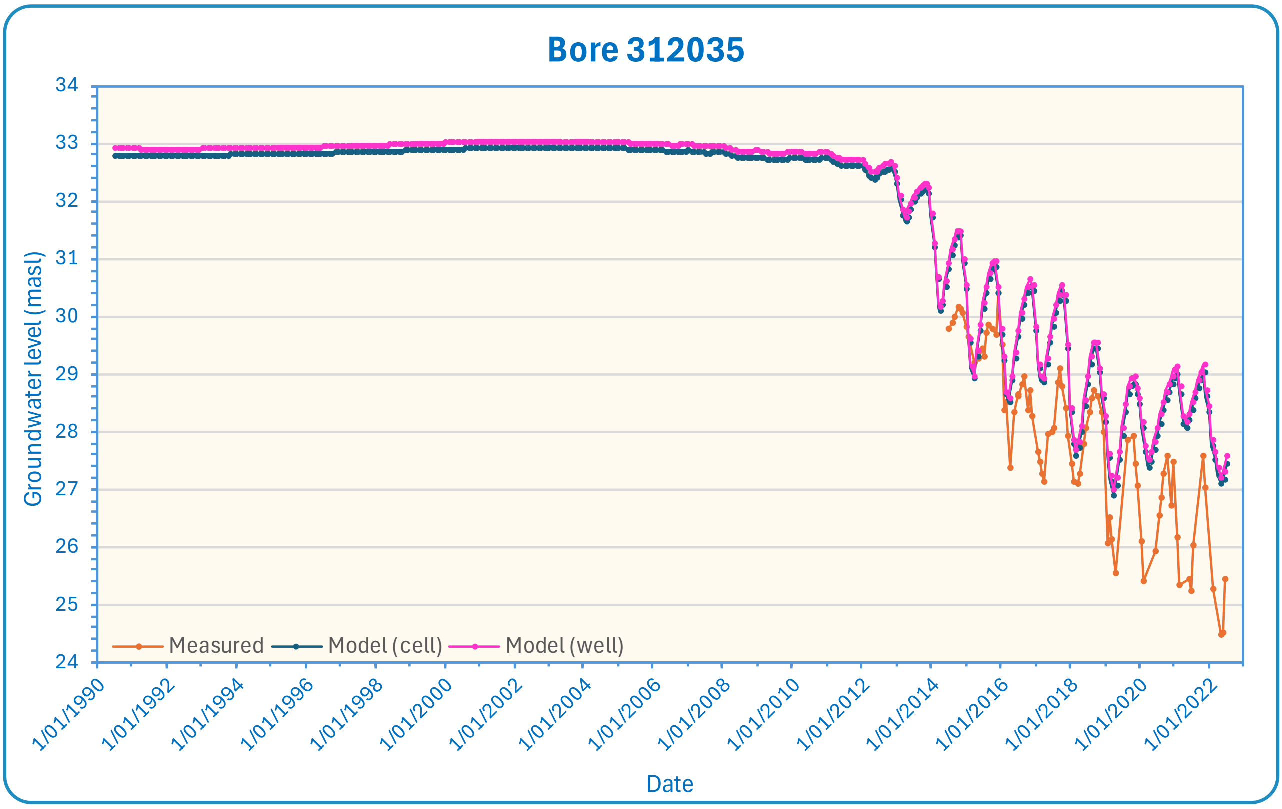


Figure D-14. Observed and model calculated groundwater level for Bore 312035.

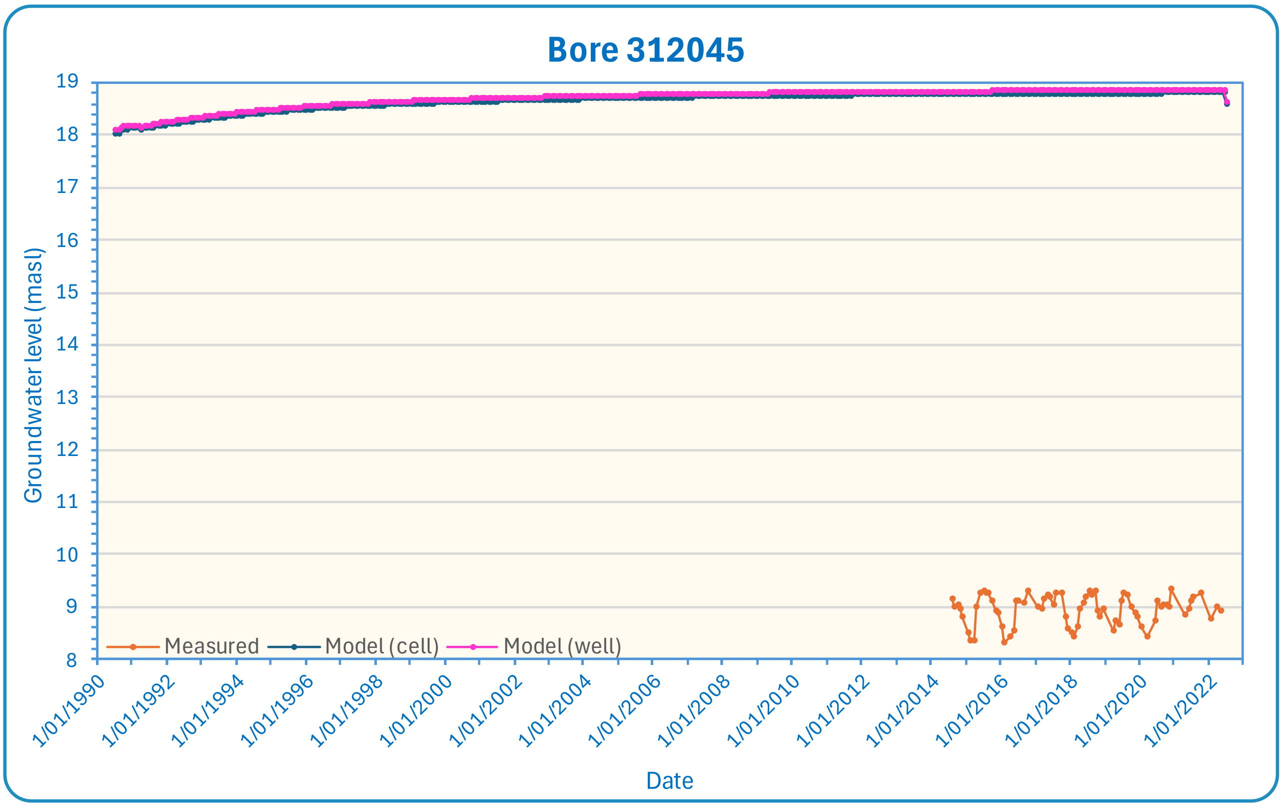


Figure D-15. Observed and model calculated groundwater level for Bore 312045.

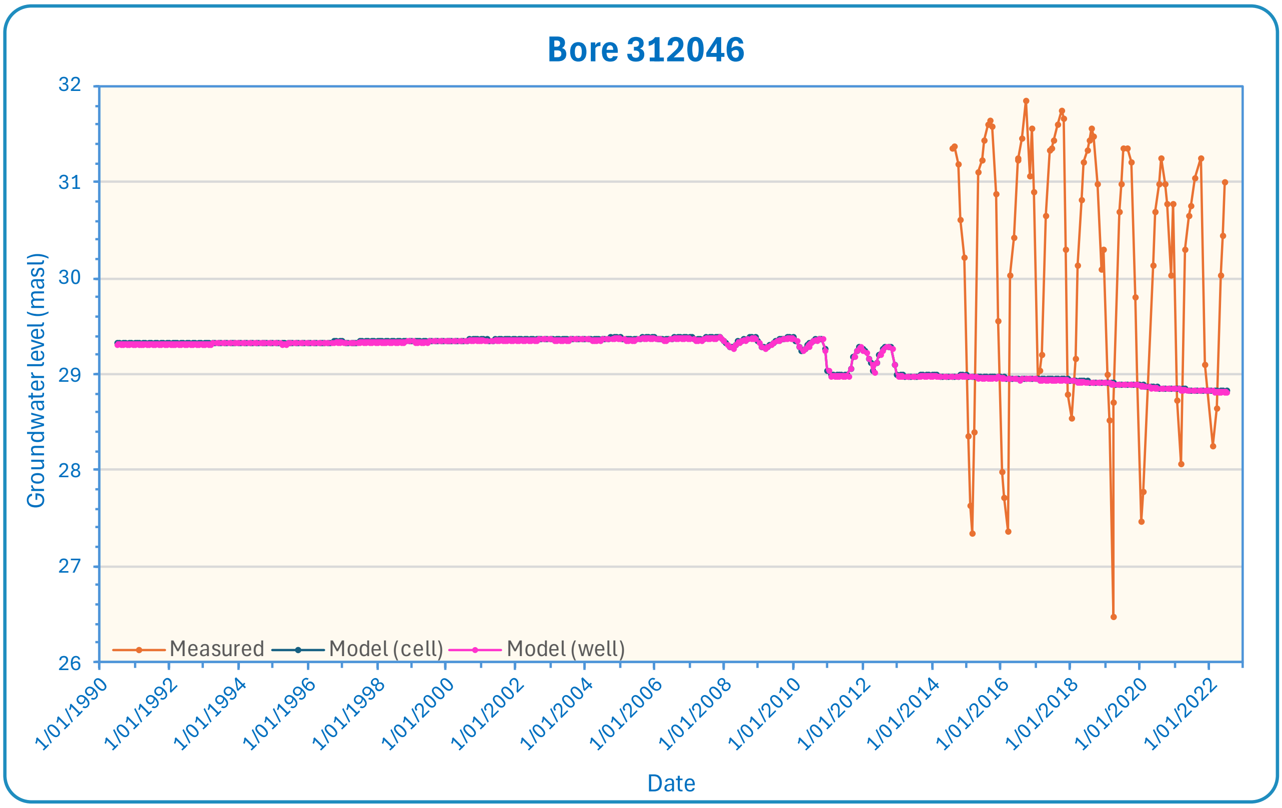


Figure D-16. Observed and model calculated groundwater level for Bore 312046.

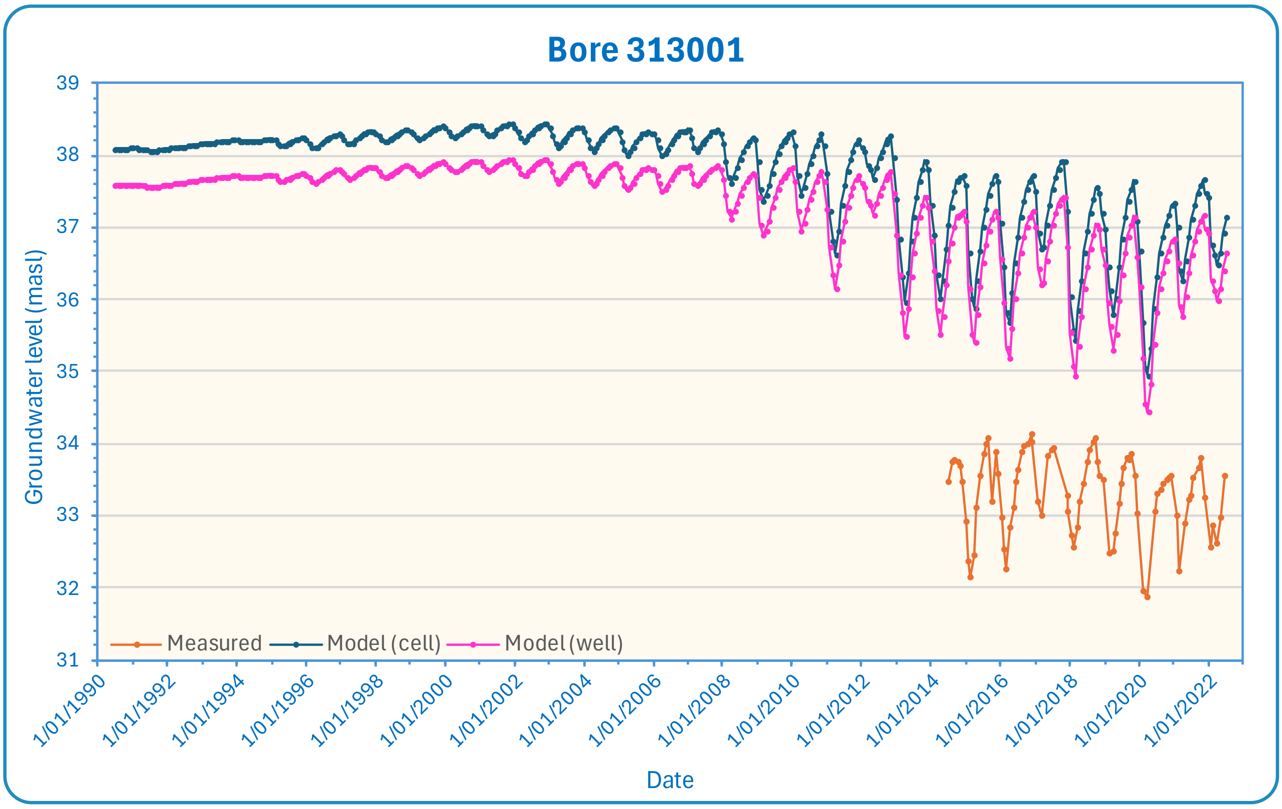


Figure D-17. Observed and model calculated groundwater level for Bore 313001.

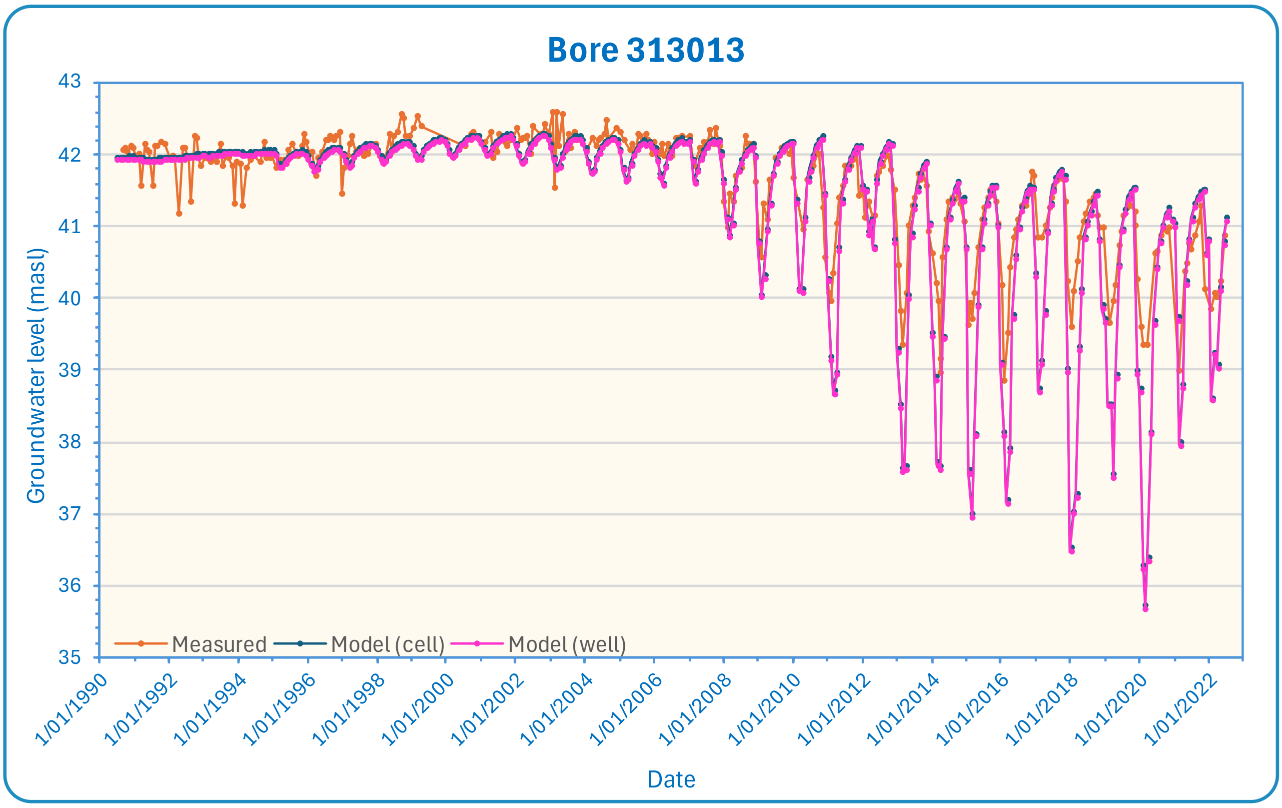


Figure D-18. Observed and model calculated groundwater level for Bore 313013.

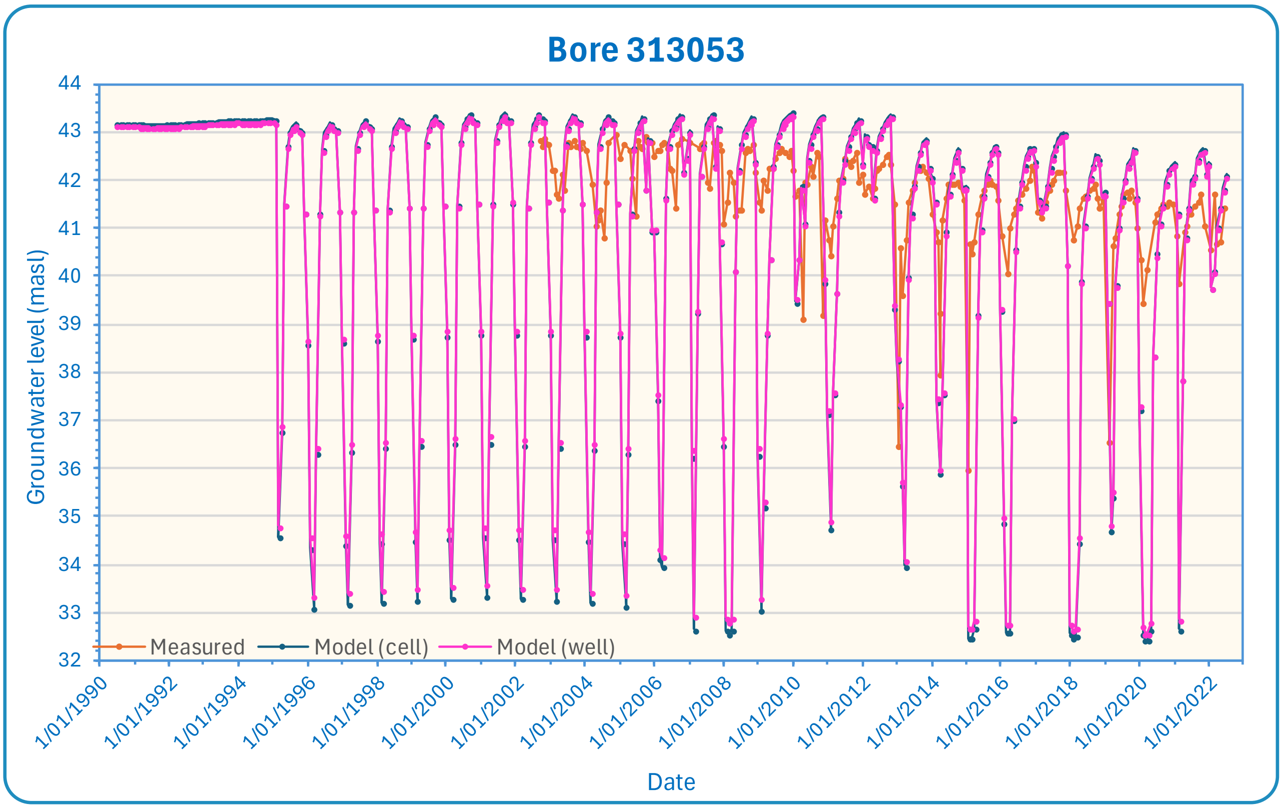


Figure D-19. Observed and model calculated groundwater level for Bore 313053.

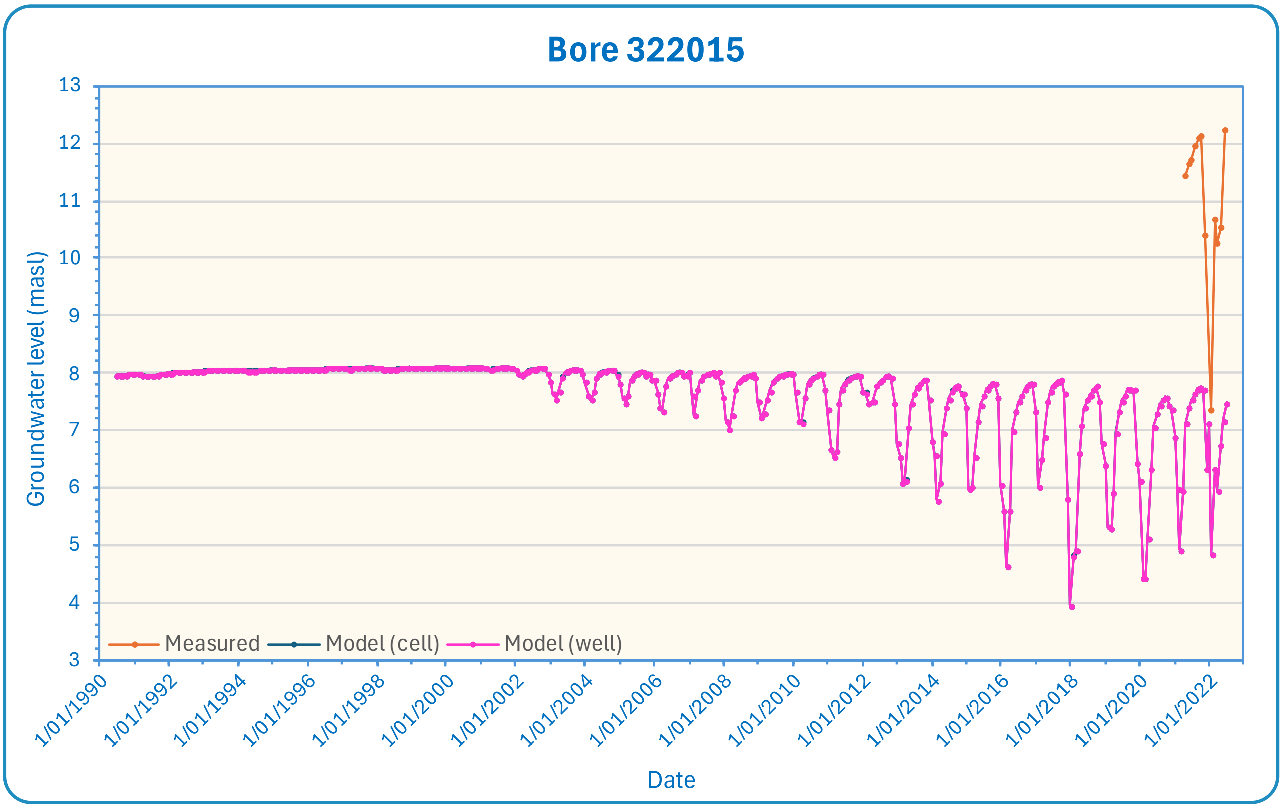


Figure D-20. Observed and model calculated groundwater level for Bore 322015.

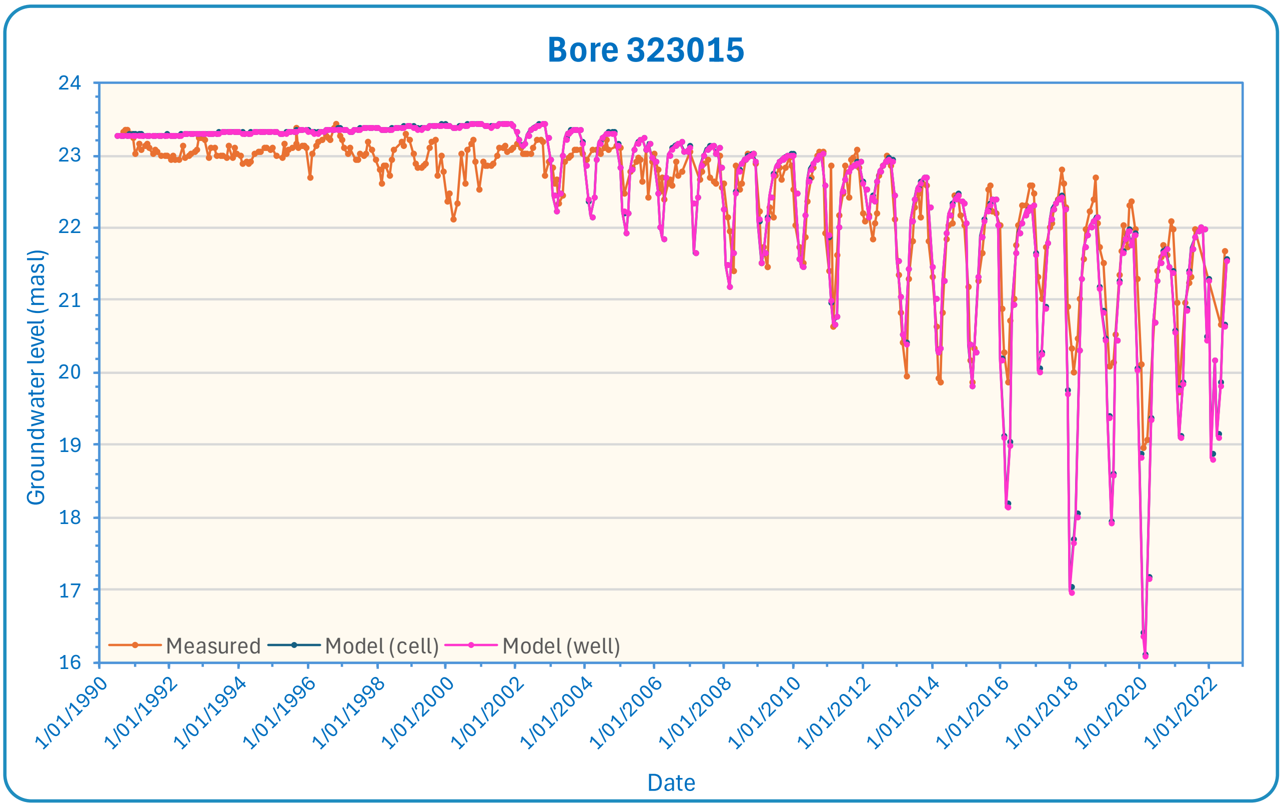


Figure D-21. Observed and model calculated groundwater level for Bore 323015.

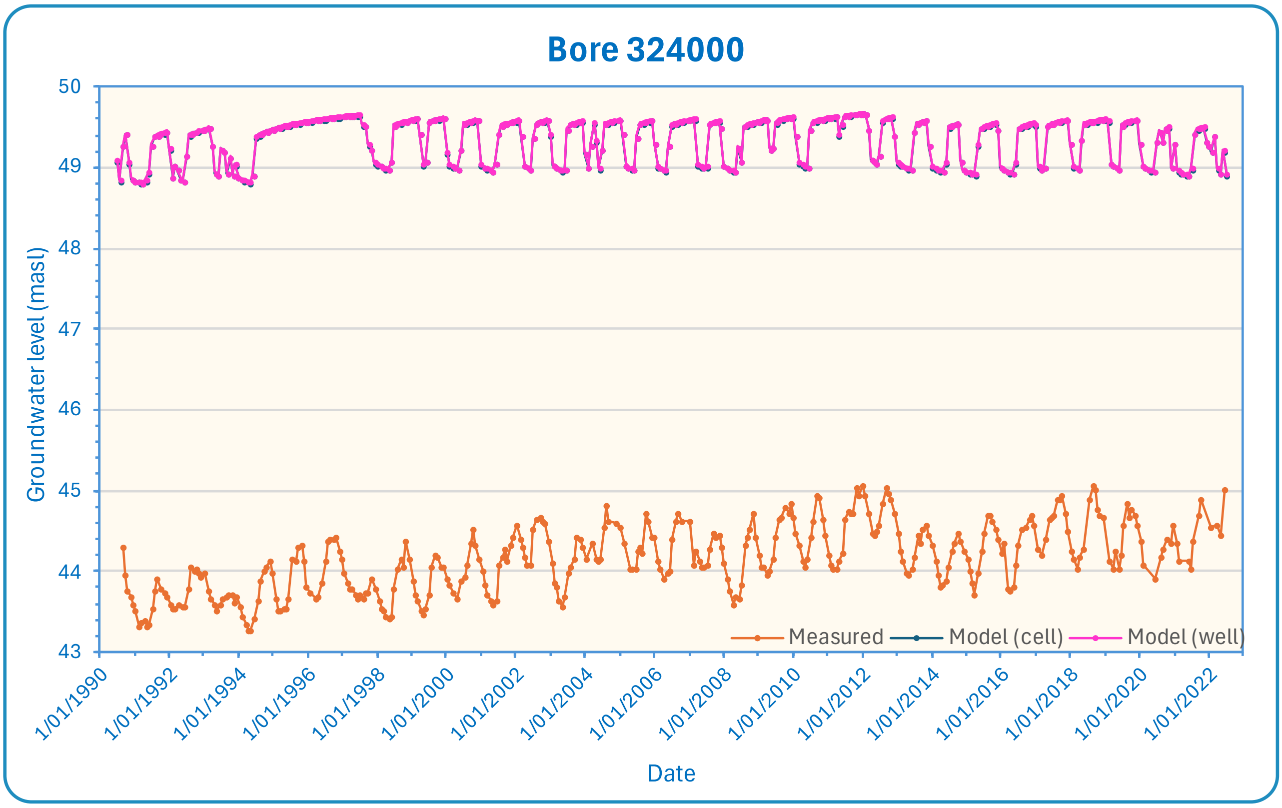


Figure D-22. Observed and model calculated groundwater level for Bore 324000.

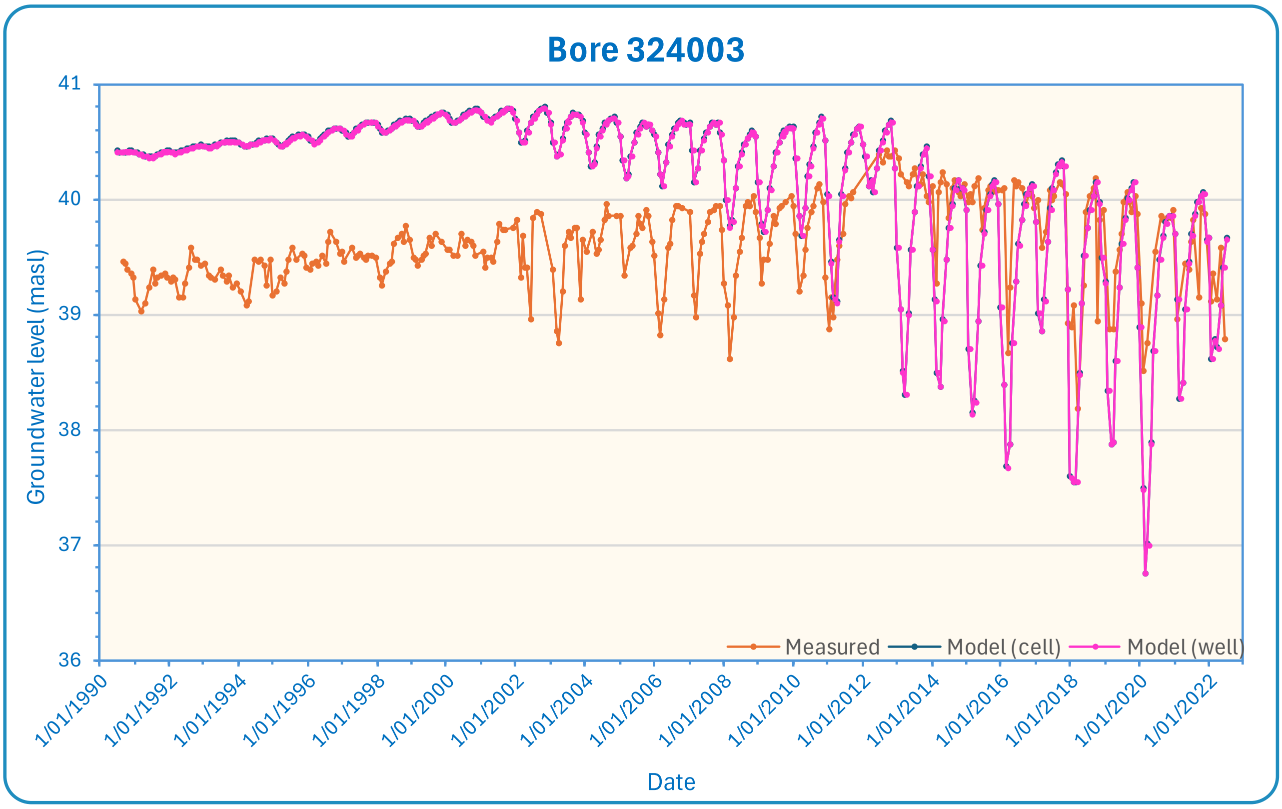


Figure D-23. Observed and model calculated groundwater level for Bore 324003.

Appendix E Observed and Modelled Groundwater Head Scatter Plots

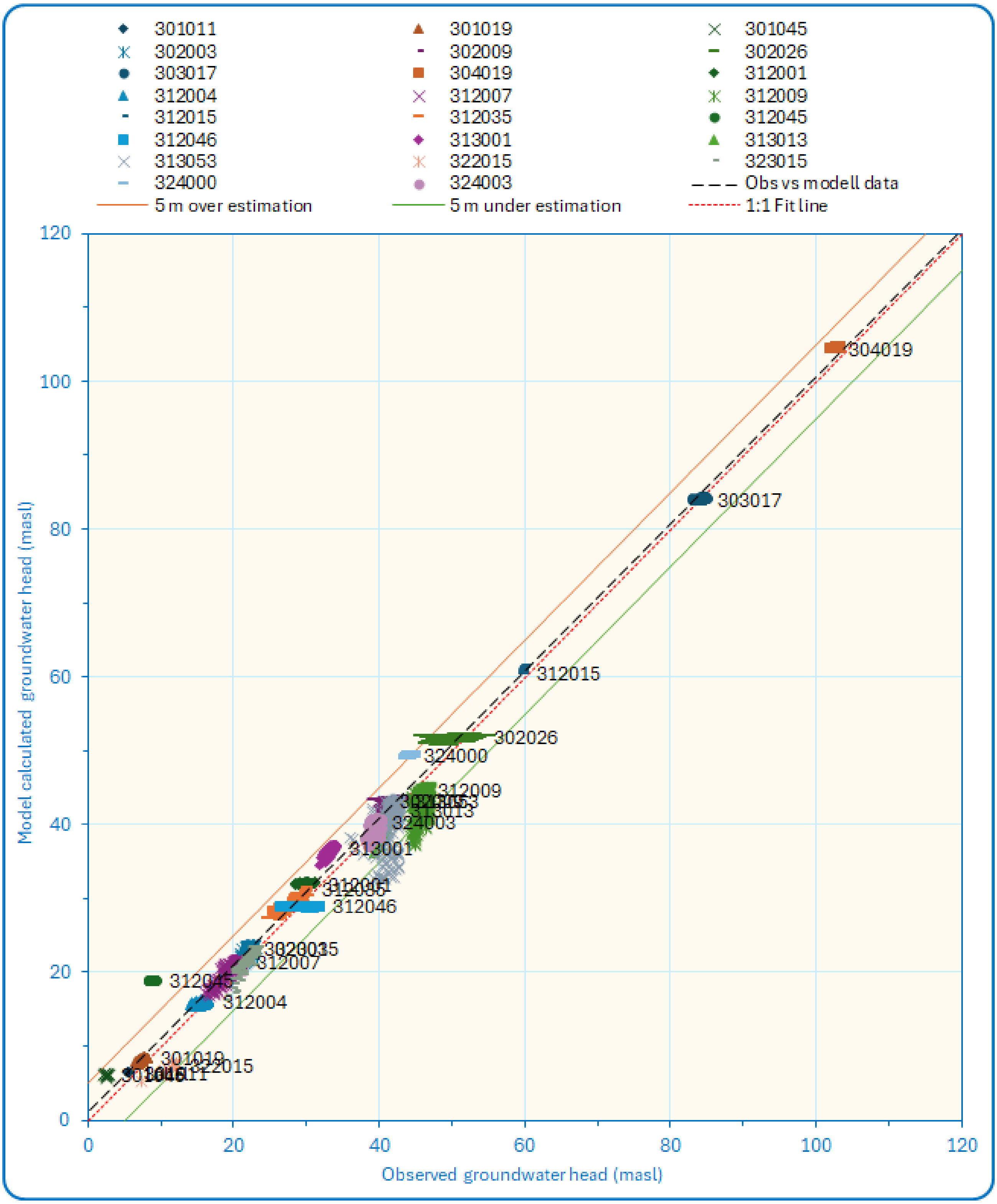


Figure E-1. Scatter plot of measured groundwater levels and model-calculated groundwater levels at the locations of bores on the dates of measurements (higher resolution plots are presented below).

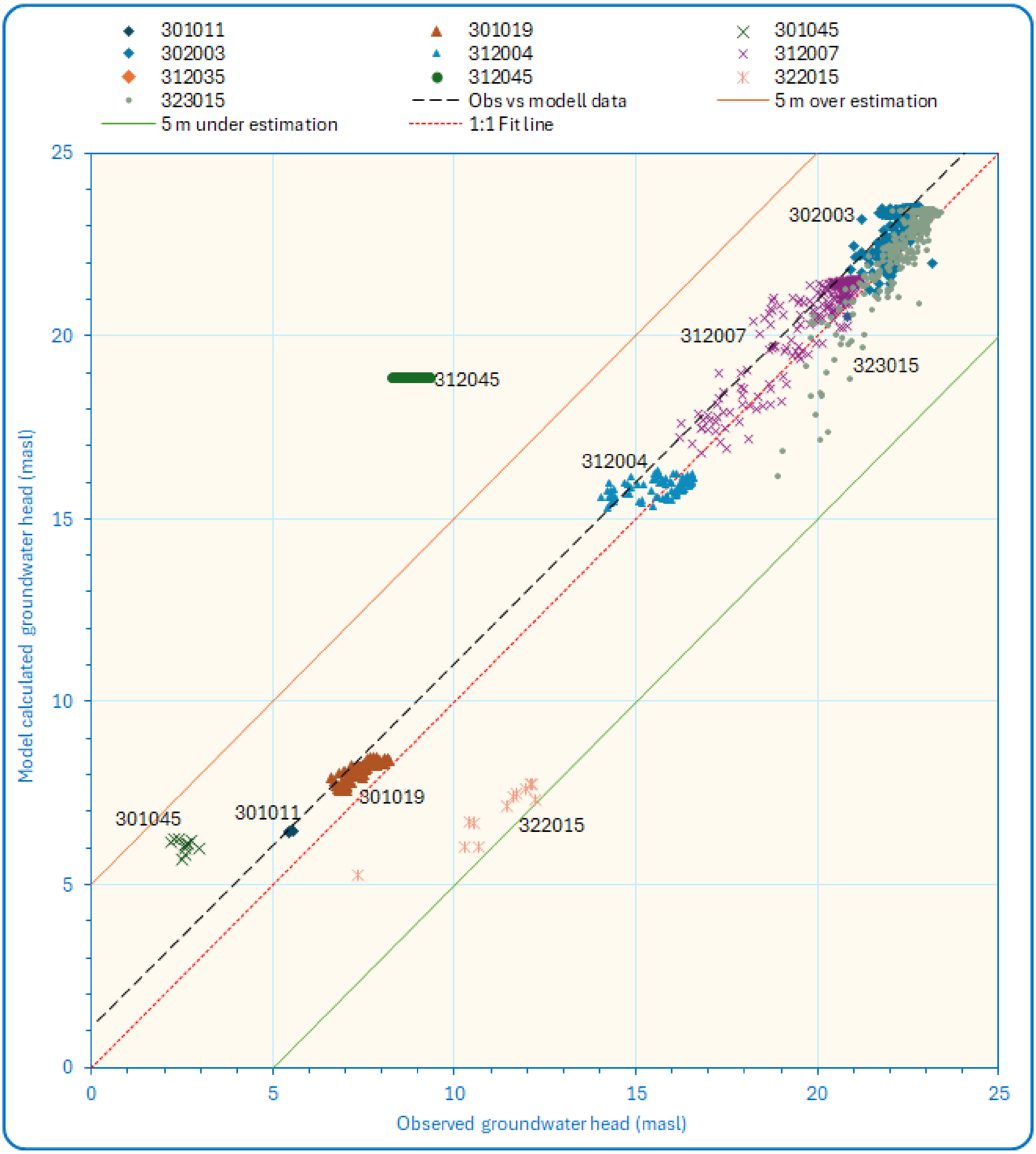


Figure E-2. Scatter plot of measured groundwater levels and model-calculated groundwater levels less than 25 masl at the locations of bores on the dates of measurements.

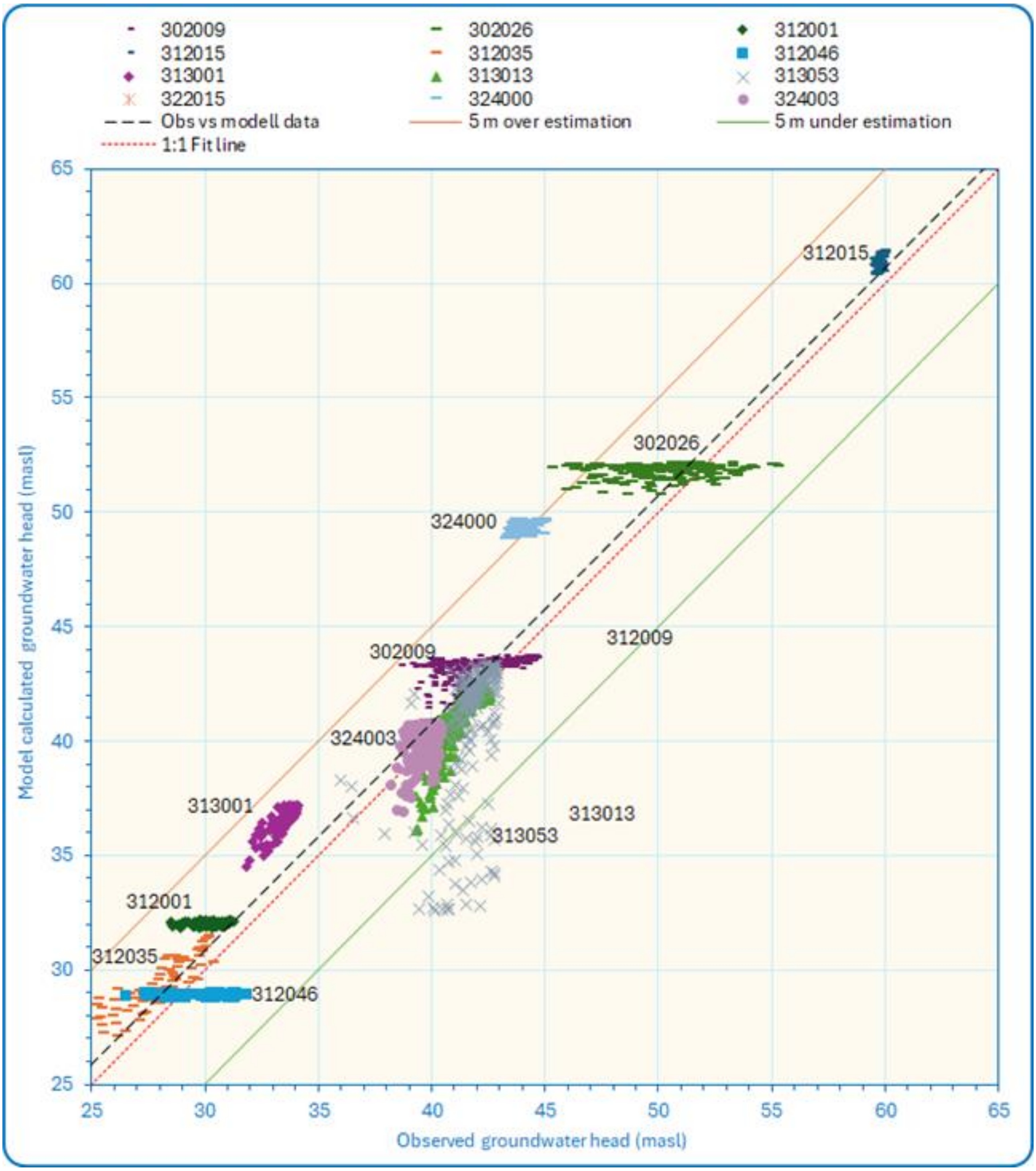


Figure E-3. Scatter plot of measured groundwater levels and model-calculated groundwater levels less than 25–65 masl at the locations of bores on the dates of measurements.

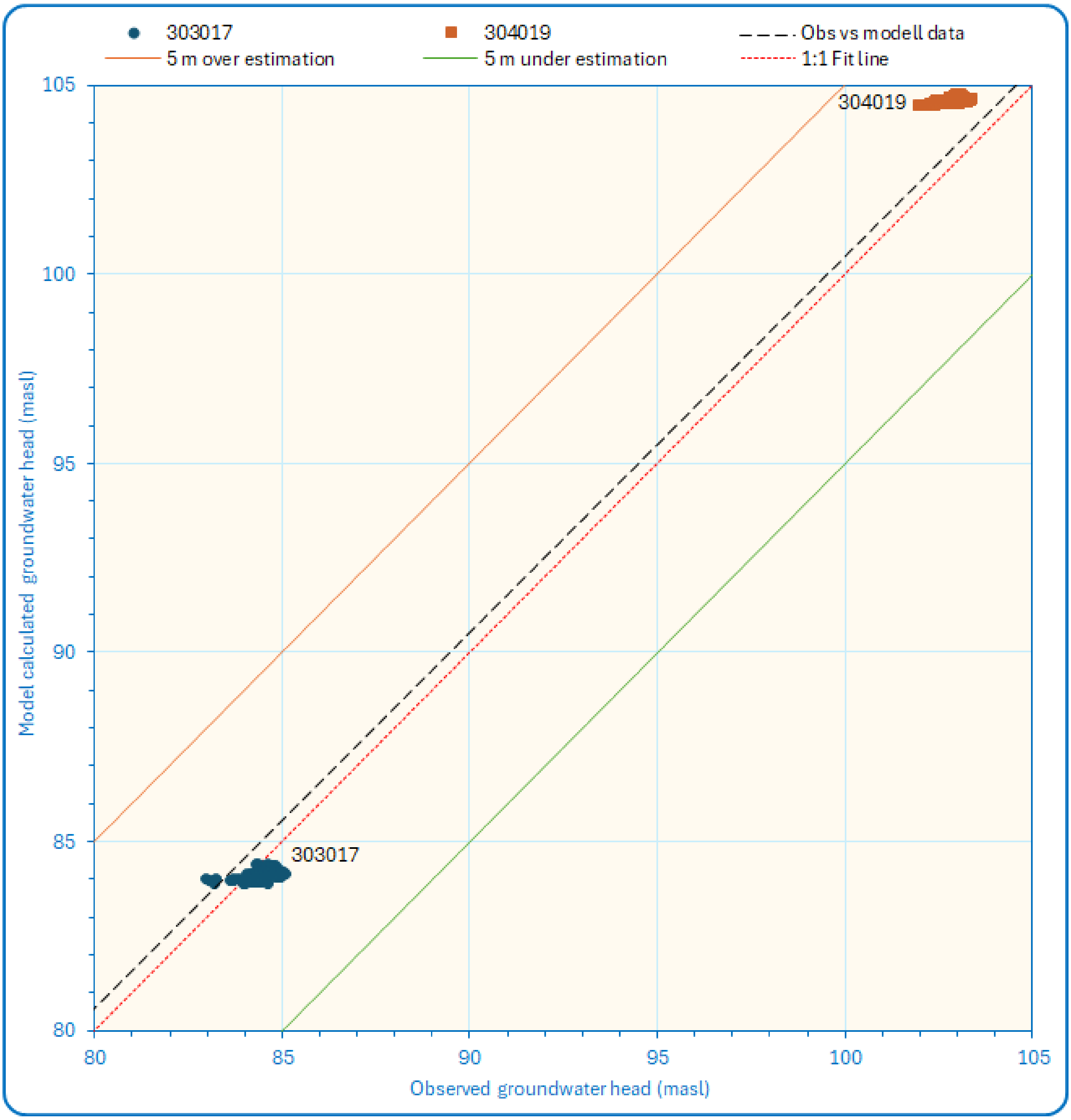


Figure E-4. Scatter plot of measured groundwater levels and model-calculated groundwater levels more than 80 masl at the locations of bores on the dates of measurements.

Appendix F Horizontal Hydraulic Conductivity Fence Diagrams

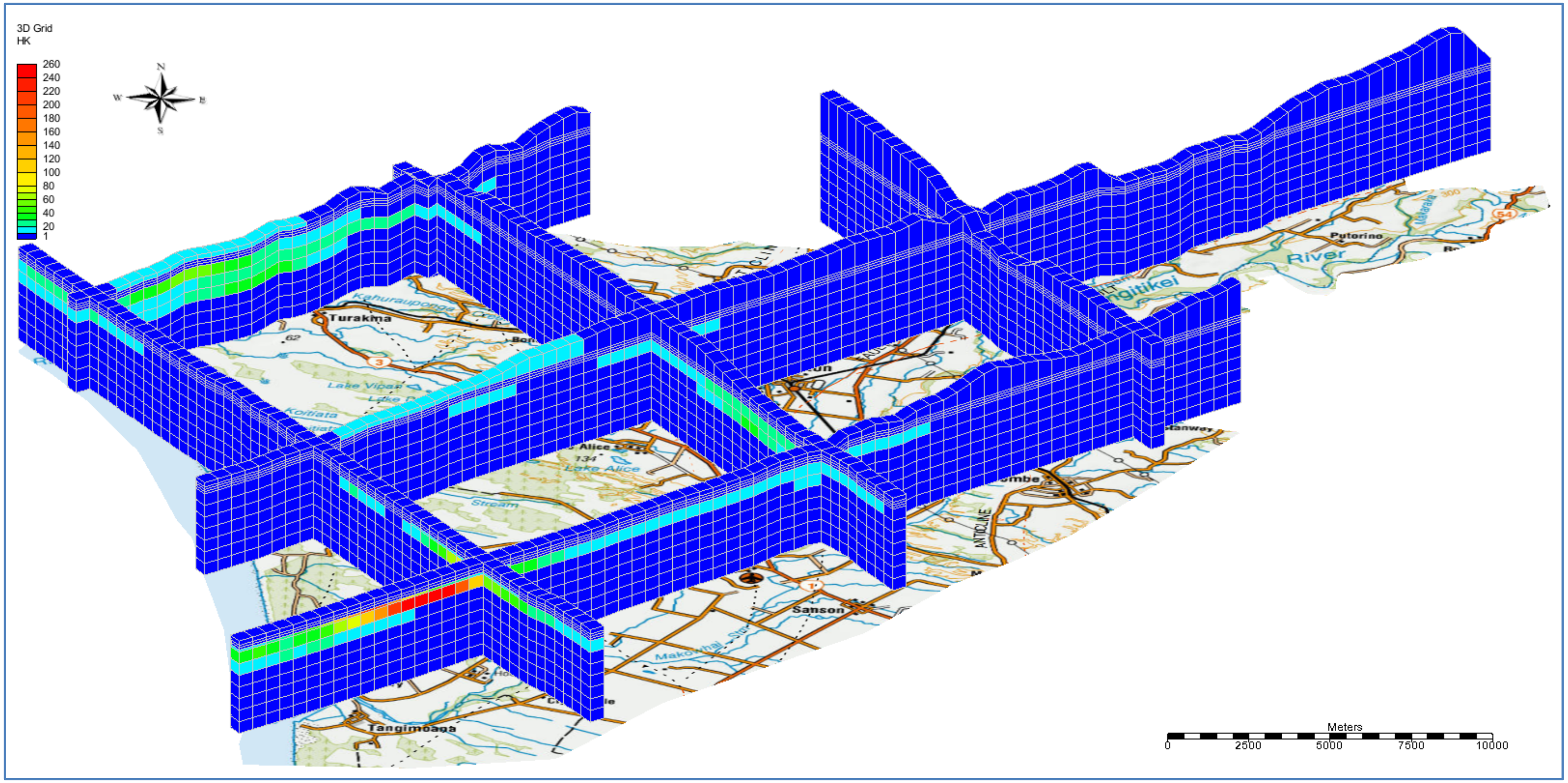


Figure F-1. Fence diagram showing calibrated model horizontal hydraulic conductivity values along selected transect on topographic basemap. For location of transcets, see Figure 4-25.

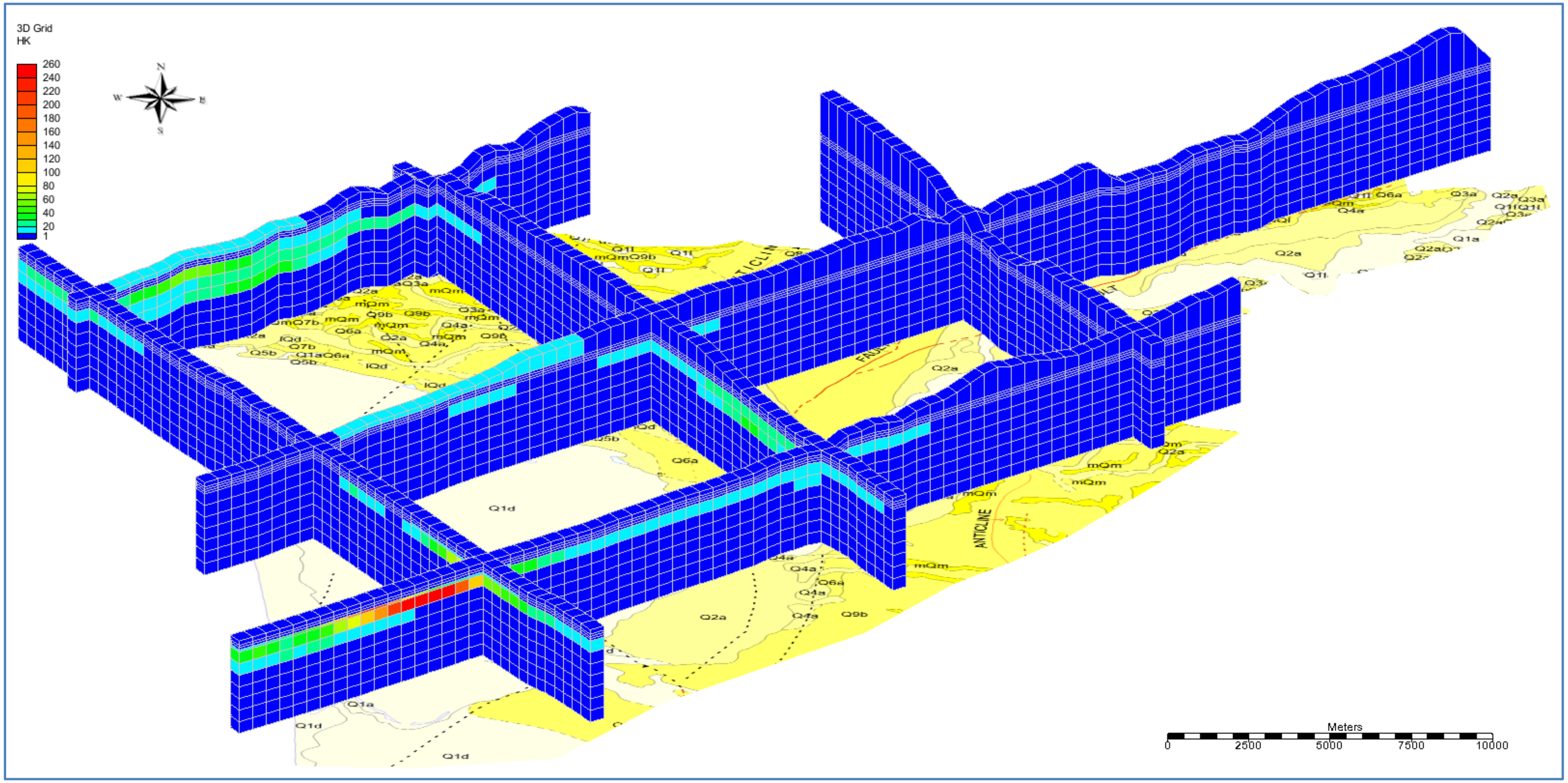


Figure F-2. Fence diagram showing calibrated model horizontal hydraulic conductivity values along selected transect on geology basemap. For location of transcets, see Figure 4-25.

Appendix G 2023 Model Calibrated Groundwater Head Sectional Views

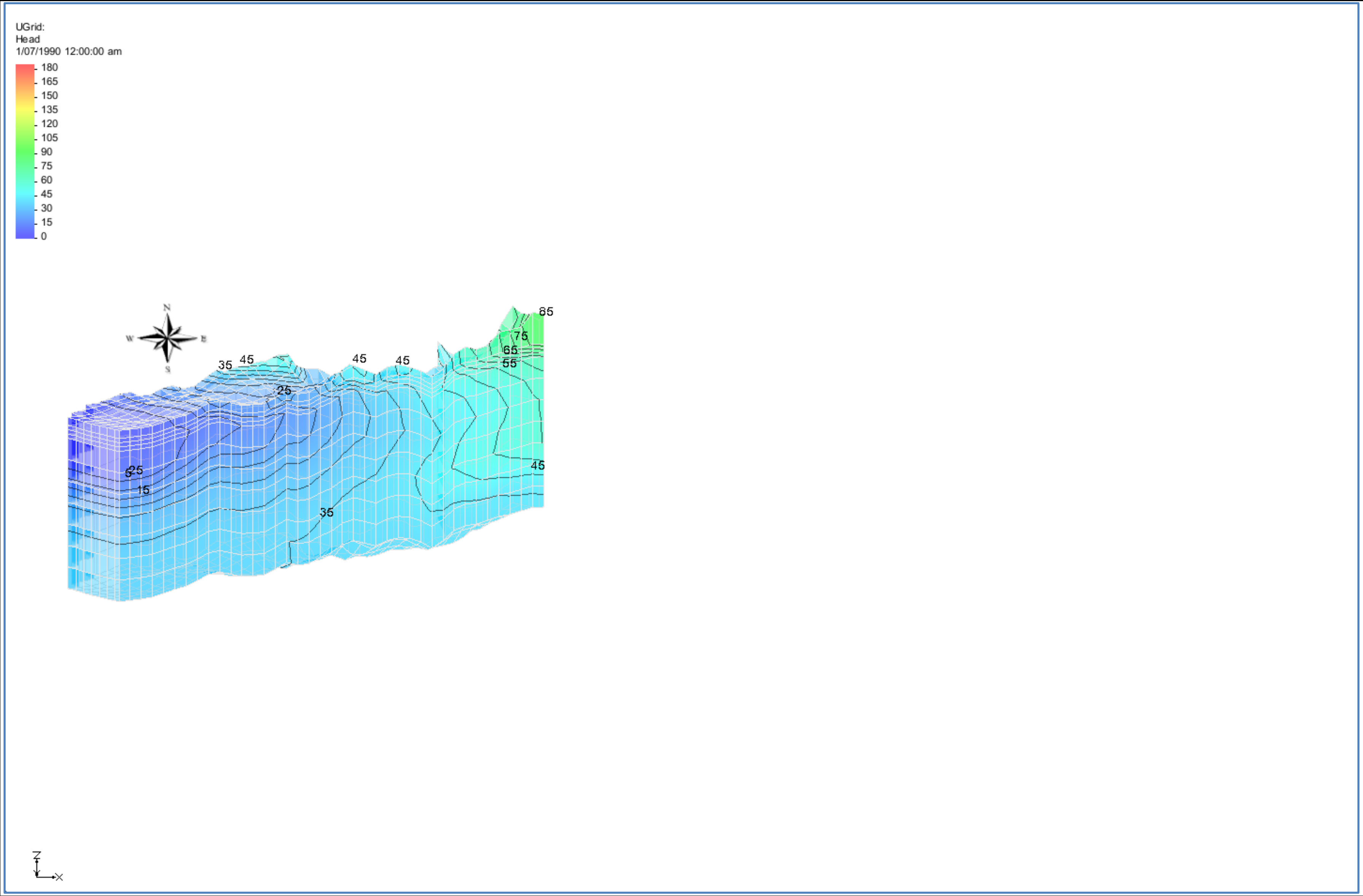


Figure G-1. SW-NE sectional view A showing groundwater heads from the calibrated model (for location, see Figure 4-25).

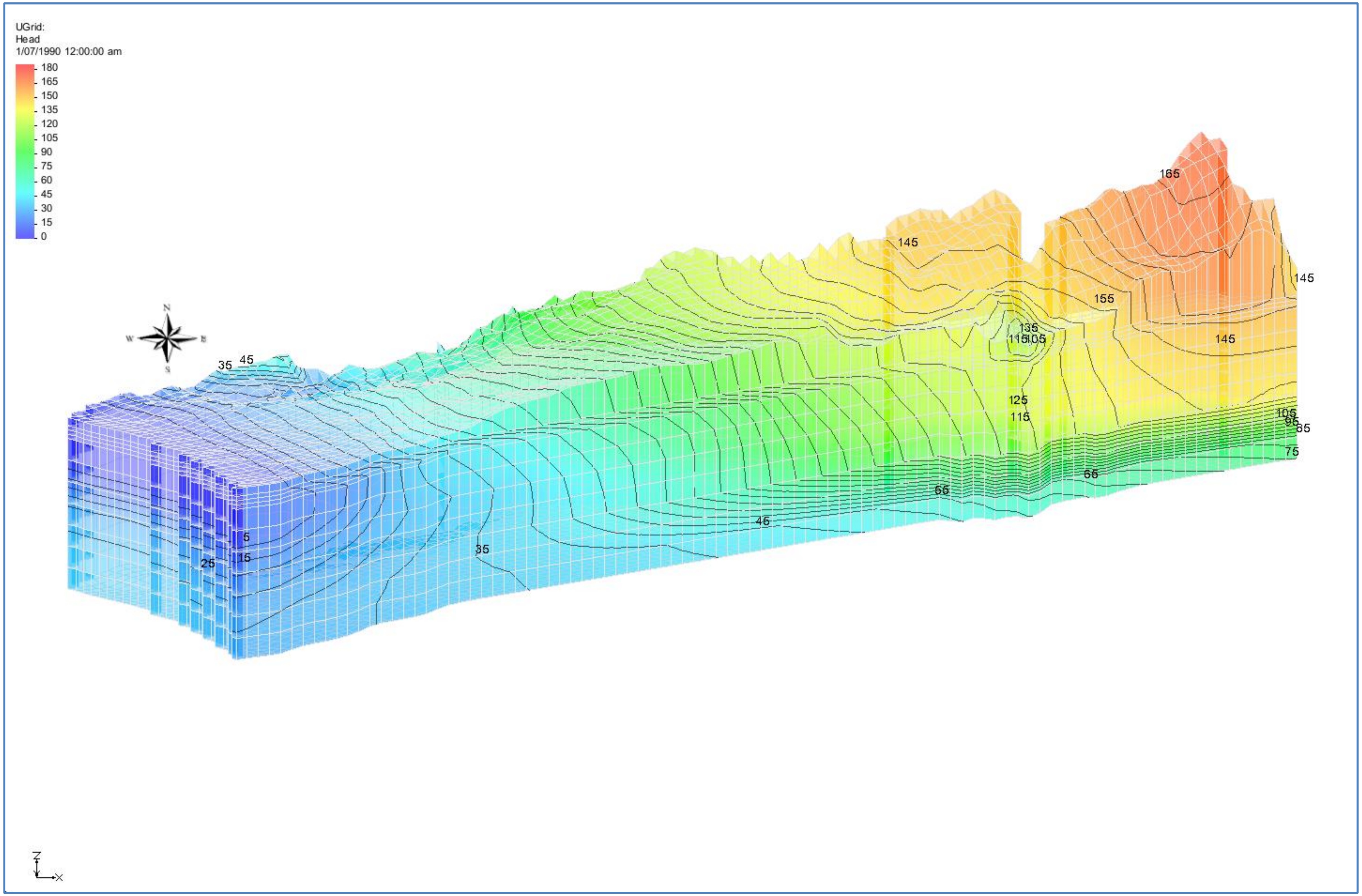


Figure G-2. SW-NE sectional view B showing groundwater heads from the calibrated model (for location, see Figure 4-25).

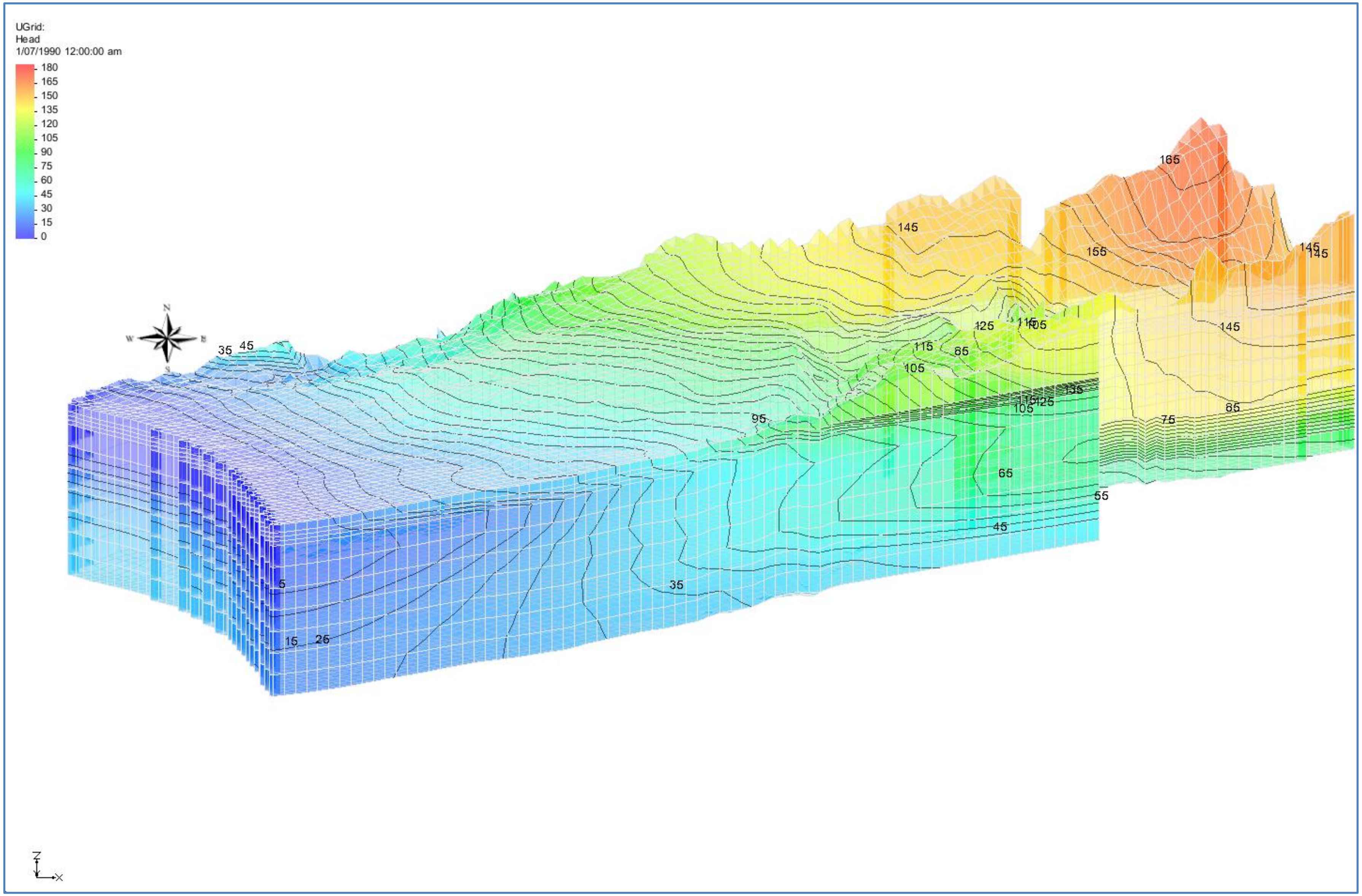


Figure G-3. SW-NE sectional view C showing groundwater heads from the calibrated model (for location, see Figure 4-25).

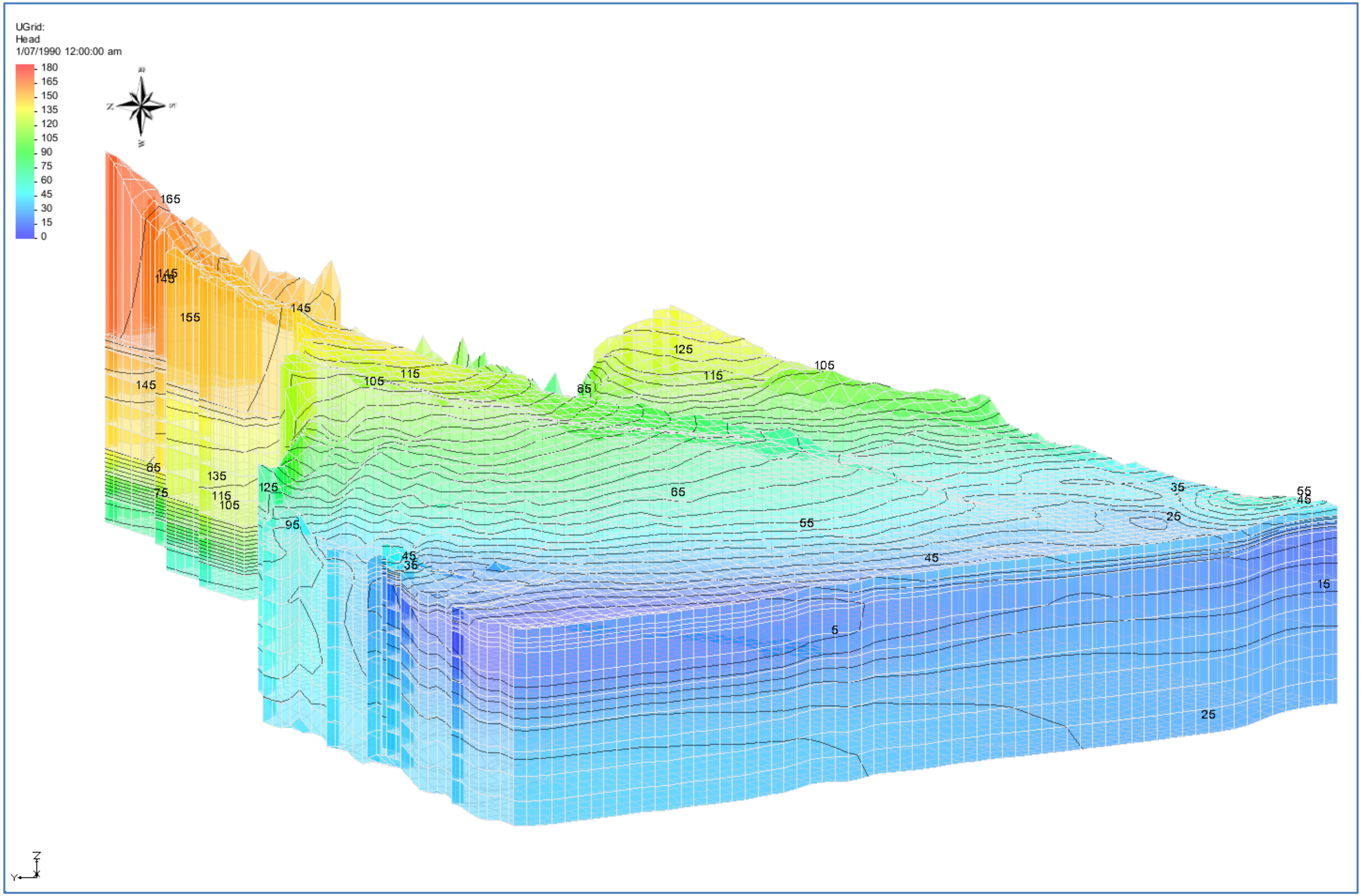


Figure G-4. NW-SE sectional view D showing groundwater heads from the calibrated model (for location, see Figure 4-25).

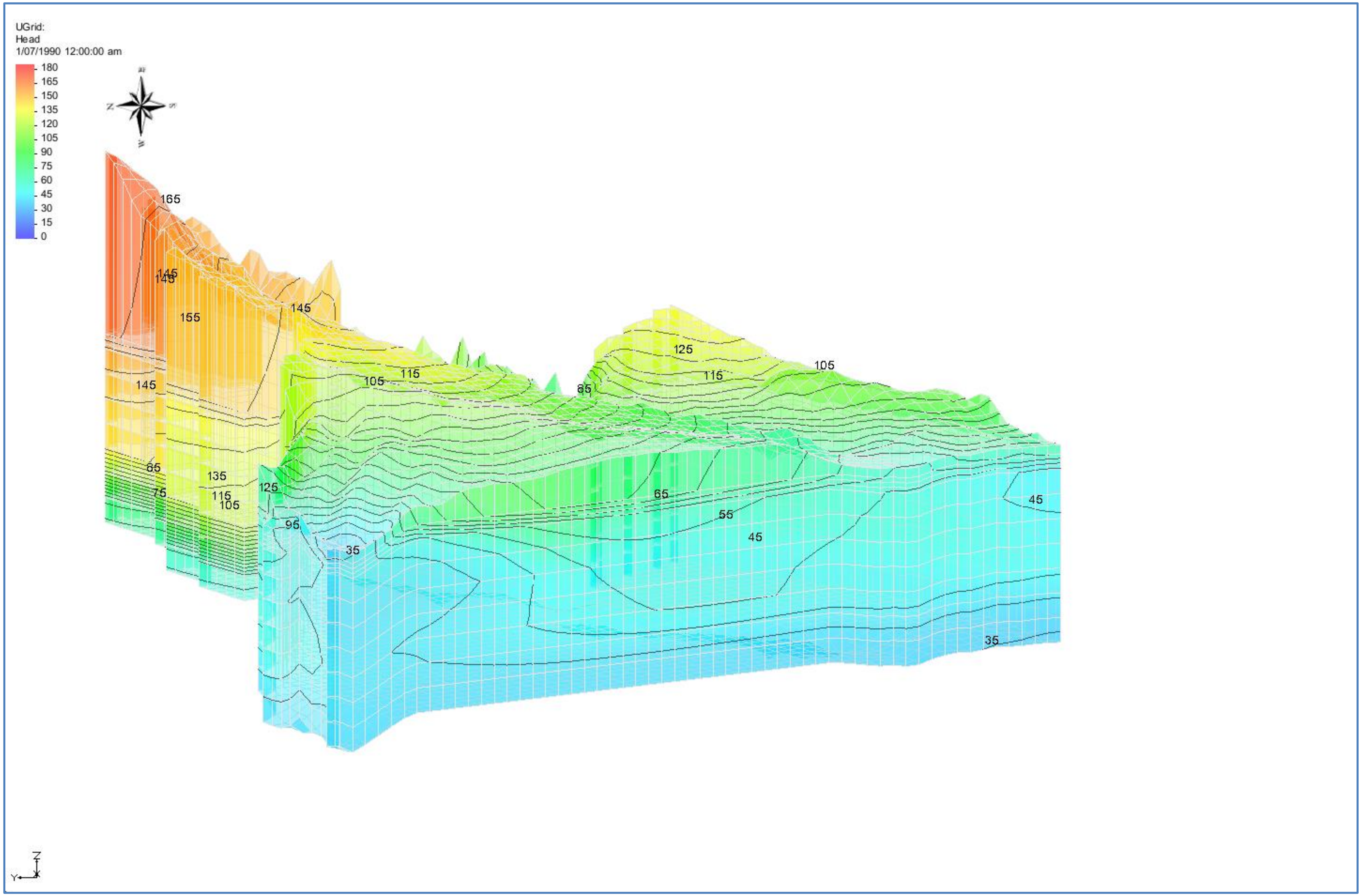


Figure G-5. NW-SE sectional view E showing groundwater heads from the calibrated model (for location, see Figure 4-25).

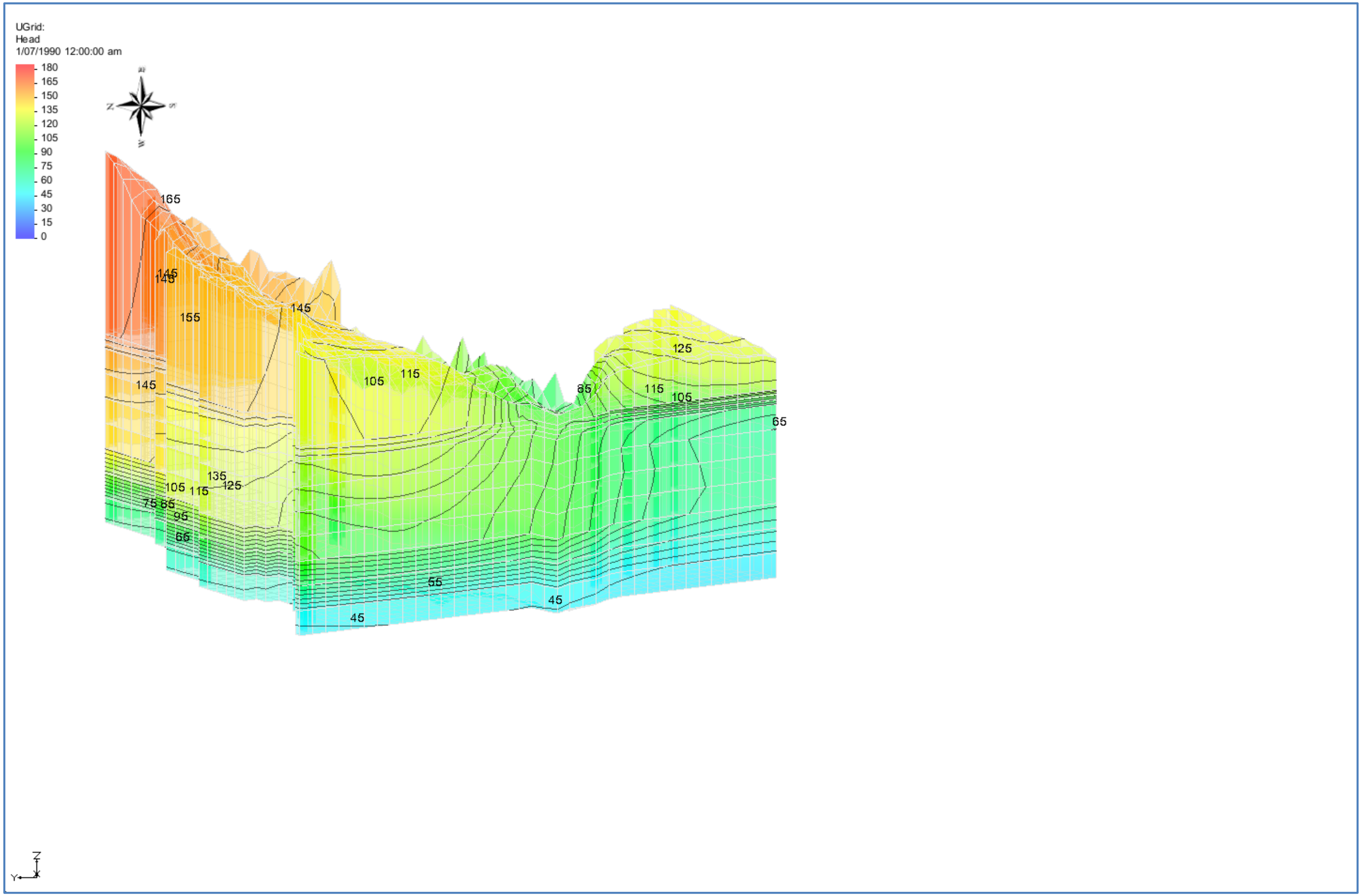


Figure G-6. NW-SE sectional view F showing groundwater heads from the calibrated model (for location, see Figure 4-25).

Appendix H Animated Groundwater Heads from the 2023 Model (digital files)

1. HistoricGroundwaterSim_2D.mp4 (video)
2. HistoricGroundwaterSim_3D.mp4 (video)

Appendix I Groundwater Model Review Checklists

The checklists in this appendix present an assessment of the latest model version, namely, the 2023 Santoft model reported on in PDP (2023). However, because PDP (2023) does not provide complete details on the model, some answers to the assessment questions are based on information from PDP (2017) and/or examination of the 2023 model files. The following two checklists are adapted from the AGMG (Barnett *et al.*, 2012):

1. Table I-1. Model [overall] compliance checklist.
2. Table I-2. Model [detailed] checklist.

In the tables below, green indicates adequacy, yellow neutrality, and red deficiency, and blue inapplicability or irrelevancy.

Table I-1. Model [overall] compliance checklist, adapted from the AGMG (Barnett *et al.*, 2012).

Question	***Answer***	***Comment***
Are the model objectives and model confidence level classification clearly stated?	Yes & No	The model **objectives** are clearly stated. The model target **confidence level classification** is not defined. Setting this target is important to help with the planning of the model, setting realistic expectations, and understand the model limitations. A water resource management model must meet Class 3 criteria—See Table 2-1 and Example 2.3 in the AGMG (Barnett *et al.*, 2012). The 2023 model is Class 1–2, with some Class 3 attributes.
Are the objectives satisfied?	No	The conceptual and numerical models fall short of adequately representing the hydrogeological system, which is critical for being able to provide credible predictions. Without adequate resemblance between real-life conditions and the model, the model cannot achieve its intended purpose/s even if it shows signs of good calibration.
Is the conceptual model consistent with objectives and confidence level classification?	No	The model is not informed by the geology. Its extent, mesh resolution, and surface water representation are not adequate. There are unrealistic parameter values. Not all processes are identified and included in the model (e.g., evapotranspiration from the water table). Overall, the model meets CLC Class X criteria.
Is the conceptual model based on all available data, presented clearly and reviewed by an appropriate reviewer?	No	The conceptual model is undermined by incomplete hydrogeological analyses. The lack of graphical presentation and the absence of an analytical water budget to guide numerical modelling are notable omissions. There is no indication that the latest version (2023) of the model was reviewed and accepted by a suitably qualified independent expert. The previous version (2017) was deemed not fit for purpose by its reviewer. The identified issues in the previous review were not addressed.
Does the model design conform to best practice?	No	There is no indication that best practice guidelines were followed in the model design. The reports reveal significant omissions and discrepancies, suggesting a deviation from accepted modelling standards. Most importantly, there is no indication that the common iterative modelling approach presented in Figure 1-4 was adopted.
Is the model calibration satisfactorily addressed?	No	The model calibration is fundamentally flawed due to a structure not informed by geology, leading to calibrated parameters that are not representative of the hydrostratigraphy. Consequently, despite a good statistical fit, the model's realism and predictive reliability are highly questionable.
Are the calibrated parameter values and estimated fluxes plausible?	No	The model files contain unrealistically high hydraulic conductivity values that are outside the reported ranges. Additionally, the model's representation of surface water is flawed. Extending the model domain 2–3 km into the sea has resulted in a critical error in the head at the coastline (6 masl), which has serious implications for heads and flows throughout the entire hydrogeological flow system, especially groundwater discharge to the ocean.
Do the model predictions conform to best practice?	Yes & No	The predictive modelling scenarios are well thought out. However, the model predictions are doubtable.
Is the uncertainty associated with the predictions reported?	No	The lack of assessment of predictive uncertainty is a serious shortcoming in the modelling exercise as it leaves decision-makers without a clear understanding of the potential range of outcomes.
Is the model fit for purpose?	No	The 2023 model is not fit for purpose. It is unreliable for informing changes in the groundwater allocation framework, and a complete revision is necessary, including enhancements in realism and a thorough uncertainty analysis, to support groundwater resource planning and management.

Table I-2. Model [detailed] checklist, adapted from the AGMG (Barnett *et al.*, 2012).

Review questions	Answer	Comment
1. Planning		
1.1 Are the Project objectives stated?	N/A	No Project is defined.
1.2 Are the model objectives stated?	Yes	Stated in the report introduction section.
1.3 Is it clear how the model will contribute to meeting the Project objectives?	Yes	By simulating different groundwater abstraction scenarios (unchanged, decreased, and increased)
1.4 Is a groundwater model the best option to address the Project and model objectives?	Yes	—
1.5 Is the target model confidence-level classification stated and justified?	No	A target model confidence level classification or equivalent is not considered.
1.6 Are the planned limitations and exclusions of the model stated?	No	PDP (2017) only indicates that the model is not intended to simulate local scale drawdown effects between abstraction bores.
2. Conceptualisation		
2.1 Has a literature review been completed, including examination of prior investigations?	No	Inadequate data compilation and literature review. This includes ignoring the review of the 2017 model during the development of the 2023 model.
2.2 Is the aquifer system adequately described?		
2.2.1 hydrostratigraphy including aquifer type (porous, Fractured Rock ...)	No	Hydrostratigraphy and lithology are not incorporated in the model.
2.2.2 lateral extent, boundaries, and significant internal features such as faults and regional folds	No	The model extent is not appropriately informed by the topography, geomorphology, geology, structure, and hydrology.
2.2.3 aquifer geometry including layer elevations and thicknesses	No	The aquifer and aquitard geometries have been set without reference to actual field data. Elevations and thicknesses do not reflect real field conditions, which undermines the model's reliability.
2.2.4 confined or unconfined flow and the variation of these conditions in space and time?	No	The hydrostratigraphy and the groundwater flow system are not adequately characterised or represented in the model.
2.3 Have data on groundwater stresses been collected and analysed?		
2.3.1 recharge from rainfall, irrigation, floods, lakes	Yes & No	Rainfall recharge data collected and analysed, but there are apparent errors. Irrigation, floods, and lakes data have not been collected or analysed.
2.3.2 river or lake stage heights	No	Lake stage heights are not included in the model. According to the report (PDP, 2017), river stage heights are assumed to be 1 m below the ground surface elevation. Inspecting the 2023 model files reveal that the river stage heights are set at 0.5 mbgl. Overall, the rivers are poorly represented in the mode, including stage and bottom heights, lateral extent, and bed conductance.
2.3.3 groundwater usage (pumping, returns etc)	Yes	Available data interpolated and extrapolated to represent abstraction reasonably. There could be other methods to do that, but the method selected in adequate. However, associated uncertainty must be factored in the modelling and subsequent decision making.
2.3.4 evapotranspiration	No	Evapotranspiration, a potentially significant stressor on the system, was neither analysed nor considered in the model, despite recommendations from the review of the 2017 model to do so.

Review questions	Answer	Comment
2.3.5 other?	—	—
2.4 Have groundwater level observations been collected and analysed?		
2.4.1 selection of representative bore hydrographs	No	All groundwater data records have been delt with equally, without assigning quality tags or selecting representative hydrographs. For example, Bore 301011 has only few measurements over a short period of time (Figure D-1) but is treated as equal to Bore 302003 which has monthly readings from the start of monitoring until now (Figure D-4).
2.4.2 comparison of hydrographs	Yes	But not adequate and not helping with understanding stress-impact relationships.
2.4.3 effect of stresses on hydrographs	No	See above.
2.4.4 water-table maps/piezometric surfaces?	Yes & No	PDP (2017) presents the results of a piezometric survey undertaken in October 2014 in the form of two piezometric maps for bores shallower and deeper than 20 mbgl. It does not present the rational for this arbitrary division of the aquifer system. The lack of a data-driven basis for this division could lead to misinterpretations of groundwater behaviour and conceptual and numerical modelling errors, including groundwater-surface water interaction and allocation limit estimates.
2.4.5 If relevant, are density and barometric effects taken into account in the interpretation of groundwater head and flow data?	N/A	Density and barometric effects are deemed irrelevant to this model due to its relatively large spatial scale and extended temporal timeframe.
2.5 Have flow observations been collected and analysed?		
2.5.1 baseflow in rivers	No	This is considered a significant omission in this model.
2.5.2 discharge in springs	N/A	There are no known springs or seepage faces within the model area.
2.5.3 location of diffuse discharge areas?	No	This is considered a significant omission in the model. Although data on surface water gains from groundwater are available, they were not utilised in the conceptual or numerical models.
2.6 Is the measurement error or data uncertainty reported?		
2.6.1 measurement error for directly measured quantities (e.g., piezometric level, concentration, flows)	No	The only 'measured' quantities are rainfall, depth to groundwater, and elevations of groundwater measurement points. The measurement accuracy of these quantities are not considered
2.6.2 spatial variability/heterogeneity of parameters	No	There is a map showing transmissivity (T) values (Figure 4-1). However, it does not reflect heterogeneity in the horizontal hydraulic conductivity (HK) as T is the product of HK by the aquifer thickness (b). The figure also presents the bore depth, but that is not enough to be able to relate HK to a hydrostratigraphic unit. Hence, HK heterogeneity remains unclear.
2.6.3 interpolation algorithm(s) and uncertainty of gridded data?	No	No information is provided on the interpolation methodologies, including the contouring method used to produce potentiometric maps like the one shown in Figure 4-14.
2.7 Have consistent data units and geometric datum been used?	No	Examination of the reports and model files reveals potential confusion between storativity (S, dimensionless) and specific storage (S_s, m^{-1}). Additionally, the units for General Head Boundary (GHB), river (RIV), and drain (DRN) conductance (C) are unclear. They could be m/d, as suggested in Table 3 of PDP (2017), or m^2/d, as commonly used in groundwater models. The choice of units reflects the calculation method used and, subsequently, the range of feasible values. As the units used in the model are unknown, the appropriateness of the values cannot be evaluated.
2.8 Is there a clear description of the conceptual model?		
2.8.1 Is there a graphical representation of the conceptual model?	No	—
2.8.2 Is the conceptual model based on all available, relevant data?	No	Surface water data not considered (location, interaction with groundwater, etc.).

Review questions	Answer	Comment
2.9 Is the conceptual model consistent with the model objectives and target model confidence level classification?		
2.9.1 Are the relevant processes identified?	No	Evapotranspiration from the water table is excluded from the analysis without sufficient data or evidence to justify this omission.
2.9.2 Is justification provided for omission or simplification of processes?	No	—
2.10 Have alternative conceptual models been investigated?	No	Alternative conceptual models have not been investigated, either as a whole or in their components. For example, alternative representations of the coastal boundary have not been discussed. This boundary could be represented as a Constant Head Boundary (CHD) or General Head Boundary (GHB). Also, its representation can consider or ignore difference in density between groundwater and oceanwater.
3. Design and Construction		
3.1 Is the design consistent with the conceptual model?	No	The numerical model boundary conditions and parameterisation are inconsistent with the conceptual model. More importantly, they do not adequately representative the modelled groundwater system.
3.2 Is the choice of numerical method and software appropriate?		
3.2.1 Are the numerical and discretisation methods appropriate?	No	The layering and grid/mesh design are inappropriate, failing to reflect the system characteristics such as surface water distribution and stratigraphy. The model design also does not accommodate the representation of receptors needed to detect the impacts of changes in stresses. There are issues with boundary condition design, including DRN cells located in the sea (see Figure 4-16) and overlapping boundaries where cells have simultaneous RCH, DRN, and RIV designations. It appears that the modellers were not aware of these problems in the 2023 model. The 2017 model files were not examined, so it is unknown whether the 2017 model suffers from the same issues.
3.2.2 Is the software reputable?	Yes	The 2023 model is built using MODFLOW 6, which comes from a reputable and proven family of modelling code. However,
3.2.3 Is the software included in the archive or are references to the software provided?	Yes	MODFLOW 6 is in the public domain and included in many commercial modelling packages.
3.3 Are the spatial domain and discretisation appropriate?		
3.3.1 1D/2D/3D	OK	The model is appropriately designed as 3D model.
3.3.2 lateral extent	No	The peripheral boundaries are assigned based on inappropriate assumptions. The assumptions are inappropriate in terms of the model objectives (checking the appropriateness of the RGMZ allocation limits) and hydrogeology.
3.3.3 layer geometry?	No	Arbitrary fixed thickness layers with the top of the grid corresponding to land surface based on poor quality data (20 m contours from 1:250 000 topographic maps).
3.3.4 Is the horizontal discretisation appropriate for the objectives, problem setting, conceptual model and target confidence level classification?	No	Fixed 500 m x 500 m cells are too big in areas of interest (rivers, streams, drains, wetlands, and coastal dune lakes).
3.3.5 Is the vertical discretisation appropriate? Are aquitards divided in multiple layers to model time lags of propagation of responses in the vertical direction?	No	The vertical discretisation is arbitrarily and does not reflect the modelled groundwater system settings.

Review questions	Answer	Comment
3.4 Are the temporal domain and discretisation appropriate?		
3.4.1 steady-state or transient	Yes	Transient model with the first step (1 year) being a transient simulation.
3.4.2 stress-periods	Yes	Monthly stress periods, enabling reasonable representation of the main stresses (climate and abstraction).
3.4.3 time steps?	Yes	One time step per stress period. Future modelling must check the need to change or keep this setting.
3.5 Are the boundary conditions plausible and sufficiently unrestrictive?		
3.5.1 Is the implementation of boundary conditions consistent with the conceptual model?	No	Attention is drawn here to inconsistencies between the model and the report. For instance, GHB conditions are reported as being assigned to the top three layers, whereas they are actually assigned to the top six layers. Additionally, surface water features are poorly represented. No-flow boundaries in the model do not correspond to groundwater divides or solid geological boundaries in the modelled system.
3.5.2 Are the boundary conditions chosen to have a minimal impact on key model outcomes? How is this ascertained?	No	—
3.5.3 Is the calculation of diffuse recharge consistent with model objectives and confidence level?	No	Not reported to be checked.
3.5.4 Are lateral boundaries time-invariant?	No	The coastal boundary is represented using a General Head Boundary (GHB), not just as a single line but as an entire zone of GHB cells extending 2–3 km into the ocean across layers 1–6. This approach is fundamentally flawed and would invalidate the model even if everything else were correct. As a result, the modelled groundwater head at the shoreline is 6 masl, which is entirely incorrect and impossible. The coastal boundary should have been simplified and represented by a Constant Head Boundary (CHD) to eliminate the need for accounting for temporal variations, as these do not occur at this boundary.
3.6 Are the initial conditions appropriate?		
3.6.1 Are the initial heads based on interpolation or on groundwater modelling?	Constant & Model	The reports do not mention initial conditions, which is a clear omission. Examination of the model files indicate: The initial head for the steady state model is a constant value of 200 masl. The initial heads for the transient model are from the steady-state model. Other options that should have been considered for the steady state model initial groundwater heads include: interpolation from observed data using the topography or top of the grid elevation.
3.6.2 Is the effect of initial conditions on key model outcomes assessed?	No	Not considered in the modelling.
3.6.3 How is the initial concentration of solutes obtained (when relevant)?	N/A	—
3.7 Is the numerical solution of the model Adequately addressed?		
3.7.1 Solution method/solver	No	Not reported.
3.7.2 Convergence criteria	No	Not reported.

Review questions	Answer	Comment
3.7.3 Numerical precision	No	Not reported.
4. Calibration and Sensitivity		
4.1 Are all available types of observations used for calibration?		
4.1.1 Groundwater head data	No	Only data from SoE monitoring network. Consent compliance, one-off surveys, studies, and other data not used.
4.1.2 Flux observations	No	This is a major deficiency in the model.
4.1.3 Other: environmental tracers, gradients, age, temperature, concentrations etc.	No	—
4.***2 Does the calibration methodology conform to best practice?***		
4.2.1 Parameterisation	No	Pilot Points are arbitrarily placed without reference to hydrostratigraphic characteristics or meaningful zoning (laterally and vertically) do not provide for meaningful modelling. Pilot Point zones are defined as whole layers except for Layer 1 where subzones are defined. The zone definition is flawed and unrealistic. It is noticed that the parameter ranges in the examined 2023 model files exceed the limits reported in PDP (2017). PDP (2023) does not report parameter ranges. Hence, it was assumed that the parameter ranges in the 2023 model are the same as the 2017 model. Additionally, it must be noticed that some of the parameters used are unrealistic and some are physically impossible.
4.2.2 Objective function	No	Not reported
4.2.3 Identifiability of parameters	No	Many important parameters have not been included in the PEST analysis and the pilot points are not grouped meaningfully. Hence, the parameter identifiability analysis is incomplete and not very useful.
4.2.4 Which methodology is used for model calibration?	PEST Pilot Points	The pilot point automated calibration method used is standard, but its application is flawed due to meaningless assignment of pilot points to arbitrary layers.
4.3 Is a sensitivity of key model outcomes assessed against?		
4.3.1 parameters	No	This is required in future versions of the model.
4.3.2 boundary conditions	No	This is required in future versions of the model.
4.3.3 initial conditions	No	This is required in future versions of the model.
4.3.4 stresses	No	This is required in future versions of the model.

Review questions	Answer	Comment
4.4 Have the calibration results been adequately reported?		
4.4.1 Are there graphs showing modelled and observed hydrographs at an appropriate scale?	Yes	PDP (2017 and 2023) present hydrographs comparing model-calculated and observed groundwater heads. However, the locations and measurement times do not exactly coincide. The model-calculated values represent the centre of the cell assigned to the observation well at the end of the stress period (end of the month). This can result in a location offset of up to 350 m and a time difference of up to 30 days between the observed measurement and the corresponding model value. These approximations can cause deviations that may make the modelled head appear closer to or farther from the measured head. Further discussion on this matter is provided in Section 4.4.4. This issue must be addressed in future models. PDP (2023) notes that the 2023 model was calibrated to annual maxima and monthly measured groundwater level data (Figure 3-3). PDP (2017) does not suggest that the 2017 model was calibrated to annual maxima groundwater level data. However, this has been picked up by the reviewer (Barnett, 2016) based on the calibration hydrographs in the report version he reviewed. The final PDP (2017) report reviewed herein does not include these plots. Barnett (2016) suggests that the 2017 model was likely calibrated against maximum annual head targets because it under-predicts the seasonal variability in groundwater heads, particularly due to increased groundwater abstraction in recent years. He notes that this approach tends to underestimate the drawdown levels indicated by the monitoring bore data. Barnett argues that if seasonal fluctuations in head are to be ignored during calibration, it would have been more appropriate to match the modelled heads to the mean annual measured head in the observation bores. In contrast, PDP (2023) argue that they placed greater emphasis on matching the annual maximum water levels rather than monthly measurements, as this approach better aligns the model with long-term trends rather than short-term seasonal changes. This review supports Barnett (2016) opinion and advice. The scale of the hydrographs in PDP (2017 and 2023) is acceptable, but better plot size and resolution are recommended for future modelling reports. More suitable y-axis value ranges must be used in future groundwater level hydrographs.
4.4.2 Is it clear whether observed or assumed vertical head gradients have been replicated by the model?	No	This analysis has not been presented in the reports.
4.4.3 Are calibration statistics reported and illustrated in a reasonable manner?	No	Some calibration statistics are reported in PDP (2017) for the 2017 model, but no calibration statistics are reported for the 2023 model, which is an updated and recalibrated version of the 2017 model.
4.5 Are multiple methods of plotting calibration results used to highlight goodness of fit robustly? Is the model sufficiently calibrated?		
4.5.1 spatially	No	This must be presented in future reports to enable checking on possible systematic errors or bias.
4.5.2 temporally	Yes	Notwithstanding the comments in question 4.4.1 above.

Review questions	Answer	Comment
4.6 Are the calibrated parameters plausible?		
4.7 Are the water volumes and fluxes in the water balance realistic?	No	Table 4 in PDP (2017) presents the 2017 model the average of the various water balance components in m³ and % of total for the years 2000, 2007, and 2012. There is no analytical water budget to compare the numerical model water budget against. In addition, the unrealistic head at the coastline (6 masl) is enough to render the entire numerical groundwater model follow budget untrustworthy.
4.8 has the model been verified?	No	The development (updating) of the 2023 model should have started with comprehensive verification (post-audit) of the 2017 model.
5. Prediction		
5.1 Are the model predictions designed in a manner that meets the model objectives?	Yes	—
5.2 Is predictive uncertainty acknowledged and addressed?	No	—
5.3 Are the assumed climatic stresses appropriate?	No	The reliability of the recharge models used in the 2017 and 2023 models cannot be verified. In addition, the model excludes evapotranspiration, which could be an important hydrogeological process in the coastal area.
5.4 Is a null scenario defined?	Yes	According to Section 3.1 in PDP (2023), '*scenario where no abstraction is simulated to indicate the effect of groundwater takes compared to a 'natural' setting.*' However, it is not called 'null scenario' in the reports.
5.5 Are the scenarios defined in accordance with the model objectives and confidence level classification?		
5.5.1 Are the pumping stresses similar in magnitude to those of the calibrated model? If not, is there reference to the associated reduction in model confidence?	Yes	—
5.5.2 Are well losses accounted for when estimating maximum pumping rates per well?	N/A	—
5.5.3 Is the temporal scale of the predictions commensurate with the calibrated model? If not, is there reference to the associated reduction in model confidence?	Yes	2017 model: 25 years calibration, 5 years predictions. 2023 model: 32 years calibration, 10 years predictions.
5.5.4 Are the assumed stresses and timescale appropriate for the stated objectives?	Yes	—
5.6 Do the prediction results meet the stated objectives?	No	Apparently, the answer to this question is yes, but the model is thought to be unreliable. Hence, the answer is 'No'.
5.7 Are the components of the predicted mass balance realistic?		
5.7.1 Are the pumping rates assigned in the input files equal to the modelled pumping rates?	Unknown	Pumping rates not provided. So, the rates in the model files could not be verified.
5.7.2 Does predicted seepage to or from a river exceed measured or expected river flow?	No	Surface water poorly and unrealistically conceptualised and simulated across the model.

Review questions	Answer	Comment
5.7.3 There are no anomalous boundary fluxes due to superposition of head dependent sinks (e.g., evapotranspiration) on head-dependent boundary cells (Type 1 or 3 boundary conditions)?	No	There are DRN cells and boundary conditions in areas they should not be in (e.g., high elevations and the sea). There are also overlapping boundary conditions at some cells (DRN, RCH, RIV).
5.7.4 Is diffuse recharge from rainfall smaller than rainfall?	Yes	—
5.7.5 Are model storage changes [not][12] dominated by anomalous head increases in isolated cells that receive recharge?	Yes	—
6. Uncertainty		
6.1 Is some qualitative or quantitative measure of uncertainty associated with the prediction reported together with the prediction?	No	Robust predictive uncertainty analysis must be included in future work.
6.2 Is the model with minimum prediction-error variance chosen for each prediction?	No	—
6.3 Are the sources of uncertainty discussed?		
6.3.1 measurement of uncertainty of observations and parameters	No	—
6.3.2 structural or model uncertainty	No	—
6.4 Is the approach to estimation of uncertainty described and appropriate?	No	—
6.5 Are there useful depictions of uncertainty?	No	—
7. Solute transport[13]		
7.1 Has particle tracking been considered as an alternative to solute transport modelling?	N/A	Particle tracking should have been used to delineate capture zones for bores to help exploring potential reasons for rising and dropping groundwater levels
7.2 Has all available data on the solute distributions, sources and transport processes been collected and analysed?	N/A	—
7.3 Has the appropriate extent of the model domain been delineated and are the adopted solute concentration boundaries defensible?	N/A	—
7.4 Is the choice of numerical method and software appropriate?	N/A	—
7.5 Is the grid design and resolution adequately addressed, and has the effect of the discretisation on the model outcomes been systematically evaluated?	N/A	—
7.6 Is there sufficient basis for the description and parameterisation of the solute transport processes?	N/A	—
7.7 Are the solver and its parameters appropriate for the problem under consideration?	N/A	—

[12] Changed from the source document (AGMG 2012).
[13] The particle tracking question has been moved under the solute transport questions. In the source document (AGMG 2012), it is included under prediction questions. This section's questions have been re-numbered accordingly.

Review questions	Answer	Comment
7.8 Has the relative importance of advection, dispersion and diffusion been assessed?	N/A	—
7.9 Has an assessment been made of the need to consider variable density conditions?	N/A	—
7.10 Is the initial solute concentration distribution sufficiently well-known for transient problems and consistent with the initial conditions for head/pressure?	N/A	—
7.11 Is the initial solute concentration distribution stable and in equilibrium with the solute boundary conditions and stresses?	N/A	—
7.12 Is the calibration based on meaningful metrics?	N/A	—
7.13 Has the effect of spatial and temporal discretisation and solution method [been] taken into account in the sensitivity analysis?	N/A	—
7.14 Has the effect of flow parameters on solute concentration predictions been evaluated, or have solute concentrations been used to constrain flow parameters?	N/A	—
7.15 Does the uncertainty analysis consider the effect of solute transport parameter uncertainty, grid design and solver selection/settings?	N/A	—
7.16 Does the report address the role of geologic heterogeneity on solute concentration distributions?	N/A	—
8. Surface Water–Groundwater Interaction		
8.1 Is the conceptualisation of surface water–groundwater interaction in accordance with the model objectives?	No	The conceptualisation of surface water–groundwater interaction in inadequate despite groundwater abstraction impacts on surface water features being a major water management consideration for Horizons and the community.
8.2 Is the implementation of surface water–groundwater interaction appropriate?	No	—
8.3 Is the groundwater model coupled with a surface water model?		
8.3.1 Is the adopted approach appropriate?	N/A	There is no definite need to couple the groundwater model with a surface water model. However, the results of hydrological investigations and modelling should inform the groundwater model setup and calibration.
8.3.2 Have appropriate time steps and stress-periods been adopted?	N/A	—
8.3.3 Are the interface fluxes consistent between the groundwater and surface water models?	N/A	The results of hydrological investigations and modelling should inform the groundwater model setup and calibration.

Stantec New Zealand
Hazeldean Business Park, Level 2,
2 Hazeldean Road, Addington 8024
PO Box 13-052, Armagh, Christchurch 8141
Tel +64 3 366 7449

horizons
REGIONAL COUNCIL

horizons.govt.nz

24 hour freephone 0508 800 800
fax 06 952 2929 | **email** help@horizons.govt.nz
Private Bag 11025, Manawatu Mail Centre, Palmerston North 4442