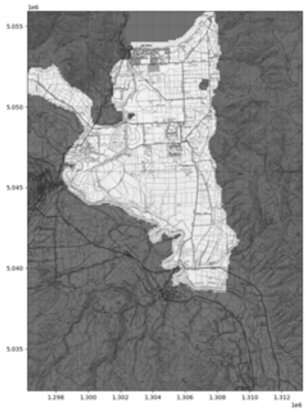
Hawea Transient groundwater model (Hawea Model)



Author:	Matt Dumont
Date:	2021-11-02
Version:	1.0.0
Status:	Final
KSL project:	Z22031HAW_hawea-model
Purpose:	This document describes the Hawea Model repo

The Hawea model domain; the inactive portions of the model are coloured dark grey. The model domain is a 3D model of the Hawea aquifer systems including the Maungawera Valley. The model domain is bounded by Lake Hawea to the North, the Clutha River to the South, and the hillslopes to the East and West. The model domain is 17 km by 23.5 km. The model cell spacing is 100 m and the model is on a regular North-South grid. The model is loosely based on the 2D model of the Hawea aquifer system developed by Wilson et al, (2011).

Index

Table of Contents

Modelling methodology and results

Rather than a traditional model report this repository serves as the detailed documentation of the modelling process. The final report with the interpretation of the modelling process and results is available in the repo here. The modelling process was broadly undertaken in the following steps; each step has its own readme document detailing its methodology and, where applicable, the results of the step:

1. Model build: build the model structure and boundary conditions
2. Model targets: define the model targets and objective function
3. Model Parameterisation: define the initial model parameters and parameterisation
4. Model Optimisation and limitations: optimise the model to the available data
5. Model Scenarios: run a series of scenarios to better understand the model behaviour and to predict the systems response to changing conditions

Subsequent Investigations

1. Historical Hawea groundwater investigation: investigate the historical data from 1976 - 1979 when the lake fell to the lowest level in the historical record (c. 327.5 m msl)
2. Quartz Creek LSR: investigate the LSR for the Quartz Creek area which was not included in the final model

Modelling Software

Most of the model was produced using open source Python packages and the MODFLOW suite. Specifically the model was built using MODFLOW NWT, optimised using PEST, and scenarios were run in MT3DMS-usgs.

Python Environment

This model was developed in Python on linux (ubuntu 20.04). The Python environment was created using the Anaconda package manager. The environment was created using the following command:

```
conda create -c conda-forge --name hawea python numpy pandas pytables openpyxl matplotlib scipy netcdf4 psutil geopandas flopy pysheds scikit-learn py7zr
conda activate hawea
pip install pyemu
pip install ppscore
pip install tabulate
pip install fpdf
pip install pdfkit
```

Versioned Conda Install

If you wish to use the exact package versions you may install with:

```
conda create -c conda-forge --name hawea python=3.10.5 numpy=1.23.4 pandas=1.4.3 pytables=3.7.0 openpyxl=3.0.9 matplotlib=3.5.2 scipy=1.9.0 netcdf4=1.6.0 psutil=5.9.1 geopandas=0.11.1 flopy=3.3.5 pysheds=0.3.3 scikit-learn=1.1.1 py7zr=0.20.0

conda activate hawea
pip install pyemu==1.2.0
pip install ppscore==1.3.0
pip install tabulate==0.9.0
pip install fpdf==1.7.2
pip install pdfkit==1.0.0
```

This install was tested successfully on windows 10 on 13/07/2023

In addition to the creation code above, the repo environment was exported in:

- environment.yml
- environment.txt

However these exports are raw and therefore may be difficult to directly install and may contain proprietary packages (e.g. kslcore) We have left them as they provide an exact copy of the development environment if future users have versioning problems with the above conda installs

Github repo structure

The full modelling process for the Hawea model was undertaken within this Github repo. The only exceptions are several large datasets (LIDAR/DEMs) which were simplified (code in repo) and then the simplified product was saved in the Github repo. This means that no external datasets are necessary to completely recreate the Hawea model and the full methodology is present in this repo.

Comment keyword standards:

We have used a number of keywords (case insensitive) to support identifying important comments within the text. These are:

- TODO: A comment that identifies a task that needs to be completed
- FIXME: A comment that identifies a problem that needs to be fixed
- KEYNOTE: A comment that identifies a key assumption or point of interest
- OPEN SOURCE IMPROVE: A comment that identifies a potential improvement to existing open source code repos

At this point only KEYNOTE and OPEN SOURCE IMPROVE should remain in the repo, however it is possible that some TODOs and FIXMEs will remain accidentally. Note that these have been dealt with, but were accidentally not removed from the code. Many IDEs have a search function that can be used to find these keywords, which we encourage you to use.

Repo index

Below is a rough guide to the repo structure. Not every file in the repo is described. Often the best way to find out what information a file contains is to look through the appropriate Python function and read the docstrings. The repo has been documented to a reasonable extent, but there is still some work that could be done to make the repo more user friendly. If you have any questions please contact Matt Dumont (matt@komanawa.com)

- README.rst: This document
- project_base.py: A script to set up the project environment and manage paths
- scott_model: A copy of the original 2D model of the Hawea aquifer system developed by Wilson et al, (2011)
- **model_build: model build process and datasets**

- README.rst: a readme file for the model build
- base_data: raw input data for the model build
- processed_input_data: processed data for the model build that was built by the scripts in this folder from the raw data in the base_data folder
- project_model_tools.py: a script to define the model tools instance, and the model structure
- get_boundary_condition_data.py: a script to get the boundary condition data
- **supporting_data_analysis: scripts to support creating the boundary condition data and structure**
    - all_wells.py: a script to get all the well location data
    - base_concept_diagram.py: a script to build a base concept diagram of the 3D model structure
    - compare_met_era5land.py: compare precipitation and PET between the available met station and the ERA5-land data
    - explore_structure.py:
    - get_era_5_land.py: script to get ERA5-land data
    - get_pumping_data.py: get and process historical pumping data
    - hillside_inflows.py: model and process estimates from the hillside inflows
    - irrigation_race_losses.py: get and process the historical race loss data
    - lake_data.py: get and process the historical lake data
    - map_flowmeter_to_wells.py: a process to map the flowmeter data to the most likely well
    - plot_borelogs.py: a process to plot the borelogs in the model
    - recharge_model.py: develop and create LSR estimates from met and ERA5-land data
    - river_data.py: : a process to get and process the river data
- modflow_model.py: a script to build a MODFLOW model instance
- utils.py: a script to define some utility functions
- zones.py: a script to define indicative model zones

- **model_parameterisation: model parameterisation and implementation**
    - README.rst: a readme file for the model parameterisation
    - base_data: raw input data for the model parameterisation
    - processed_data: processed data for the model parameterisation that was built by the scripts in this folder from the raw data in the base_data folder
    - static_params.py: a script to define the static model parameters
    - pilot_points.py: a script to create, define, and interpolate pilot points for kh and sy
    - inital_parametersiation.py: a script to define the initial model parameters (before optimisation)
    - plot_parameter_names.py: a script to plot the parameter names generates parameter_map.png
    - optimised_parameterisation.py: a script to easily access optimised parameter sets
    - **optimised_parameter_sets: optimised parameter sets**
        - 3d_v1a_opt.par: optimised parameter set for the 3D model version 1a

        - 3d_v1b_opt.par: optimised parameter set for the 3D model version 1b
        - 3d_v1d_opt.par: optimised parameter set for the 3D model version 1d
    - parameter_map.png: a map of the model parameters
- **targets_and_sensitive_sites: target development and data**
    - README.rst : readme document detailing the methods and data used to develop the targets
    - model_output.py: script to extract consistent model outputs and plots
    - get_raw_target_data.py: ingest raw target data
    - get_indicative_times.py: get indicative times for the targets that fall outside of the optimisation period
    - head_targets.py: definition of the head targets
    - riv_gain_loss_targets.py: definition of the river gain and loss targets
    - senstive_sites.py: identification of sensitive sites
    - target_structure_checks.py: checks to ensure that the targets and the model structure were not mutually exclusive
    - base_data: base input data for the targets
    - processed_data: processed target data, this was developed from the raw data in the base_data folder
- **optimisation: optimisation code and results**
    - README.rst: readme document detailing the optimisation process and methodology
    - **PEST optimisation build, run, and post processing scripts**
        - build_optimisation.py: script and functions to build the PEST files
        - a_build_run_optimisation_version.py: build and run a PEST optimisation
        - run_opt_step_models.py: run the step models from a PEST optimisation
        - manual_optimisations: manual optimisations that were run, in the end these never contributed more than some information to the modeller
        - model_utils_for_forward_run.py: functions to build and run a model from PEST parameter files
        - compare_parameterisations.py: script to compare parameters across multiple parameter files
        - hawea_plot_optimisation.py: script to plot the optimisation results
        - plot_multiple_high_freq.py: script to plot multiple high frequency observations for given PEST obs files
    - **Manage optimisation period:**
        - determine_opt_start.py: script to determine the start and end of the optimisation period
        - optimisation_period.py: script to manage and hold the information about the optimisation period
    - **Optimisation Results**
        - **optimisation_results: results for the optimisation holding all of the pest input and output files**

- [3d_v1a](): optimisation results for the 3D model version 1a
- 3d_v1b: optimisation results for the 3D model version 1b
- 3d_v1d: optimisation results for the 3D model version 1d (final model)

- **final_opt_models: the final optimised model files**
  - 3d_v1a: final optimised model files for the 3D model version 1a
  - 3d_v1b: final optimised model files for the 3D model version 1b
  - 3d_v1d: final optimised model files for the 3D model version 1d (final model)
  - compress_uncompress_model.py: utilities to compress and uncompress the model files so they could be included in the Git repo (50mb limit)
- **Computational support files**
  - compile_pest: compile PEST for linux
  - pest_run_data: static data needed by PEST to run the model
  - git_setup.sh: script to setup the Git repo for the optimisation on a machine
- **Model overview**
  - pre_optimisation_overview.py: make pre optimisation overview plots
  - make_preopt_slideshow.py: make a pre optimisation slideshow
  - pre_optimisation_plots_png: pre optimisation plots of boundary conditions, targets, parameterisation, and other supporting work, many of these figures are referenced in the various readme.rst files
  - make_opt_presentation.py: make a presentation of the optimisation results for a meeting

- **quartz_creek_lsr: modelling of LSR for the Quartz Creek area see the Scenarios readme for more information**
  - results: results from the LSR modelling
  - model_qtz_ck_lsr.py: script for LSR modelling
- **Scenarios: scenario modelling code and results**
  - README.rst: document describing the scenario modelling methods and results
  - **Scenario development and supporting scripts**
    - scen_period.py: script to handle the scenario period
    - boundary_condition_plots: plots of the scenarios boundary conditions
    - base_data: base input data for the scenarios
    - processed_input_data: processed input data for the scenarios, these files were all developed from the base data
    - boundary_conditions.py: develop the input boundary conditions for the scenarios
    - supporting_data_analysis: additional data analysis scripts to support creating boundary conditions
    - scenario_outputs.py: script to make consistent scenario outputs
    - run_flow_scenario.py: script to run a flow scenario
    - run_scenario.py: script to run a scenario (in multiprocessing)

- **Model information and MT3D indicator modelling**
    - run_mt3d_scenario.py: script to support running MT3D
    - mt3d_indicator_scens.py: script to run MT3D indicator scenarios
    - compare_boundary_sensitivity.py: compare the results of the boundary condition sensitivity analysis
    - model_info_scenarios.py: script to run model information scenarios
    - **model_info_scen_results: model results and plots for model information scenarios**
        - 0_results: plots for model information scenarios
        - {scenario name}: Model results for model information scenarios: input and output data for the scenario
    - mt3d_indicator_scenarios: model results and plots for the MT3D scenarios
- **Low Lake Hawea level scenarios**
    - low_lake_scenario_data.py: script to develop typological lake levels and perturbations
    - low_lake_scenarios.py: script to run low lake scenarios
    - compare_low_lake.py: script to compare low lake scenarios
    - **low_lake_scenarios: model results and plots for low lake scenarios**
        - 0_results: plots for low lake scenarios
        - {scenario name}: Model results for low lake scenarios: input and output data for the lake scenario
- **Allocation modelling**
    - allocation_zones.py: get and plot allocation zones
    - allo_rch_hillside.py: scripts to get and compare the allocation, hillside recharge, and LSR for each zone
    - allocation_scenarios.py: script to develop all allocation scenarios and to run the non-gridded allocation scenarios
    - run_grid_allocation.py: script to run the gridded allocation scenarios
    - compare_allocation_scens.py: script to compare allocation scenarios
    - allocation_scenarios: model results for allocation scenarios
    - **allocation_results: plots of allocation results**
        - old_allo_zones.png: figure of the old allocation zones (Wilson et al., 2012)
        - new_allo_zones.png: figure of the new allocation zones
        - Hawea Flat_results: results for the gridded Hawea Flat allocation scenarios
        - Maungawera Flat_results: results for the gridded Maungawera Flat allocation scenarios
        - Terrace-Hill_results: results for the gridded Terrace-Hill allocation scenarios
        - nat_current_full: results for the naturalised, current allocation, and full allocation scenarios

      - Te Awa_results: results for the gridded Te Awa allocation scenarios
      - Terrace-River_results: results for the gridded Terrace-River allocation scenarios
      - mangawera_valley: results for the Maungawera Valley allocation reduction scenarios
      - allo_zone_rch: results comparing LSR, hillside inflows, and allocation for each zone
      - example_quantile_plots: example quantile plots for the allocation scenarios to support presentations
  - **Wetland Setback Modelling**
    - wetland_setback_campbells: wetland setback modelling for Campbells wetland scripts and results
    - wetland_setback_butterfield: wetland setback modelling for Butterfield wetland scripts and results
- support_figures: supporting figures for this and other README.rst documents
- dummy_packages: dummy packages for the proprietary packages used in the model, these packages have some, but not all of functionality of the original packages
- **historical_investigation: historical Hawea groundwater investigation of historical data from 1976 - 1979 where the lake fell to the lowest level in the historical record (c. 327.5 m msl)**
  - README.rst: readme document detailing the historical investigation methods, results, and conclusions.
  - **base_data: The raw input data for the historical analysis. For more info see the "Dataset and Resources" section at the end of this document.**
    - MWD_Hawea_Flats_Groundwater_1984.pdf: The original PDF of the Ministry of Works and Development report, which contains the historical data.
  - current_model_prediction.py: run the naturalised 3d_v1d model for the historical period with low lake levels
  - figures: output figures from the historical analysis. For more info see the "Dataset and Resources" section at the end of this document.
  - generated_data: data generated by the analysis. For more info see the "Dataset and Resources" section at the end of this document.
  - get_historical_data.py: read in and access the historical data
  - lake_drop_scenarios.py: run and compare the lake drop scenarios to the historical data
  - mt3d_indicator_scenarios: results for the MT3d component analysis
  - mt3d_indicator_scens.py: run the MT3d component analysis on a steady state model with low lake levels
  - plot_historical_data.py: plot the historical data
  - plot_historical_natualised_model.py: plot the historical data and the naturalised model results
  - shift_diff.py: compare lake and historical data, and calculate the shift between the two
  - simple_smoothing_model.py: develop and apply a simple smoothing model to the historical data

Supporting data index

This repository contains all of the input and processed data needed to build and run the model. There are two exceptions to this; the 1 m LIDAR dem for the Clutha and Hawea rivers and the 15 m DEM used for the model top. Both DEMs are too large to store in the repo, so they have been simplified and the simplified versions are stored in the repo. There are a number of directories in the repo that contain the input and processed data each of these directories contains a readme.rst file that briefly describes the data in the directory. The directories are:

- model_build/base_data: contains the base data used to build the model
- model_build/processed_input_data: contains the processed data used to build the model
- model_parameterisation/base_data: contains the base data used to parameterise the model
- model_parameterisation/processed_data: contains the processed data used to parameterise the model
- Scenarios/wetland_setback_butterfield/base_input_data: contains the base data used to run the butterfield wetland setback scenario
- Scenarios/wetland_setback_butterfield/processed_input_data: contains the processed data used to run the butterfield wetland setback scenario
- Scenarios/base_data: contains the base data used to run the scenarios
- Scenarios/processed_input_data: contains the processed data used to run the scenarios
- Scenarios/wetland_setback_campbells/base_input_data: contains the base data used to run the campbells wetland setback scenario
- Scenarios/wetland_setback_campbells/processed_input_data: contains the processed data used to run the campbells wetland setback scenario
- targets_and_sensitive_sites/base_data: contains the base data used to define the model targets and objective function
- targets_and_sensitive_sites/processed_data: contains the processed data used to define the model targets and objective function
- historical_investigation datasets: contains the datasets used in the historical investigation

Proprietary packages

For the most part we relied on open source packages to create the Hawea model, but we did use some proprietary in house packages. These packages are not included in this repository, but we have included dummy packages that contain the same structure as the original packages and replicates some of the functionality. These dummy packages are located in the dummy_packages folder in the model repo and the python scripts have all been adjusted to load the dummy package version if the original version is not available.

Additionally, to ensure future use of this model we have included outputs of the data which necessitated the use of the proprietary packages. These outputs are located in the *processed_input_data* folders. The functions that use these packages to develop the outputs tend to follow are "recalc" structure, that is:

```
def get_data(*args, **kwargs, recalc=False):
    save_path = processed_data_dir.joinpath('data.csv') # path in the processed data folder where the outputs are saved
    if save_path.exists() and not recalc:
        # read the data from the saved path and return it
        # sometimes additional processing (e.g. other args) is done after loading the data
        return pd.read_csv(save_path)
    else:
        # the process by which the data was generated
        outdata = None
        # save the data to the save_path
        outdata.to_csv(save_path)
        return outdata
```

This structure allows the user to run the model without the proprietary packages, but also allows the user to see the full methodology used to generate the outputs. This also keeps the links between the data generation and the data use (e.g. in a model) explicit. This prevents the 'black box' problem that can occur when the data is generated by a different process and then ingested into the model.

Excluding a full model re-build, these proprietary packages should not be needed; however, if the user wishes to run the model with the proprietary packages or to generate a next generation with the proprietary packages they are encouraged to contact the author of this model: matt@komanawa.com

The proprietary packages used in this model are:

- **Dummy packages provided:**
    - **from model_tools.time_discretization import TimeDis**
        - mange the human time to model time
    - **from model_tools.regular_modeltools import ModelTools_RegularGrid**
        - manage the model structure and real world coordinates to model coordinates
- **No Dummy packages provided**
    - **kslcore**
        - an internal package used to ensure consistant access to our computational resources (google drive, NAS, etc.) across multiple machines
    - **from rushton_model.rushton import Rushton**
        - land surface recharge model
    - **from run_managers.beopest_manager import BeopestManager**
        - Manage Beopest across multiple linux machines
    - **from run_managers.ssh_distributor import SshDist**
        - Distribute a list of model runs across linux machines
    - **from model_tools.util_functions.list_file_utils import ListSolverInfo**
        - extract solver information from the list file
    - **from model_tools.plot_borelogs import plot_borelogs, plot_single_log, make_single_log_handles**
        - plot bore logs
    - **from model_tools.model_plotting import plot_spd, first, last, FakePath**
        - plot model results
    - **from model_tools.plot_optimisation import plot_optimisation_and_extract_info**
        - plot optimisation results

Dead links

We have made a substantial effort to ensure that all links in the model are valid. However, there are likely some links that return a 404 error. If you come across this, then please contact the author of this model: Matt@komanawa.com so that he can fix the links. Typically the links are relative to the repository. if the link is broken you can likely infer the correct location by looking at the link and the repo structure.

Branches and releases

The process of the model optimisation required multiple structural changes to the model as well as changes to the objective function to attain a satisfactory history match. These different structures and changes were all set up as unique branches within the repo. For more information on branches see github's explanation of branches. At the end of the calibration process there were 24 unique branches, most of which were abandoned. These branches were issued as pre-production releases (More information about releases). Only the key structures were retained and the "final" model was merged back to the main branch.

Active Branches

Main (3d_v1d)

- The 'final' optimised model.
- Contains 3D structure around the Lake Hawea Moraine
- Best fits for the high frequency targets.
- Bund elevation set to 335 msl
- NGMP well head observations removed from objective function as there is significant tension between these records and the high frequency observations. The NGMP wells are pumped irrigation bores and the primary purpose for sampling was water quality monitoring.

3d_v1a

- Identical to "Main (3d_v1d)" except that the NGMP wells were included in the objective function
- decent history matching; however "Main (3d_v1d)" provides better results
- retained as active branch for comparison to "3d_v1b"

3d_v1b

- Identical to "3d_v1a" except that the bund elevation was set to 333 MSL.
- history matching results were similar to "3d_v1a" suggesting that the bund elevation is largely non-unique
- retained to demonstrate the non-uniqueness of the 3D structure

terrace_only

- This model structure only includes the High Terrace (south of Hawea Flat) to the Clutha river
- this optimisation was undertaken to see if the High Terrace could be history matched (within the accepted parameter ranges) in isolation from the rest of the Hawea aquifer system.
- History matching was not achieved.

3d_v10a

- Identical to "Main (3d_v1d)" including parameterisation, but the bund elevation was set to 330 m MSL to investigate the predictions of the historical investigation.

previous branches (releases)

There are many previous branches that were issued as pre releases and then deleted (effectively archived). There should be no reason for other users to delve into these previous branches as they ended up with unsatisfactory history matching; however, they are available and briefly described below (working notes) for completeness.

1. Main (before 2/11/22) The main build branch. First structural version
2. Structure v2, Changes:
    - Increase parameterisation via pilot points to Maungawera
    - Add recharge multiplier pilot points across model (NI)
    - Remove sandy point from model
    - abandoned but retained
3. Structure v3,Changes:
    - Set ss=sy
    - Set the model to confined to reduce computational burden
    - This helped but the model preformed poorly,
    - Error did not reduce saturated thickness.
    - abandoned and deleted
4. Structure_v4:
    - From structure v2
    - Add new mean annual head targets from regular
    - Increase steps to 7 in transient
    - Expand hillside streams to all adjacent cells (up to 9 cells per hill)
    - Optimisation never run here, just saved to version structural changes
5. Structure_v5
    - From structure_v4
    - Remove near river pumping wells.
6. Structure_v6
    - From structure 5
    - Add a 1m confined layer below the bottom of layer 1 (may improve stability)
7. Structure_v6a
    - From v6, but set ss to sy
8. Structure_v7 (built but not run)
    - From structure 5
    - Reduce thickness to reasonable pumped thickness and then Maximum 30m sat thickness
    - Set ss = sy
    - run as a confined model
9. Structure_v8
    - From structure_v6a
    - increase initial conductivity (to 50, 100 and 70 was too unstable)

- rch multiplier only by irrigated not irrigated bounds of multiplier 0.5-1.2

10. Structure_v9
    - Fix river targets (they were backwards!)
    - Implement grandview and john creek (+Hawea and Clutha) as str package
    - Lake stage vs g40_0415
    - Looks fine, honestly the fact that them model isn't matching it suggests some sort of structural error. Reworked transport in grandview stream?/ water through grandview stream??? Likely the problem google maps shows water in grandview to the lake (and in john creek (to the north), all other creeks are probably fine.
    - Lower basement around g40_0366
11. Structure_v10
    - Set weight of regular year targets to 0
    - set each of the 'h_hf' targets equal weights despite different data lengths
    - look/lower basement in dry cells near model boundaries
    - NE hillside area (done)
    - Near clutha river (done)
    - I think I need some more pilot points
    - Near pt 402 on camp hill moraine (move Maungawera south?) () and another in the moraine (to interpolate with other river group
    - To stop dry cells south of camp hill moraine
    - Significant number in the hillslope area just off the bounds to allow conductivity to fall there if needed for stability. And to manage the change in geologic setting near hillslope
    - Adjust some locations based on the new pilot point locations
    - New rivergroup south of Maungawera valley entrance to allow for the difference between the two settings
    - Additional point in the middle of the terrace to manage near hillside environment.
    - Try lowering hillside conductance → set to 100 vs 1000 for Hawea/Clutha, which means much of the peak flow does does not make it into the model.
12. Structure_v11
    - Move to 1 global recharge modifier (done)
    - Much higher initial kh (lake=5, rest = 300) (in progress
    - Lower sy, and lower sy bounds
    - Change weights (lower low frequency targets)
    - Bit of a hail mary before the weekend
    - retired (even though I'm happy with the parameterisation. If I want to change back to vll parameters do it from v12
13. Structure_v12
    - Increase kh/sy parameterisation in the near lake environment
14. p_lake
    - As per structure_v11 but with a single additive parameter for lake heads (e.g. lake hds = lake hds + mod

- A test to see if the lake levels problems are sorted everything else works great?
- Note the parameter is offset by 100m as pyemu has bugs!

15. lake_bar
    - Add a 1 cell thick barrier for kh
    - Remove additional v12 parameterisation
16. cond_int
    - Try to fit the heads by simply setting lake conductance (1 cell width lake)
17. 3d_v1
    - Address the 3D moraine issues in structure
    - 3 layers the bottom two pinch out against the bottom of the model.
    - well management
    - target management
    - other structural pieces
    - Add abrupt parameter change at terrace interface
    - Remove from dam to “dam control” road from model (e.g. no flow)
    - Re-run pre_optimisation_overview.py
    - remove the slope fixer on the east side
    - remove additional parameterization of v12
18. 3d_v2
    - As per v1 but fully confined (to increase stability)
    - Ss[0] = sy[0]
    - Initial parameters do not manage the drop quite so well. This may really need the unconfined aspects of the model.
    - Bit of a hail mary over xmas. Really need the unconfined action to make the ‘waterfall happen’
19. 3d_v1c
    - As 3d_v1a but with top of bund set to 337
    - great difficulty getting this to converge
    - abandoned
20. 3d_v4
    - As 3d_v1a, but top of bund is set to 340m MSL instead of 335
    - Difficult to get model to converge
    - abandoned
21. 3d_v5
    - As 3d_v1a, but top of bund is parameterised
    - Largely unstable

References

- Wilson, J., 2012. Hawea Basin Groundwater Review. prepared by Resouce Science Unit of Otago Regional Council, June 2012, Dunedin.

Hawea Transient groundwater model (Hawea Model) build methods and results

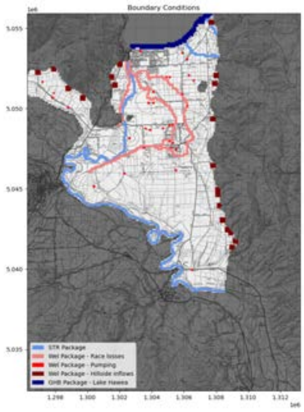


Figure: Overview of model boundary conditions

Author:	Matt Dumont
Date:	2021-11-02
Version:	1.0.0
Status:	Final
KSL project:	Z22031HAW_hawea-model
Purpose:	This document provides the methodology and results for the model build process

Index

Table of Contents

Module Index

- README.rst: this document
- base_data: raw input data for the model build

- processed_input_data: processed data for the model build that was built by the scripts in this folder from the raw data in the base_data folder
- project_model_tools.py: a script to define the model tools instance, and the model structure
- get_boundary_condition_data.py: a script to get the boundary condition data
- **supporting_data_analysis: scripts to support creating the boundary condition data and structure**
    - all_wells.py: a script to get all the well location data
    - base_concept_diagram.py: a script to build a base concept diagram of the 3D model structure
    - compare_met_era5land.py: compare precipitation and PET between the available met station and the ERA5-land data
    - explore_structure.py:
    - get_era_5_land.py: script to get ERA5-land data
    - get_pumping_data.py: get and process historical pumping data
    - hillside_inflows.py: model and process estimates from the hillside inflows
    - irrigation_race_losses.py: get and process the historical race loss data
    - lake_data.py: get and process the historical lake data
    - map_flowmeter_to_wells.py: a process to map the flowmeter data to the most likely well
    - plot_borelogs.py: a process to plot the borelogs in the model
    - recharge_model.py: develop and create LSR estimates from met and ERA5-land data
    - river_data.py: : a process to get and process the river data
- modflow_model.py: a script to build a MODFLOW model instance
- utils.py: a script to define some utility functions
- zones.py: a script to define indicative model zones

Model boundaries

The model domain (see figure below) was initially defined to include the following aquifers:

- **The main Hawea flat aquifer** stretching from Lake Hawea in the North to the base of the High Terrace in the South. This aquifer is bounded by the Hawea River on the West and the Grandview Ridge on the East
- **The High Terrace aquifer** stretching from the base of the High Terrace in the North to Clutha River in the South. This aquifer is also bounded by the Hawea River on the West and the Grandview Ridge on the East
- **Aquifers near the Hawea river** including Te Awa, Maungawera Flat, and river adjacent aquifers to the south of Maungawera flat and East of the High terrace.
- **The Maungawera Valley aquifer** including the Maungawera Valley aquifer from the approximate Hawea River/ Lake Wanaka flow divide in the Northwest to the Maungawera Flat aquifer
- **The Sandy Point Aquifer** which is to the East of the Clutha River to the South of the High Terrace aquifer. This aquifer is also bounded by the Grandview Ridge on the East

During the model build the steep topography of the Sandy Point aquifer caused model convergence issues. The Sandy Point has minimal data available (only one historical groundwater measurement). Therefore we resolved the convergence issue by removing the Sandy Point aquifer from the model domain. We still produced estimates of land surface recharge (LSR) and Hillside inflows to this aquifer, which were used to inform groundwater allocation decisions.

The boundaries of the model domain were all defined by no-flow boundary conditions. In addition, at the Lake Hawea Dam, Camp Hill, and Cameron Hill bedrock is exposed. Therefore these outcrops were also defined as no-flow boundaries. Finally, the Camp Hill Medial Moraine, located between Te Awa and the Maungawera Flat aquifers, is comprised of poorly sorted and unworked moraine sediments. While there are a few domestic supply bores in this area, the groundwater system is likely minimal, particularly in comparison with the other outwash dominated aquifers. We therefore chose to define this area as a no-flow boundary.

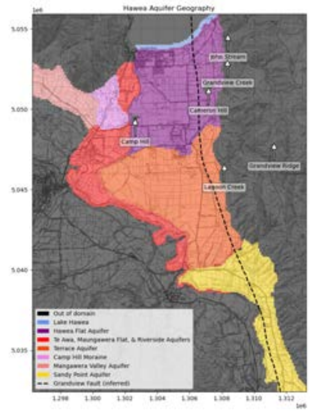


Figure: Map of the model domain with key features labelled

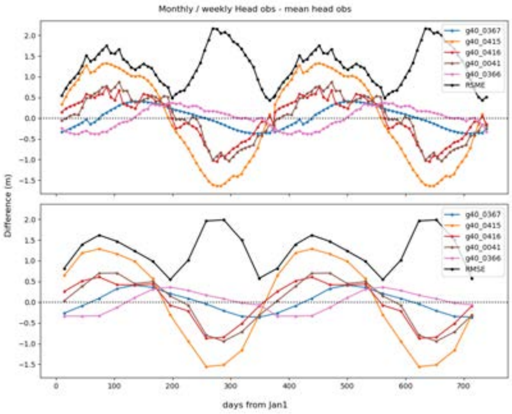
Model Time period

This model is a transient groundwater model with the first period defined as a steady state period. For the purposes of boundary conditions we defined two time periods for the model:

- **Optimisation period: 2015-07-18 to 2020-06-27:**
    - the period where we have the most data available across boundary conditions, observations (targets).
    - we defined the start year of the optimisation period to 2015 as this was the first year that we had reliable groundwater abstraction data were available.
    - weekly MODFLOW stress periods were used.
- **Scenario Period: 1980-07-18 to 2020-12-01**
    - the period where we have reasonable data available across boundary conditions, but minimal observations (targets)
    - weekly MODFLOW stress periods were used.

Model Starting heads

The challenge of defining starting heads for a transient groundwater model is that any choice of starting heads can then impact the subsequent model results. For this model we chose to use the top of cell as the starting heads and then run a steady state model period. To identify the best month to transition between the steady state model and the transient model we calculate the difference between the monthly and weekly high frequency data (see target readme for more details). and the mean of the full dataset. The results are shown in the figure below, but there were obvious local minimums in the RMSE at 200 and 365 days from January 1 or approximately start of January and mid July. We chose a start date of in mid July (2015-07-18) as this will minimize the variability in heads associated with the irrigation season.



Model Structure

The model structure was initially created as a 1 layer model, but during the course of the optimisation it became clear that the model could not reproduce the data without additional structure and layering. For more information on the optimisation process see the optimisation readme.

1 layer model structure

The 1 layer model was largely based on Wilson et al. (2012). The model top was defined based on a 15 m DEM (from NZWaM - Hydro), and the model bottom was initially set from the model bottom used in Wilson et al. (2012). The model bottom and top were then adjusted as follows:

- All cells with stream package cells with the stream rbot parameter below the model bottom were set as 0.5 m below rbot.
- A number of cells which routinely caused dry cells (and instability in the model) had the bottom gradient reduced
- The model top was adjusted so that the tops were always at least 0.5 m above the rbot of the stream package cells.
- There were a number of cells near the Clutha River that caused dry cells due to the incised nature of the river. For these cells the bottom was set to the bottom of the nearby river cells
- The model bottom was adjusted to ensure that the model thickness was at least 2 m

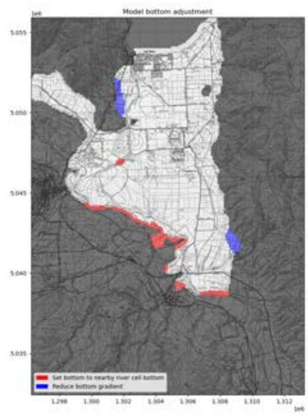


Figure: location where the gradient of the bottom or the absolute bottom elevations were reduced

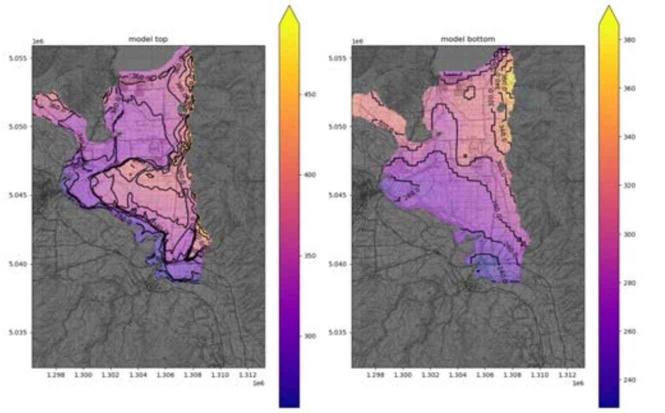


Figure: Model top and bottom

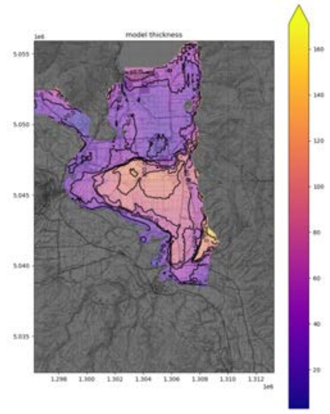


Figure: Model thickness

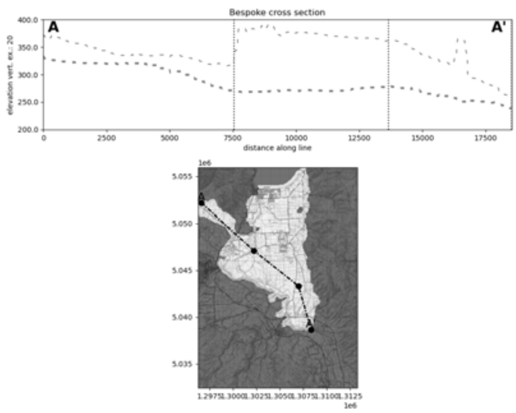


Figure: Example model cross-section 1

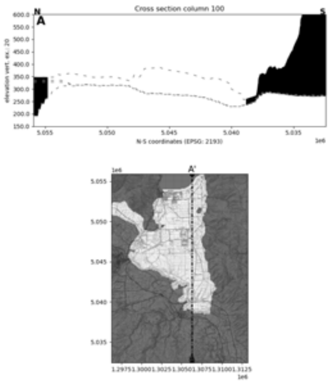


Figure: Example model cross-section 2

multi-layer (3d) model structure

The multi-layer model structure that was created better represents the complex geology in and around the Southern edge of Lake Hawea. There is likely to be other areas of the model domain that have more complex geology; however, excluding the structure at the Lake Hawea moraine precluded our model from fitting the observed data. For more information on the optimisation process see the optimisation readme.

Lake Hawea Moraine Conceptual Model

In the 1 layer model structure the Lake Hawea moraine was represented as a single layer and its impact on the groundwater system was parameterised as a single parameter -- hydraulic conductivity. However, in reality the Lake Hawea moraine is a complex geological structure. From a groundwater perspective the key observations that precluded the 1 layer model from fitting the data were the high frequency measurements at well G40/0415 (roughly at the intersection of Cemetery Road and Gladstone Road). These observations showed that the groundwater levels in this well are highly correlated with the lake levels, but with approximately 10 m of vertical displacement. We developed and fitted a very simple numerical model to the groundwater levels at G40/0415 to better understand the relationship between the lake levels and the groundwater levels in this well. The model parameterised the groundwater levels as:

$$h_{gw}(t) = \sum_{n=t+l}^{t+l+s}h_{lake mod(n)} / s$$

$$h_{lake mod}(t) = ((h_{lake}(t) - h_{lakeMean}) * a) + h_{lakeMean} + \Delta_{h}$$

where:

- h_{gw} is the groundwater level,
- h_{lake} is the lake level,
- t is the time (day),
- l is the lag parameter (days),
- s is the number of days to smooth the lake levels,
- $h_{lakeMean}$ is the mean lake level,
- Δ_{h} is the vertical step parameter,
- a is the lake level amplitude modifier,

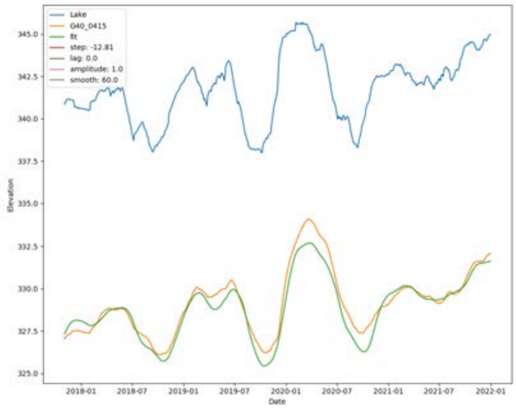


Figure: results of the simple fit to Lake Hawea levels

The simple numerical model provides a good fit to the groundwater levels at G40/0415 with a step change of 12.81 m and a 60 day smoothing period.

With the 1D model we were unable to fit the water levels at G40/0415. We could either fit the shape of the groundwater levels but there was substantial bias in the mean groundwater level (too high) or we could fit the mean groundwater level but the shape of the groundwater levels was lost.

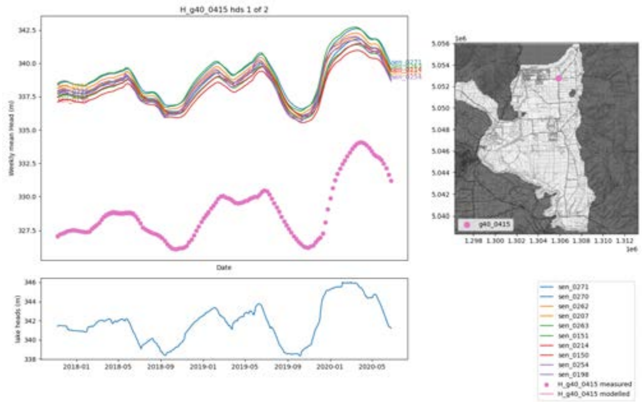


Figure: the results of the 1 layer model which fit the shape of the groundwater levels, but not the mean

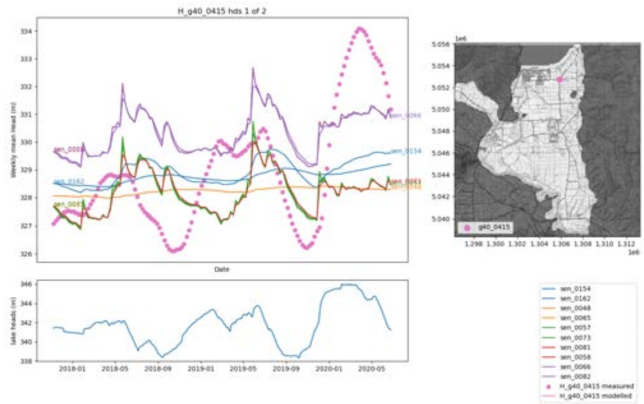


Figure: the results of the 1 layer model which fit the mean of the groundwater levels, but not the shape

We do not have other high frequency observations of groundwater levels near the Lake Hawea moraine. However, we do have a number of static water levels that were measured shortly after drilling the bore. These water levels, relative to the lake level at the time of measurement are shown in the figure below. This figure shows a constant vertical offset between the groundwater levels and the lake levels of approximately 10 m. Some of these boreholes are located in the moraine less than 200 m from the lake. Many of the water supply wells near the lake (within the mapped moraine) are relatively deep (e.g. 50+ m).

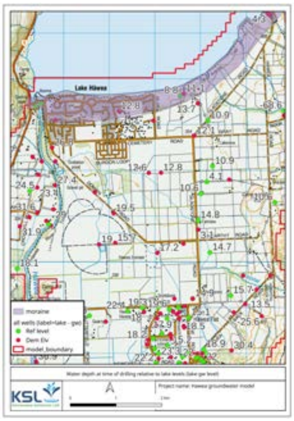


Figure: all groundwater levels relative to the Lake Hawea level on the sampling date (positive values are groundwater levels below the lake)

Several of the bores have recorded bore logs which are shown in the figure below. Drillers logs can be imprecise; however, the logs demonstrate lower conductivity sediments overlying more conductive sediments at ~ 320 m msl. This elevation is also typically where drilling stopped, suggesting they had finally found conductive material. Some logs also show more conductive sediments overlying the less conductive sediments, but this is spatially variable.

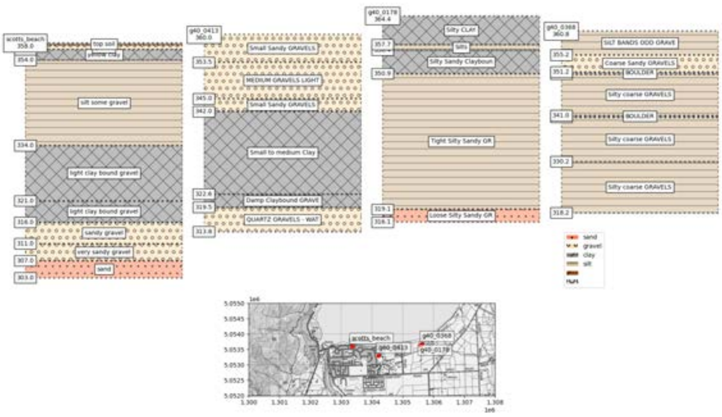


Figure: bore logs from the Lake Hawea Moraine

The Quaternary geological history of the Lake Hawea area is made up of a number of glacial advances and retreats described below and show in the figure below.

1. After the Q6 Luggate Advance relatively conductive glacial out-wash filled the basin and deposited the sediments of the high terrace
2. The Q4 Albert Town Advance scoured many of the previous sediments and deposited relatively impermeable, glacial moraine at the Camp Hill Moraine and the moraine at the northern base of the High Terrace.
3. The subsequent Q3 retreat meant relatively conductive outwash sediments were deposited between the High Terrace and Lake Hawea.
4. The Q2 Hawea Advance scoured the previous sediments (producing Lake Hawea) and deposited a relatively impermeable moraine at the southern edge of Lake Hawea.

Figure: Quaternary geological history of the Lake Hawea area from Wilson(2012)

Based on the borelog information, the geological history of the area, and the groundwater levels we propose a conceptual model of the Lake Hawea Moraine. The conceptual model is shown in the figures and described below:

1. During the Q2 Hawea Advance the glacier did not fully scour the previous outwash sediments and the new Lake Hawea moraine was deposited on a thin wedge of Q3+ outwash.
2. The moraine forms a relatively impermeable barrier to groundwater flow from Lake Hawea to the Hawea Flat aquifer system.
3. During and after the Q2 Lake Hawea Advance the moraine was locally reworked and eroded by some combination of the Hawea river, Grandview and John Creek periglacial activity (e.g. local reworking of the moraine by surface water on the top of the glacier). Regardless, this reworking produced a locally continuous cap of relatively conductive material on top of the moraine.
4. After the completion of the Lake Hawea dam in the 1950s the Lake Hawea level was raised by approximately 20 m, which allowed Lake Hawea water to seep through the permeable cap of the moraine and into the Hawea Flat aquifer system. This is anecdotally supported by reports of relatively low groundwater levels prior to the completion of the dam.
5. The permeable cap of the moraine is relatively thin and is almost certainly not continuous across the moraine. This forms a small, possibly perched, aquifer system on top of the previous moraine. This aquifer system then effectively spills over the moraine into the Hawea Flat aquifer system with very steep localised groundwater gradients.
6. The Hawea Flat aquifer system is a relatively conductive system and has relatively low groundwater gradients. Locally, where the Lake Hawea moraine overlies more conductive material, groundwater will flow from this conceptual pour point back towards Lake Hawea, providing the groundwater in the aforementioned groundwater bores (Scott's Beach, G40/0413, G40/0368, G40/0178)

The main significance of this conceptual model is that at some point the groundwater system could become disconnected from Lake Hawea. If this were to happen then groundwater levels could significantly decline in the Hawea Flat aquifer. More details on these scenarios are discussed in the scenarios readme file..

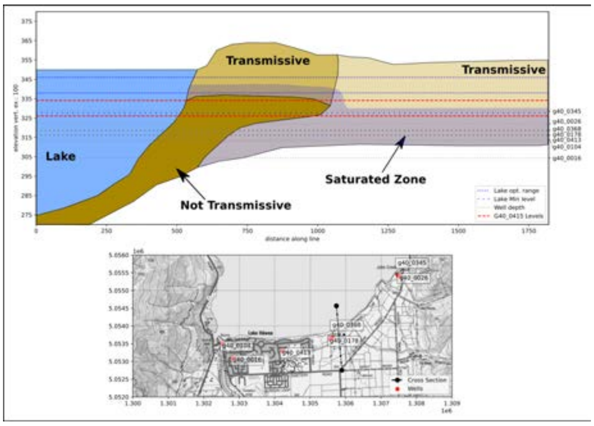


Figure: conceptual model of the Lake Hawea Moraine (across the moraine)

The location and elevation of the conceptual pour point is fundamentally unknown. If we assume a perched aquifer (rather than a local penetration, discussed below) then, based on groundwater levels, and the bore logs, we can estimate the bottom elevation of the impermeable moraine to be approximately 320 m msl. The top of the impermeable moraine is likely is more difficult to estimate. Based on the behaviour of the water levels in G40/0415, we can suggest it is likely below the typical minimum operating level of Lake Hawea (338 m msl). The observed water levels in G40/0415 do not become disconnected (e.g. variations in Lake levels which are not seen in the measured bore levels) and the rise and fall of the groundwater levels are relatively symmetrical. If the lake was disconnected we would expect to see a period of groundwater decline with a sudden increase in groundwater levels (as the lake became re-connected to the groundwater system). There is some anecdotal evidence that the groundwater levels in the Hawea Flat aquifer system have declined significantly when the lake levels reached their historical minimum (327.6 m msl in 1976 & 1977). We could interpret this as evidence of a lake disconnect, suggesting that the top of the moraine is likely above 327.6 m msl, but there are not sufficient records from this event to be certain.

It is also possible that the moraine is not continuous across the Lake Hawea foreshore, but is instead locally penetrated by a paleo-channel of either the Hawea River, John Creek, or Grandview Creek. If this is the case the penetration would likely be within the spatial extent of the Grandview Creek and John Creek alluvial fans. The response of the groundwater system to a local penetration of the moraine is likely to be dependent on the shape of the paleo-channel. The response would likely be non-linear as the transmissivity of the paleo channel would decline rapidly with decreasing water levels This scenario is fundamentally uncertain and very difficult to include in a groundwater model without additional information. Therefore we chose to assume a perched aquifer for the conceptual model used in the multi-layer model structural design.

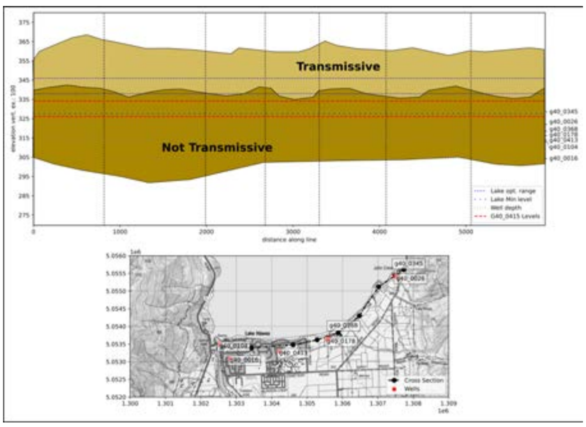


Figure: indicative conceptual model of the Lake Hawea Moraine (along the moraine) assuming a perched aquifer

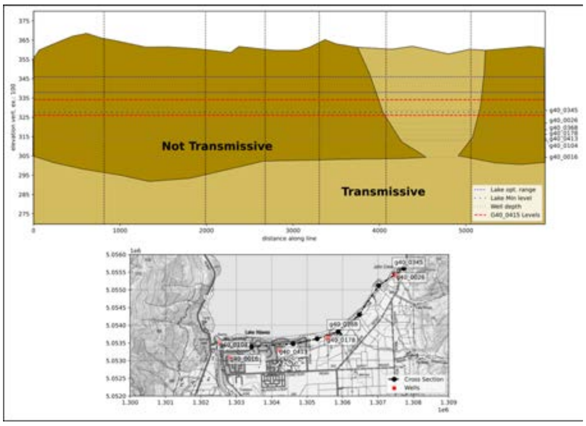


Figure: indicative conceptual model of the Lake Hawea Moraine (along the moraine) assuming local penetration of the moraine with a paleo-channel

Implementation of the Lake Hawea Moraine Conceptual Model into the groundwater model

We implemented a very simple version of the conceptual model described above, assuming a perched aquifer system. We:

1. implemented a 3 layer system. Layer number follows Python indexing (i.e. layer 0 is the top layer)
2. defined four new zones for the model (see figure below)
    1. The moraine zone: Layer 0 represents the permeable cap of the moraine, layer 1 represents the impermeable moraine, and layer 2 represents the conductive out-wash left by the Q4+ glaciations of the main Hawea Flat aquifer system.
    2. The Lake bar: This zone does not occur in layer 0, but in layer 1 and 2 it represents the impermeable material that separates Lake Hawea from the Hawea Flat aquifer system.
    3. The Lake: This zone occurs in layer 0, 1, 2; it is conductive material and contains GHB boundary conditions (see more below)
    4. The Layer pinch out area: This zone occurs in layer 0, 1, 2 and is used to pinch out the three layer system so that for the rest of the model domain, layer 0 is the main aquifer system (e.g. like the 1 layer model) layers 1 and 2 consist of 1 m thick layers
3. set the top of layer 2 (bottom of the moraine) to be 328 m msl.
4. set top of layer 1 (top of the moraine) to be 335 m msl. note in some of the branches of this repo we used different "bund_top" elevations. the layer 1 elevation is always specified in model_build.project_model_tools.bund_top object

Several cross sections are presented below; additional crossections and spatial figures about the multi layer model structure are available in the support figures folder.

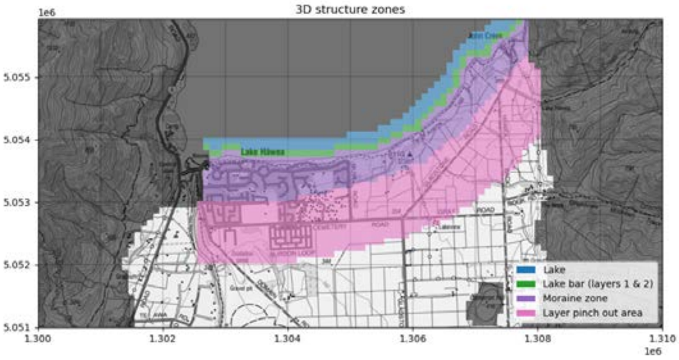


Figure: 3D spatial view of the multi-layer model structure zones

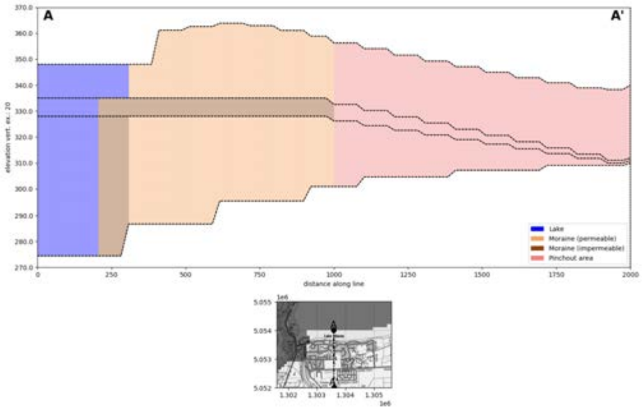


Figure: 3D cross section view 2 of the multi-layer model

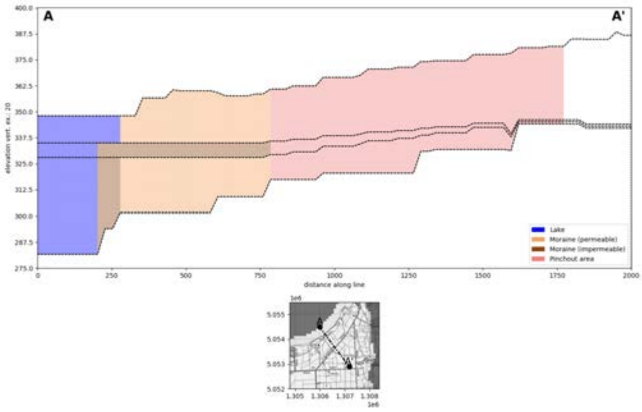


Figure: 3D cross section view 6 of the multi-layer model

Model boundary conditions

Land surface recharge (LSR)

LSR model

We chose to use the Rushton model to estimate LSR. The Rushton model is simple, easy to implement, and has been used in a number of other studies. In general the Rushton model uses the following methods to estimate soil moisture balance:

1. **Calculation of infiltration to the soil zone (In), and near surface soil storage for the end of the current day**

    (SOILSTOR). Note that Infiltration (In) as specified by the Rushton algorithms is not just infiltration (Rainfall-Runoff). It also includes SOILSTOR from the previous day.

2. **Estimation of Actual ET**

    The spreadsheet calculates TAW and RAW from field capacity, wilting point, and rooting depth data. Typical values for field capacity and wilting point are given in Table 19 of Allen et al. (1998). Rooting Depth changes with the season, and is typically 0.5-1m for grass (Table 22 of Allen et al.,1998) A depletion Factor, p, needs to be estimated for the calculation of RAW. p is the average fraction of TAW that can be depleted from the root zone before moisture stress (reduction in ET) For NZ conditions p should be around 0.4-0.6, typically 0.5 for grass. See Table 22 of Allen et al. (1998) for more values Fracstor (near surface soil retention) needs to be estimated. Typical values are 0 for a coarse sandy soil, 0.4 for a sandy loam, 0.75 for a clay loam (Rushton, 2006, pg 388)

3. **Calculation of Soil Moisture Deficit and recharge.**

    Note that the Soil Moisture Deficit equation, section (d) of Rushton, is ambiguous. SURFSTOR for this equation should be for the end of the current day, as calculated in section (b). The three steps outlined above partition near surface soil storage between near surface soil storage for the following day, AET, and the soil moisture deficit/reservoir respectively

Groundwater recharge occurs only when the soil moisture deficit is negative, i.e. there is surplus water in the soil moisture reservoir.

We also added an irrigation component to the Rushton model as follows:

1. Natural irrigation demand (before irrigation is applied) is calculated to reach the target value (taw * self.irrig_targ) if Irrigate (bool parameter):

    1. define the irrigation index (those cells with soil moisture < trig (taw* irrig_trig) AND which have not been irrigated more recently than the minimum number of days between irrigation (min_irrig_return))
    2. calculate used irrigation demand * if date is not in the irrigation days (between irrig start and stop) then use demand = 0 * else use demand = max(max_irrigation applied, irrigation demand + irrigation inefficiency)
    3. irrigate from the scheme (irrig_available)
    4. where excess demand remains irrigate from storage
    5. where excess water from the scheme is available add it to storage up to maximum storage
    6. add irrigation water to use_rain and recalculate the soil moisture balance note that irrigation will only be allowed to runoff if allow_irrigation_to_runoff = True
    7. calculate remaining irrigation demand (after irrigation is applied)

2. next day

LSR model inputs -> Precip and PET

We used two sets of inputs for meteorological data to estimate LSR:

- **ERA5-land**: a global reanalysis dataset of meteorological data (1950 - 2020) accessed here
- **Met station data**: Hawea met station data provided by ORC (2012-2021)

We chose to use these two datasets as the met station data is measured data and is therefore more accurate and covers the full optimisation period. For the longer scenario period we relied on the ERA5-land data as it is an available, well documented and validated reanalysis that is available for the full scenario period.

LSR model inputs -> Irrigated area and efficiency

The Rushton model accounts requires irrigation efficiency and irrigation area to be specified. The irrigation area is from MFE's national irrigated land spatial dataset. The irrigation efficiency, triggers, return frequencies and application rates are all specified in the recharge modelling script and are largely informed from McIndoe (2002)

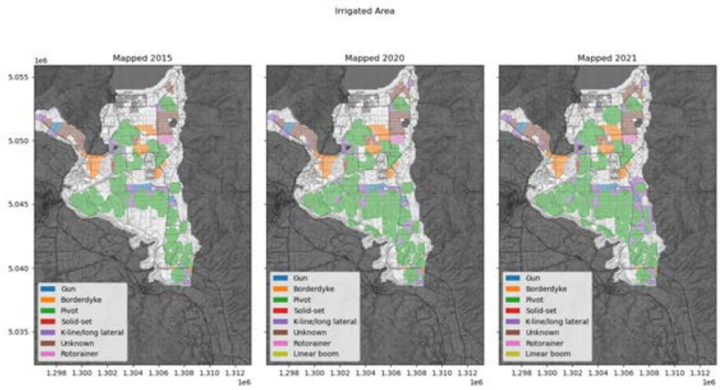


Figure: irrigated area and irrigation types

Correcting ERA5-land data

Unsurprisingly, the ERA5-Land has biases and unit conversion issues. We corrected the ERA5-land data by simple multi-linear regression. For the PET we used the daily ERA5-land PET and the season as the predictor variables and daily met PET. For the precipitation we used the weekly mean ERA5-land precipitation as the predictor variable and the weekly mean met precipitation as the dependent variable. The results of the regression are shown in the figure below.

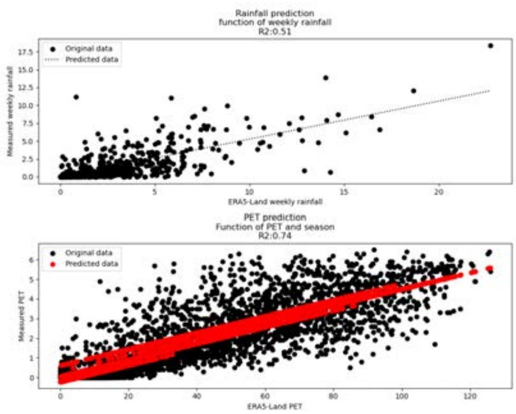


Figure: Era5-land vs met data and regressions

Met station based LSR

The weekly mean met station based recharge and spatial mean recharge are presented in the figures below.

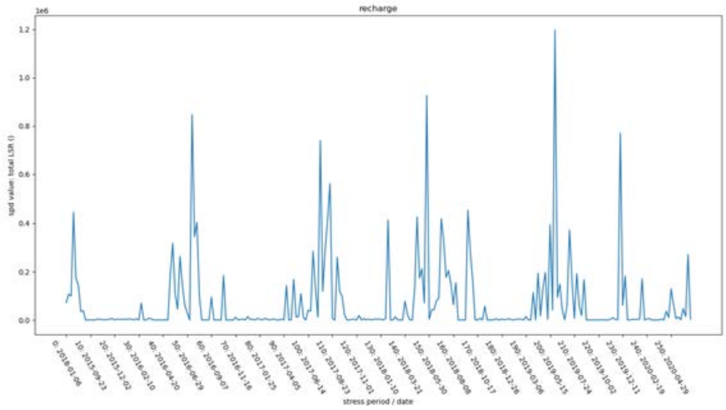


Figure: Weekly mean met data recharge

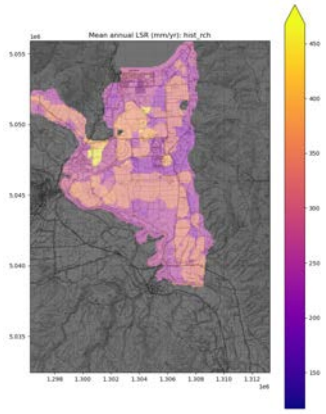


Figure: spatial variation of mean met data recharge

Correcting ERA5-land based LSR

The ERA5-land based recharge was biased relative to the met station based recharge despite the corrections applied to the meteorological data. We corrected the ERA5-land based recharge by two simple multilinear regressions one for irrigated sites and another for dryland sites based on the weekly mean LSR. The regressions and the results are shown in the figures below.

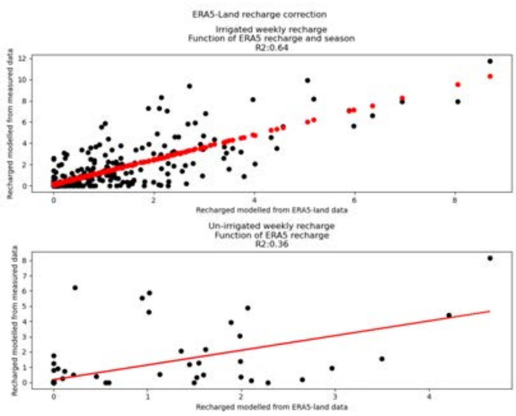


Figure: regressions for ERA5-land and metdata recharge

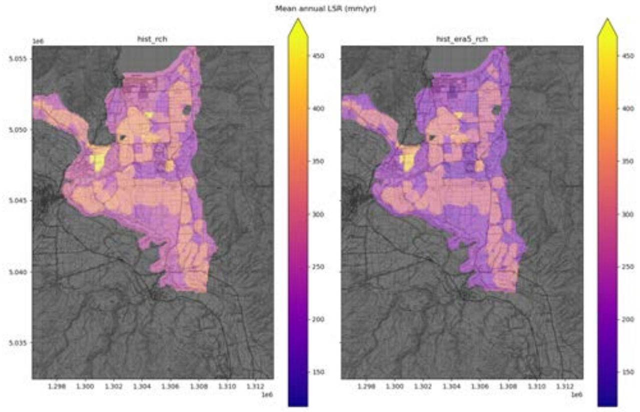


Figure: comparison for the spatially distributed mean recharge, Note that hist_rch is the metdata based recharge and hist_era5_rch is the same period, but using the ERA5-land data

While the regressions are not perfect, they do improve the large scale bias between the ERA5-land based recharge and the met station based recharge. We use the met station based recharge for the optimisation and the ERA5-land based recharge for the scenarios. While this does introduce some bias in our scenarios we analyse the results of the scenarios relative to the optimisation period run with the ERA5-land based recharge, which should mitigate the bias.

Generating a Long record of LSR

The advantage of using the ERA5-land data is that it is available for the full scenario period. We generated several long records of LSR for the full scenario period. The records are defined as follows and are show in the figure below.

- **dryland_rch**: recharge calculated from ERA5-land assuming this is no irrigation in the catchment (e.g. no irrigation losses)
- **irr_rch**: recharge calculated from ERA5-land assuming that irrigation in the catchment maintains the spatial coverage from 2021, but all irrigation is applied via pivot irrigators (e.g. 85% irrigation efficiency)
- **hist_rch**: recharge calculated from the met station data for the optimisation period (2015-2020)
- **hist_era5_rch**: recharge calculated from ERA5-land for the optimisation period (2015-2020)

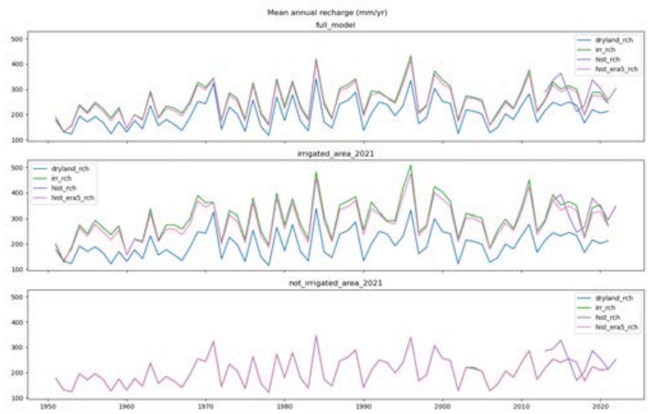


Figure: comparison of the temporal recharge

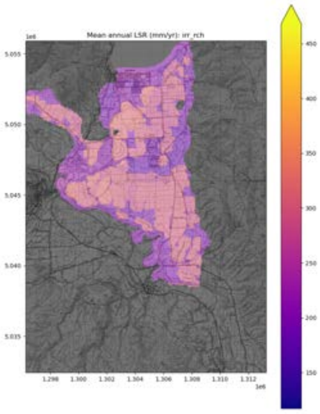


Figure: spatially distributed mean recharge for the irr_rch scenario

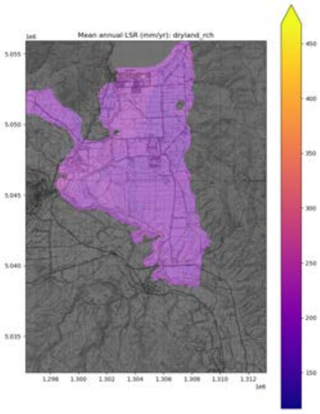


Figure: spatially distributed mean recharge for the dryland_rch scenario

Groundwater Abstraction (pumping)

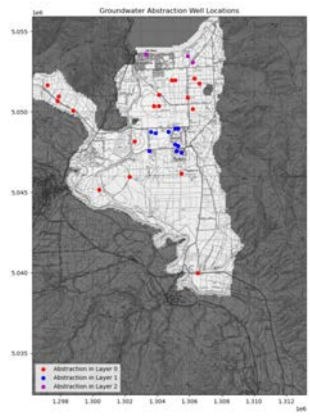


Figure: Groundwater abstraction locations

Groundwater abstraction was defined from the ORC usage data. The water use data was provided by the ORC and further interpreted in Kitteridge (2022). Good metering data is available from 2015 to 2020. The linkage between water metering data and the water abstraction point is complex with multiple abstraction points using 1 meter and multiple meters service 1 abstraction point. Where possible we matched the abstraction to the metering data, and where this was not possible we simply used the location data associated with the meter for the groundwater abstraction points.

The layering of the model is sufficiently simple that we made simple assumptions to define the layer for the abstraction points. In general all abstraction points were assumed to be in the top layer with the following exceptions:

- The abstraction points within the moraine and the layer pinch out zone (see multi-layer (3D) model structure above) were assumed to be in layer 2 (recalling that layers follow Python indexing standards and layer 0 is the top layer). This prevented any abstraction to be misplaced into low conductivity units or in cells that may become dry (i.e. while layer 1 and 2 are being pinched out).
- A number of abstraction bores in and around the Hawea Flat township were placed in layer 1 (the middle layer) as the top model cell would become dry during some periods. This is likely due to model structural error (there are no reports of these bores routinely going dry) and is a limitation of the model. By placing the bores in layer 1 we ensure that the abstraction occurs and is consistent with the model water balance.

The abstraction points included within the model and the temporal variation in groundwater abstraction are shown in the figures below.

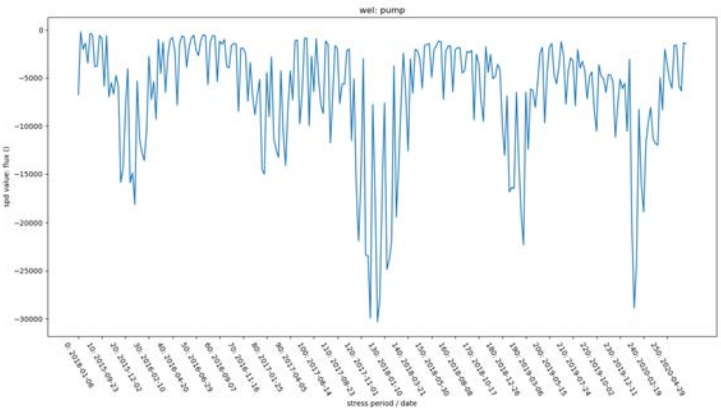


Figure: Total model groundwater abstraction (m/day)

Near river bores

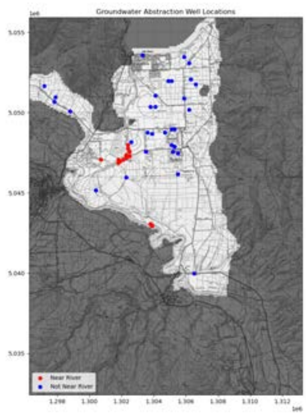


Figure: Groundwater abstraction locations including the abstraction near the Hawea and Clutha Rivers

The bulk of the "groundwater" abstraction in the model domain comes from two locations adjacent to the Hawea and Clutha Rivers. These abstraction bores occur in river proximal gravels which likely have a significantly higher hydraulic conductivity than the rest of the aquifer system. We initially attempted to include these bores in the model, but our model structure was not sufficiently resolved to include this river proximal aquifer. The very high localised abstraction caused dry cells and significant model instability. We therefore considered these river proximal wells as surface water abstraction (e.g. via a gallery) and removed these abstraction points from the model. We did use the river proximal abstraction data to adjust our river gain and loss targets and therefore conserve the water budget.

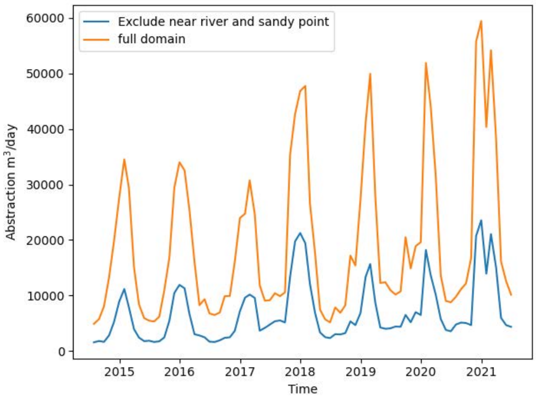


Figure: Comparisons of abstraction including and excluding the near river bores

Major Rivers (Hawea river and Clutha River)

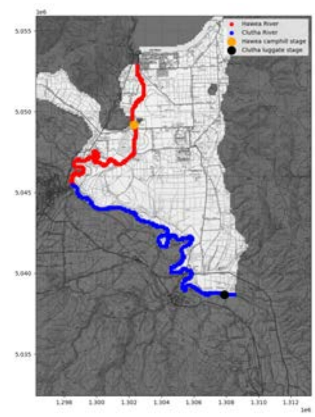
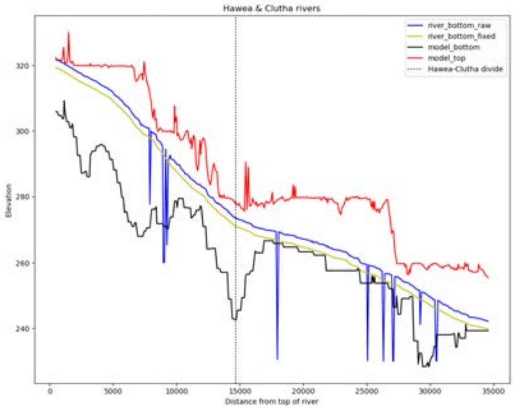


Figure: Major Rivers and Stage monitoring locations

The Hawea and Clutha rivers were included in the model using the stream boundary condition package.. The stream boundary condition package models both stream flow and surface-ground water interactions. While the package allows for modelling of stream stage, for this model we specified the stream stage. The package requires the following inputs:

- Stream location and riverbed elevation
- Stream stage
- Stream flow (at the top segment of each stream)
- The stream bed conductance factor

We defined the stream location with a carefully drawn line along the riverbed informed by a LiDAR dataset provided by Otago Regional Council. The raw riverbed elevation was defined as the minimum LiDAR elevation in each river model cell. This left a river profile that was not consistently decreasing downstream. To correct this we used a rolling mean to define the river bed elevation. Finally we inset the river bottom by 2.5 m so that the riverbed elevation was always below the river stage.



river bed elevations

The streambed conductance factor was a parameter in the model inversion. See the model parameterisation readme for more information The steam flow did not need to be particularly precise as the river would never come close to losing all of its water to the aquifer system. Therefore we set the Hawea River flow to the historical flow measured at Camp Hill. The Clutha River flow was arbitrarily set to 10 * the Hawea River flow. We prescribed the river stage for both the Hawea and Clutha rivers by interpolating historical river stage data at Camp Hill (Hawea River) and at a point on the Clutha River 200 m downstream of Luggate Confluence. The Clutha stage data did not cover the full optimisation period; therefore, we used the ISO-weekly mean river stage for the missing data. The Hawea River stage data was temporally complete. To interpolate the river stage spatially we simply applied the stage measured at Camphill relative to the riverbed elevation to the riverbed elevation in all other Hawea River model cells. The same approach was used for the Clutha River; however where the Clutha River joined the Hawea River there was an offset. To avoid this offset causing model convergence issues we linearly interpolated the stage at the end of the Hawea River to the stage on the Clutha River 200 m downstream of Luggate Confluence. The river stages generated this way do not cover the full scenario period. Therefore we used the ISO-weekly mean river stage for the scenario period.

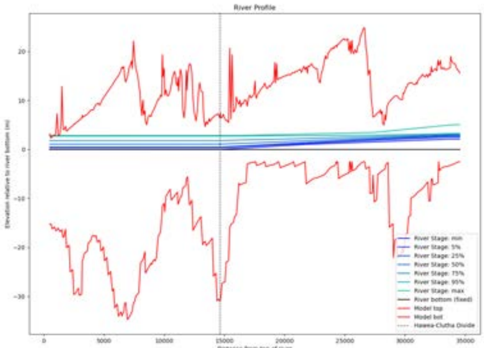


Figure: river stage relative to riverbed elevation

Lake Hawea

Lake Hawea was modelled with the General Head Boundary Package, which allows for time variant heads to be set. The package requires the following inputs:

- Location
- Head
- Conductance

For this model the lake locations were defined as all layers where the model cells that intersected the lake polygon. The lake conductance was set to a very high value (1e10) so that the only parameter defining the lake - model interaction was the cell's hydraulic properties (e.g. hydraulic conductivity). The lake head was set based on the historical lake stage measured at the dam. The historical lake stage covered both the full optimisation period and the full scenario period.

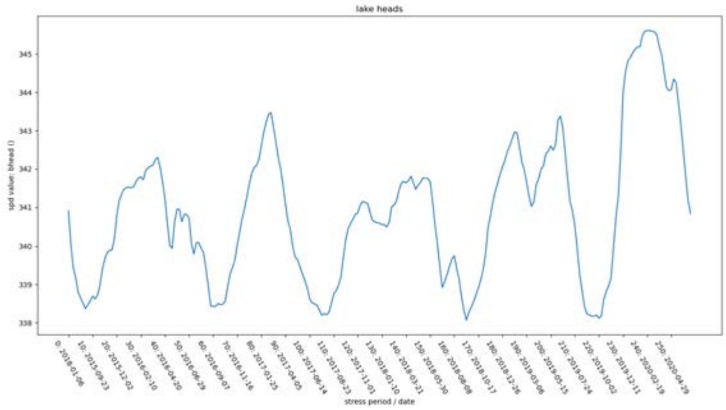


Figure: Lake levels for the optimisation period

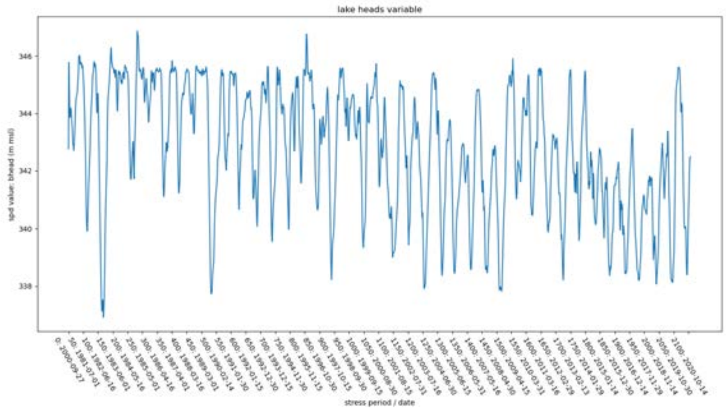


Figure: Lake levels for the Scenario period

Irrigation Supply Race Losses (race losses)

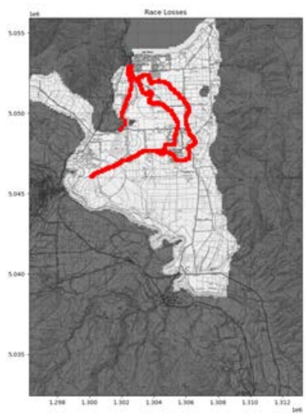


Figure: Race boundary condition locations

There are a number of irrigation supply races across the model domain. Estimates of race water losses are uncertain, however McIndoe (2002) suggests that approximately 10% of the race flows are lost to groundwater. We have access to records of daily race takes from the Hawea Irrigation Co. from 2012-01-01 to 2021-12-31, which covers the full optimisation period. For the scenario period we simply used the ISO weekly mean race losses.

Race losses were implemented as well boundary conditions using the Wel package. Well boundary conditions were placed in every model cell that intersected the race shapefiles and the flux was specified as 10% of the daily race flows spread evenly across every 'race' boundary condition.

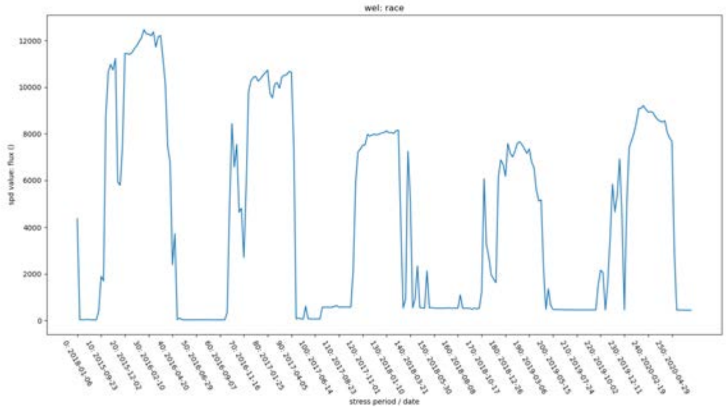


Figure: race losses (m/day) during the optimisation period

Hillside stream inflows (hillside inflows)

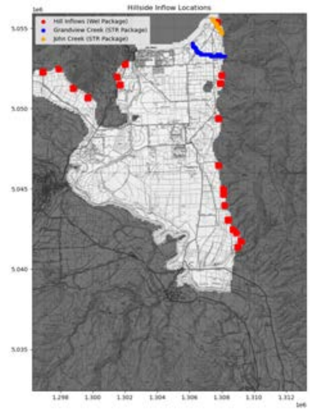


Figure: Hillside inflow locations

Method to estimate hillside inflows

There is has been rather minimal gauging data for the various hillside creeks that flow into the model domain, but it likely that these creeks contribute significantly to the groundwater budget. Recorders were put into Grandview and Lagoon Creek during the winter of 2017, so we have a daily data flow record for the period 2017-08-21 to 2021-02-09. This period is insufficient for even the optimisation period, let alone the scenario period. In addition there are another 19 hillside creeks that have not been gauged. We chose to estimate the hillside inflows based on the long term record of nearby Lindis River. The Lindis River is a much larger river that drains the mountains to the east of Lake Hawea. While the Lindis River catchment is much larger than the hillside inflows, it drains areas with similar geography and climate and has a historical high frequency gauging record at Lindis Peak starting in 1976-09-23. To estimate the hillside inflows we used the following methodology:

1. We estimated the catchment area (CA) for each of the hillside catchments that flow into the model domain using pysheds

2. We manually estimated the Lindis River Catchment above the Lindis Peak recorder (by drawing a shapefile). Note we did not use pysheds here as the lower gradient topography in the Lindis River created complications with the precision of the available DEM
3. We normalised the daily flows of the Lindis River, Lagoon Creek, and Grandview Creek to their respective catchment areas
4. We calculated the mean annual low flow (MALF) normalised to the catchment area for each of the hillside creeks and the Lindis River
5. We then conducted a logarithmic regression of the MALF/CA against catchment area (see figure below). Note that our regression predicted a MALF of zero at a catchment area of 0.14 km^2, which is consistent with the behaviour we would likely expect.
6. We then conducted a multiple linear regression of daily flows of the hillside creeks against the independent variables of Lindis River Flow/CA and the predicted MALF/CA. (see figure below) The Root Mean Squared errors for the daily and monthly flows at Lagoon Creek and Grandview Creek are shown in the table below.
7. We then used both of these regressions to predict the daily flows of the hillside creeks for the period of 1976-09-23 to 2021-06-30. Where the prediction was negative we set the flow to zero.
8. Finally, to reduce the impact of very high flows (where overland flow may not be inconsequential) we set any daily flows greater than the 98th percentile of the daily flows to the 98th percentile.

This methodology certainly has its limitations. Regression scores are not as high as we would like, but given the minimal data this was one of the very few options available. Other options could be based on rainfall-runoff modelling, but this would be very complex, and would introduce additional biases associated with the meteorological data and other modelling parameters. The root mean squared error of the daily flows at Lagoon Creek and Grandview Creek are presented in the table below. Note that the monthly mean flows are much better predicted than the daily flows. Given these RSME values we would consider our predictions to be good enough for the modelling process. In addition, we added a parameterised multiplier to the hillside inflows during our model inversion.

Creek	rsme_daily (m3/s)	rsme_monthly(m3/s)
Grandview	0.057	0.036
Lagoon	0.024	0.014

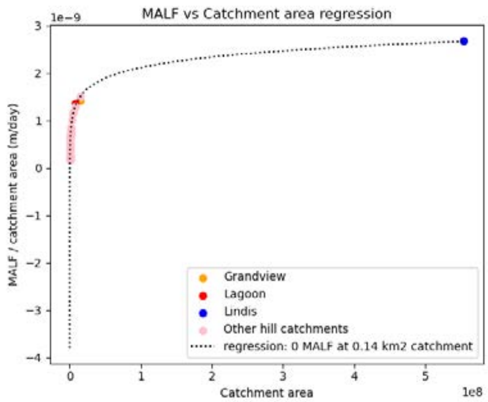


Figure: The relationship used to predict the catchment area normalised MALFs

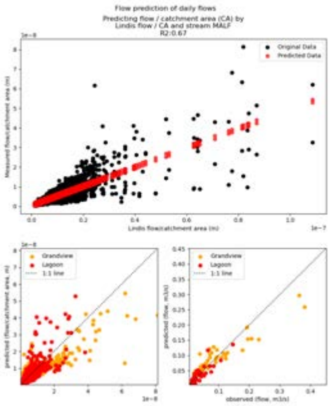


Figure: The relationship used to predict daily hillside creek flows

Large Hillside Inflows (Grandview and John Creek) implementation

Both John Creek and Grandview Creek can have significant flows, flow directly into Lake Hawea, and sometimes do not lose all of their water to groundwater. Therefore we implemented these using the stream boundary condition package.. This allowed the model to partition the groundwater losses across the length of the stream. The stream bottom was set to 2 m below the model top. The stream bottoms were then adjusted so that they were continuously decreasing downstream. The conductance factor was parameterised. The stream flow at the top of the stream was set using the inflow estimates described above and the stream stage was set at the smoothed model top (i.e. 2 m above the stream bottom).

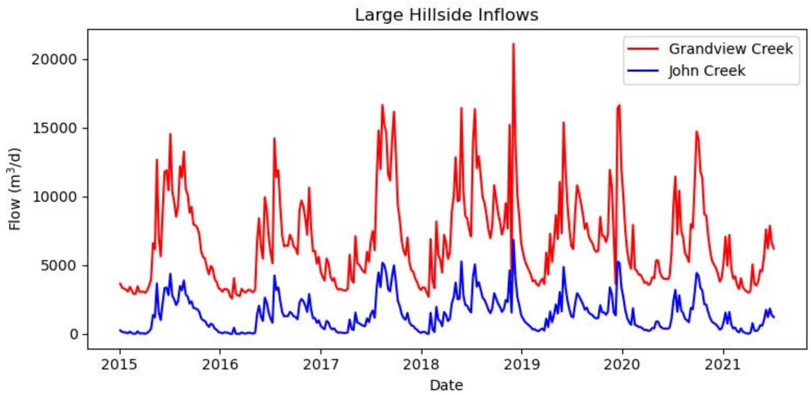


Figure: Large hillside inflows rates

Smaller Hillside inflows (other hillside inflows) implementation

All of the smaller inflows were implemented using the Well package. A series of 9 well boundary conditions were placed, centered on model cells that intersected the hillside inflow shapefiles. The flux was set to the daily hillside inflow estimate divided by 9 and spread evenly across the 9 well boundary conditions.

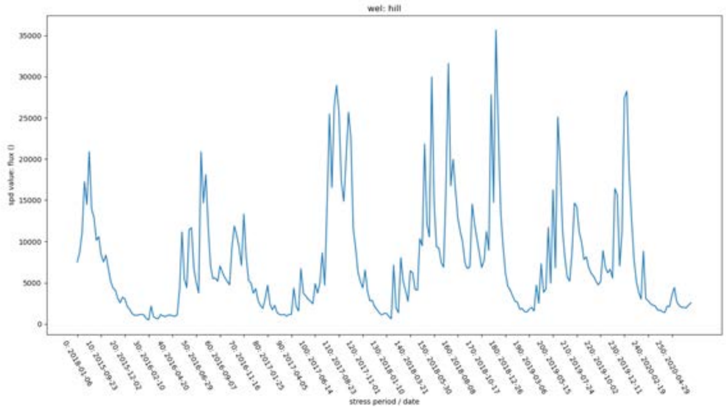


Figure: Total inflow for the smaller hillside catchments (m/day)

Model Zones

A number of model zones were generated to more easily visualise the model results. The generated zones are shown below.

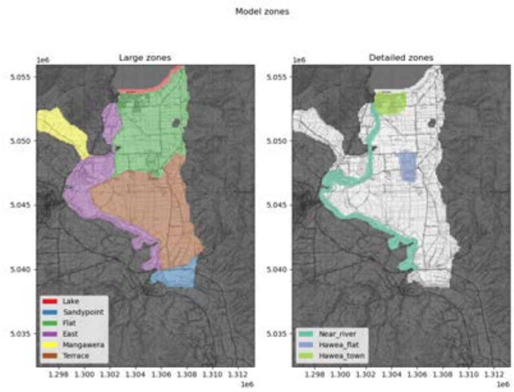


Figure: helpful model zones

References

- Wilson, J., 2012. Hawea Basin Groundwater Review. prepared by Resouce Science Unit of Otago Regional Council, June 2012, Dunedin.
- McIndoe, I., 2002. Efficient and reasonable use of water for irrigation.
- Rushton, K.R., Eilers, V.H.M., Carter, R.C., 2006. Improved soil moisture balance methodology for recharge estimation. Journal of Hydrology 318, 379-399.

Hawea Transient groundwater model (Hawea Model) parameterization

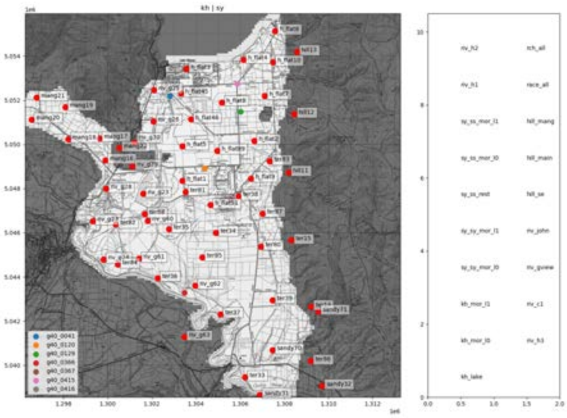


Figure: All Hawea Model parameters and their location in the model domain.

Author: Matt Dumont

Date: 2021-11-02

Version: 1.0.0

Status: Final

KSL project: Z22031HAW_hawea-model

Purpose: This document describes the development of model parameters

Index

Table of Contents

Module Index

- README.rst: this document
- base_data: raw input data for the model parameterisation
- processed_data: processed data for the model parameterisation that was built by the scripts in this folder from the raw data in the base_data folder
- static_params.py: a script to define the static model parameters
- pilot_points.py: a script to create, define, and interpolate pilot points for kh and sy

- inital_parametersiation.py: a script to define the initial model parameters (before optimisation)
- plot_parameter_names.py: a script to plot the parameter names generates parameter_map.png
- optimised_parameterisation.py: a script to easily access optimised parameter sets
- **optimised_parameter_sets: optimised parameter sets**
    - 3d_v1a_opt.par: optimised parameter set for the 3D model version 1a
    - 3d_v1b_opt.par: optimised parameter set for the 3D model version 1b
    - 3d_v1d_opt.par: optimised parameter set for the 3D model version 1d
- parameter_map.png: a map of the model parameters

Static parameters

There are four static parameters in the model. They are:

Parameter Name	Parameter type	Value	Comment
lake_ss	specific storage in the lake zone	1e-10	Minimal storage in the lake zone
lake_sy	specific yield in the lake zone	1e-10	
vka	vertical conductivity	1	Set to 1 as we are not modelling vertical gradients (no z data for targets)
lake_conduct	lake conductance	1e10	Set high so that lake hydraulic conductivity is the only parameter bounding the lake-groundwater interaction

Spatial parameters

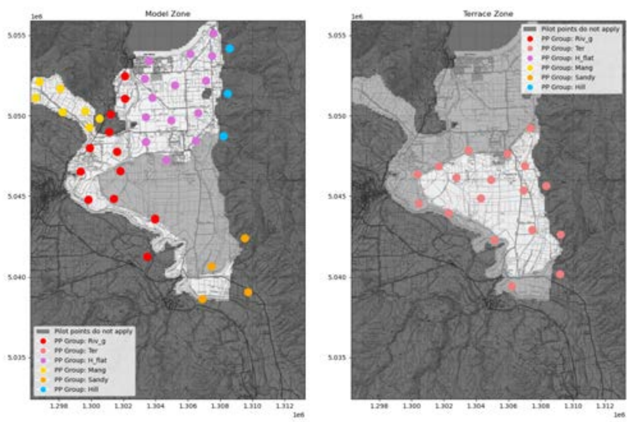


Figure: Spatial parameters for hydraulic conductivity (kh) and specific yield (sy) in the model domain.

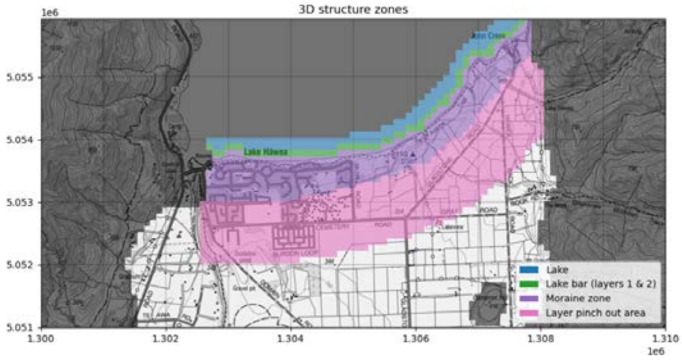


Figure: 3D spatial view of the multi-layer model structure zones

Hydraulic conductivity (kh) and specific yield (sy) were parameterised using pilot points in two discrete zones; one for the high terrace and another for the rest of the model. We chose to divide the model into two zones because it would be reasonable that the significantly older high terrace sediments could have different hydraulic properties to the younger sediments in the rest of the model. In addition this sharp parameter change could allow for the impacts of the Q4 Albert Town Advance moraine which we did not explicitly represent in the model. Pilot points extend beyond the active model domain to allow interpolated values to be calculated at the boundary of the active model domain. Note that specific storage, and parameters for the moraine zones are described in *other parameters* below. The interpolated parameters apply to all layers in the model (inc. the layer pinch out zone) except in the moraine zone where these parameters apply to only layer 2 (the bottom layer recalling that layering follows python indexing format layer 0 = top layer). These parameters do not apply to the lake zone or the lake bar zone.

We initially parameterised the model as a relatively homogeneous system. We considered leveraging the parameters generated by Wilson et al. (2012), but decided against is as it could introduce bias from the previous model. Parameter ranges and initial values are defined in the table below.

Parameter group	Initial	Min	Max
Main Kh	3e2	1e-2	1e4
Terrace Kh	5e1		
Main Sy	1e-2	1e-4	3e-1
Terrace Sy			

Wilson et al. (2012) assessed the annual lake fluctuation in Lake Hawea using the Jacob tidal equation. They found values of specific yield = 0.012 and a transmissivity estimate of 1300 m^2/d. We used the specific yield value to set the initial values for the specific yield. The tidal estimates of transmissivity do not include the complex three dimensional structure which almost certainly limits the transmissivity. We therefore set the initial value to a higher value than the tidal estimate. We set the upper bound based on the highest recorded transmissivity from pump tests, though as noted in Wilson et al. (2012) these tests are somewhat suspect. We set the lower bound at an arbitrary, but very low value allowing kh to span six orders of magnitude.

Interpolating spatial parameters

We interpolated the pilot points to continuous values using Radial basis function (RBF) interpolation with a multiquadric kernel. The RBF interpolation was performed using the *scipy.interpolate.Rbf* function. The RBF interpolation was preformed on log (base 10) transformed values. An example of the RBF interpolation is shown below.

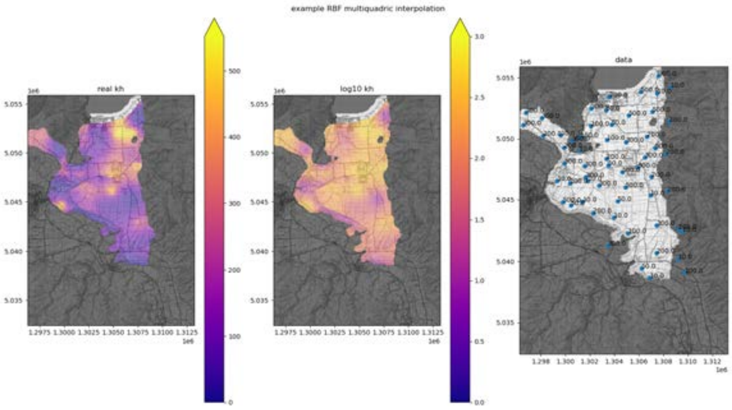


Figure: Example of the RBF interpolation of the spatial parameters. The interpolated values are shown for the bottom layer of the model.

Other parameters

A number of other parameters were included in the model. These parameters are listed in the table below. Recall layering follows Python indexing format e.g. layer 0 = top layer.

Parameter type	Package	Name	Initial	Min	Max	Applies to
Hillside inflow multiplier	Wel	hill_se	1	0.8	1.2	see figure below
Hillside inflow multiplier	Wel	hill_main	1	0.8	1.2	
Hillside inflow multiplier	Wel	hill_mang	1	0.8	1.2	
Recharge Multiplier	Rch	rch_all	1	0.5	1.2	full model domain
Race loss multiplier	Wel	race_all	1	0.8	1.2	
River Conductance	Riv	riv_h1	1000	100	10000	see figure below
River Conductance	Riv	riv_h2	1000	100	10000	
River Conductance	Riv	riv_h3	1000	100	10000	
River Conductance	Riv	riv_c1	1000	100	10000	
River Conductance	Riv	riv_gview	100	50	5000	
River Conductance	Riv	riv_john	100	50	5000	
Conductivity	Upw	kh_mor_l0	300	0.001	1000	moraine zone layer 0 & lake bar layer 0
Specific Yield	Upw	sy_sy_mor_l0	0.01	0.0001	0.3	
Specific Storage	Upw	sy_ss_mor_l0	0.0001	0.000001	0.001	

Conductivity	Upw	kh_mor_l1	0.0001	0.0000001	1	moraine zone layer 0
Specific Yield	Upw	sy_sy_mor_l1	0.001	0.0001	0.3	& lake bar layers 1 & 2
Specific Storage	Upw	sy_ss_mor_l1	0.0001	0.000001	0.001	
Specific Storage	Upw	sy_ss_rest	0.0001	0.000001	0.001	area of pilot points
Conductivity	Upw	kh_lake	300	0.001	1000	lake zone all layers

We set the initial parameters and ranges for the multipliers (hill_se, hill_main, hill_mang, rch_all, race_all) to allow a 20% change around the predicted in the inflow values. The initial values for the Hawea and Clutha River conductance (riv_h1, riv_h2, riv_h3, riv_c1) were roughly pulled from the model developed in Wilson et al. (2012). The initial values for the smaller river conductance (riv_gview, riv_john) were set as an order of magnitude lower than the Hawea and Clutha River conductance. The ranges for the river conductance are somewhat arbitrary but act to allow the model to explore a range of values. The initial and ranges of kh_lake, and kh_mor_l0 were set to the values used for the main pilot points. The specific storage values (sy_ss_rest, sy_ss_mor_l0, sy_ss_mor_l1) were set to typical specific storage values. The specific yield parameters were set to the range used for the main pilot points. The initial value for sy_sy_mor_l0 was set to the starting pilot point values while the initial value for sy_sy_mor_l1 was set an order of magnitude lower. kh_mor_l1 (the moraine low conductivity unit) was set to the typical values for glacial till.

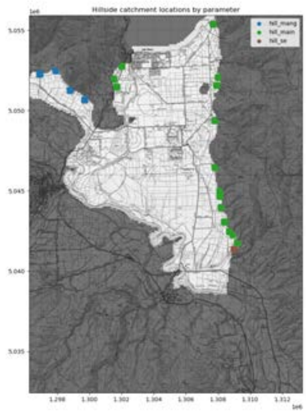


Figure: Spatial location of the hillside inflow multiplier parameters.

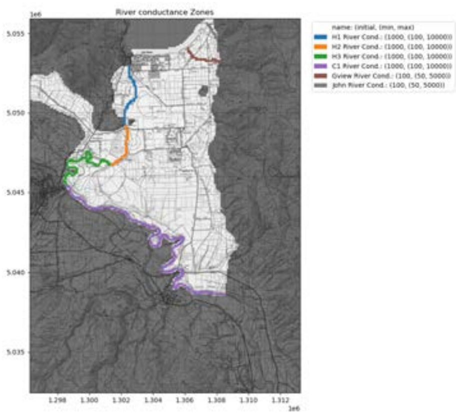


Figure: Spatial location of the river conductance parameters.

References

- Wilson, J., 2012. Hawea Basin Groundwater Review. prepared by Resouce Science Unit of Otago Regional Council, June 2012, Dunedin.

Hawea Transient groundwater model (Hawea Model) Targets and Sensitive sites

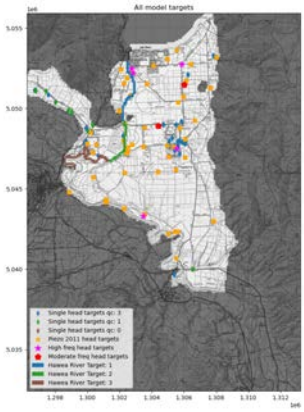


Figure: All Hawea Model targets

Author:	Matt Dumont
Date:	2021-11-02
Version:	1.0.0
Status:	Final
KSL project:	Z22031HAW_hawea-model
Purpose:	This document describes the development of model targets and the identification of sensitive sites

Index

Table of Contents

Module Index

- README.rst : readme document detailing the methods and data used to develop the targets
- model_output.py: script to extract consistent model outputs and plots
- get_raw_target_data.py: ingest raw target data
- get_indicative_times.py: get indicative times for the targets that fall outside of the optimisation period
- head_targets.py: definition of the head targets
- riv_gain_loss_targets.py: definition of the river gain and loss targets
- senstive_sites.py: identification of sensitive sites
- target_structure_checks.py: checks to ensure that the targets and the model structure were not mutually exclusive
- base_data: base input data for the targets
- processed_data: processed target data, this was developed from the raw data in the base_data folder

Groundwater head targets

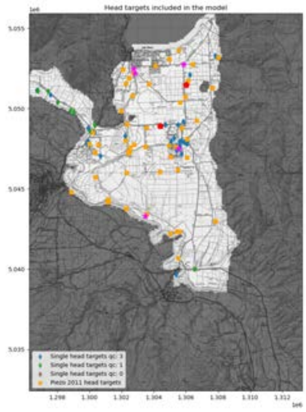


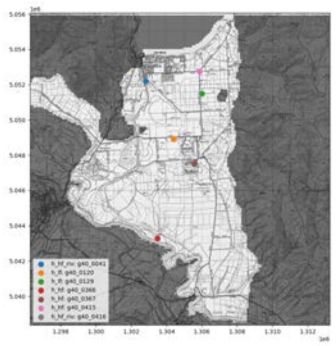
Figure: Spatial distribution of groundwater head targets

High and moderate frequency targets

High and moderate frequency targets were provided by Otago Regional Council (ORC) and included the following wells. Note that in the code these targets are often referred to as "regular" targets:

Well name	Group	Description
G40/0041	h_hf_riv	High frequency loggers installed after 2014, near Hawea river
G40/0416	h_hf_riv	High frequency loggers installed after 2014, near Hawea river
G40/0366	h_hf	High frequency loggers installed after 2014
G40/0367	h_hf	High frequency loggers installed after 2014

G40/0415	h_hf	High frequency loggers installed after 2014
G40/0120	h_lf	NGMP wells, which have a longer record than the others, but less frequent sampling
G40/0129	h_lf	NGMP wells, which have a longer record than the others, but less frequent sampling



Targets from the 2011 Piezometric Survey

The 2011 piezometric survey was conducted on 21-sept-2011 and is detailed in Wilson et al., (2012). The survey provides a significant spatial distribution of targets across the model domain and is highly reliable. These targets make up much of the information used to constrain the model in the optimisation away from the high frequency targets.

Single targets

Often when drillers are installing wells, they will take a single, static water level measurement at the time of installation. This information is less reliable than a piezometric survey, but it is still useful to constrain areas of the model where there is little other information. There is often missing information for these records, as such we created four quality categories for these targets:

- 0: no date data for the depth to water field -- Not included in the model
- 1: no elevation data present (read from DEM) -- Only included in the model in the Sandy Point and Maungawera Valley, where there is a relative dearth of other information.
- 2: no depth data for the well -- No wells matched this category
- 3: as good as it gets -- included in all parts of the model

River gain and loss targets

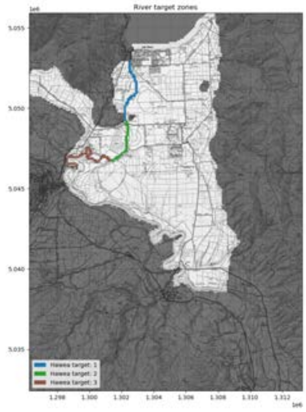


Figure: Spatial distribution of river gain/loss targets

Measured data

Two sets of four concurrent river gaugings on the Hawea River were used to develop the river gain/loss targets. The first set of gaugings were taken on 2017-09-29 and the second set were taken in on 2018-02-07. These targets are inherently uncertain as the gauging error is typically >= 10% of the river discharge and in braided river systems such as the Hawea River, the river discharge can vary significantly over short distances as water travels in and out of the river proximal and riverbed gravels. Nevertheless, the river gain/loss targets are the only measured constraint on the model and are therefore used in the optimisation.

Expert judgement

In addition to the measured datasets described above, the expert judgement of the Hawea River is that it is largely a gaining reach from the Hawea Dam to Camp Hill. After Camp Hill, the river loses a significant amount of water as the river turns west against the high terrace. The lower reaches are gaining and losing until it reaches the Clutha River. The Clutha River is exclusively thought of as a losing reach. This expert judgement is in agreement with the measured data and while it is not explicitly included within the model optimisation targets, it was used to qualitatively assess the performance of the model.

Managing targets outside of the optimisation period

Many of the targets used in the model fall outside of the optimisation period. If we only included information from within the optimisation period we would be left with only the "regular" high and moderate frequency targets -- that is seven spatial targets across the model domain and no targets in the Maungawera Valley. Therefore we needed to apply the targets out of the optimisation period to the most appropriate time within the optimisation period. This is done by:

1. Calculating the last 12 months normalised average recharge and hillslope inflow for each month in the scenario period (1980-07-18 to 2020-12-01). Any targets outside of this period were excluded from the model. The choice to use the last 12 months was based off of the annual cycle, and confirmed by calculating the predictive power of multiple different time periods (e.g. 6 months, etc.). The annual data provided the best predictive power.
2. The target dates outside of the optimisation period were then assigned to the month within the optimisation period that was the closest (cartesian distance of the normalised 12 previous month recharge and hillslope inflows) to the target date and had the closest normalised recharge and hillslope. Targets were allowed to shift up to 1 month (e.g. a target measured in September could was assigned to the closest August, September or October month). This was done to maintain any seasonal effects while allowing a larger potential pool of matches.
3. Targets which were temporally shifted in this way were assigned to all stress periods in the month.

The following figures show the results of this process. Figures of each target month are available here

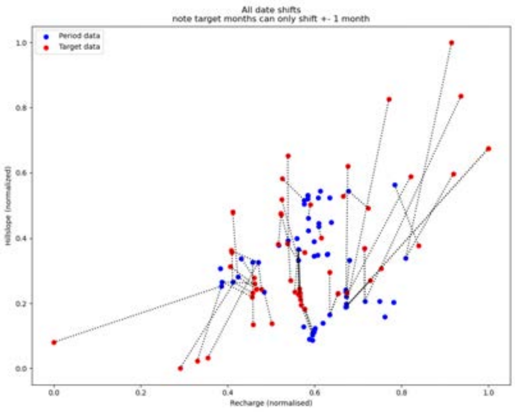


Figure: Target period shifts for all targets

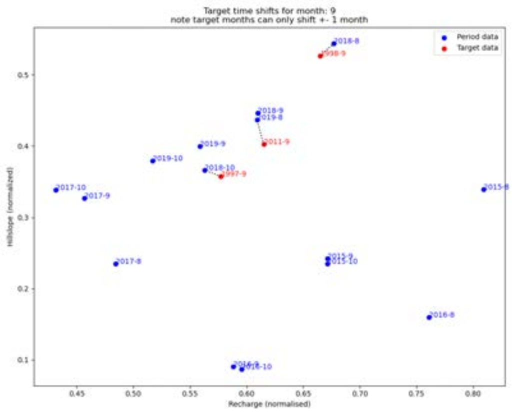


Figure: Target period shifts for the targets measured in September, recall that the 2011-9 targets include the piezometric study

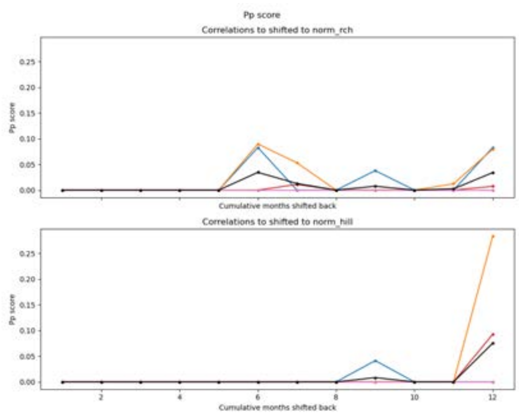


Figure: The predictive power score of different monthly aggregations of the normalised recharge and hillslope inflows

Temporal distribution of targets

The final temporal distribution of targets in the model is shown in the following figures. Recall that the targets which were measured outside of the optimisation period were assigned an indicative time during the optimisation period.

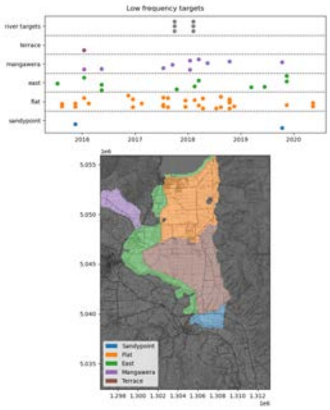


Figure: Temporal distribution of low frequency groundwater head and river gain/loss targets

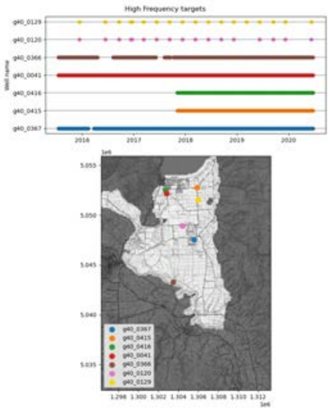


Figure: Temporal distribution of high frequency groundwater head targets

Model Objective Function and target weighting

The objective function is at a high level simply the weighted sum of squared errors between the modelled values and target. The weighting strategy is often adjusted during the course of the model optimisation (for more info on the final weighting scheme see the optimisation readme); however, the initial strategy for the weighting of the targets is described below.

1. **We developed a hierarchy of target groups as follows:**

    1. **h_hf:** High frequency groundwater head targets: these are high quality data with a high temporal resolution and a moderate spatial resolution. They are by far the most important targets within the model.
    2. **h_hf_riv:** High frequency groundwater head targets near the Hawea River: these are high quality data with a high temporal resolution, but they are adjacent to the Hawea River so are more susceptible to structural bias based on our implementation of the Hawea River in the model.
    3. **h_lf:** Moderate frequency groundwater head targets: these targets provide two additional sites with a number of samples across the optimisation period.
    4. **h_piezo:** Scott 2011 piezo survey: these targets provide a significant spatial distribution of targets across the model domain and are highly reliable; however, they fall outside of the optimisation period and therefore likely have some bias.
    5. **h_single_1:** Single targets Q3: generally lower quality targets but they provide some additional data.

6. **h_single_3:** Single Targets Q1: generally very low quality targets, but in areas they are the only data available.
7. **riv:** River targets: useful and the only way to constrain the river gain/loss, but they are not very reliable (as described above).
8. **There were two additional target groups that were never weighted:**
    1. **rwh_hf:** ISO weekly mean for each high frequency target (e.g. for the "typical" water year)
    2. **rwh_hf_riv:** ISO weekly mean for each high frequency near river target (e.g. for the "typical" water year)

2. The initial weights were set so that all targets weights were proportional to their value, that is that the expected value * weight = 1
3. The weights were then adjusted so that single_3 had twice the impact relative to single_1
4. The weights were then normalised so that the total impact of each group was equal regardless of the number of targets within the group.
5. The groups were then manually weighted with a multiplier to adjust the relative impact of each group in accordance with the hierarchy described above. This weight factor was adjusted during the course of the optimisation. For more info on the final weighting scheme see the optimisation readme.

Note that when a target occurred in a dry model cell the modelled value was set to the head value in the cell in the layer below, or if that cell was also dry, the head value in the cell below that. This was done to ensure that the modelled values did not get impacted by the dry cell flag.

Other sensitive sites

There are two sensitive wetlands within the model domain -- Butterfield Reserve and Campbell's Reserve. These wetlands were not included within the model as boundary conditions, but they are indexed here.

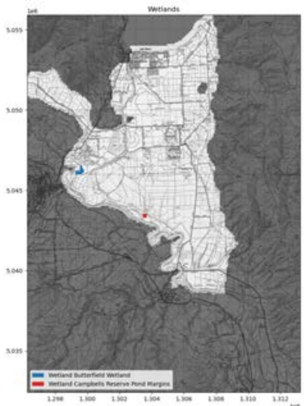


Figure: Wetlands in the Hawea model domain

References

- Wilson, J., 2012. Hawea Basin Groundwater Review. prepared by Resouce Science Unit of Otago Regional Council, June 2012, Dunedin.

Hawea Transient groundwater model (Hawea Model) optimisation and results

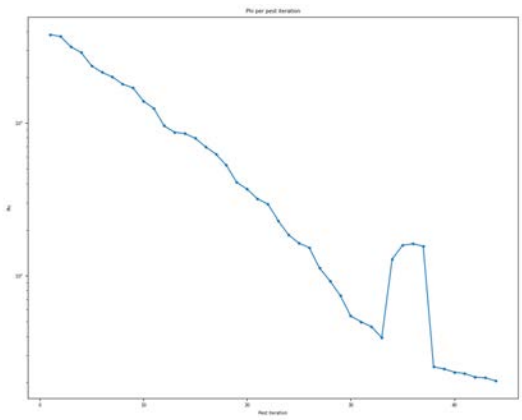


Figure: Model objective function over the optimisation

Author:	Matt Dumont
Date:	2021-11-02
Version:	1.0.0
Status:	Final
KSL project:	Z22031HAW_hawea-model
Purpose:	This document provides the methodology and results for the model optimisation process and discusses the model limitations

Index

Table of Contents

Module Index

- README.rst: readme document detailing the optimisation process and methodology
- **PEST optimisation build, run, and post processing scripts**
    - build_optimisation.py: script and functions to build the PEST files
    - a_build_run_optimisation_version.py: build and run a PEST optimisation
    - run_opt_step_models.py: run the step models from a PEST optimisation
    - manual_optimisations: manual optimisations that were run, in the end these never contributed more than some information to the modeller
    - model_utils_for_forward_run.py: functions to build and run a model from PEST parameter files
    - compare_parameterisations.py: script to compare parameters across multiple parameter files
    - hawea_plot_optimisation.py: script to plot the optimisation results
    - plot_multiple_high_freq.py: script to plot multiple high frequency observations for given PEST obs files
- **Manage optimisation period:**
    - determine_opt_start.py: script to determine the start and end of the optimisation period
    - optimisation_period.py: script to manage and hold the information about the optimisation period
- **Optimisation Results**
    - **optimisation_results: results for the optimisation holding all of the PEST input and output files**
        - 3d_v1a: optimisation results for the 3D model version 1a
        - 3d_v1b: optimisation results for the 3D model version 1b
        - 3d_v1d: optimisation results for the 3D model version 1d (final model)
    - **final_opt_models: The final optimised model files**
        - 3d_v1a: final optimised model files for the 3D model version 1a
        - 3d_v1b: final optimised model files for the 3D model version 1b
        - 3d_v1d: final optimised model files for the 3D model version 1d (final model)
        - compress_uncompress_model.py: utilities to compress and uncompress the model files so they could be included in the Git repo (50mb limit)
- **Computational support files**
    - compile_pest: compile PEST for linux
    - pest_run_data: static data needed by PEST to run the model
    - git_setup.sh: script to setup the Git repo for the optimisation on a machine
- **Model overview**
    - pre_optimisation_overview.py: make pre optimisation overview plots
    - make_preopt_slideshow.py: make a pre optimisation slideshow

- pre_optimisation_plots_png: pre optimisation plots of boundary conditions, targets, parameterisation, and other supporting work, many of these figures are referenced in the various readme.rst files
- make_opt_presentation.py: make a presentation of the optimisation results for a meeting

Optimisation overview

The optimisation process involved changing group weightings, parameter bounds, and the model structure. This chaotic process is typical in model optimisation and is typically very difficult to follow for anyone other than the primary modeller. To try and make this process more transparent and to provide a record of the process, we created a new Git branch for each optimisation version. This allowed us to track the changes to the model and reproduce and archive all of the key input data with any changes. To reduce the size of the Github repo we deleted all of the abandoned branches, but we produced a Github release for each abandoned branch so that the data could be recovered if needed.

Broadly speaking there were 3 main stages to the optimisation process:

1. A 1 layer model (2D)
2. Specific sub model (2D) to test whether or not the terrace observations could be fit by disconnecting the terrace from the main model
3. A 3 layer model (3D)

Optimisation setup / PEST structure

The model was optimised via PEST which is a model calibration and optimisation package. The interface to the model was handled via flopy which is a Python package for working with MODFLOW models and pyemu which is a Python package for working with PEST models. The PEST iterations were run in parallel on a cluster of linux machines using Beopest which is a subpackage of PEST. Beopest was managed via an in house class called BeopestManager. In addition, some manual optimisation was undertaken during the optimisation process to better understand the limits of specific model structures. These manual optimisations were undertaken using another in house class called SshDist. The main optimisation script was a_build_run_optimisation_version.py

The build of the PEST runfile was undertaken in build_optimisation.py. which has a number of component functions:

- raw_pest: - Overarching function to build the PEST runfile (calls the following functions). - Also handles the singular value decomposition (SVD) parameters.
- make_template_and_infiles: - Make the template and infiles for PEST to interact with the model parameter inputs.
- make_ins_and_output_files: - Make an example output files (model outputs) and the PEST instruction files to read the model output data (targets).
- set_control_data: - Set the control data for the PEST runfile.
- set_parameter_data_groups: - Set parameter data groups, limits, transformations, and derivative handling.
- set_obs_data: - Set the observation data, weights, and group weightings.

While the full specification for our PEST optimisation is available in the code the following relevant key parameters are listed below:

- Kh and river conductance parameters were varied on a log transform (*partrans*)
- All other parameters were varied with no transform (*partrans*)
- PEST was run in estimation mode (*pestmode = 'estimation'*)
- PEST allowed model failures in lamda calculation (*lamforgive = 'lamforgive'*)

- PEST allowed model failures in derivative calculation (*derforgive = 'derforgive'*)
- PEST was run with singular value decomposition (*svdmode = 1*)
- The eigenvalue threshold for svd was set to 5e-7 (*svd_dataeigthresh = 5e-7*)

Standard optimisation outputs

For each optimisation version we produced a number of standard outputs that are consistent across each optimisation result. We detail them here so that individuals can have easy access to the key outputs.

The output structure is as follows (links to the files are provided to 3d_v1d which is the final optimised model):

- **Base_Optimisation_plots:**

    - final_opt_model: plots of the final optimised model including the parameterisation, model heads, and the model fits to the observations - cross_sections: plots of heads in model cross sections - spatial_hds: spatial plots of head target residuals - spatial_riv: spatial plots of the river gain and losses - str_flow: plots of the stream flow in the river boundary conditions - Max_heads_(Hawea_aquifer).png: minimum heads across the model time steps - Min_heads_(Hawea_aquifer).png: maximum heads across the model time steps - Range_of_Heads_(Hawea_aquifer).png: range of heads across the model time steps - Steady_state_heads_(Hawea_aquifer).png: plot of the steady state heads (in layer 0 for most of the model, but layer 2 for the moraine areas) - 3d_hds.png: plot of heads in the 3D zone - SS_budget.png: plot of the steady state water budget - all_riv_targets_mes_mod.png: plot of the measured vs modelled river targets - all_riv_targets_residual.png: plot of the river target residuals - all_river_fluxes_hill.png: total river fluxes for each conductance parameter zone for John and Grandview Creek - all_river_fluxes_large.png: total river fluxes for each conductance parameter zone for the Hawea and Clutha rivers - dry_cells_l0.png: number of time steps with dry cells for layer 0 - dry_cells_l1.png: number of time steps with dry cells for layer 1 - dry_cells_l2.png: number of time steps with dry cells for layer 2 - flooded_cells_l0.png: number of time steps with flooded cells for layer 0 - flooded_cells_l1.png: number of time steps with flooded cells for layer 1 - flooded_cells_l2.png: number of time steps with flooded cells for layer 2 - hds_all_mod_v_meas.png: plot of the measured vs modelled heads - hds_all_residual_time.png: plot of the head target residuals - hds_closeup_h_g40_0041.png: plot of measured and modelled heads for well G40/0041 - hds_closeup_h_g40_0120.png: plot of measured and modelled heads for well G40/0120 - hds_closeup_h_g40_0129.png: plot of measured and modelled heads for well G40/0129 - hds_closeup_h_g40_0366.png: plot of measured and modelled heads for well G40/0366 - hds_closeup_h_g40_0367.png: plot of measured and modelled heads for well G40/0367 - hds_closeup_h_g40_0415.png: plot of measured and modelled heads for well G40/0415 - hds_closeup_h_g40_0416.png: plot of measured and modelled heads for well G40/0416 - hds_h_piezo_mod_v_meas.png: plot of measured and modelled heads for the piezo survey - hds_h_piezo_residual_time.png: plot of the piezo survey head residuals - hds_h_single_1_mod_v_meas.png: plot of measured and modelled heads for the single head targets Q1 - hds_h_single_1_residual_time.png: plot of the single head targets Q1 residuals - hds_h_single_3_mod_v_meas.png: plot of measured and modelled heads for the single head targets Q3 - hds_h_single_3_residual_time.png: plot of the single head targets Q3 residuals - hds_normal_year_all.png: plot of the measured and modelled heads for the normal water year (ISO week mean) - hds_normal_year_mod_v_meas.png: plot of the measured vs modelled heads for the normal water year (ISO week mean) - hds_regular_mod_v_meas.png: plot of the measured and modelled heads for the regular observations (e.g. high frequency) - hds_regular_residual_time.png: plot of the regular observation head residuals - hds_regyear_h_g40_0041.png: plot of the regular water year heads (ISO weekly mean) for well, G40/0041 - hds_regyear_h_g40_0366.png: plot of the regular water year heads (ISO weekly mean) for well, G40/0366 - hds_regyear_h_g40_0367.png: plot of the regular water year heads (ISO weekly mean) for well, G40/0367 - hds_regyear_h_g40_0415.png: plot of the regular water year heads (ISO weekly mean) for well, G40/0415 - hds_regyear_h_g40_0416.png: plot of the regular water year heads (ISO weekly mean) for well, G40/0416 - more_than_50_outers.png: areas in the model that required more than outer 50 iterations to converge - more_than_100_outers.png: areas in the model that required more than outer 100 iterations to converge - more_than_300_outers.png: areas in the model that required more than outer 300 iterations to converge - more_than_500_outers.png: areas in the model that required more than outer 500 iterations to converge - more_than_800_outers.png: areas in the model that required more than outer 800 iterations to converge

- obs_plots: plots of the model objective function and target residuals through the optimisation process (e.g. at each optimisation step))
- regular_hds_closeup: plots of changes in the fit to the regular observations ((e.g. high frequency) through each optimisation step)
- param_plots: plots of parameter values through the optimisation process (e.g. at each optimisation step)
- param_fail_plots: plots of parameter values that failed to converge vs those that did not
- param_sen_plots: plots of parameter sensitivity through the optimisation process (e.g. at each optimisation step)
- parameters_norm_to_bounds.txt: a text file of the parameter values of the final model normalised to the parameter bounds
- parameters_norm_to_bounds_close.txt: as above, but only those that are close to their bounds
- parameter_norm_sy_kh.png: ignore, bug in plot
- jacobian_filled_0_of_1.png: plots of whether or not the Jacobian was filled (red values had model failure)
- hk_values.png: plot of the final hk parameter values at the pilot points
- kh_array.png: plot of the interpolated final kh parameter values
- sy_values.png: plot of the final sy parameter values at the pilot points
- sy_array.png: plot of the interpolated final sy parameter values

1 layer model (2D) optimisation results

With the 2D model we were able to fit many of the targets within the model, but despite numerous (we ran 17 unique optimisations) parameterisations, observation weighting schemes, and change to the model structure we were unable to replicate the water levels at G40/0415. The model could either fit the shape of the groundwater levels but there was substantial bias in the mean groundwater level (too high) or we could fit the mean groundwater level but the shape of the groundwater levels was lost.

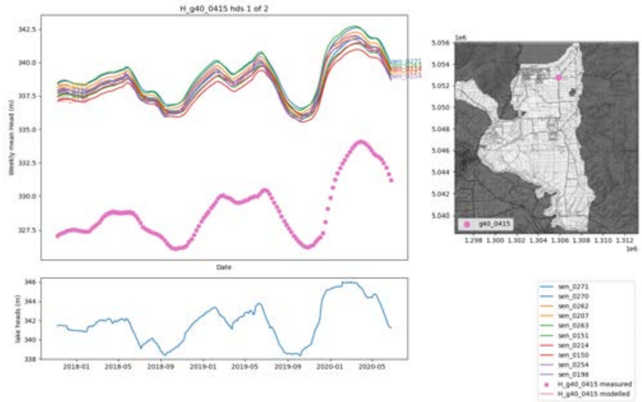


Figure: the results of the 1 layer model which fit the shape of the groundwater levels, but not the mean

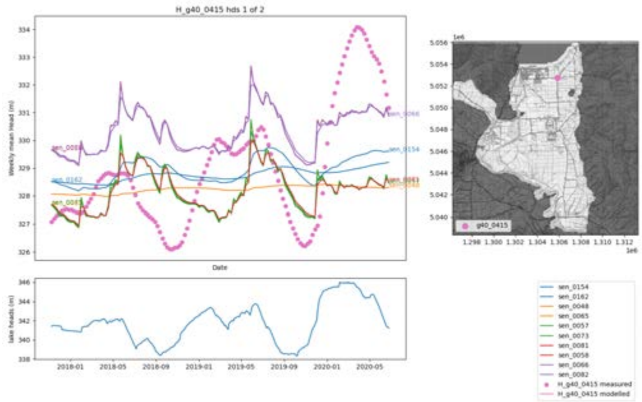


Figure: the results of the 1 layer model which fit the mean of the groundwater levels, but not the shape

We do not have other high frequency observations of groundwater levels near the Lake Hawea moraine. However, we do have a number of static water levels that were measured shortly after drilling the bore. These water levels, relative to the lake level at the time of measurement are shown in the figure below. This figure shows that a constant vertical offset between the groundwater levels and the lake levels of approximately 10 m. Some of these boreholes are located in the moraine less than 200 m from the lake. Many of the water supply wells near the lake (within the mapped moraine) are relatively deep (e.g. 50+ m).

Figure: all groundwater levels relative to the Lake Hawea level on the sampling date (positive values are groundwater levels below the lake)

Because we could not reproduce the groundwater levels at G40/0415, we deemed that the we could reject the hypothesis that a 1 layer model could reproduce the groundwater levels the Hawea system with confidence. This is an essential outcome from the Hawea groundwater model as the complex three dimensional structure has a key implications for the management of the groundwater system; the groundwater system is likely to be either disconnected or have other non-linear responses to Lake Hawea level if the lake falls below a threshold value.

Terrace only model (2D) optimisation results

The terrace only model had 1 target to fit, G40/0366. The model was unable to fit this target. The heads are higher than the measured data, and theape of the curve does not match the observed data. The full model was able to fit this target significantly better. From these data we can reject the hypothesis that the terrace only model can reproduce the observed groundwater levels in G40/0366. This suggests that there is indeed some connection between the High Terrace aquifer and the Hawea Flat aquifer.

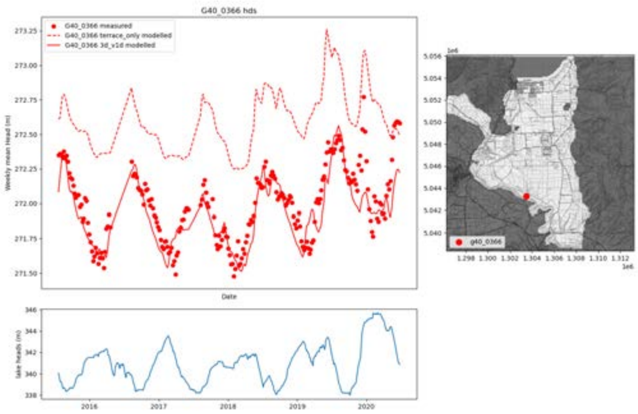


Figure: Comparison of the terrace only model and the 3d_v1a model at the south end of the High Terrace

3 layer model (3D) optimisation results

We produced 3 final 3 layer models. The differences (relative to the final model 3D_v1d) are listed below.

Model | "bund_top"*	NGMP wells included in objective function
3d_v1a | 335m msl	yes
3d_v1b | 333m msl	yes
3d_v1d | 335m msl	no

*the "bund_top" is the elevation of the top of the low conductivity layer in the moraine zone, which is also the threshold value for the non-linear response of the groundwater system to the lake level.

Final observation weightings:

Parameter group	Weighting
'rwh_hf'	0
'rwh_hf_riv'	0
'h_hf'	150
'h_hf_riv'	50
'h_lf'	{0, 10} {v1d, (v1a, v1b)}
'riv'	1e-3
'h_piezo'	10
'h_single_1'	5

'h_single_3'	5

3D_v1d (final model)

The model 3d_v1d is the final model that we used for all of the scenarios. This model did an excellent job reproducing the groundwater levels in our high frequency monitoring points across all of the historical data. The figures for these high frequency observation and a discussion of the results are provided in the *Comparison of 3d_v1a, 3d_v1b, and 3d_v1d* section below. The full set of optimisation plots for this model are available in the 3d_v1d optimisation results plots folder.

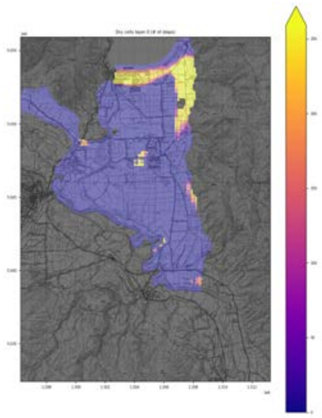


Figure: the number of model stress periods with dry cells in the final model

There are a number of areas in the model with consistent dry cells. There are a number dry cells directly south of the moraine. These are not a concern and are instead simply an artefact of the complex 3D structure in the area. Many more of these persistent dry cells are relatively isolated and occur in areas of steep topographical gradients (e.g. near the Clutha River, Camp Hill Moraine, or just adjacent to the Grandview Ridge). These cells are likely caused by structural error in topographical data and are of little concern.

More concerning are the dry cells in and around the Hawea Flat township. These dry cells are likely caused by the relative thinning of the model in this area, local abstraction and parameter structural errors, and/or a missing structure in the model. It is quite possible that there is a lower conductivity layer to the west of the Hawea Flat township from the Q4 Albert Town Advance. There was not enough information to justify adding this structure to the model, but it may warrant further investigation.

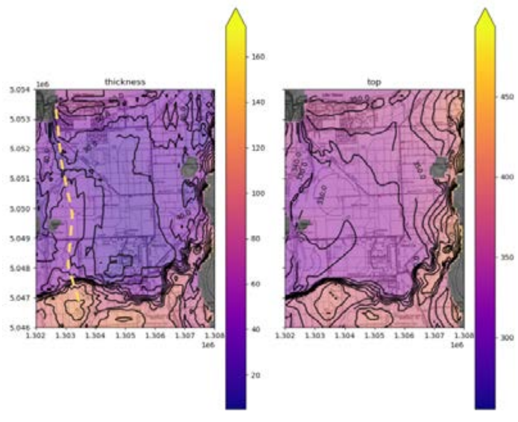


Figure: the possible indicative location (yellow dashed line) of a moraine structure to the east of the Hawea Flat township

Finally the model does a very poor job of reproducing the groundwater levels in the area to the east of the inferred Grandview Fault. There is very limited information in this area (3 single observations) most of which are near boundary conditions (e.g. Grandview Creek), so it is difficult to draw any conclusions about the cause of the poor model performance in this area. Instead, we have to accept that the model is not suitable in this area and the model results should not be used.

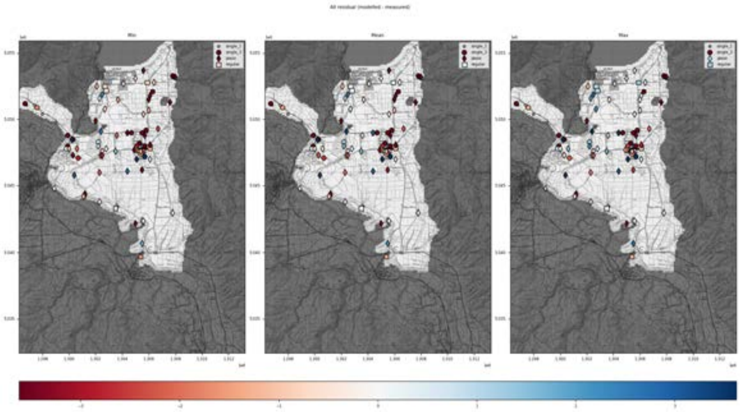


Figure: the groundwater levels residuals plotted spatially for the full model domain. Note that where a target has multiple temporal observations the min, mean and max residuals are shown. The color bar units are m

In general the model does a good job of replicating the groundwater levels across the model domain. There are some areas where the model over or under predicts the groundwater levels. As discussed above the areas to the east of the inferred Grandview Fault are significant under estimates and the model does not preform well in these areas. There are multiple targets in and around the Hawea flat township which are underestimated by the model, but given the close proximity of the high frequency observations at bore G40/0367, we believe that these misfits are most likely due to either poor data quality of the targets or problems arising from applying the historical measured water levels to the time period within the optimisation period. The latter is the most likely as there is a significant amount of abstraction in and around the Hawea Flat Township that may not have been present when the historical water levels were measured.

The two figures below show the modelled groundwater-surface water interaction for the Hawea and Clutha Rivers. While there are some target misfits, the model does a very good job of reproducing the expected interaction between the groundwater and surface water systems. The misfits occur in areas which both lose and gain water across the model period. The model does a good job reproducing the expected behaviour (gain/loss), but underestimates the total losses relative to the measured data. As discussed in the model target readme there is significant uncertainty in the measured gauging, therefore, we feel that the model is performing well in this area. In addition, the model does an excellent job of reproducing the expert judgment of the surface water and groundwater interaction. The Hawea is gaining below the dam to approximately Camp Hill, and then loses a significant amount of water between Camp Hill and sharp westward bend.

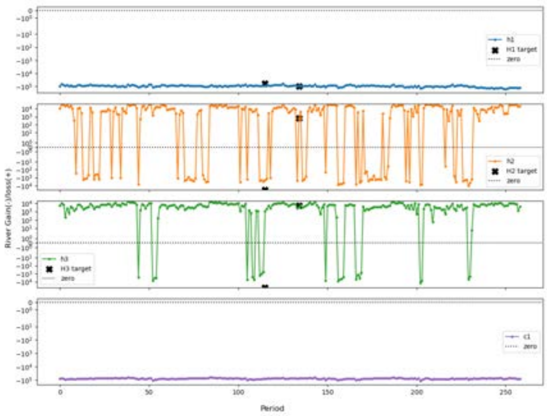


Figure: the river fluxes for the final model at each of the parameter zones

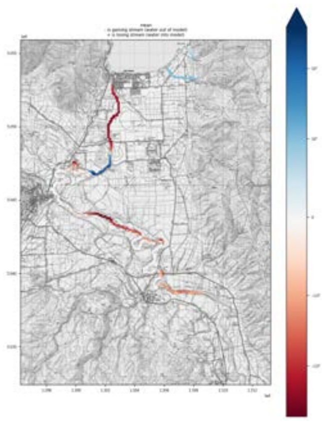


Figure: the mean river fluxes spatially for the final model

The figure below shows the steady state groundwater heads around the moraine zone in all three layers. We don't have any targets to inform this data, but it does produce a key prediction that could be tested in the future. The model predicts that the groundwater levels are significantly higher in the northeastern edge of the moraine and that this is the area which ultimately controls flow between the lake and the groundwater system. This is consistent with either the perched aquifer conceptual model or the local penetration conceptual model (see the Lake Hawea Moraine Conceptual Model section of the model build readme).

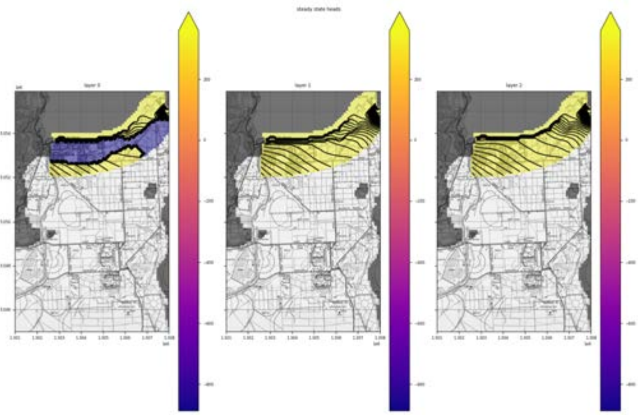


Figure: the steady state groundwater heads around the moraine zone in all three layers

3D_v1a

We will not independently discuss the results for 3d_v1a here, but the full set of optimisation plots for this model are available in the 3d_v1a optimisation results plots folder. A discussion of the differences between the three models is provided below.

3D_v1b

We will not independently discuss the results for 3d_v1a here, but the full set of optimisation plots for this model are available in the 3d_v1b optimisation results plots folder. A discussion of the differences between the three models is provided below.

Comparison of 3d_v1a, 3d_v1b, and 3d_v1d

The figures below show the results for all three of the 3D models for the high frequency groundwater levels.

Fit to higher frequency groundwater levels at G40/0415

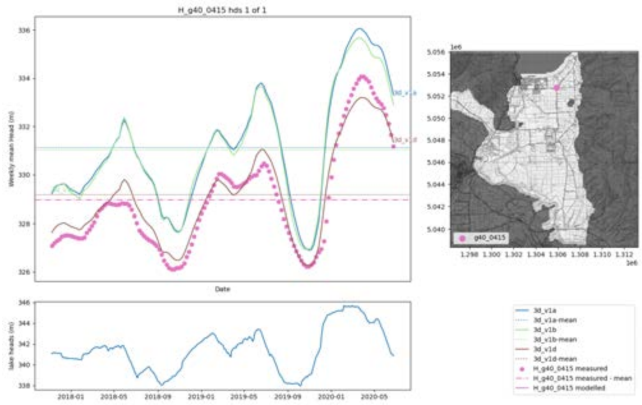


Figure: the groundwater levels for the 3d_v1a/b/d models at bore G40/0415

Fit to higher frequency groundwater levels at G40/0416

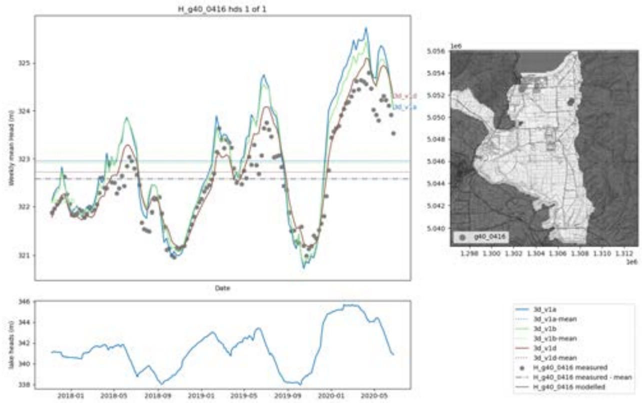


Figure: the groundwater levels for the 3d_v1a/b/d models at bore G40/0416

Fit to higher frequency groundwater levels at G40/0041

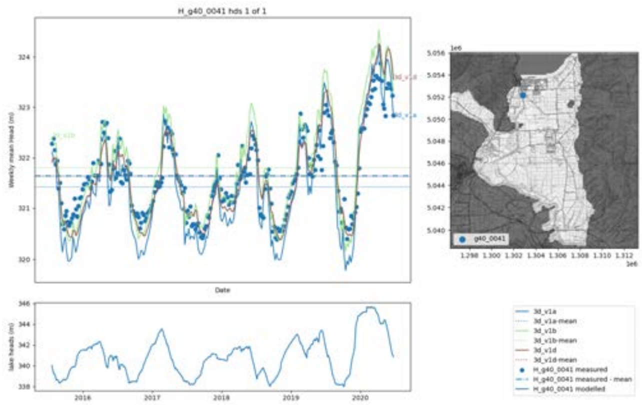


Figure: the groundwater levels for the 3d_v1a/b/d models at bore G40/0041

Fit to higher frequency groundwater levels at G40/0129

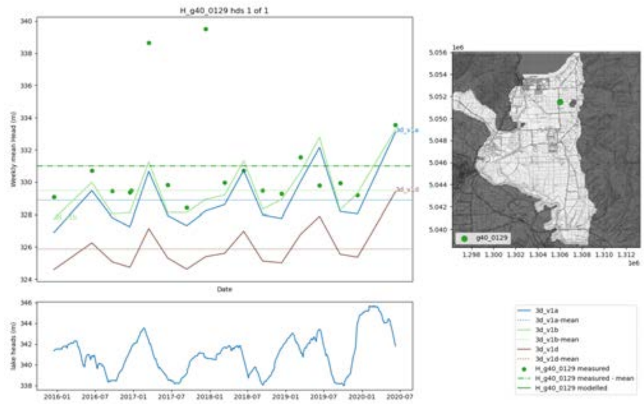


Figure: the groundwater levels for the 3d_v1a/b/d models at bore G40/0129 (NGMP bore)

Fit to higher frequency groundwater levels at G40/0120

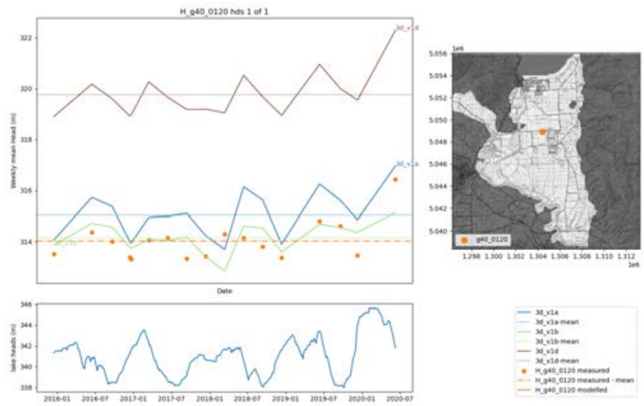


Figure: the groundwater levels for the 3d_v1a/b/d models at bore G40/0120 (NGMP bore)

Fit to higher frequency groundwater levels at G40/0367

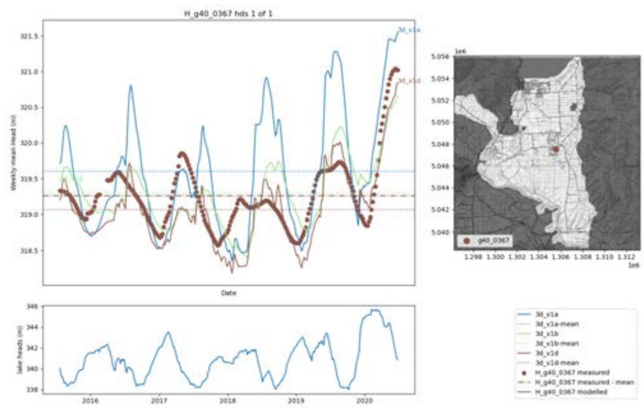


Figure: the groundwater levels for the 3d_v1a/b/d models at bore G40/0367

Fit to higher frequency groundwater levels at G40/0366

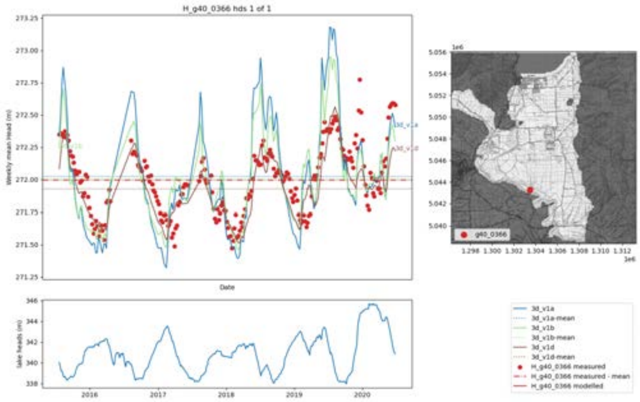


Figure: the groundwater levels for the 3d_v1a/b/d models at bore G40/0366

Discussion and implications

Model versions 3D_v1a, 3D_v1b, and 3D_v1d all do an adequate job of replicating the high frequency groundwater level observations. 3D_v1d does a slightly better job of replicating the high frequency groundwater levels than 3D_v1a and 3D_v1b, but it does this at the expense of the moderate frequency groundwater observations (i.e. the NGMP wells G40/0129 & G40/0120). This prioritisation was intentional as the NGMP bores are designed to monitor contaminants rather than static water levels. These bores are often pumped irrigation bores and structural error in the pumping is likely to impact the observation fits. Therefore we suggest that the 3D_v1d model is the best model for predicting the likely impacts of alternative management conditions.

Model versions 3D_v1a and 3D_v1b produce very similar results. The key outcome of the similar results from the 3d_v1a and 3d_v1b models is that the current observations do not constrain the threshold value, below which the groundwater levels exhibits a non-linear response to the lake level. We discuss plausible ranges for the threshold value in the Lake Hawea Moraine Conceptual Model section of the model build readme, but we are not able to further constrain this range. More discussions of the impacts of this threshold are discussed in the scenarios readme file.. Note that we did attempt to model the threshold value as 337 m msl, but this optimisation did not converge. However, we did not spend a significant amount of resources trying to get this optimisation to converge, so we do not believe that the lack of convergence here indicates that the threshold value cannot be as high as 337 m msl.

Steady State Model Water Budget (3D_v1d)

The table below provides the steady state water budget for the 3D_v1d model. Note that breaking the model boundary conditions budget down to the individual components (e.g. hillside inflow vs abstraction) components which introduces some dependencies due to double counting (e.g. where multiple boundary conditions exist in a single cell). This discrepancy is c. 1700 m^3/d; the discrepancy for the model (e.g. all_well/all_riv) is c. 0.3 m^3/d.

Boundary Condition	Steady State Flux
Lake	110384.9
Recharge	36216.2
Hawea1_flux	-96472.1
Hawea2_flux	10649.3
Hawea3_flux	5536.4
Clutha1_flux	-78458.9
Grandview_flux	3100.0

John_flux	1500.0
all River	-154145.3
Race_flux	4605.7
Abstraction_flux	-7365.3
hill_maungawera_flux	850.2
hill_flat_west_flux	259.8
hill_flat_east_flux	1537.8
hill_terrace_east_flux	3357.6
hill_south_east_flux	6083.2
all well	7544.0
discrepancy	-0.3

Access to final optimised parameter sets and models

The final optimised parameter sets for the 3d_v1a, 3d_v1b, and 3d_v1d models are available in this repository in the optimised_parameter_sets directory and are accessible via Python by the model_parameterisation.optimised_parameterisation.get_3d_v1{a|b|d}_params method. The final models are available in the final_opt_models directory. Due to the limit on file sizes that Github implements, the final models have been compressed and some files have been split into multiple parts with the 7zip library. To uncompress these models you must use the optimisation.final_opt_models.compress_uncompress_model.uncompress_model function.

To uncompress the 3d_v1d model to your downloads folder can use the following code:

```
from optimisation.final_opt_models.compress_uncompress_model import
from project_base import proj_root
from pathlib
# proj_root is the path to the root of the repo
# path to the model in the repo, you can substitute an absolute path
compressed_path = proj_root.joinpath('optimisation/final_opt_models/3d_v1d')
# path to save the uncompressed model to (currently set to 3d_v1d in your downloads folder)
out_path = Path.home().joinpath('Downloads', '3d_v1d')
uncompress_model(compressed_path, out_path)
```

Model limitations

There are a number of limitations to this model and the model optimisation. The main limitations are:

- **A non-unique model structure**: Because the complex structure in the moraine zone is not well constrained by the data we have assumed a very simple model structure that almost certainly introduces structural error. It is likely that some of the parameters in the model are compensating for this model structural error, which may have flow on effects, particularly for scenarios that are well outside the model optimisation conditions.
- **A non-unique parameterisation**: the PEST optimisation process is a poorly posed problem (that is there is not enough observations to calculate a unique solution to the model parameters). This means that there are multiple solutions to the model parameters that represent a good fit to the observations. This model has not undergone a parameter uncertainty process so we cannot predict the likely range or implications of the parameter uncertainty. In addition, the uncertainty of the model parameters is compounded by the uncertainty in the model structure. There are likely many other model structure/parameter sets that would fit the observations as well as this model.

- **Area to the east of grandview fault**: The model has persistent dry cells to the east of the inferred location of the Grandview fault. This is likely due a combination of model structural and parameterisation errors. The results from this area should not be used and there is currently insufficient data to produce a trustworthy model in this area.
- **Limited data for hillside streams**: The hillside streams are a major source of water to the Hawea aquifer systems, but we only have observations from two gauging sites for a limited period of time. We estimated the inflows from a a correlation with the Lindis River; however, both of the gauged streams are of similar size so the correlation may not hold for the smaller hillside streams.
- **Piezometric survey date**: the only piezometric survey conducted for the Hawea aquifer system was conducted in Sept of 2011. This survey was conducted outside our optimisation period (we did not have adequate abstraction information to conduct the model optimisation during 2011). A significant portion of the model domain therefore does not have any observations taken during the optimisation period, which adds another source of uncertainty to these areas of the model.
- **Parameterisation near Butterfield Reserve**: Butterfield Reserve is a sensitive wetland in an old oxbow of the Hawea River. The final parameterisation in this area is for a very low hydraulic conductivity which likely underestimates the degree of connectivity between Butterfield Reserve and the Hawea River. Model results here should be used with caution.

Recommended additional data

Over the course of our optimisation process we identified a number of additional data sources that would be useful. We have listed them here, with a discussion of why the information would be useful, but we have not included any feasibility assessments or costings to acquire these data sources. We recognise that some of these data sources may not be feasible but we have included them here so that decision makers can consider their relative value. We have not ranked these additional data sources in any way as any prioritisation is an intersection of priories (which we cannot address) and scientific merit.

- **A high frequency groundwater record near the Northeast Corner of the Hawea Flat aquifer:** One of the key model predictions for the complex moraine structure is that groundwater levels (impacted by the lake) should be elevated in and around the Grandview/John Creek alluvial fans. Testing this prediction would require a high frequency groundwater level record in this area of at least a couple of years in length. The exact location of such a bore would need more detailed consideration.
- **A high frequency groundwater record near the exit of the Maungawera Valley:** The Maungawera Valley has a relative paucity of data which makes predictions regarding the sustainable use of groundwater uncertain. A high frequency monitoring bore near the exit of the valley (e.g. up valley of the Maungawera Valley Road and Lake Hawea Albert Town Road intersection) would act as an integrator for the up valley groundwater system and would provide significantly more information about the local groundwater system. The exact location of such a bore would need more detailed consideration.
- **A high frequency bore near the Hawea domain and/or Butterfield Road**: Water from Lake Hawea can flow either toward Hawea Flat township or it can flow back towards the Hawea River. Understanding the piezometric surface in the aforementioned area would help constrain that flow. The exact location of such a bore would need more detailed consideration.
- **A detailed investigation of moraine structure**: As mentioned multiple times within this repository, the moraine structure is not well constrained by the data. A detailed investigation of the moraine structure would help constrain the model structure and reduce the uncertainty in the model predictions. The method of investigation would need significant consideration and would likely require a combination of geophysical and drilling investigations.
- **An investigation of structure to West of Hawea flat Township**: As discussed above, glacial geomorphology suggests that there could be a low permeability structure to the west of Hawea Flat township associated with a potential lateral moraine of the Albert Town advance. Further investigation of this possible structure would help constrain our understanding of the groundwater system in this area.

- **multiple concurrent gauging of multiple hillside streams:** As described in the model build readme the hillside streams are a major source of water to the Hawea aquifer system. However, we only have observations from two gauging sites for a limited period of time. We estimated the inflows from a a correlation with the Lindis River; however, both of the gauged streams are of similar size so the correlation may not hold for the smaller hillside streams. Multiple concurrent gaugings (at high and low flows) of multiple hillside streams (both large and small catchment areas would help constrain the predictions of inflows from the hillside streams.
- **Additional Piezometric surveys:** At present the only piezometric survey completed in the Hawea region was conducted in Sept of 2011. This survey was conducted we had adequate abstraction information and so we had to transpose these groundwater head targets to dates inside the optimisation period, which adds error. One or more additional piezometric surveys (e.g. at high and low water levels) would help constrain the model predictions in the areas that are only informed by the piezometric survey.
- **Model parameter / structural uncertainty analysis**: In the absence of additional data collection more information about the uncertainty of the model predictions could be obtained by conducting a parameter / structural uncertainty analysis. Given the significant structural uncertainty in the moraine zone we would recommend calibrating and conducting parameter uncertainty analysis on many different model structures. This would be a significant undertaking, but could easily build on the work and data analysis from this project and contained within this repo.

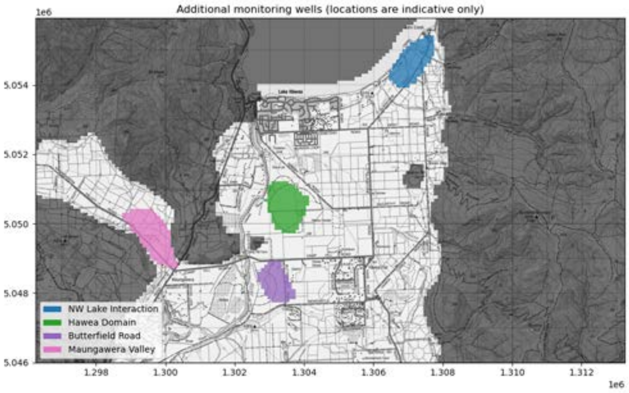


Figure: Areas for possible additional high frequency monitoring bores

Optimisation working notes

These working notes are largely verbatim except that the active branches are elevated above the abandoned branches. The number indicates the order in which the branches were developed.

Active Branches (optimisation versions)

25. Main (3d_v1d)

- The 'final' optimised model
- Contains 3D structure around the Lake Hawea Moraine
- Best fits for the high frequency targets
- Bund elevation set to 335 msl
- NGMP well head observations removed from objective function as there is significant tension between these records and the high frequency observations. The NGMP wells are pumped irrigation bores and the primary purpose for sampling was water quality monitoring

20. 3d_v1a

- Identical to “Main (3d_v1d)” except that the NGMP wells were included in the objective function
- decent history matching; however “Main (3d_v1d)” provides better results
- retained as active branch for comparison to “3d_v1b”

21. 3d_v1b

- Identical to “3d_v1a” except that the bund elevation was set to 333 MSL.
- history matching results were similar to “3d_v1a” suggesting that the bund elevation is largely non-unique
- retained to demonstrate the non-uniqueness of the 3D structure

14. terrace_only

- This model structure only includes the High Terrace (south of Hawea Flat) to the Clutha River
- this optimisation was undertaken to see if the High Terrace could be history matched (within the accepted parameter ranges) in isolation from the rest of the Hawea aquifer system.
- History matching was not achieved

Abandoned branches (releases, optimisation versions)

There are many previous branches that were issued as pre releases and then deleted (effectively archived). There should be no reason for other users to delve into these previous branches as they ended up with unsatisfactory history matching; however they are available and briefly described below (working notes) for completeness.

1. Main (before 2/11/22) The main build branch. First structural version
2. Structure v2, Changes:
    - Increase parameterization via pilot points to Maungawera
    - Add recharge multiplier pilot points across model (NI)
    - Remove sandy point from model
    - abandoned but retained
3. Structure v3,Changes:
    - Set ss=sy
    - Set the model to confined to reduce computational burden
    - This helped but the model preformed poorly,
    - Error did not reduce saturated thickness.
    - abandoned and deleted
4. Structure_v4:
    - From structure v2
    - Add new mean annual head targets from regular
    - Increase steps to 7 in transient
    - Expand hillside streams to all adjacent cells (up to 9 cells per hill)
    - Optimisation never run here, just saved to version structural changes

5. Structure_v5
    - From structure_v4
    - Remove near river pumping wells.
6. Structure_v6
    - From structure 5
    - Add a 1 m confined layer below the bottom of layer 1 (may improve stability)
7. Structure_v6a
    - From v6, but set ss to sy
8. Structure_v7 (built but not run)
    - From structure 5
    - Reduce thickness to reasonable pumped thickness and then maximum 30 m sat thickness
    - Set ss = sy
    - run as a confined model
9. Structure_v8
    - From structure_v6a
    - increase initial conductivity (to 50, 100 and 70 was too unstable)
    - rch multiplier only by irrigated not irrigated bounds of multiplier 0.5-1.2
10. Structure_v9
    - Fix river targets (they were backwards!)
    - Implement Grandview and John Creek (+Hawea and Clutha) as str package
    - Lake stage vs g40_0415
    - Looks fine, honestly the fact that them model isn't matching it suggests some sort of structural error. Reworked transport in grandview stream?/ water through grandview stream??? Likely the problem google maps shows water in grandview to the lake (and in john creek (to the north), all other creeks are probably fine.
    - Lower basement around g40_0366

11. Structure_v10
    - Set weight of regular year targets to 0
    - set each of the ‘h_hf’ targets equal weights despite different data lengths
    - look/lower basement in dry cells near model boundaries
    - NE hillside area (done)
    - Near Clutha River (done)
    - I think I need some more pilot points
    - Near pt 402 on camp hill moraine (move Maungawera south?) () and another in the moraine (to interpolate with other river group
    - To stop dry cells south of camp hill moraine
    - Significant number in the hillslope area just off the bounds to allow conductivity to fall there if needed for stability. And to manage the change in geologic setting near hillslope
    - Adjust some locations based on the new pilot point locations
    - New rivergroup south of Maungawera valley entrance to allow for the difference between the two settings
    - Additional point in the middle of the terrace to manage near hillside environment.
    - Try lowering hillside conductance → set to 100 vs 1000 for Hawea/Clutha, which means much of the peak flow does does not make it into the model.
12. Structure_v11
    - Move to 1 global recharge modifier (done)
    - Much higher initial kh (lake=5, rest = 300) (in progress
    - Lower sy, and lower sy bounds
    - Change weights (lower low frequency targets)
    - Bit of a hail mary before the weekend
    - retired (even though I’m happy with the parameterisation. If I want to change back to vll parameters do it from v12
13. Structure_v12
    - Increase kh/sy parameterisation in the near lake environment
14. terrace only
    - see above
15. p_lake
    - As per structure_v11 but with a single additive parameter for lake heads (e.g. lake hds = lake hds + mod
    - A test to see if the lake levels problems are sorted everything else works great?
    - Note the parameter is offset by 100 m as pyemu has bugs!
16. lake_bar
    - Add a 1 cell thick barrier for kh
    - Remove additional v12 parameterisation
17. cond_int
    - Try to fit the heads by simply setting lake conductance (1 cell width lake)

18. 3d_v1
    - Address the 3D moraine issues in structure
    - 3 layers the bottom two pinch out against the bottom of the model.
    - well management
    - target management
    - other structural pieces
    - Add abrupt parameter change at terrace interface
    - Remove from dam to “dam control” road from model (e.g. no flow)
    - Re-run pre_optimisation_overview.py
    - remove the slope fixer on the east side
    - remove additional parameterisation of v12
19. 3d_v2
    - As per v1 but fully confined (to increase stability)
    - Ss[0] = sy[0]
    - Initial parameters do not manage the drop quite so well. This may really need the unconfined aspects of the model.
    - Bit of a hail mary over xmas. Really need the unconfined action to make the ‘waterfall happen’
20. 3d_v1a
    - see above
21. 3d_v1b
    - see above
22. 3d_v1c
    - As 3d_v1a but with top of bund set to 337
    - great difficulty getting this to converge
    - abandoned
23. 3d_v4
    - As 3d_v1a, but top of bund is set to 340 m MSL instead of 335
    - Difficult to get model to converge
    - abandoned
24. 3d_v5
    - As 3d_v1a, but top of bund is parameterised
    - Largely unstable
25. 3d_v1d
    - see above

References

- Wilson, J., 2012. Hawea Basin Groundwater Review. prepared by Resource Science Unit of Otago Regional Council, June 2012, Dunedin.

Hawea Transient groundwater model (Hawea Model) scenarios methods and results

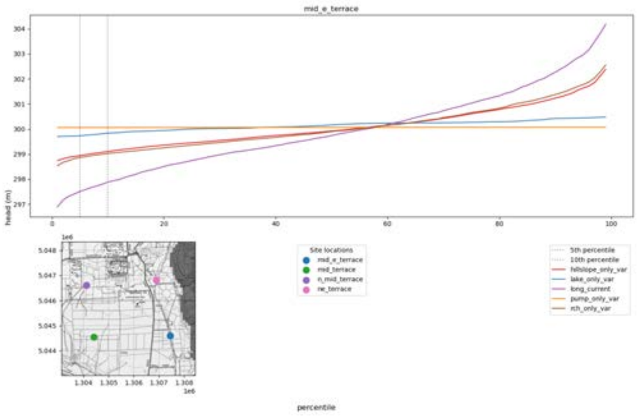


Figure: results of the boundary condition sensitivity analysis at the mid_e_terrace indicator well

Author:	Matt Dumont
Date:	2021-11-02
Version:	1.0.0
Status:	Final
KSL project:	Z22031HAW_hawea-model
Purpose:	This document provides the methodology and results for the model scenario process

Index

Table of Contents

Module Index

- README.rst: document describing the scenario modelling methods and results
- **Scenario development and supporting scripts**
  - scen_period.py: script to handle the scenario period
  - boundary_condition_plots: plots of the scenarios boundary conditions
  - base_data: base input data for the scenarios
  - processed_input_data: processed input data for the scenarios, these files were all developed from the base data
  - boundary_conditions.py: develop the input boundary conditions for the scenarios
  - supporting_data_analysis: additional data analysis scripts to support creating boundary conditions
  - scenario_outputs.py: script to make consistent scenario outputs
  - run_flow_scenario.py: script to run a flow scenario
  - run_scenario.py: script to run a scenario (in multiprocessing)
- **Model information and MT3D indicator modelling**
  - run_mt3d_scenario.py: script to support running MT3D
  - mt3d_indicator_scens.py: script to run MT3D indicator scenarios
  - compare_boundary_sensitivity.py: compare the results of the boundary condition sensitivity analysis
  - model_info_scenarios.py: script to run model information scenarios
  - **model_info_scen_results: model results and plots for model information scenarios**

        - 0_results: plots for model information scenarios
        - {scenario name}: Model results for model information scenarios: input and output data for the scenario
    - mt3d_indicator_scenarios: model results and plots for the MT3D scenarios
- **Low Lake Hawea level scenarios**
    - low_lake_scenario_data.py: script to develop typological lake levels and perturbations
    - low_lake_scenarios.py: script to run low lake scenarios
    - compare_low_lake.py: script to compare low lake scenarios
    - **low_lake_scenarios: model results and plots for low lake scenarios**
        - 0_results: plots for low lake scenarios
        - {scenario name}: Model results for low lake scenarios: input and output data for the lake scenario
- **Allocation modelling**
    - allocation_zones.py: get and plot allocation zones
    - allo_rch_hillside.py: scripts to get and compare the allocation, hillside recharge, and LSR for each zone
    - allocation_scenarios.py: script to develop all allocation scenarios and to run the non-gridded allocation scenarios
    - run_grid_allocation.py: script to run the gridded allocation scenarios
    - compare_allocation_scens.py: script to compare allocation scenarios
    - allocation_scenarios: model results for allocation scenarios
    - **allocation_results: plots of allocation results**
        - old_allo_zones.png: figure of the old allocation zones (Wilson et al., 2012)
        - new_allo_zones.png: figure of the new allocation zones
        - Hawea Flat_results: results for the gridded Hawea Flat allocation scenarios
        - Maungawera Flat_results: results for the gridded Maungawera Flat allocation scenarios
        - Terrace-Hill_results: results for the gridded Terrace-Hill allocation scenarios
        - nat_current_full: results for the naturalised, current allocation, and full allocation scenarios
        - Te Awa_results: results for the gridded Te Awa allocation scenarios
        - Terrace-River_results: results for the gridded Terrace-River allocation scenarios
        - mangawera_valley: results for the Maungawera Valley allocation reduction scenarios
        - allo_zone_rch: results comparing LSR, hillside inflows, and allocation for each zone
        - example_quantile_plots: example quantile plots for the allocation scenarios to support presentations
- **Wetland Setback Modelling**

- wetland_setback_campbells: wetland setback modelling for Campbells Wetland scripts and results
- wetland_setback_butterfield: wetland setback modelling for Butterfield Wetland scripts and results

Scenarios Overview

Scenario Type	Key Questions
Model information Scenarios	How does the model behave, and what impacts groundwater levels?
Low lake scenarios	What happens if the management of Lake Hawea changes significantly?
MT3D indicator scenarios	Where is the water sourced from?
Allocation Scenarios	What is a sustainable level of abstraction?
Wetland setback scenarios	Where, and to what extent, does abstraction impact significant wetlands?

Boundary Conditions Overview

The table below has an overview of the different possible boundary conditions for the model Scenarios. Some of these boundary conditions were defined in the model_build readme and other are defined more fully below. Note that we also used static recharge, hill inflows, lake levels, and river flows for the scenarios. These were simply the steady state component of each boundary condition (i.e. the mean value).

Boundary condition name	Package	Water component	Overview	Reference
dryland_rch	Rch	LSR	scenario period ERA5 dry-land recharge	../model_build readme
irr_rch	Rch	LSR	scenario period ERA5 irrigated recharge	../model_build readme
hist_rch	Rch	LSR	opt period met recharge	../model_build readme
hist_era5_rch	Rch	LSR	opt period ERA5 recharge	../model_build readme
large rivers	Str	Hawea and Clutha R.	ISO weekly mean river flows / stage	../model_build readme

hillside flows	Str/Wel	Hillside inflows	long record of hillside inflows	../model_build readme
race losses	Wel	race losses	ISO weekly mean race losses	../model_build readme
no_pump	Wel	GW abstraction	no abstraction	n/a
pump curve	Wel	GW abstraction	typological annual pumping curve (0-1)	below
static_pump	Wel	GW abstraction	steady state optimisation pumping	below
extended_pump	Wel	GW abstraction	ISO weekly mean pumping	below
extended_full_allo	Wel	GW abstraction	ISO weekly mean of maximum daily allocation normalised to historical pumping record	below
extended_max_allo	Wel	GW abstraction	maximum daily allocation applied to every day of the year	below
extended_max_allo_pc	Wel	GW abstraction	maximum daily allocation applied to pump curve	below
reduced abstraction	Wel	GW abstraction	allocation reduction (fraction of extended_pump)	below
grid_pump	Wel	GW abstraction	gridded abstraction (additional allocation)	below
lake	Ghb	Lake Hawea	long record of Lake Hawea levels	../model_build readme
low lake levels	Ghb	Lake Hawea	typological low Lake Hawea levels	below

Boundary conditions Methodology

The sections below provide additional documentation for the boundary condition options that have not been sufficiently described in the model build readme.

Groundwater abstraction

Development of the pumping curve

In order to apply additional levels of groundwater abstraction in a sensible way that is constant with the annual usage patterns we developed a typological pumping curve. This was developed by analysing the ISO weekly mean pumping data. The ISO weekly data still has some variation so we then applied a centered moving window mean of 9 weeks to smooth the data. The data was then transformed via min/max normalisation to a range of 0-1. For increased abstraction scenarios we could then apply a maximum daily take rate to the pumping curve to get the daily abstraction. For reference the integral of the pumping curve is c. 135 suggesting that on average the annual usage is 135 times the mean annual maximum daily take.

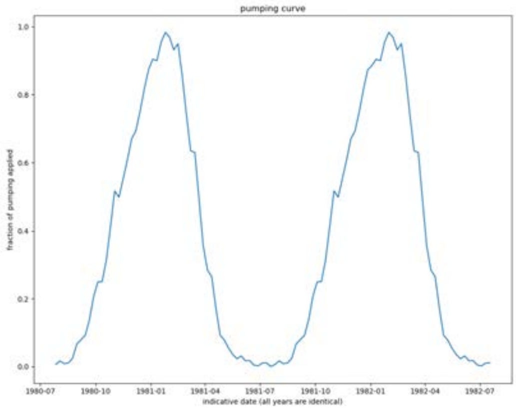


Figure: typological pumping curve

extended_pump

We produced the extended_pump time series to apply the current level of abstraction to the full scenario period. This was done by simply repeating the ISO weekly mean historical pumping data to all weeks of the scenario period.

extended_full_allo

The extended_full_allo pumping record is maximum daily allocation applied to the min/max normalised historical pumping record (2015-2020). This likely underestimates the full allocation scenario as most users do not regularly take their maximum daily usage every year.

extended_max_allo

The extended_max_allo pumping record is maximum daily allocation applied to every day of the year. This is almost certainly not attainable for 2 reasons. 1) no water users are likely to use their allocation every day of the year (e.g. irrigators do not irrigate in winter) and 2) many consents have maximum annual allocations as well as maximum daily allocations; however, analysis of these data were not included in the usage analysis preformed by Kitteridge (2022).

extended_max_allo_pc

The extended_max_allo_pc pumping record is maximum daily allocation applied to the typological pumping curve. This is likely the most realistic scenario of the level of abstraction that could be achieved with the current consented activities.

Additional abstraction (grid_pump)

One of our goals was to explore the impacts of additional water allocation. Introducing new allocation is challenging because the spatial locations of these abstractions are unknowable. To address this we developed a 500 m grid of abstraction points for each of the allocation zones. To assess the impact of additional abstraction we then applied an additional maximum daily take to the typological pumping curve and then evenly distributed this abstraction across the grid points within the allocation zone. This approach will underestimate the local impacts of abstraction, but it should represent the impacts to the allocation zone as a whole. We anticipate that other approaches (e.g. well interference and/or stream depletion assessments) would be used to limit the local impacts of any additional consented abstraction.

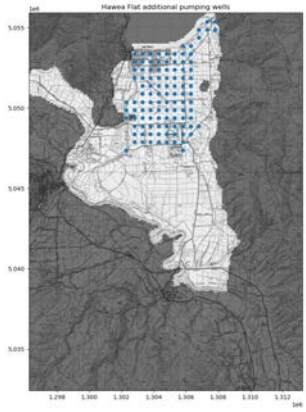


Figure: grid of abstraction points for the Hawea Flat allocation zone

Reduced abstraction (reduced allocation scenarios)

Where we needed to assess the impacts of reduced abstraction (either from the current usage or one of the aforementioned pumping records), we simply applied a percentage reduction to each Wel boundary condition.

Plots of the abstractions

It is beyond the scope of this document to provide and interpret all of the plots of the boundary conditions. However, these plots are all available in the following directories:

- Scenarios/boundary_condition_plots/pumping: full record of the pumping records.
- Scenarios/boundary_condition_plots/pumping_use_allo_difs: comparisons of the different pumping records.

Lake Hawea Levels for low lake scenarios

In order to assess the impacts of heretofore unseen lake levels we developed a typological annual lake level variation. This was done by creating the best fit between the ISO weekly mean lake levels to a modified sin wave function:

$$l = a * \sin((t - d) / 52 * 2 \pi) + b$$

where:

- l is the lake level
- a is the amplitude of the lake level variation
- t is the ISO week
- d is the phase shift of the lake level variation
- b is the mean lake level

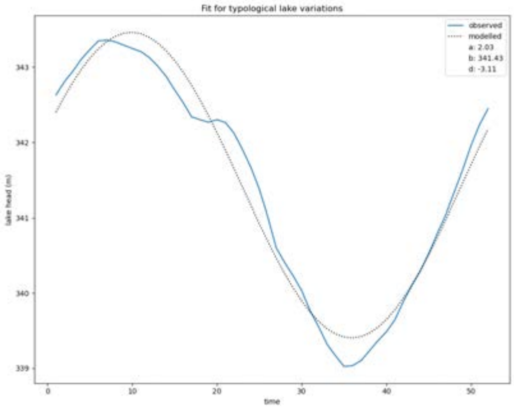


Figure: The fit of the typological lake levels to ISO weekly lake levels for the low lake scenarios

We then used these parameters as a base value for the following equation to perturb the historical lake levels:

Where the lake levels were greater than the annual mean lake level:

$$lh = a_{high} (| sin((t - d) / 52 * 2pi) |)^{k_{high}} + b_{high}$$

Where the lake levels were less than the annual mean lake level:

$$ll = -a_{low} (| sin((t - d) / 52 * 2pi) |)^{k_{low}} + b_{high}$$

Where:

- lh is the high lake level
- ll is the low lake level
- a_{high} is the amplitude of the high lake level variation
- a_{low} is the amplitude of the low lake level variation
- b_{high} is the mean lake level for the high lake level variation
- b_{low} is the mean lake level for the low lake level variation
- k_{high} is the width parameter of the high lake level variation
- k_{low} is the width parameter of the low lake level variation
- t is the ISO week
- d is the phase shift of the lake level variation (not modified)

This allows us to modify the lake amplitude, mean, and the width of the sin wave for both the high and low lake levels.

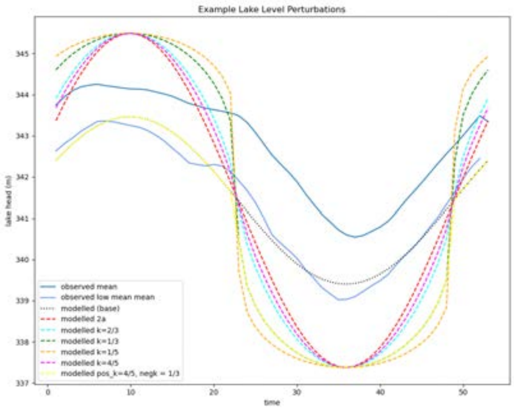


Figure: example of the perturbations to the lake levels for the low lake scenarios

Standard Scenario Outputs

Indicator monitoring points

The Hawea model is a 4 dimensional model (3 spatial dimensions and 1 temporal dimension). In order to visualise the impacts of scenarios we have developed a series of indicator monitoring points. These points are spaced through the model domain and were developed by expert judgement.

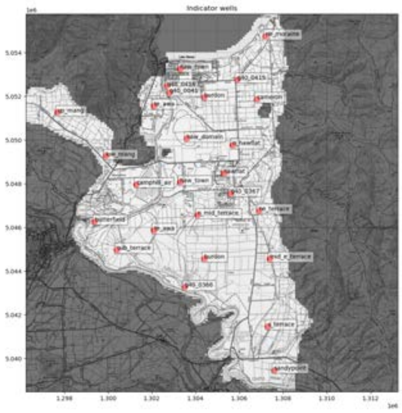


Figure: The indicator monitoring points for the Hawea model

Data outputs

We have developed standard data outputs for the scenarios, they are as follows:

- all_well_output_dataset.csv: model heads at every known well in the model domain.
- converged.txt: text file with a boolean value indicating whether the model converged.
- key_input_data.csv: a record of key input data including recharge, hill inflows, and lake levels.
- output_dataset.csv: the key output data for each time step which is comprised of: - model heads at the indicator monitoring points and at the high frequency monitoring bores. - river gain/losses for each conductance parameter zone. - key extracts from the zone budget.
- zone_budget.csv: the full zone budget.

In addition, some scenarios also contain the following files:

- {}.list: list file from the model runs.
- {}_hds: compressed and split model heads
- plots: plots of the model results (similar to the plots in the Standard optimisation outputs in the model_optimisation)

Figure outputs

The standard output for comparing multiple scenarios is a series of plots including:

- comp_plots: direct comparison of the model results for the full time period
- quantile_plots: comparison of the model quantile data
- qq_plots: comparison of the model quantile to a base model scenario's quantile data

Reading quantile plots

An example of a quantile plot is shown below. This plots the calculated percentile of the model heads at the indicator well. In the figure below the point at (10, 342.5) on the long current line indicates that 10% of the model heads for the long_current scenario at the upper Maungawera indicator well are less than or equal to 342.5 m msl. Because the long_current scenario is a reasonable sample of the historical record we can infer that if the future record is similar to the historical record (e.g. weather/climate) and the boundary conditions are similar to the long_current scenario then there is a 10% chance that the model heads at the upper Maungawera indicator well will be less than or equal to the modelled value (342.5 m msl). Note that this does not account for potential biases in the model results.

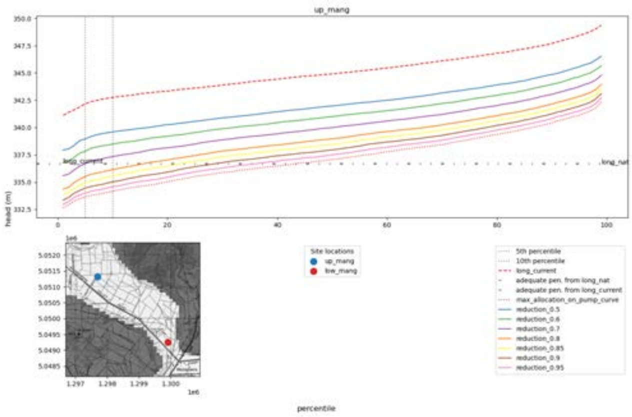


Figure: example of a quantile plot

Reading q-q plots

An example of a quantile-quantile plots is shown below. Note that the data is the same a the example quantile plot described in the *reading quantile plots* section above. This figure compares the quantile data from a base scenario (in this case the long_current scenario) to the quantile data from the scenario of interest. So the point at approximately (10, 60) on the reduction_0.5 line can be interpreted as the modelled 10th percentile heads for the long current scenario (342.5 m msl,from the quantile plot in the section above) would be the 60th percentile heads for the reduction_0.5 scenario. From this if we assume that the long_current scenario is a reasonable representation of the historical record then we can infer that if we transitioned to the reduction_0.5 scenario then we would expect that the current low levels that were experienced 10 percent of the time would now be experienced 60 percent of the time. This addresses the potential bias problems (e.g. if the model over/under estimates the groundwater levels) that are present in the quantile plots, but does make it harder to interpret.

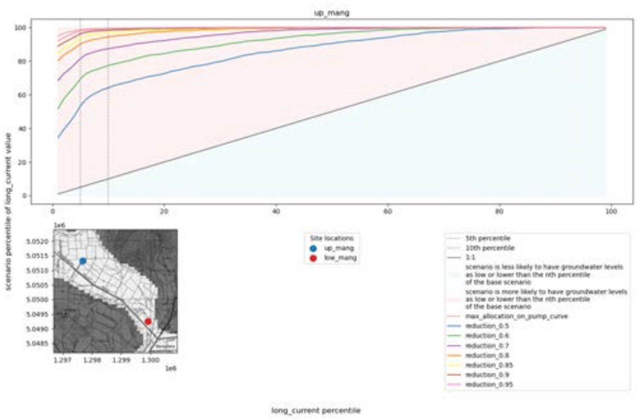


Figure: example of a quantile-quantile plot

Adequate penetration

In addition to the standard outputs we need to estimate an adequate penetration level for each indicator well in the model. The adequate penetration level is the level at which we would expect water supply wells to be screened to maintain reliability. That is if an individual owns a water supply well that is screened above the adequate penetration level then they would expect to have interruptions in their water supply, while if the well is screened below the adequate penetration level then they would expect to have a reliable water supply. The adequate penetration level becomes a key constraint for the model scenarios (e.g. if the model scenario results are lower than the adequate penetration level we would expect significant water supply issues). Adequate penetration levels can be set based on an analysis of current well screens and historical water levels; however, this approach is difficult in the Hawea Area as many bores do not have detailed well construction data. In addition this approach can cause challenges when there are biases in the model results (e.g. if the model over/under estimates the groundwater levels). To address these issues we have calculated the adequate penetration level for a given scenario based on the model results as follows:

$$apl = h_{mean} - 3\sigma$$

where:

- apl is the adequate penetration level
- h_{mean} is the mean of the model heads at the indicator well
- σ is the standard deviation of the model heads at the indicator well

For these models we typically compare results to the adequate penetration level for the long_current scenario and long_nat scenario. The long_current scenario is a reasonable representation of the historical record, while the long_nat scenario is a reasonable representation of the natural record (i.e. no pumping).

Scenario Methods & Results

The following sections describe the methods and results for each scenario set. Note that we typically will only include the results for an individual indicator well. A link to the location for the full sets of model results is provided in the section. We include some discussion and analysis of these results, but our primary discussion is in final report. Regardless, a full discussion of all the results is beyond the scope of this document and this project, but below we provide some example discussion of the results, so that these results could be used to address a future specific question.

Model information scenarios

Scenario Name	Purpose/comment
optimised	Optimised model results
long_current	Long scenario with long_current abstraction and irr_LSR
long_nat	Long current, but with dryland recharge, and no pumping (races left on)
no_pumping	Long current, but with no pumping
hillslope_only_var	What extent does the hillslope inflow variation influence total model variation
lake_only_var	What extent does the Lake Hawea level variation influence total model variation
pump_only_var	What extent does the groundwater abstraction influence total model variation
rch_only_var	What extent does the LSR variation influence total model variation
static_pumping	What variation exists with only pumping held static

Boundary condition sensitivity

To conduct the boundary condition sensitivity we compared the model scenarios where all but one boundary condition was held static to the long current scenario. We could then qualitatively assess the contribution of each boundary condition to the total model variation. An analytical approach is also possible, but we feel it would suffer from false precision, that is that while it would perfectly represent the contribution of each boundary condition to the total model variation the contribution to the experienced real world variation would likely have a significant and unspecified level of error. The results shown in the table below are the results of the qualitative assessment, which we believe should hold true in the real world.

Zone	Recharge	Hill inflows	Lake levels	Pumping
Hawea Flat	Moderate	Low	High	Low
Te Awa	Low	Low	High	N/A
Terrace-River	Moderate	Moderate	Moderate	Localised high
Terrace-Hill	High	High	Moderate	Low
Maungawera Flat	Low	Low	High	N/A
Maungawera Valley	High	High	N/A	Moderate
Sandy Point	High	Moderate	Low	N/A

The figure below provides an example of the results of the boundary condition sensitivity analysis. The full set of results are available at Scenarios/model_info_scen_results/0_results/boundary_sense. In the figure below the pump_only_var scenario has minimal variance across the quantiles relative to the long_current scenario. Conversely, the hillslope_only_var and rch_only_var scenario has a large amount of variance across the quantiles compared to the long_current scenario. This indicates that the hillslope inflows and LSR are a significant contributor to the total water level variation at this indicator well, while the pumping is a relatively minor contributor. Lake levels are minor contributor to the total water level variation at this indicator well, but are still noticeable, particularly at high and low groundwater levels (0-20th and 80-100th percentiles).

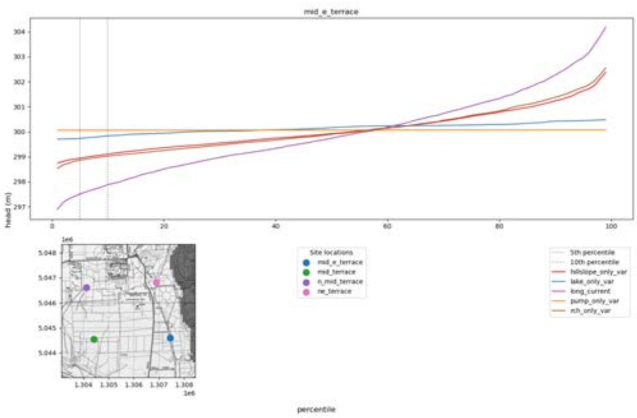


Figure: example results of the boundary condition sensitivity analysis at the mid_e_terrace indicator well

Naturalised vs current without abstraction

These scenarios compare the results of the long_current, long_nat, and no_pump scenarios. In the example figure below the naturalised scenario (long_nat) has a significantly lower mean water level than the no_pump scenario. This is likely because there is significant irrigation on the high terrace, which is not represented in the naturalised scenario. The long_current scenario has a mean water level between the long_nat and no_pump scenarios, which shows that while the current level of abstraction has a significant impact on the water levels, it does not reduce the water levels below the naturalised state. This is consistent with our understanding of the area as the irrigation on the High Terrace is primarily supplied by abstraction from the Hawea River, thus development in this area has shown a net increase in groundwater levels. The full set of results are available at Scenarios/model_info_scen_results/0_results/nat_long_nopump.

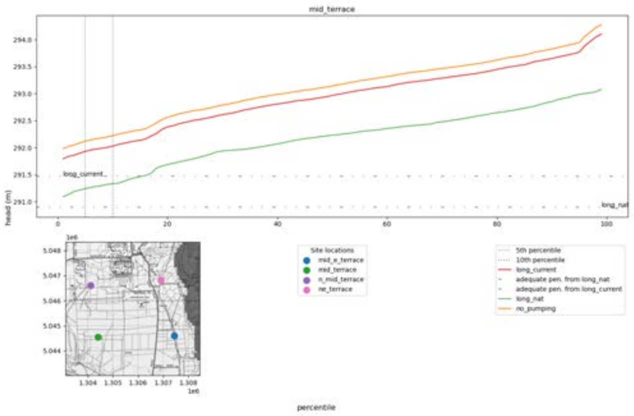


Figure: example results of the comparison of the naturalised, long current, and no abstraction scenarios at the mid_terrace indicator well

Naturalised vs current vs long current

These scenarios compare the results of the long_current, long_nat, and optimised scenarios. We do not include an example figure here as the results are very similar to the naturalised vs current without abstraction scenarios. The full set of results are available at Scenarios/model_info_scen_results/0_results/nat_long_opt.

Naturalised vs current vs long current (opt period only)

These scenarios compare the results of the long_current, long_nat, and optimised scenarios but only for the optimisation period. We do not include an example figure here as the results are very similar to the naturalised vs current without abstraction scenarios. The full set of results are available at Scenarios/model_info_scen_results/0_results/nat_opt_long_opt_per_only.

Low Lake Hawea Level Scenarios

The Low Lake Hawea Level scenarios are designed to test the sensitivity of the model to the Lake Hawea head boundary conditions and to test the impacts of the complex moraine structure specifically. These scenarios involve creating a synthetic lake level boundary conditions and then ascertaining how the model responds to these conditions. The figures below show all of the synthetic lake level scenarios that were tested.

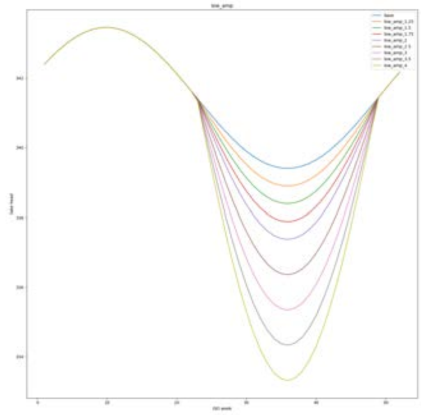


Figure: Lake Hawea head boundary conditions for the low_amp scenarios

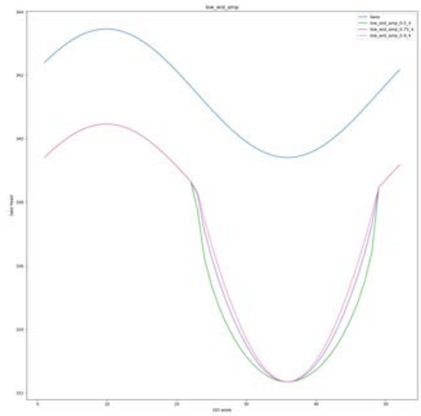


Figure: Lake Hawea head boundary conditions for the low_wid_amp scenarios

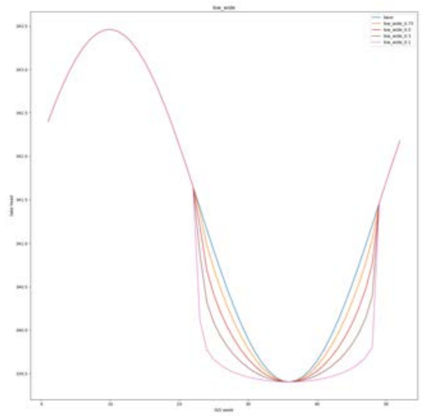


Figure: Lake Hawea head boundary conditions for the low_wide scenarios

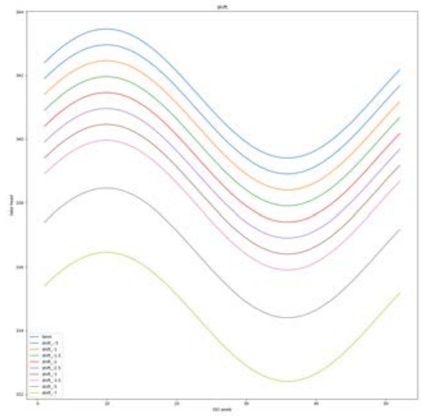


Figure: Lake Hawea head boundary conditions for the shift scenarios

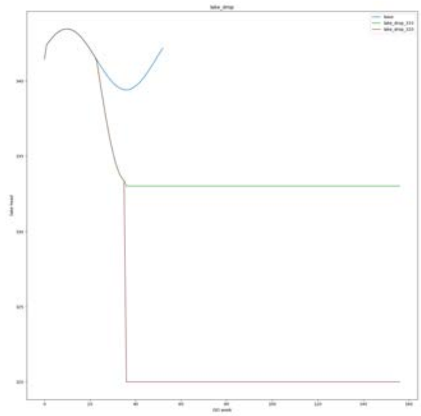


Figure: Lake Hawea head boundary conditions for the lake_drop scenarios.

Methods and results

The synthetic lake levels were created for a three year period. All other boundary conditions were set at the steady state (mean of optimised period) values. The synthetic lake levels were then applied to the model and the results were extracted at the indicator boundary points. Finally we compared these scenarios to the base sin fit to the observed ISO weekly mean lake levels

The sections below show some key figures for the various lake level scenarios. Many more figures for each scenario are available in the low lake level results folder. Scenarios/low_lake_scenarios/0_results/lake_drop/comp_plots/hds_monitoring.png

The changes in lake levels clearly propagate directly to the groundwater levels. However the lake_drop and the low_wide_amp scenarios clearly show the predicted impacts of Lake levels falling below the threshold value.

Lake_drop Scenario results

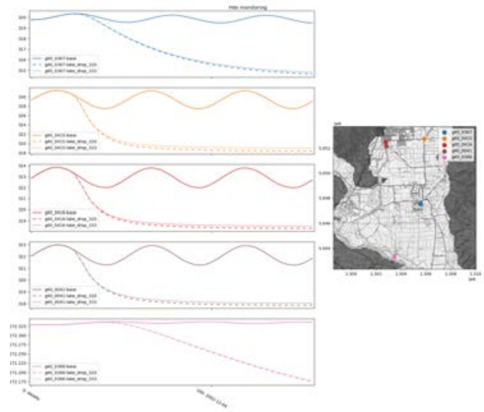


Figure: Responses at the high frequency monitoring bores for the lake drop scenarios.

Low_amp Scenario results

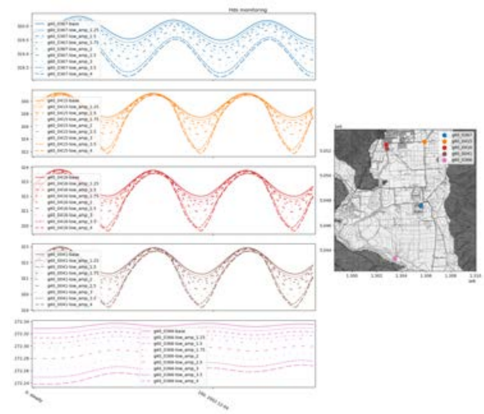


Figure: Responses at the high frequency monitoring bores for the low_amp scenarios.

Low_wid_amp Scenario results

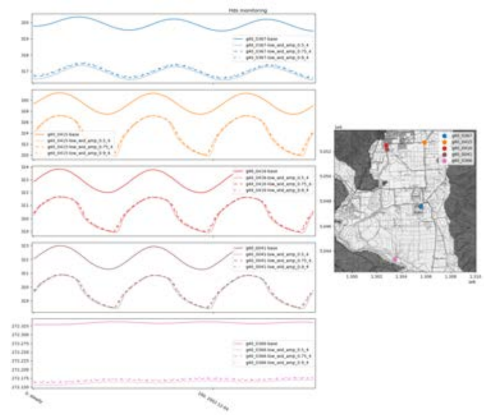


Figure: Responses at the high frequency monitoring bores for the low_wide_amp scenarios.

Low_wide Scenario results

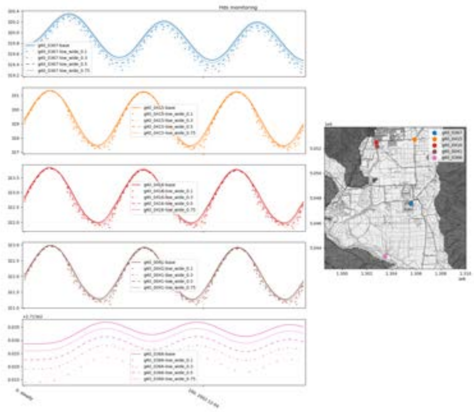


Figure: Responses at the high frequency monitoring bores for the low_wide scenarios.

Shift results

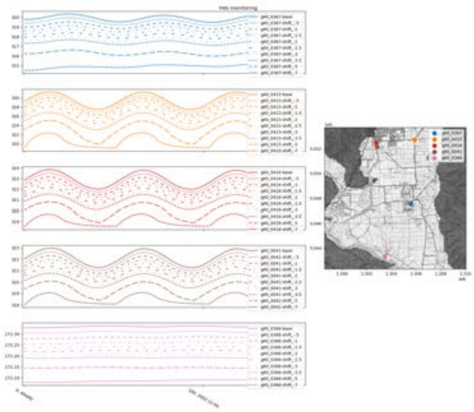


Figure: Responses at the high frequency monitoring bores for the shift scenarios.

MT3D Indicator Scenarios

Scenario name	Boundary condition concentration
all_any	all boundary conditions set to 1 (process check)
all_hill_indicator	all hillside inflows (excluding John and Grandview creeks) set to 1
all_str	all stream boundary conditions set to 1
hill_rch_indicator	all hillside inflows (excluding John and Grandview Creeks) and recharge set to 1
lake_con_indicator	all lake boundary conditions set to 1
not_any	all boundary conditions set to 0 (process check)
not_str	all boundary conditions (except str package)
race_con_indicator	all race cells set to 1
rch_indicator	recharge concentration set to 1

Methods

We ran a steady state model with the mean optimised period boundary conditions and then ran the MT3D model with the concentrations for each boundary condition set to 1 and all others set to 0. The MT3D model was run for an arbitrary period (7.305 E5 years) to ensure the concentrations were at pseudo steady state. The final concentrations for these scenarios were saved and plotted. They are available in the mt3d_indicator_scenarios folder.

Results

Interpretation of these results are relatively trivial a value of 0 means no water in the cell is modelled to originate from the boundary condition, while a value of 1 means all water in the cell originated from the boundary condition. Below we include the figures for the key boundary conditions. These results are useful to determine the relative importance of each boundary condition and help set allocation zone boundaries.

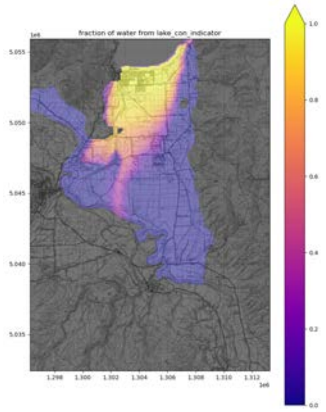


Figure: fraction of the water in the model sourced from Lake Hawea at steady state conditions.

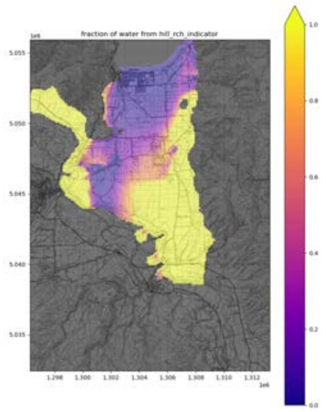


Figure: fraction of the water in the model sourced from hillside inflow or LSR at steady state conditions. Note this excludes John and Grandview Creeks.

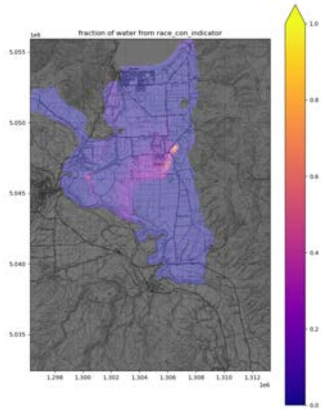


Figure: fraction of the water in the model sourced from water race leakage at steady state conditions.

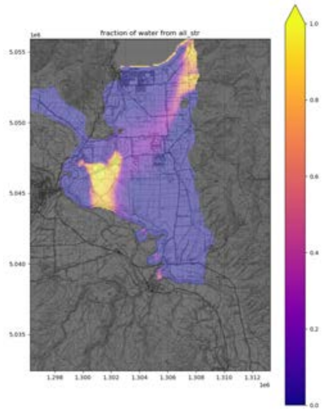


Figure: fraction of the water in the model sourced from the Hawea and Clutha Rivers and John and Grandview Creeks at steady state conditions.

Recommended Allocation Zones

Wilson (2012) recommended a set of allocation zones for the hawea area. These are shown in the figure below.

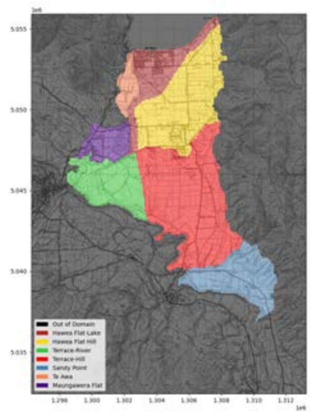


Figure: Allocation zones recommended by Wilson (2012) for the Hawea area. Note the allocation zones are not recommended for the entire model area, only the area of the Hawea Basin Groundwater Review (2012) study.

From the current modelling we have recommended some slight changes to the allocation zones. These are shown in the figure below. The main differences are:

1. The Hawea Flat-hillside and Hawea Flat-lake allocation zones have been rearranged into the Hawea Flat allocation zone and Grandview allocation zone. This is based on the new information that Lake Hawea levels impact the groundwater levels in most of the Hawea Flat area. The Grandview allocation zone is defined as the area to the west of the Hawea Flat allocation zone which is not impacted by Lake Hawea levels. The location of the Grandview-Hawea Flat boundary was approximately defined by the inferred Grandview Fault location. The Grandview Fault is thought to uplift basement rock above the Hawea Flat groundwater levels and therefore prevent the lake levels from impacting the groundwater in this area. Because the location of the Grandview Fault is not well defined we suggest that areas in the Grandview allocation zone could be reassigned to the Hawea Flat allocation zone if further investigation shows that water levels at that location are impacted by the Lake Hawea levels and that it would be reasonable to infer that groundwater flows from the lake to the area of interest.
2. We explicitly included the Camp Hill Moraine, but we suggest that very limited water is likely to be available in this zone.
3. We extend the sandy point zone to the Northern border of the next allocation zone.
4. We explicitly model and include the Maungawera Valley allocation zone, which was not addressed in the Wilson (2012) report.
5. Here we have suggested possible setback areas for the Butterfield and Campbell wetlands, note these location are suggestions however, the actual location of the setback areas requires planning consideration and is beyond the scope of this project. We have provided scientific guidance to support the planning process.

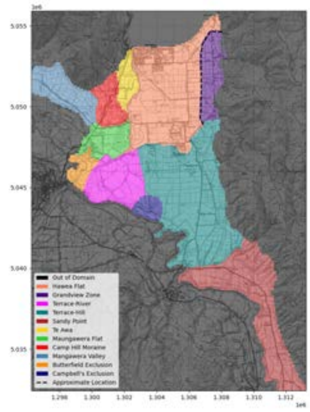


Figure: Recommended allocation zones for the Hawea area from this work.

Allocation Scenarios

Allocation Scenario overview

Scenario Name	Abstraction	LSR	Other comments
optimised	optimisation period	optimisation period	final optimised model
long_current	extended_pump: ISO weekly mean pumping	irr_rch	

long_nat	None	dryland_rch	race losses still included
no_pumping	None	irr_rch	
full_allocation	extended_full_allo: maximum daily allocation multiplied by the min/max normalised to historical pumping record	irr_rch	
max_allocation_on_pump_curve	extended_max_allo_pc: maximum daily allocation applied to pump curve	irr_rch	
{zone} MAPC + {rate} $ m^3/day $	extended_max_allo_pc + {rate} applied to the grid_pump wells for {zone}. Increased allocation scenarios.	irr_rch	
reduction _ {faction}	extended_max_allo_pc, wells in the Maungawera Valley allocation zone are multiplied by {fraction} e.g. .9 = 90% of maximum allocation on pumping curve for the maungawera valley	irr_rch	only for the Maungawera Valley allocation zone
max_allocation	extended_max_allo: maximum daily allocation applied to every day of the year	irr_rch	not realistic

Full List of Zone Specific Scenarios

Scenario Name	Applicable Zone	Pumping Increase (+) / Decrease(-)	Percent Increase (+) / Decrease (-)
Hawea Flat MAPC + 3424 m^3/day	Hawea Flat	3424	0.05
Hawea Flat MAPC + 6847 m^3/day	Hawea Flat	6847	0.1
Hawea Flat MAPC + 13694 m^3/day	Hawea Flat	13694	0.2
Hawea Flat MAPC + 20542 m^3/day	Hawea Flat	20542	0.3
Hawea Flat MAPC + 34236 m^3/day	Hawea Flat	34236	0.5
Hawea Flat MAPC + 51354 m^3/day	Hawea Flat	51354	0.75
Hawea Flat MAPC + 68472 m^3/day	Hawea Flat	68472	1
Hawea Flat MAPC + 102708 m^3/day	Hawea Flat	102708	1.5
Maungawera Flat MAPC + 500 m^3/day	Maungawera Flat	500	
Maungawera Flat MAPC + 1000 m^3/day	Maungawera Flat	1000	

Maungawera Flat MAPC + 2500 m^3/day	Maungawera Flat	2500	
Maungawera Flat MAPC + 5000 m^3/day	Maungawera Flat	5000	
Maungawera Flat MAPC + 7500 m^3/day	Maungawera Flat	7500	
Maungawera Flat MAPC + 10000 m^3/day	Maungawera Flat	10000	
reduction_0.5	Maungawera Valley	-2348	-0.5
reduction_0.6	Maungawera Valley	-1878.4	-0.4
reduction_0.7	Maungawera Valley	-1408.8	-0.3
reduction_0.8	Maungawera Valley	-939.2	-0.2
reduction_0.85	Maungawera Valley	-704.4	-0.15
reduction_0.9	Maungawera Valley	-469.6	-0.1
reduction_0.95	Maungawera Valley	-234.8	-0.05
Te Awa MAPC + 500 m^3/day	Te Awa	500	
Te Awa MAPC + 1000 m^3/day	Te Awa	1000	
Te Awa MAPC + 2500 m^3/day	Te Awa	2500	
Te Awa MAPC + 5000 m^3/day	Te Awa	5000	
Te Awa MAPC + 7500 m^3/day	Te Awa	7500	
Te Awa MAPC + 10000 m^3/day	Te Awa	10000	
Terrace-Hill MAPC + 135 m^3/day	Terrace-Hill	135	0.1
Terrace-Hill MAPC + 336 m^3/day	Terrace-Hill	336	0.25
Terrace-Hill MAPC + 673 m^3/day	Terrace-Hill	673	0.5
Terrace-Hill MAPC + 1346 m^3/day	Terrace-Hill	1346	1
Terrace-Hill MAPC + 2019 m^3/day	Terrace-Hill	2019	1.5

Terrace-River MAPC + 1011 m^3/day	Terrace-River	1011	0.1
Terrace-River MAPC + 2527 m^3/day	Terrace-River	2527	0.25
Terrace-River MAPC + 5054 m^3/day	Terrace-River	5054	0.5
Terrace-River MAPC + 10109 m^3/day	Terrace-River	10109	1
Terrace-River MAPC + 15164 m^3/day	Terrace-River	15164	1.5

Allocation via Zonal recharge

For each of the allocation zones we extracted the range of hillside inflows and LSR. An example figure is shown below. It is a violin plot, so the width of the violin represents the probability of the value occurring. The full results, including tabular results are available in the allo_zone_rch folder.

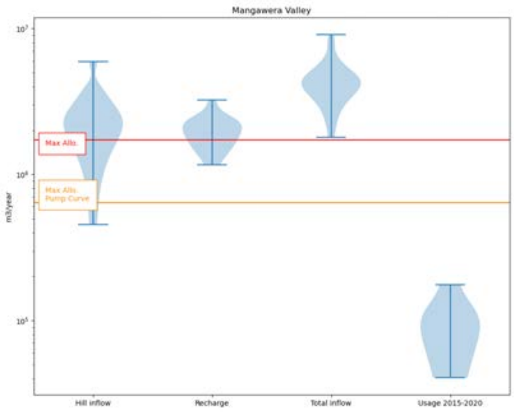


Figure: Example violin plot of the range of hillside inflows and LSR for the Maungawera Valley allocation zone.

Naturalised vs current vs full

In addition to the zonal recharge modelling we also compared the impact of the long current scenario with various full allocation and naturalised scenarios. The full results are available in the nat_current_full folder.

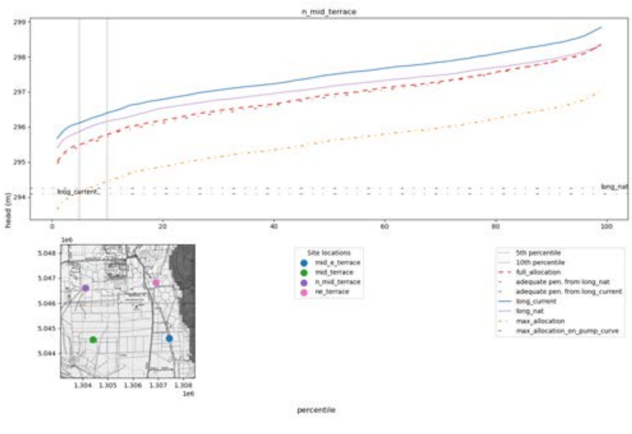


Figure: An example of the comparison between the long_current, naturalised, and various full allocation scenarios.

Zone specific Scenario Methods and Results

We modelled a number of scenarios to determine the impact of potential changes to zonal allocation on the groundwater levels for most allocation zones. An overview of these results are shown in the table below. However, there were some allocation zones that we did not conduct scenario modelling for. These zones and the rational for not modelling them are described below and we suggest any allocation in these zones should be based on the results of the zonal recharge modelling.

- **Grandview Zone:** The model does not provide a good representation of the groundwater levels in this area.
- **Camp Hill Moraine:** We suggest that very limited water is likely to be available in this zone and it was not modelled.
- **Sandy Point:** The model does not provide a good representation of the groundwater levels in this area as there is very limited data available to constrain the model.

Allocation Zone	Scenarios modelled allocation:	Path to results
Hawea Flat	Increases	Scenarios/allocation_results/Hawea Flat_results
mangawera _valley	Decreases	Scenarios/allocation_results/mangawera_valley
Maungawer a Flat	Increases	Scenarios/allocation_results/Maungawera Flat_results
Te Awa	Increases	Scenarios/allocation_results/Te Awa_results
Terrace-Hill	Increases	Scenarios/allocation_results/Terrace-Hill_results
Terrace-Riv er	Increases	Scenarios/allocation_results/Terrace-River_results

Full discussion of potential allocation levels are reserved for the final report. However we want to highlight several key points:

1. Changes in allocation in one zone can impact the groundwater levels in other zones. Therefore, we suggest that allocation in all zones should be considered collectively.
2. These scenarios do not discriminate between groundwater and induced stream depletion. Therefore, we suggest that allocation should be considered in the context of the stream depletion that is acceptable for the Hawea River and Clutha River.
3. Where the results suggest that a significant increase in allocation is possible, particularly where there is limited or no present allocation we suggest a very conservative approach. The lack of current abstraction means that the model has minimal to no information about the impact of abstraction to local groundwater levels. Any increase in allocation in these areas should be phased in over time to allow for monitoring and revised assessment of the impact of the allocation on the groundwater levels.

Quartz Creek Allocation Zone LSR analysis

The Quartz Creek Allocation Zone is a small area to the west of the Maungawera Valley in the catchment and alluvial fan of Quartz Creek. The allocation zone is shown in the figure below. Water in this area drains towards Lake Wanaka and was therefore out of scope of this project. However, we have included the results of LSR modelling for this area in this repository for completeness and because it required minimal additional effort. The LSR results are located in the quartz_creek_lsr/results folder and are shown in the figure below. Note that allocation in some of these areas have the potential to impact the groundwater levels in the Maungawera Valley and therefore should be considered in the context of the Maungawera Valley allocation zone.

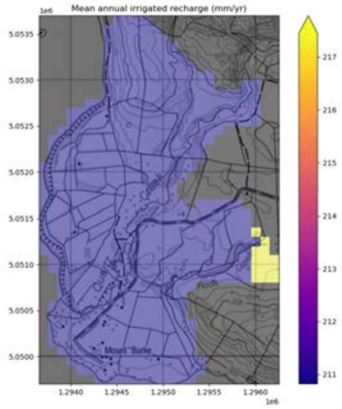


Figure: Locations of the Quartz Creek allocation zone and the LSR results.

Wetland Setback Scenarios

The purpose of the wetland setback modelling was to assess the impact of abstraction on the two sensitive wetland in the Hawea Basin.

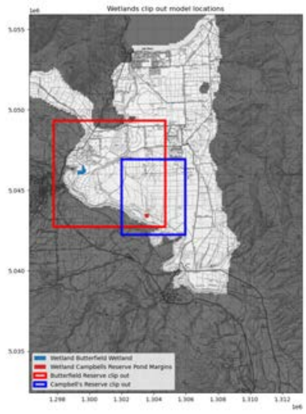


Figure: Locations of the two wetlands in the Hawea Basin and the boundaries for the clip out models.

Methodology

We wanted to examine a larger range of parameters than the optimised model parameter set. Therefore we made two clip out models; one for each wetland. The clip out models were made by clipping out the wetland and an approximately 5 km up gradient buffer. We set the boundary conditions of these models as follows:

- No existing abstraction was included in the model.
- Steady state LSR
- Steady state conditions for the Hawea and Clutha Rivers.
- For the Campbell's wetland model we set the North, East, and West boundaries as constant head boundaries with the head set from the optimised model.

We then created a semi-regular polar grid of test abstraction well locations. To identify the impact of abstraction on the wetland we ran a base scenario (no abstraction) and then ran a series of scenarios. The head difference between the base and new scenarios at the wetland location was extracted. We varied the following parameters:

- pumping rate
- hydraulic conductivity

- specific yield
- river conductance

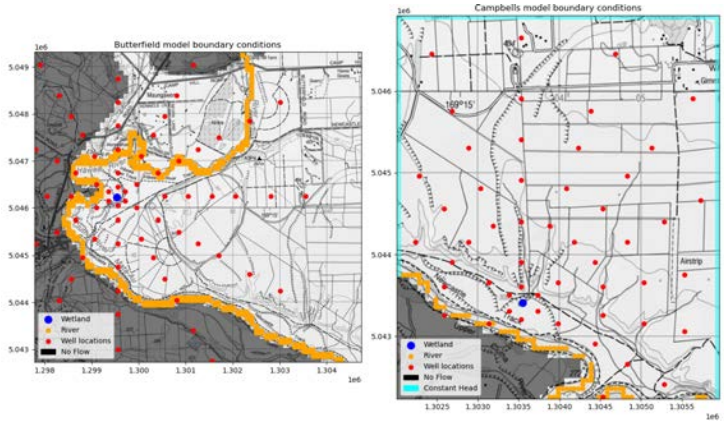


Figure: boundary conditions for the clip out models.

To visualise the results we assigned the head difference at the wetland location to the well location and then interpolated the results for each unique combination of parameters. Finally we qualitatively assessed the various parameterisations to determine an area where abstraction would likely impact the wetland.

Results

The full suite of results are available in:

- Campbell's: Scenarios/wetland_setback_campbells/results
- Butterfield: Scenarios/wetland_setback_butterfield/results

An example of the results for the Campbell's Wetland are shown below. The results for the Butterfield Wetland are similar. This specific example is for a pumping rate of 500 m^3/d, hydraulic conductivity of 0.316 m/d, specific yield of 0.316, and river conductance of 750 m^2/d. This specific example shows that abstraction within 1 km of the wetland is likely to cause c. 0.5 m drawdown at the wetland location. However, abstraction 2 km or further from the wetland is unlikely to cause drawdown at the wetland location. Note that several of the example bore locations were not able to sustain abstraction at this rate. The drawdown at these locations are excluded from the interpolation.

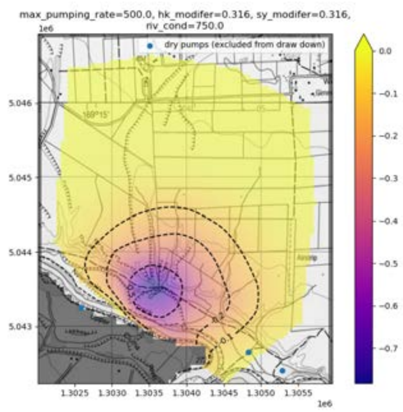


Figure: An example of the wetland setback modelling results for Campbell's Wetland.

References

- Wilson, J., 2012. Hawea Basin Groundwater Review. prepared by Resouce Science Unit of Otago Regional Council, June 2012, Dunedin.

Analysis of Historical Lake H■wea Low Lake levels corresponding groundwater levels)

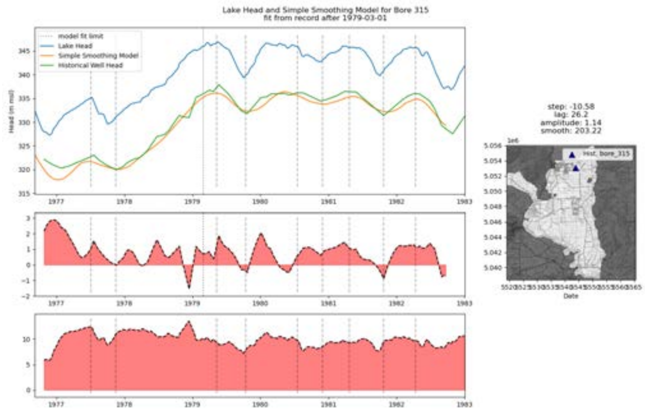


Figure: Historical Low level levels

Author: Matt Dumont

Date: 2023-11-14

Version: 1.0.0

Status: Draft

KSL project: Z22031HAW_hawea-model - supplemental

Purpose: This document provides the results of a re-analysis of the historical low lake levels at Lake H■wea (1976-1979). The analysis was conducted to determine the corresponding groundwater levels at the time of the low lake levels.

Index

Table of Contents

Module Index

- README.rst: This document
- **base_data: The raw input data for the historical analysis. For more info see the "Dataset and Resources" section at the end of this document.**
    - MWD_Hawea_Flats_Groundwater_1984.pdf: The original PDF of the Ministry of Works and Development report, which contains the historical data.
- current_model_prediction.py: run the naturalised 3d_v1d model for the historical period with low lake levels
- figures: output figures from the historical analysis. For more info see the "Dataset and Resources" section at the end of this document.

- generated_data: data generated by the analysis. For more info see the "Dataset and Resources" section at the end of this document.
- get_historical_data.py: read in and access the historical data
- lake_drop_scenarios.py: run and compare the lake drop scenarios to the historical data
- mt3d_indicator_scenarios: results for the MT3d component analysis
- mt3d_indicator_scens.py: run the MT3d component analysis on a steady state model with low lake levels
- plot_historical_data.py: plot the historical data
- plot_historical_natualised_model.py: plot the historical data and the naturalised model results
- shift_diff.py: compare lake and historical data, and calculate the shift between the two
- simple_smoothing_model.py: develop and apply a simple smoothing model to the historical data

Investigation Context and Objectives

A key prediction of the Lake H■wea groundwater model hosted in this repository was that there was some threshold lake level below which the lake would become disconnected from the groundwater system. Subsequently groundwater levels could fall significantly. These predictions were based on the model structure being required to match observed groundwater levels, particularly in Bore G40/0415. This repository holds more information on the model structure and predictions at low lake levels.

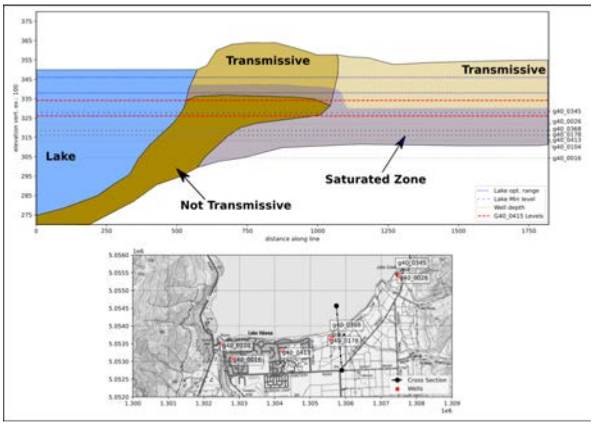


Figure: conceptual model of the Lake H■wea Moraine (across the moraine)

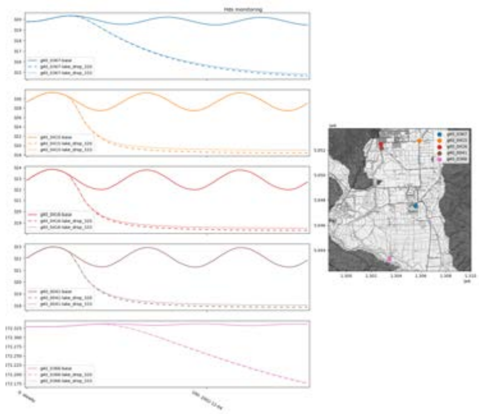


Figure: Responses at the high frequency monitoring bores for the lake drop scenarios.

Modern Lake H■wea levels are strictly controlled to be above 338 m msl. Historically, between 1976 to 1979, Lake H■wea levels fell to their lowest recorded level of c. 327.5 m MSL. This was reportedly because of exceptionally high energy demands.

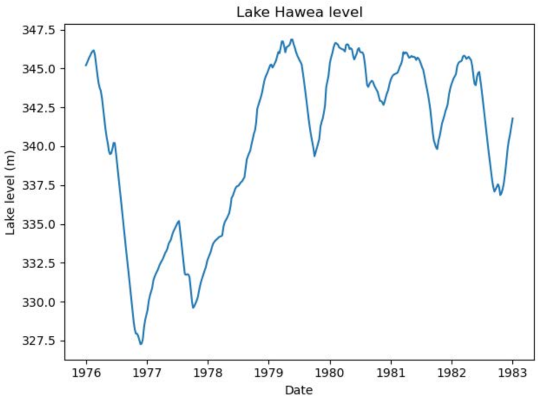


Figure: Lake H■wea levels from 1976 to 1979.

As part of the Lake H■wea modelling project the authors requested all information on historic lake levels and groundwater levels. It was understood that there were no regular groundwater monitoring records concurrent with the historic low lake levels. This was incorrect. At the end of the modelling project one of the modelling team found a record of an appendix to a 1984 Ministry of Works and Development report which contained plots of high frequency groundwater level data for bores during the historic low lake levels. A copy of this appendix is included in this repository and is accessible here. This appendix was previously unknown to both the modelling team and the science team at the Otago Regional Council. Its discovery was too late to be included in the modelling project; however the Otago Regional Council commissioned this addendum to the modelling project to digitise the data, investigate the implications of the historic low lake levels on groundwater levels, and to ascertain if the model predictions were consistent with the newly available historic groundwater level data.

Digitization of Historical data and discussion of their results

Summary of the historical data

The historical appendix contains:

- A map of the historic monitoring bores (figure reproduced below)
- **Groundwater level records at 5 bores during the historic low lake levels**
    - Bore 13
    - Bore 315
    - Bore 513
    - Bore 515
    - Bore Butterfields
- A contoured map of the change in groundwater level for every 1 meter drop in lake level
- A contoured map which identifies the lake level where the groundwater level is unlikely to be affected by the lake level
- Discussion and conclusions

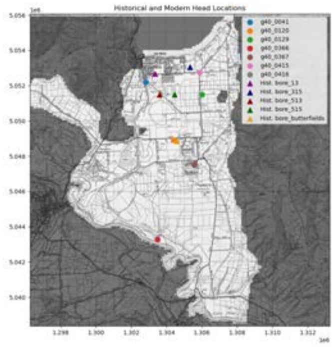


Figure: Map of the historic groundwater monitoring bores.

The key results from the appendix are:

- Variation in Lake H■wea levels below 329.66 m msl (330 m Dunedin 1958) are unlikely to have any impact on the groundwater levels.
- Groundwater levels further from the lake edge are less affected by variations in lake level.
- **The threshold for lake level variations to affect groundwater levels varies by distance from the lake edge:**
    - "A lake level of between 330 and 333 m will influence the area within 2.8 km of the lake but no further"
    - "A lake level of between 333 and 336 m will influence the area within 2.8 to 5.3 km of the lake but no further"

Digitization of the historical data and data access

All historical data was digitised and is included in the repository. The geospatial data (contours and bore locations) were geo-referenced using Qgis. The maps of bore locations appear to have been distorted (either while printing or being scanned), therefore we manually adjusted the locations of the reported bores based on the available landmarks (road intersections etc.). The contour maps contained less distortion, but should be considered indicative only. The digitized geospatial data available in the base_data/georeferenced folder.

The historic groundwater levels were digitised using the WebPlotDigitizer software. The raw digitised data is available in the base_data folder. The digitised data was then processed to convert to groundwater levels to meters above sea level (msl) from meters Dunedin 1958. The digitisation was then checked by comparing the digitised historic lake levels with the available lake level record.

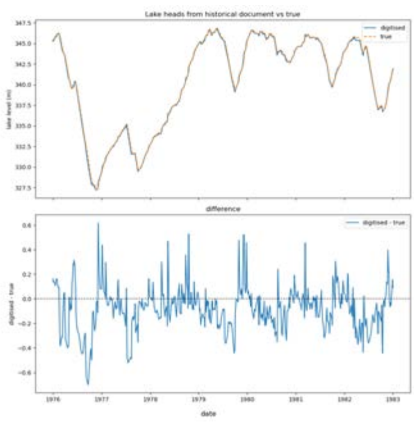


Figure: Comparison of the digitised lake levels with the available lake level record.

The digitised groundwater levels are available in the generated_data/historical_data.hdf file. The data is also available as csv files in the generated_data/csv_archive folder. For more information see the "Dataset and Resources" section at the end of this document.

Discussion of the historical results

A full re-analysis of the data is described below, but as a summary:

1. We generally agree with the level of the lake at which groundwater levels are affected.
2. We disagree with the interpretation that the threshold for lake level variations to affect groundwater levels varies by distance from the lake edge. We believe that the observed insensitivities are better accounted for by the smoothing of the groundwater level response to lake level variations.

Re-analysis of Historical Lake H■wea Low Lake levels & corresponding groundwater levels

Observed groundwater level response to historic lake level variations

The digitised groundwater levels were plotted against the available lake level record. The results are shown below.

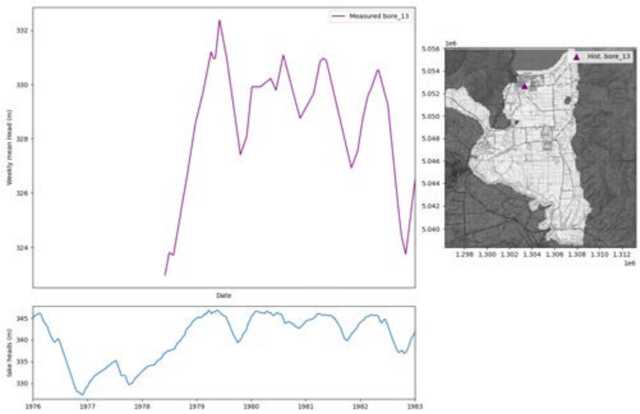


Figure: Observed groundwater level response to historic lake level variations at Bore 13.

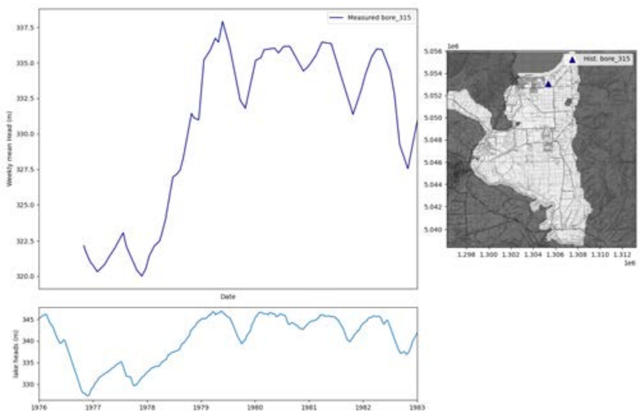


Figure: Observed groundwater level response to historic lake level variations at Bore 315.

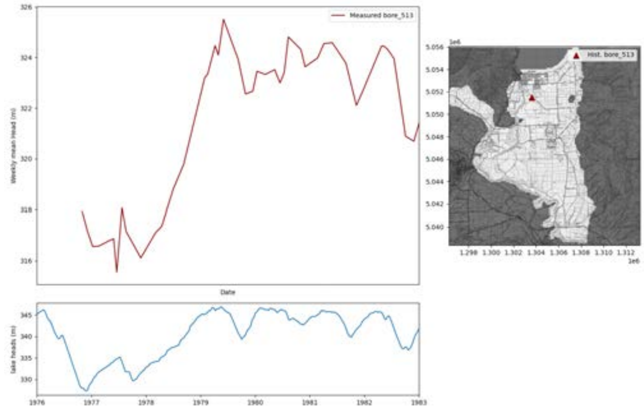


Figure: Observed groundwater level response to historic lake level variations at Bore 513.

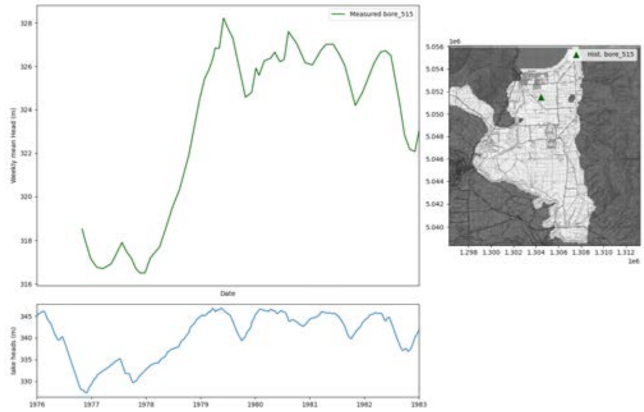


Figure: Observed groundwater level response to historic lake level variations at Bore 515.

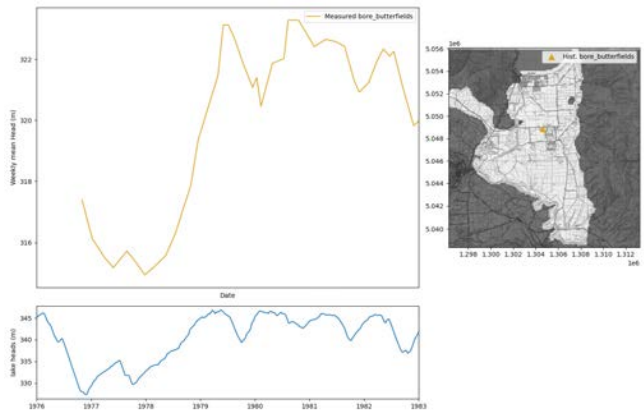


Figure: Observed groundwater level response to historic lake level variations at Bore Butterfields.

Smoothed Extrema matching

To understand the shift between the observed groundwater level local minima and maxima and the observed lake level local minima and maxima we conducted an extrema matching process. We first smoothed the observed groundwater and lake level data using a centered 100 day moving window. This process removes much of the noise while leaving the system relative minima/maxima intact. We then calculated the period between each lake level local minima and maxima and the next nearest groundwater level local minima and maxima, respectively. The results are shown below. In most signals there are minima and maxima that do not correlate well, likely due to other influences in the groundwater system. Never the less we identified the likely lag between changes in lake level and each historic bore as approximately:

- bore_13: 26 days
- bore_315: 21 days
- bore_513: 45 days
- bore_515: 45 days
- bore_butterfields: 71 days

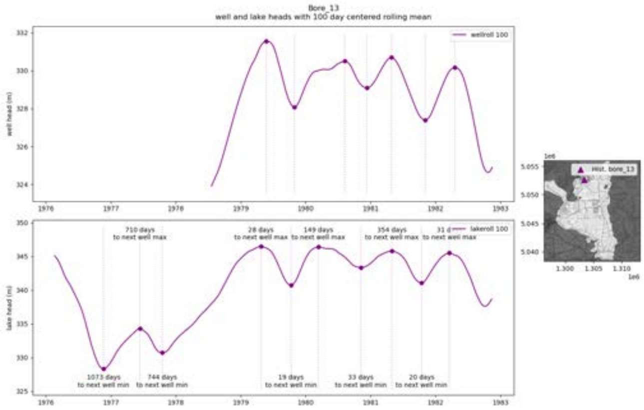


Figure: Observed groundwater level response to historic lake level variations at Bore 13.

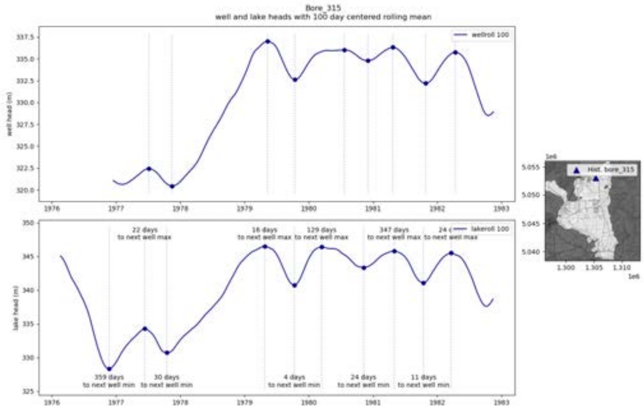


Figure: Observed groundwater level response to historic lake level variations at Bore 315.

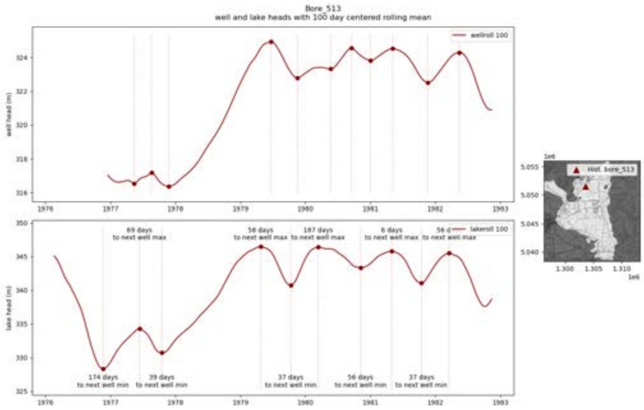


Figure: Observed groundwater level response to historic lake level variations at Bore 513.

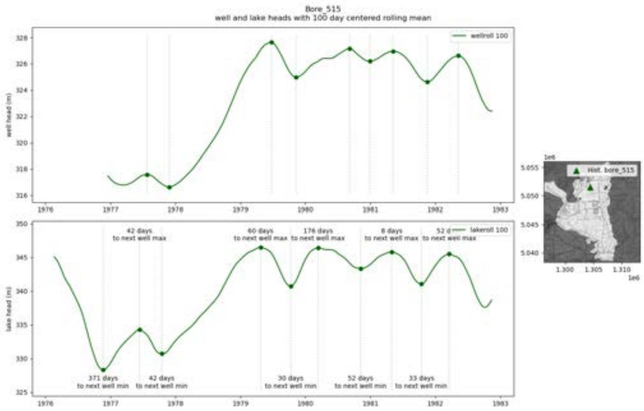


Figure: Observed groundwater level response to historic lake level variations at Bore 515.

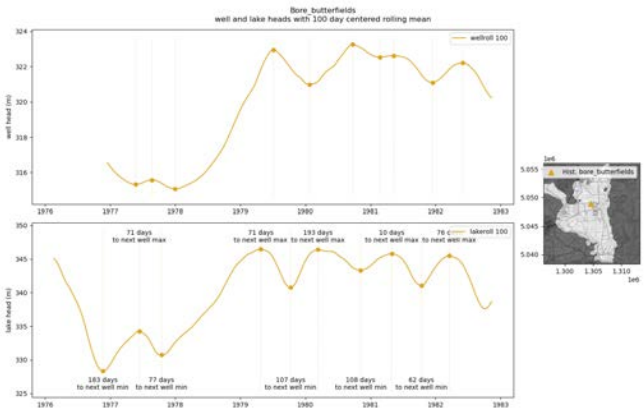


Figure: Observed groundwater level response to historic lake level variations at Bore Butterfields.

Lake and Groundwater level data comparison

We compared the shifted (see above) observed lake levels and groundwater level data for the recovery (only where the groundwater levels are increasing) periods in 1977 and 1978. The results at Bore 315 are the most useful and are shown below. In the first recovery there is a clear slope change when the shifted Lake H■wea level reaches 332 m msl. This slope change is also evident in the second recovery, but is less clear. We suggest that the increases in groundwater levels while the shifted lake levels are below 332 m msl are due to other influences in the groundwater system. We suggest that the groundwater level response to lake level variations is only evident when the shifted lake levels are above 332 m msl.

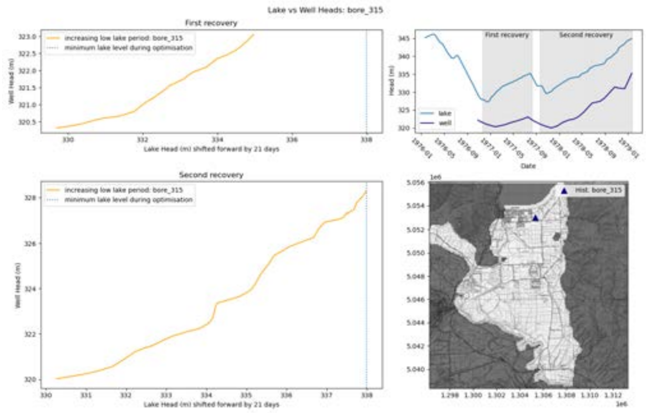


Figure: Observed groundwater level response to historic lake level variations at Bore 315.

Simple smoothing model

In the model build a simple smoothing model was developed to match the observed groundwater level response to lake level variations. The details for this model are described in the model build documentation. The model was developed to match the observed groundwater level response to lake level variations at Bore G40/0415. The purpose of this model was to define the model structure required to match the observed groundwater level response to lake level variations. Here we apply the model to the historical Bores to see where it deviates from the observed groundwater level response to lake level variations.

Simple smoothing model from Bore G40/0415

We applied the simple smoothing model developed for Bore G40/0415 to the nearby historical Bore 315. The results are shown below.

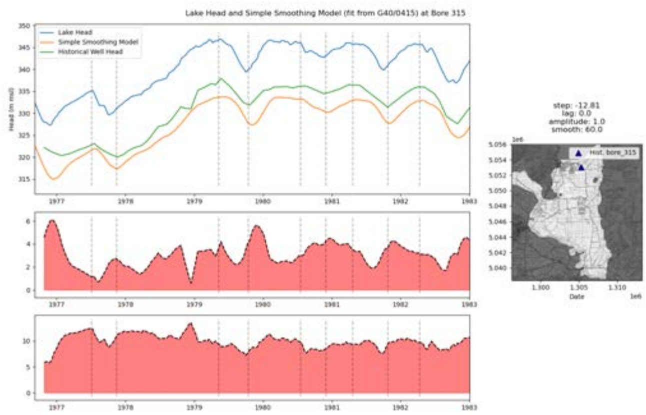


Figure: Simple smoothing model trained on Bore G40/0415 applied to the historical Bore 315.

Bespoke simple smoothing models

In addition, we trained the simple smoothing model on the historical bore data (Bores 315, 515, Butterfields) after 1979-03-01 (once Lake H■wea levels returned to their normal operational range). The results are shown below.

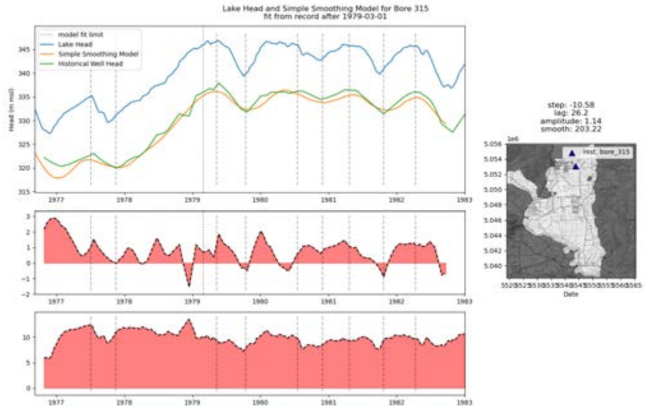


Figure: Bespoke simple smoothing model trained on Bore 315.

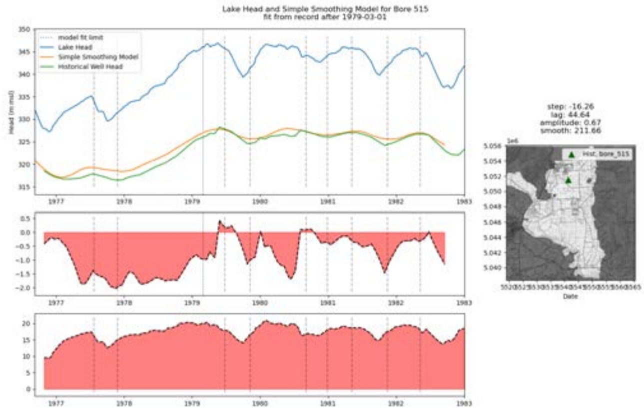


Figure: Bespoke simple smoothing model trained on Bore 515.

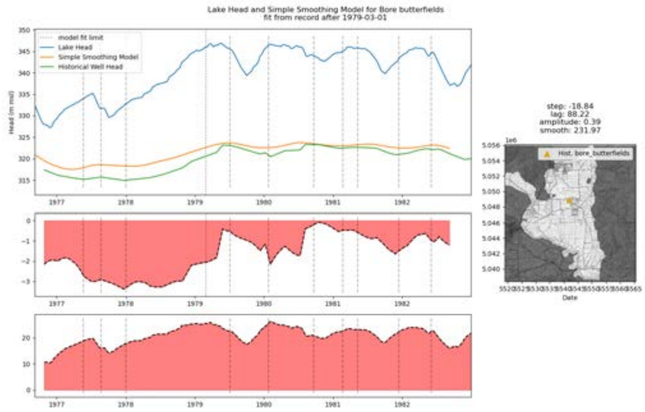


Figure: Bespoke simple smoothing model trained on Bore Butterfields.

Re-analysis discussion

Our re-analysis finds:

- **That the lake level impacts in the historical bores are likely lagged behind the observed lake level variations by:**
    - Bore 13: 26 days
    - Bore 315: 21 days

    - Bore 513: 45 days
    - Bore 515: 45 days
    - Bore Butterfields: 71 days
- Shifted lake levels in comparison to Bore 315 groundwater levels show a clear slope change when the shifted lake levels reach 332 m msl. This slope change is also evident in the second recovery, but is less clear. We suggest that the increases in groundwater levels while the shifted lake levels are below 332 m msl are due to other influences in the groundwater system. This suggests that the critical level for disconnection between the groundwater and lake system is 332 m msl. Note that this is higher than the historical analysis suggests (329.6 m msl).
- The simple smoothing model developed for Bore G40/0415 is able to match the observed groundwater level response to lake level variations at the nearby historical Bore 315, but includes some bias.
- The bespoke simple smoothing models do a good job matching the groundwater elevations at the historical bores after 1979-03-01 (once Lake H■wea levels returned to their normal operational range).
- The bespoke simple smoothing model for Bore 315 underestimates the groundwater elevation significantly during the first low lake period (1976-1977), but adequately matches the groundwater elevation during the second low lake period (1977-1979). In the first period the lake was below 332m msl for 7 months, and below 330m msl for 4 months. In the second low period the lake was below 332m for 5 months and below 330m msl for just over 3 weeks. The second low period is shorter than the fit model's smoothing period of 203 days (c. 6 months), which explains why the simple smoothing model does not show deviation too far from the observed groundwater levels despite lake levels that obviously fell below the cutoff elevation. This analysis does not disagree with an cutoff level between 330 and 332 m msl.
- The bespoke simple smoothing models at bores 515 and Butterfields overestimate the groundwater elevation rather than underestimate as we would expect. This suggests that the discontinuity between the lake and the groundwater levels may be obscured by the smoothing of lake levels (e.g. via storage) and/or other processes. Therefore we disagree with the Ministry of Works and Development conclusion that different elements of the system become disconnected from Lake H■wea at different lake levels. That conclusion is not uniquely supported by the data and is conceptually/structurally difficult to explain.

Model performance during Historic Low Lake levels

To assess the current model performance during the historic low lake levels we:

1. Ran the existing optimised model (3d_v1d) for the period 1976-01-01 to 1983-01-01, extracted the predicted groundwater levels at the historical bore locations, and compared the results to the observed groundwater levels.
2. Developed a new model (3d_v10a) which set the invert of the bund (see model build documentation) to 330 m msl and ran the new model for the period 1976-01-01 to 1983-01-01. We then extracted the predicted groundwater levels at the historical bore locations and compared the results to the observed groundwater levels.
3. Extracted the lake drop scenario results (see Scenarios documentation) at the historical bore locations, matched the results so that the time that the lake drop scenario went below the 3d_v1a bund elevation (335m) matched the time that the observed lake levels went below the historically observed limit (330m msl), and compared the results to the observed groundwater levels.

3d_v1d model results

The modelled vs observed groundwater levels for the 3d_v1d model are shown below.

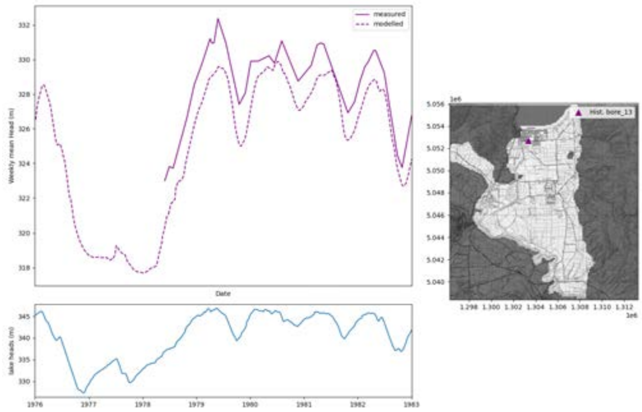


Figure: 3d_v1d model results for Bore 13.

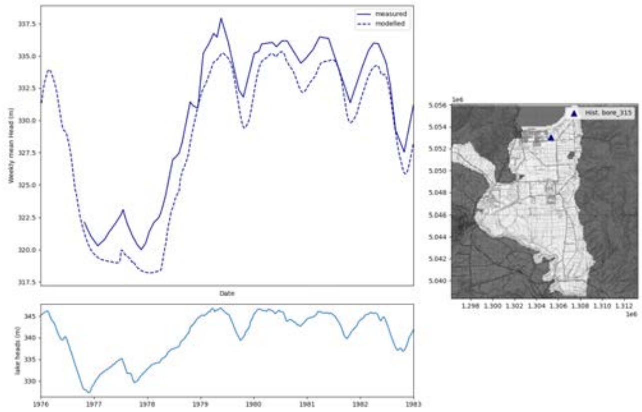


Figure: 3d_v1d model results for Bore 315.

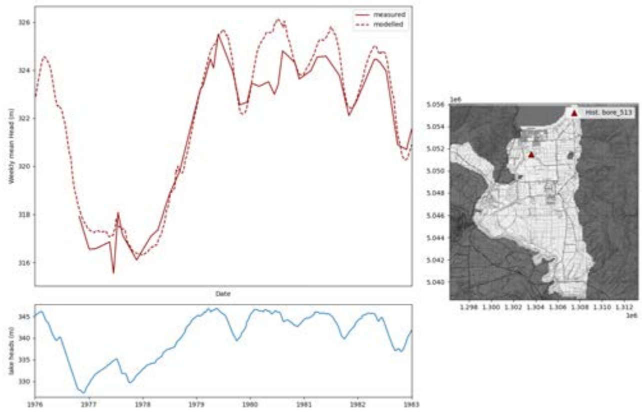


Figure: 3d_v1d model results for Bore 513.

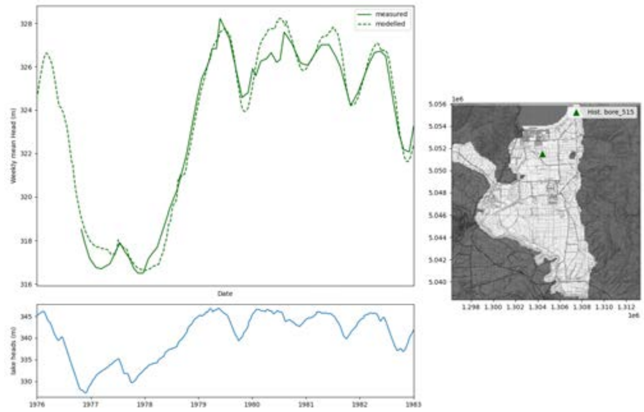


Figure: 3d_v1d model results for Bore 515.

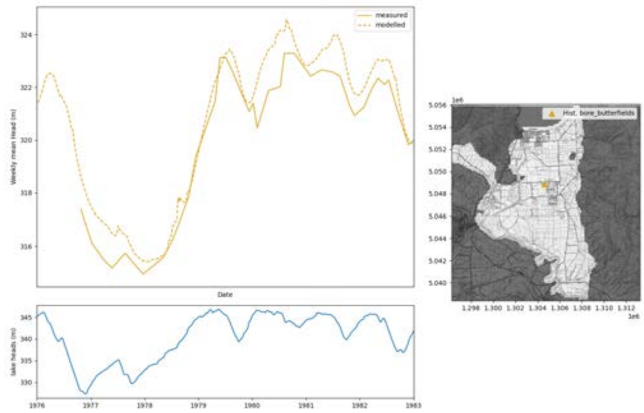


Figure: 3d_v1d model results for Bore Butterfields.

3d_v10a historical period model results

The modelled vs observed groundwater levels for the 3d_v10a model are shown below.

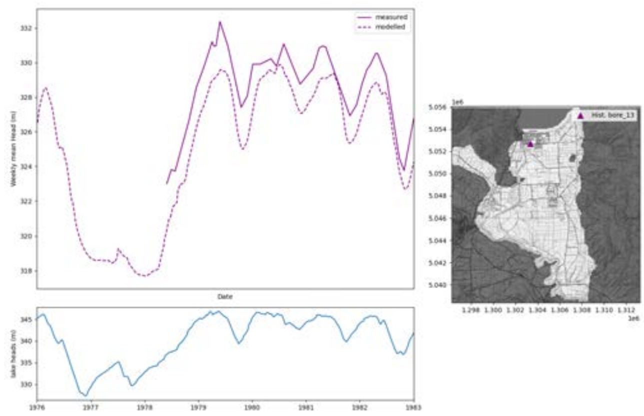


Figure: 3d_v10a model results for Bore 13.

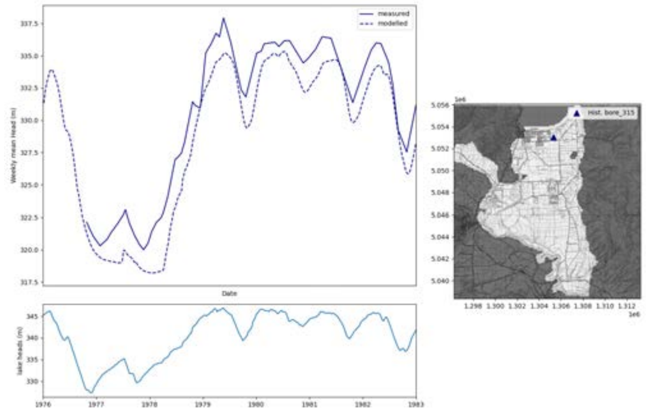


Figure: 3d_v10a model results for Bore 315.

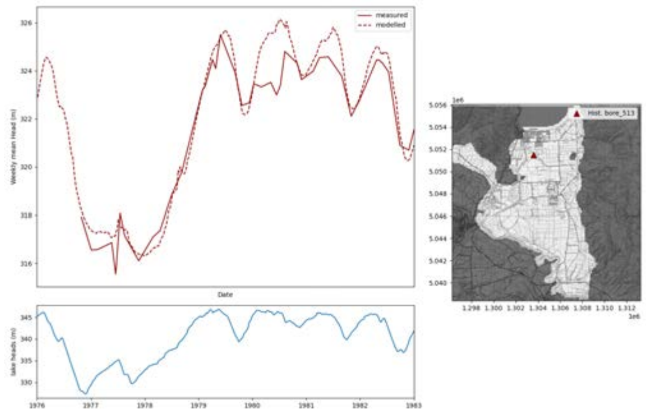


Figure: 3d_v10a model results for Bore 513.

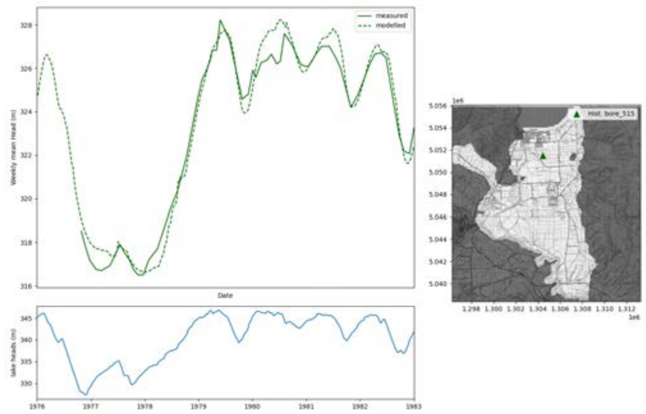


Figure: 3d_v10a model results for Bore 515.

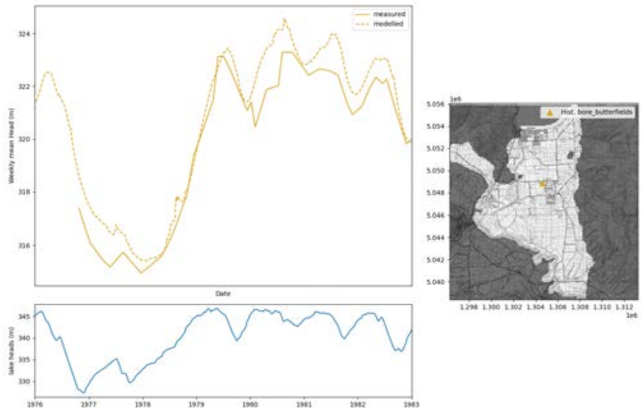


Figure: 3d_v10a model results for Bore Butterfields.

3d_v10a model period results

As the 3d_v10a model was not optimised for the model period we also present the modelled vs observed groundwater level for the high frequency targets (see the target documentation for more information) for the 3d_v10a model. The results are shown below.

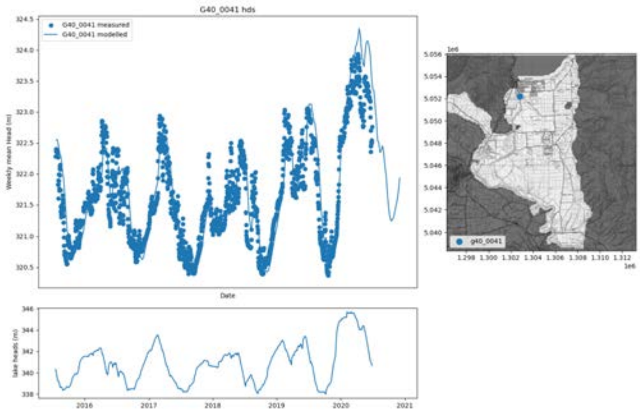


Figure: 3d_v10a model results for Bore G40/0041.

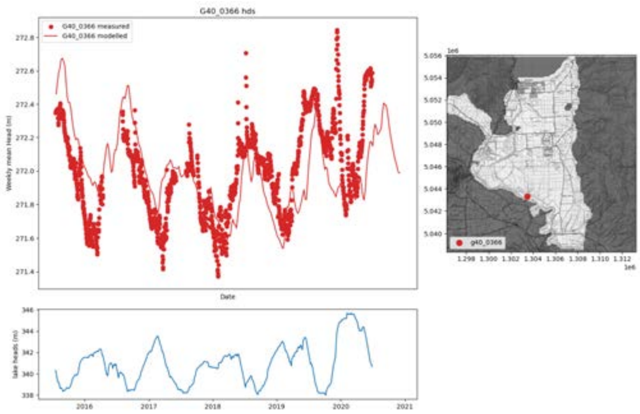


Figure: 3d_v10a model results for Bore G40/0366.

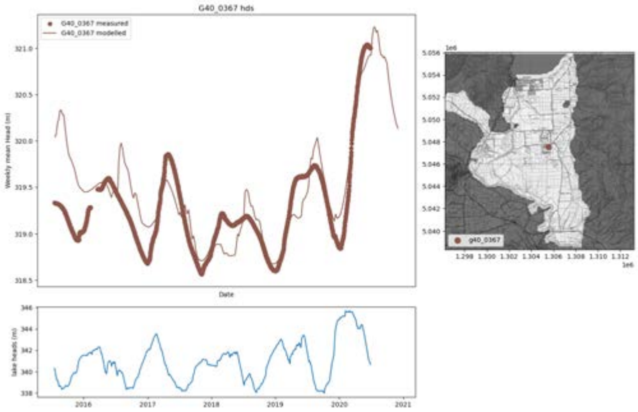


Figure: 3d_v10a model results for Bore G40/0367.

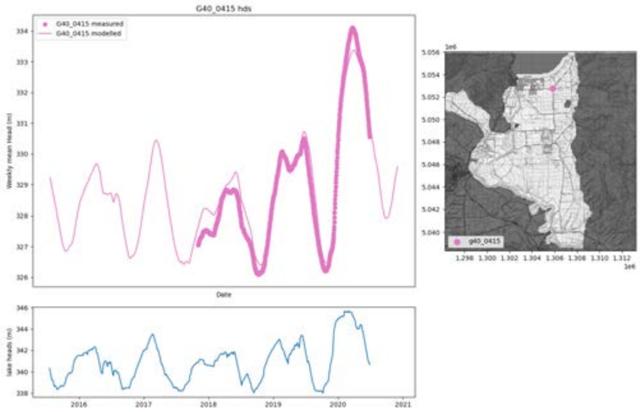


Figure: 3d_v10a model results for Bore G40/0415.

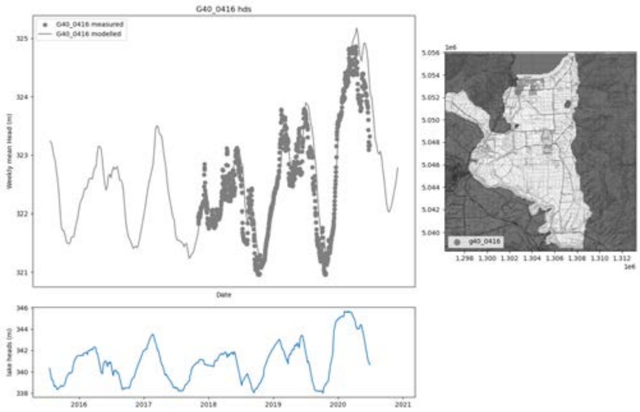


Figure: 3d_v10a model results for Bore G40/0416.

Lake Drop scenarios (3d_v1d model)

Note we have only presented the lake drop 320 (where lake levels were set to 320m msl) results here, but the results for the other lake drop scenarios are available in the figures/lake_drop_scenarios folder.

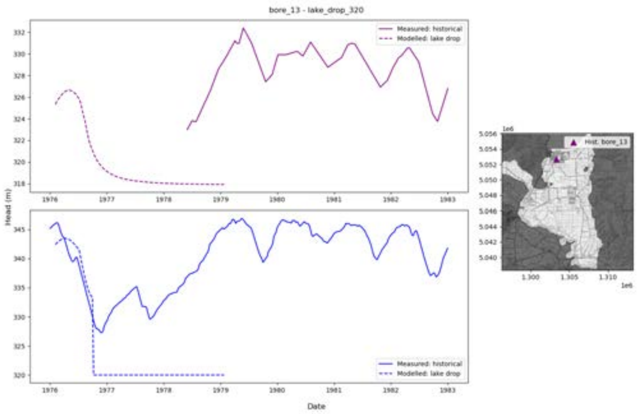


Figure: Lake drop scenario results for Bore 13.

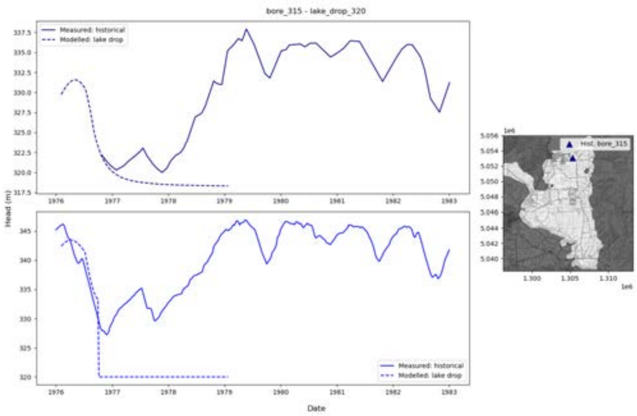


Figure: Lake drop scenario results for Bore 315.

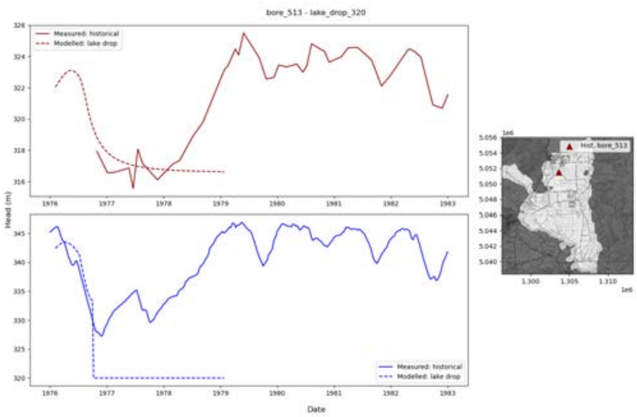


Figure: Lake drop scenario results for Bore 513.

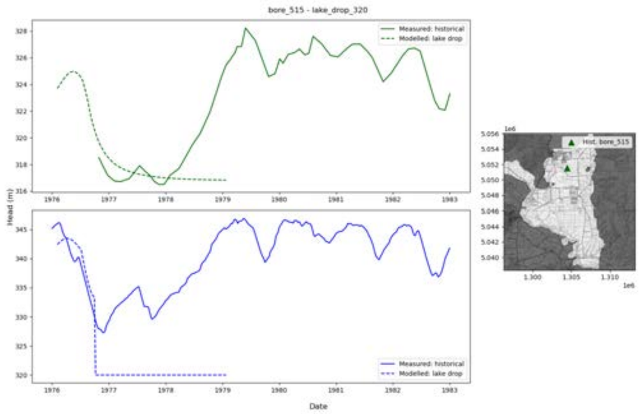


Figure: Lake drop scenario results for Bore 515.

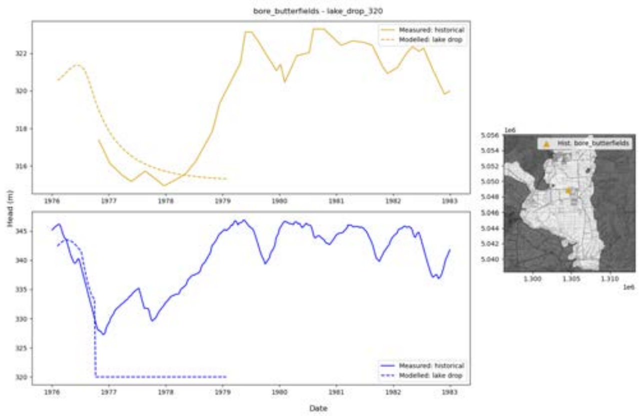


Figure: Lake drop scenario results for Bore Butterfields.

MT3d component analysis during low lake levels (3d_v1d model)

Given the performance of the model during the historic low lake levels, we also present the MT3d component analysis for the steady state 3d_v1d model where Lake H■wea levels are below the invert level. The results are shown below.

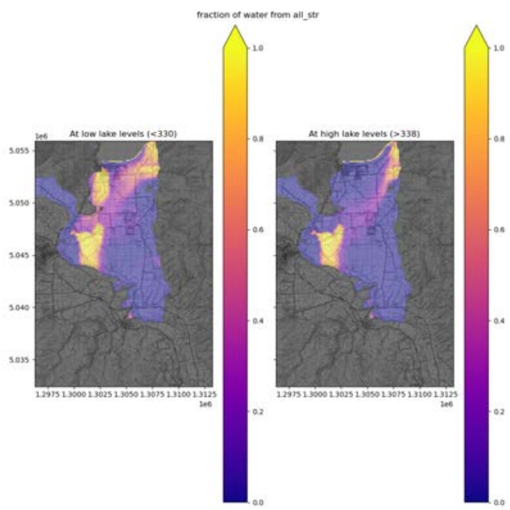


Figure: MT3d component analysis for Stream boundary conditions.

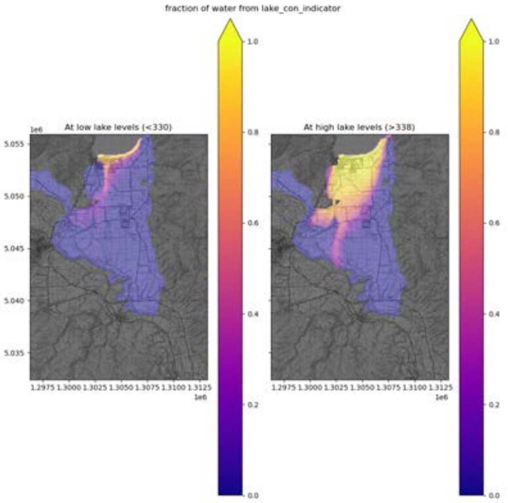


Figure: MT3d component analysis for Lake boundary conditions.

Model performance discussion

Our model performance analysis finds:

- The model preforms surprisingly well during the historic low lake levels. The modelled groundwater levels are within c. 2 m of the observed groundwater levels for the historical lake lows.
- Most wells show very little bias, but the model consistently underestimates the groundwater levels at Bore 13 and 315. The underestimation occurs during both the high and low lake levels (note that there is no data for Bore 13 during the historical low levels).
- The lake drop scenarios are broadly consistent with the observed data.
- Lake drop results at Bore 315 suggest that the observed historical groundwater levels did not reach their equilibrium state without Lake H■wea influences.
- The low lake levels associated with the lake drop scenarios at bores: 513, 515, and Butterfields propagate to the bore levels more slowly than the observed historical low lake levels. This suggests that the model may have lower than expected transmissivity in the main aquifer system. The lower transmissivity was necessary to fit the high frequency monitoring records; therefore this misfit may provide additional evidence of a second complex moraine structure in the main aquifer system. For more information see: model optimisation readme.
- In combination the misfit of the low lake levels provide evidence that there is likely information in the historical data that could contribute to a better groundwater model with re-calibration.
- The results for the 3d_v10a model (bund top set at 330 m msl) are very similar to the results of the 3d_v1d model (bund top set at 335 m msl). This suggests that the model is relatively insensitive to the exact elevation of the impermeable moraine. Further analysis, particularly formal structural and parameter uncertainty analysis, is required to better understand this perceived insensitivity to the impermeable moraine top elevation.
- The MT3D component analysis suggests that at low Lake H■wea levels the groundwater system in the Northern portion of the Lake H■wea aquifer system becomes dominated by the Grandview and John Creek inflows. At times of higher groundwater flow much of the losses from these creeks are diverted into Lake Hawea. This conclusion matches the conventional understanding of the system but assumes that most of the Grandview and John Creek flow is lost to groundwater. If these creeks are less strongly losing then groundwater levels would likely be over predicted by the model during low lake levels.

Conclusions and Recommendations

Our conclusions are:

- The Ministry of Works and Development data provides essential information on the Lake H■wea system.
- The new data confirms the key modelling prediction that there is some level below which the lake becomes disconnected from the groundwater system.
- The prior information suggested that the bottom elevation of the impermeable moraine was likely to be approximately 320 m msl and that the top of the impermeable moraine was likely between 327.6 - 338 m msl. The new data further constrains the likely elevation of the top of the impermeable moraine to be between 330 and 332 m msl.
- Given the impact of the lake level on the groundwater system we recommend that the lake level is maintained above a minimum elevation of 333 m msl to ensure that the groundwater system is not disconnected from the lake. This suggestion should not be seen as conclusive evidence to reset the current Lake H■wea consent conditions, but rather as an absolute minimum level below which the lake should not be allowed to fall even in extraordinary circumstances. As always, should lake levels fall below the current minimum level (338 m msl) the groundwater system should be closely monitored.
- These results do not discredit the recommended additional datasets (see model optimisation readme) as they are required to further constrain the model structure and to improve the model performance; however the results do increase the confidence in the implemented model and conceptual structure.
- The model generally performs well during the historic low lake levels. This result is surprising given there was no prior data to inform the model optimisation under these conditions.
- Despite the good fit there is likely more information to be gleemed from the historical data via model re-calibration.
- The model predicts slower than observed low head propagation to the more distant historical bores. This observation may provide additional evidence for a second complex moraine structure in the main aquifer system. For more information see: model optimisation readme.
- The model is remarkably insensitive to the level of the top of the impermeable moraine. A drop of 5m causes very little change in both the calibration and historical data matches. Further analysis, particularly formal structural and parameter uncertainty analysis, is required to better understand this perceived insensitivity to the impermeable moraine top elevation.
- Comparison of the "lake drop" scenarios with the observed historical data suggests that the groundwater levels during this historical period of low lake levels did not reach steady state conditions / equilibrium state of the aquifer when disconnected to the lake.
- MT3D component analysis suggests that as lake levels fall below the impermeable moraine, the groundwater in the northern Lake H■wea aquifer system becomes increasingly dominated by losses from Grandview and John Creeks. If these creeks are less well connected to the groundwater system than assumed, the model may significantly over predict steady state groundwater levels during low lake levels.

Our recommendations from this work are:

- Despite the adequate performance of the model, we would strongly suggest re-optimising the model and including targets for the period 1976-01-01 to 1983-01-01.
- Given the relative insensitivity of the invert level, we would suggest conducting formal parameter and structural uncertainly analysis to better understand the range of predictions.
- These results further support the need for a high frequency monitoring well near the Northeast Corner of the H■wea Flat aquifer (see model optimisation readme). One of the key model predictions for the complex moraine structure is that groundwater levels (impacted by the lake)

should be elevated in and around the Grandview/John Creek alluvial fans. Testing this prediction would require a high frequency groundwater level record in this area of at least a couple of years in length.

Dataset and Resources

- **base_data: This folder contains all the data used in this analysis.**
    - Bore_13: This folder contains the raw digitised data for Bore 13.
    - Bore_315: This folder contains the raw digitised data for Bore 315.
    - Bore_513: This folder contains the raw digitised data for Bore 513.
    - Bore_515: This folder contains the raw digitised data for Bore 515.
    - Butterfield_bore: This folder contains the raw digitised data for the Butterfield bore.
    - MWD_Hawea_Flats_Groundwater_1984.pdf: The original PDF of the Ministry of Works and Development report, which contains the historical data.
    - **georeferenced: This folder contains the georeferenced data.**
        - boreholes_georef.shp: The georeferenced borehole locations.
        - georef_map1_modified.tif: The georeferenced map 1 from the report
        - georef_map2_modified.tif: The georeferenced map 2 from the report
        - map_3_md.tif: The georeferenced map 3 from the report
        - map_4_md.tif: The georeferenced map 4 from the report
    - lake_plot.csv: The digitised lake level data, which is used to confirm the digitised data.
- **figures: This folder contains all the figures generated in this analysis.**
    - 3d_v10a_historical: comparison of the 3d_v10a model results with the observed data.
    - 3d_v1d_historical: comparison of the 3d_v1d model results with the observed data.
    - historic_period_hds: raw plots of the observed groundwater levels during the historic low lake levels.
    - historical_head_locs.png: The map of the historical bore locations.
    - lake_drop_scenarios: comparison of the lake drop (330|320) scenario results with the observed data.
    - lake_heads_from_hist_document.png: A comparison of the digitised lake levels with the available lake level record.
    - lake_v_hds_historic_only: comparison of the shifted lake levels with the observed data.
    - lake_v_hds_modelled: comparison of the shifted lake levels with the modelled data.
    - lake_v_hds_nearest_modern: comparison of the shifted lake levels with the observed data and the nearest modern bore.
    - model_period_hds: comparison of the modelled groundwater levels during the optimisation period.
    - simple_smoothing_model: comparison of the simple smoothing model results with the observed data.
    - well_lake_delta: plots of the difference between the observed groundwater levels and the observed lake levels.
    - well_lake_extrema: plots of the extrema matching process.

- **generated_data: This folder contains all the generated data for this investigation.**
    - csv_archive: a CSV archive of the historical_data.hdf file. For ease of access
    - historical_data.hdf: All historical data digitised and processed to meters above sea level (msl) from meters Dunedin 1958. This includes the bore records, the lake level records, the bore locations, and the model period heads.
    - long_nat_v10a.7z: standard outputs of the 3d_v10a model for the period 1976 to 2020.
    - min_fit_lake_bore_315_curve.p: The simple smoothing model parameters for Bore 315.
    - min_fit_lake_bore_515_curve.p: The simple smoothing model parameters for Bore 515.
    - min_fit_lake_bore_butterfields_curve.p: The simple smoothing model parameters for Bore Butterfields.
    - nat_historical_data.hdf: the modelled heads for the historical low lake period Model 3d_v1d.
    - nat_historical_data_v10a.hdf: the modelled heads for the historical low lake period Model 3d_v10a.
- **mt3d_indicator_scenarios: This folder contains the MT3d indicator scenarios for the 3d_v1d model.**
    - plots: plots of the MT3d indicator scenarios.
    - ucn_data: the MT3d indicator scenario data (arrays from 0-1 where 0 means no water from the component and 1 means all water from the component).