# CSG-induced groundwater impacts in the Pilliga region: prediction uncertainty, data-worth and optimal monitoring strategies

Sreekanth Janardhanan, Dan Gladish, Dennis Gonzalez, Dan Pagendam, Trevor Pickett, Tao Cui
GISERA – W8
31st March 2018

Citation

Copyright

Important disclaimer

Contents

Figures

Tables

Acknowledgments

Executive Summary

The proposal for development of coal seam gas from the Gunnedah Basin in northern NSW is currently being considered by the NSW Government for environmental approval. The Great Artesian Basin (GAB) aquifer(s) in this region are important sources of freshwater and the community has raised concerns about gas development potentially affecting quantity and quality of groundwater currently available for beneficial uses like irrigation, stock and domestic use. In data-poor areas like deep sedimentary basins, assessment of groundwater impacts from resource development should be an iterative procedure to inform adaptive management. The essential steps in such a management framework include predictive analysis of impacts using models informed by existing datasets, identifying the type and locations for new monitoring data that can inform the predictions of interest and refining the predictions and management actions including make-good arrangements where needed. Such an approach ensures that potential water quantity and quality changes of a proposed development are quantified in advance and verified by collecting compliance monitoring data and managed with an appropriate level of certainty so that regulators can ensure that the resource development projects are managed with acceptable level of impacts to environment and water resources. In this study we undertook one iteration of such a risk assessment-monitoring-management approach for predictive analysis, monitoring and management of drawdown impacts for a generalized case of coal seam depressurization in the Narrabri Gas Project area.

The groundwater model developed as part of the Bioregional Assessments for the Namoi subregion was used and modified for the modelling analysis in this study. Two alternative approaches were used for conceptualizing uncertain hydraulic property fields in the model. The first used a parsimoniously parameterized modelling approach (previously developed as part of groundwater modelling in the Bioregional Assessment Programme) that helps to quantify potential extreme drawdown impacts using limited computational resources. The second approach used highly parameterized modelling to characterize the spatial variability of heterogeneous hydraulic properties underpinned by measurements obtained from numerous cores in key aquitard sequences in the Surat and Gunnedah basins in the region (Turnadge et al., 2017). Both the modelling approaches considered wide range of variability and uncertainty in the volume of water extraction by CSG project. The analysis resulted in quantifying the drawdown impacts from CSG water extractions in the range 4.4 GL to 107.1 GL over life time of the gas project. Geostatistical models underpinned by these datasets were used to constrain the prior parameter and prediction uncertainties of drawdown impacts. A comparison of the 95th percentile drawdown impacts predicted under two approaches showed that incorporating spatial trends in aquitard hydraulic characteristics helped to reduce prediction uncertainty in the estimates for the median, 5th and 95th percentile values of predicted impacts at various receptor locations. This uncertainty reduction is attributable to increased level of information about the aquitard properties obtained from measured hydraulic property datasets.

The predictive analysis of drawdown estimated the median drawdown for the important GAB aquifer, the Pilliga Sandstone, to be less than 0.2 m for most of the areas within and beyond the

project area. The model simulations and uncertainty analysis indicated that it is highly likely that CSG-induced drawdown in the GAB aquifer, the Pilliga Sandstone, would be less than 1.2 m if CSG development similar to the generic scenario considered in this study were to go ahead.

We also undertook a probabilistic analysis of travel times and the distances over which water quality changes might travel, should any contaminant particles migrate to the GAB aquifer because of well integrity loss or accidental damage. Probabilistic simulation of travel times using forward particle tracking indicated that the velocity of flow of groundwater in the GAB aquifer is very slow and particles are likely to travel on the order of hundreds of metres within a design period of 100 years. Independent analysis of groundwater flow velocities using environmental tracers in the companion GISERA study (Raiber and Suckow, 2018) also inferred that the groundwater flows very slowly in this region.

One simulation for a very long period (3000 years) indicated that the maximum distance travelled by particles in this long period is about 6.5 km. This analysis indicated that, should any detrimental water quality changes be induced around the CSG well interface with the GAB aquifers, the risk of the resulting poor-quality water flowing travelling towards and reaching farmers bores and other receptors in the wider region beyond the gas project area within the 100-year simulation period is very unlikely. Probabilistic particle tracking, and spatial proximity analysis was conducted to identify the number of receptor bores which have a likelihood of intercepting particles flowing from the CSG bore locations. The analysis indicated that none of the receptor bores in the region have a likelihood of intercepting particles flowing from the CSG bore locations within a 100-year design period considered for monitoring network design. Probabilistic simulations of reverse particle tracking were undertaken to investigate if the flow-lines arriving at the risk receptors would be coming from the CSG wells, and thus providing an indication of whether the development posed any risk of contamination. The reverse particle tracking analysis showed similar results to the forward particle tracking with very small travel distances observed.

Spatial analysis was conducted to identify the number and types of risk receptors that would be located within the extents of different percentiles of drawdown and particle travel distances. Predictive analysis was undertaken to obtain the distribution of predicted drawdown change at 57 receptor nodes that had at least a 5% chance of a non-zero drawdown. The analysis indicated that pressure/water level changes in the GAB aquifers will be small and monitoring should also focus on deeper formations to identify CSG-induced depressurization propagating upwards. Currently the density of existing monitoring bores is 4.2 monitoring bores/ 100 km^2 for the GAB aquifers within 95th percentile of predicted drawdown.

The spatial patterns of predicted drawdown and particle travel distances were used in a mathematical model reduction framework that enables this knowledge to be used within an optimization routine to identify optimal locations for monitoring. The value of the design of the monitoring bores as a network for minimising prediction uncertainty across the region is demonstrated by designing a 10-bore monitoring network that collects drawdown data from multiple depths at each location. Geostatistical analysis of baseline water quality data collected from the GAB aquifer, the Pilliga Sandstone, was undertaken to identify areas where estimates of water quality are uncertain. This knowledge was used together with a Differential Evolution algorithm to identify (globally) optimal locations for monitoring baseline water quality.

The worth of monitoring data collected from these locations was further evaluated using the data-worth analysis techniques. Data worth analysis was undertaken to analyse the parameter contributions to predictive uncertainty and relative worth of monitoring data in reducing the prediction uncertainty of CSG-induced drawdown at these locations. Since the predicted level of drawdown is small for the considered generic scenario of coal seam gas development, the drawdown simulated in the GAB and water table aquifers were found to be more dependent on how the CSG-induced flux losses are compensated by releases of water from storages and/or river or recharge processes.

The analysis indicated that hydraulic characteristics of deeper formations in the Surat and Gunnedah basins govern the likelihood of propagation of CSG-induced drawdown into the GAB and other aquifers closer to the surface. For example, it was found that the data worth for informing drawdown predictions in the Pilliga Sandstone was most for monitoring data collected from the model layer immediately below the Pilliga Sandstone – i.e. the Purlawaugh formation. The highly parameterized modelling approach also helps to identify the type and location of the parameter to be measured to inform prediction of drawdown at each location and target formation of interest. While the monitoring network design presented in our study used the spatial patterns of predicted drawdown in multiple model layers, the data worth analysis was independent of the absolute values of predictions and used only the sensitivity of predictions to observations and parameters. Thus, data worth analysis can be used as an independent method for evaluating the veracity of the design. In general, the predictive analysis, monitoring network design and data worth analysis, combine to indicate that having multi-level piezometers in and around the 95th percentile drawdown extent can help to significantly reduce uncertainty in CSG-induced drawdown predictions. The probabilistic particle tracking and CSG well – receptor proximity analysis helped to identify areas of relatively higher importance for monitoring water quality impacts, while the analysis indicated that water quality risks at the regional scale considered in this study is very low.

The focus of our study was scientific analysis of predictive modelling results undertaken in a probabilistic framework to provide practical insights for optimising future monitoring in the Pilliga forest region under a generic CSG development scenario. Whilst the utility of a co-ordinated network of bores in reducing prediction uncertainty is demonstrated, the study does not recommend any single monitoring network design or density of monitoring bores for the study area. Those are policy decisions and are made by the regulators considering a wide range of factors often including the effects of multiple water users and industries and implications of predicted impacts in a water resources management context.

1 Introduction

1.1 Overview

The proposal for development of coal seam gas from the Gunnedah Basin in northern NSW is currently being considered by the NSW Government for environmental approval. The Great Artesian Basin (GAB) aquifer(s) in this region are important sources of freshwater and the community has raised concerns about gas development potentially affecting the quantity and quality of groundwater currently available for beneficial uses like irrigation, stock and domestic use. Concerns have also been raised about CSG development potentially affecting groundwater recharge into the broader GAB as the Pilliga forest is part of the intake bed of the GAB in northern NSW. The GISERA research into the "Impacts of CSG development on GAB fluxes" undertakes modelling and field studies to assess potential extremes of CSG-induced flux losses from the CSG development area and intake beds and to improve the conceptual understanding of recharge in this area. This report presents the results from the study undertaken as part of the GISERA project on "Spatial design of monitoring networks".

This study demonstrates an iterative approach for risk assessment and monitoring for adaptive management of the impacts of onshore gas development on groundwater resources. The risks to groundwater quantity and quality are initially quantified with a groundwater model that is underpinned with limited available datasets. The models are then used to inform the relative worth of existing and potential future datasets, their types and locations that can be most useful in informing and reducing uncertainty in the predictions of interest. The predictive analysis is then used for designing optimal monitoring strategies.

Specifically, this study integrated the knowledge generated from: 1) probabilistic modelling of CSG-induced pressure/water level changes; and 2) particle travel times from CSG wells, to important risk receptors for informing future groundwater monitoring strategies. The predictive modelling of drawdown (pressure/ water level) changes and particle tracking analysis were undertaken considering a generalized scenario of CSG depressurization and accounting for the substantial uncertainty in the understanding of the hydraulic characteristics of deep aquifer and aquitard formations in the Gunnedah and Surat basins. This approach enabled the use of knowledge emerging from the predictive analysis to inform the type, location and timing of data to be collected for monitoring compliance, improving our general understanding if the system and minimizing uncertainty in the prediction of potential impacts. Undertaking this type of analysis prior to the start of the CSG project, as in this case, provides valuable information for the regulators, groundwater managers and community stakeholders about the nature and probable extent of changes, and how they may undertake monitoring to evaluate compliance and adapt mitigation strategies, if required in advance, during or after the life of the CSG project.

1.2　Objectives

This report presents scientific analysis of potential drawdown impacts, and particle travel times and distances from CSG wells induced by a generic coal seam gas development scenario for the Pilliga forest region of the Gunnedah Basin. The CSG-well distribution proposed by the Narrabri Gas Project was used as a basis for defining this generic scenario. This knowledge was then integrated into statistical and optimization analysis frameworks to inform data-worth and spatial design of monitoring networks. The data-worth and spatial design of monitoring networks presented in this study pertains to a generic scenario of depressurization induced by CSG development and demonstrates how the design of a monitoring network can inform compliance monitoring and reduce prediction uncertainties. Multiple monitoring designs are presented, and these consider different monitoring objectives. While the scientific analysis presented can provide useful insight for the industry, regulators and other stakeholders for informing monitoring decisions and relative data-worth, this study does not recommend any specific monitoring density or strategy for existing or proposed coal seam gas projects. Such decisions are to be governed by the regulators based on the relevant water monitoring policies and requirements. Groundwater monitoring implemented by the industry must be designed for the planned (or most likely) scenario of coal seam gas development. The methods proposed in by this study are generic and can be adapted for such purposes. The specific objectives of this study are:

- Undertake probabilistic simulation of groundwater pressure and water level drawdown changes induced by a generic CSG depressurization scenario.

- Undertake probabilistic particle tracking from CSG wells as a screening analysis for water quality risks posed to important receptors if well integrity loss acts as a potential contamination pathway.

- Develop a generic method for optimal spatial design of groundwater monitoring network for early detection of groundwater system changes and reducing prediction uncertainty.

- Undertake modelling and uncertainty analysis to inform data-worth for improving prediction accuracy of pressure/ water level changes at selected risk receptors.

- Integrate the knowledge from the analyses above and demonstrate optimal design of a monitoring network for a generic scenario of CSG depressurization in the Pilliga forest in northern NSW.

1.3　Overview of the methodology

This study uses an optimization framework to inform future optimal strategies for groundwater monitoring and worth of data collected at optimal location. This is achieved by combining the data and knowledge emerging from predictive analysis of CSG-induced drawdown and particle travel distances and times, together with spatial and statistical analysis of trends. A numerical groundwater model based on MODFLOW-USG was developed as part of the Bioregional Assessments Programme (Sreekanth et al., 2017a) and subsequently used in the GISERA GAB flux project for predictive analysis of CSG-induced flux changes in the Pilliga forest area of the GAB aquifer. That model was used in this study as the basis for undertaking calibration and data-worth analysis.

The flow model was also used together with a particle path and travel time simulator for unstructured grids called MODPATH-3DU (Papadopoulos, 2016). This allowed for undertaking probabilistic predictive analysis of travel paths, times and distances from CSG wells and receptors. Spatial trends in CSG-induced drawdown in multiple model layers and depths from the surface were analysed. Sensitive ecological and economic risk receptors that were hydrogeologically connected to potentially impacted zones were evaluated with respect to the likelihood and magnitude of CSG development impacting these receptors.

This analysis considered the likelihood of changes in pressure or water level near a risk receptor as well as the likelihood of particles released from the CSG wells reaching these receptors within a 100-year timeframe. Uncertainty in prediction of these changes were quantified. This information was then used to inform how monitoring could reduce these uncertainties and inform management decisions. The details of these components are described in the following sections. An overview of the work flow used for predictive analysis, data worth and monitoring network design is shown in Figure 1. Details of the methods for each of these components and relevant theory are briefly described in the sub-sections.
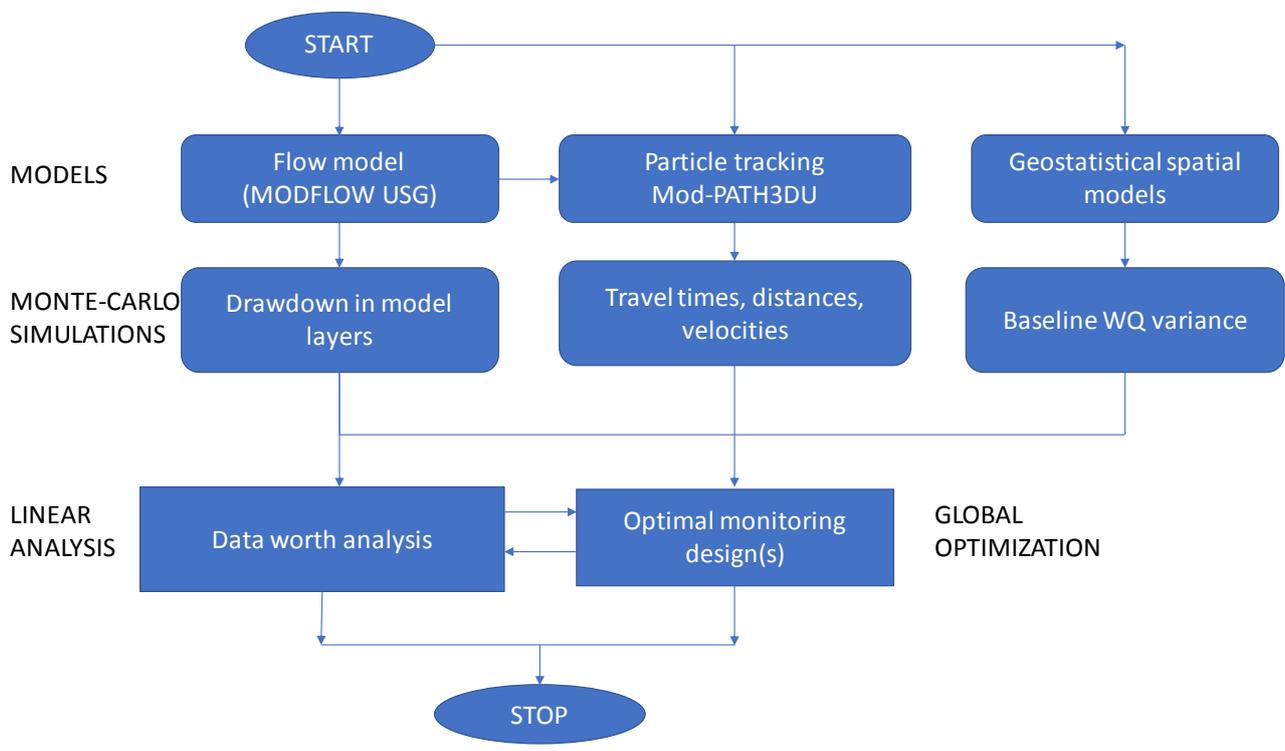
Figure 1: Workflow for predictive analysis, data-worth and monitoring network design

2 Probabilistic modelling of CSG-induced drawdown

2.1 Overview

One of the two most important prediction variables used in the data-worth analysis and spatial design of monitoring networks is the CSG-induced pressure/water level changes (drawdown) in aquifers and aquitards that overlie the coal formations. CSG-induced drawdown in each model layer was computed as the difference between model predictions of two possible states of GAB groundwater resource – one corresponding to no depressurisation and the other corresponding to depressurisation based on a generic case of gas development. These are respectively referred to as the baseline and the CSG development cases. This approach assumes that all other uses of groundwater remain unchanged over this period and quantifies only the changes in pressure and/or water level resulting solely from CSG-induced depressurisation. The details of the modelling methodology are described in detail in Crosbie et al. (2016). Development of the regional scale groundwater model for the Namoi subregion that we used for this study is described in detail in Sreekanth et al. (2017a). Interested readers are referred to those reports for details of the model and prediction of cumulative impacts of CSG and coal mine developments in the Namoi subregion. In the current report, we present a brief overview of the methods and results from applying that model, but with a different parameterization scheme for the predictive analysis of drawdown. The parameterisation is intended to reflect a generic case of coal seam gas development from the Pilliga forest area. This generic case CSG development is also reported in Sreekanth et al (2017c) that reported the GAB flux impacts and considers large variability in CSG-induced water production.

2.2 The groundwater model

A 15-layer numerical groundwater model that was developed for the Namoi sub-region in the Bioregional Assessment Programme was adopted with modified and improved parameterisation scheme in this study. Considering the complexity of the geology comprising two distinct sedimentary basins, MODFLOW-USG (Panday et al., 2013) was adopted as the preferred modelling platform as it permits pinching out layers and allows model layers to follow the geometry of the geological formations and boundaries.

CSG development targets the Hoskissons Coal and Maules Creek Formation in the Gunnedah Basin. Important assets including farmers' bores and groundwater-dependent ecosystems in the Namoi subregion rely on water from the alluvial formations, Pilliga Sandstone or in the outcrop of other formations. Therefore, these formations were represented by independent layers in the numerical groundwater model. The Namoi alluvium was vertically discretised into two different model layers corresponding to the upper Narrabri and lower Gunnedah formations. Other formations that are present between the alluvium and Pilliga Sandstone were conceptualised as an inter-burden layer with depth-dependent hydraulic characteristics. Similarly, the formations

between the Pilliga Sandstone and the Hoskissons Coal and between the Hoskissons Coal and Maules Creek Formation are represented in the numerical groundwater model by means of inter-burden layers with distinct effective hydraulic characteristics that vary with depth. The basement rock under the Maules Creek Formation is represented by means of another layer in the numerical model layer. This resulted in a numerical model architecture with nine layers to represent the hydrostratigraphy. Each inter-burden layers in the hydrostratigraphy was further divided into three layers in the numerical model grid to account for variable hydraulic properties at varying depths. This resulted in a numerical model grid with 15 layers as described in Table 1.

MODFLOW-USG was used to build the model using unstructured Voronoi grids. Model cells with a size of 300 m was used near the CSG project area and river nodes (and coal mines), and up to 3 km elsewhere. The finer mesh clearly identifies the areas of mining and coal seam gas development within the Namoi subregion. The resulting mesh has 58,649 Voronoi cells in plain view, covering an area of approximately 59,000 km^2. The number of cells in each model layer may be less than this number depending on the extent of each model layer which can be smaller than the entire model domain where layers are absent. Layers are absent in the numerical groundwater model where they do not exist, the most obvious example is the alluvium which only covers a fraction of the model domain.

The eastern boundary of the Namoi subregion groundwater model was defined by a fault system called the Hunter- Mooki Thrust Fault and was assumed to be impermeable. The northern boundary was also assumed to be a no-flow boundary as it is parallel to groundwater-flow direction in the Great Artesian Basin. Toward west and south the model domain extends beyond the subregion boundary to minimise boundary effects in the model results general-head boundary conditions were used along these boundaries.

The model's land surface was subject to recharge due to rainfall, overbank flooding and irrigation. They were implemented in the model using the Recharge (RCH) package for MODFLOW-USG. Rainfall recharge is spatially and temporally varying, reflecting spatial differences in near-surface geology and temporal variation in rainfall. A mean annual recharge surface for the Namoi subregion was obtained using a chloride mass balance approach (Aryal et al., 2017a). The temporal variation of rainfall recharge was provided by the Australian Water Resources Assessment (AWRA) landscape model (AWRA-L). Spatially varying evapotranspiration based on vegetation-dependent root zone depths was represented using the EVT package of MODFLOW-USG.

The major rivers and selected minor reaches, were represented in the model using MODFLOW-USG river package. These represented the points at which the groundwater model interacts with the surface water model. The river package enabled two-way exchange of water between groundwater and river: when the river stage is above the water table the river can leak to the groundwater, when the water table is above the river stage the groundwater model will contribute baseflow to the river. A total of 11,785 groundwater extraction bores are included in the model. Each is assumed to extract water according to its full entitlement. A detailed description of the development of this groundwater model is available in Sreekanth et al (2017a)

Table 1: Conceptualisation of hydrostratigraphy units and numerical model layers for the formations of the Gunnedah and Surat basins

Province	Period	Formation	Layer in geologic model	Layer in GW model	Hydrostratigraphic unit
Namoi Alluvium	Pleistocene	Narrabri Formation	1	1	aquifer
Namoi Alluvium	Pliocene	Gunnedah Formation	2	2	aquifer
Namoi alluvium	Miocene	Cubbaroo Formation	2	2	aquifer
Surat Basin	Cretaceous	Rolling Downs Group and Liverpool Range Volcanics	3	3-5	Inter-burden
Surat Basin	Creteceous	Blythsdale Group	3	3-5	Inter-burden
Surat Basin	Jurassic	Pilliga Sandstone	4	6	aquifer
Surat Basin	Jurassic	Purlawaugh Formation	5	7-9	Inter-burden
Surat Basin	Jurassic	Garrawilla Volcanics	6	7-9	Inter-burden
Gunnedah Basin	Triassic	Napperby and Deriah formations	7	7-9	Inter-burden
Gunnedah Basin	Triassic	Digby Formation	7	7-9	Inter-burden
Gunnedah Basin	Permian	Black Jack Group – Coogal and Nea Subgroup	7	7-9	Inter-burden
Gunnedah Basin	Permian	Hoskissons Coal	8	10	Coal
Gunnedah Basin	Permian	Black Jack Gropu – Brothers Subgroup	9	11-13	Inter-burden
Gunnedah Basin	Permian	Watermark Formation	9	11-13	Inter-burden
Gunnedah Basin	Permian	Porcupine Formation	9	11-13	Inter-burden
Gunnedah Basin	Permian	Upper Maules Creek Formation	10	11-13	Inter-burden
Gunnedah Basin	Permian	Maules Creek coal seams	10	14	Coal
Gunnedah Basin	Pemian	Lower Maules Creek Formation	10	14	Coal
Gunnedah Basin	Permian	Goonbri Formation	NA	15	Basement
Gunnedah Basin	Permian	Leard Formation	NA	15	Basement

Two model parameterization and predictive analysis approaches were tested in this project using the numerical groundwater model developed for the Namoi sub-region. These are named: 1) the parsimoniously parameterized modelling approach and 2) highly parameterized modelling approach. In the current study for data-worth analysis and spatial design of monitoring networks we chose the highly parameterized approach for undertaking predictive analysis of CSG-induced drawdown changes and particle tracking. This is because, the highly parameterized approach enabled a better description of the spatial heterogeneity in the hydraulic characteristics of the aquitard and aquifer sequences in the Gunnedah and Surat basins underpinned by measured data available from core samples (Turnadge et al., 2017) and down hole porosity logs and geostatistical models that conform to this dataset. The highly parameterized modelling approach was also more tenable for explicit analysis the model sensitivity and data-worth of observations and parameters to many spatially explicit predictions.

2.3   Modelling approaches

The parsimoniously parameterized approach was found to be useful in quantifying potential extreme changes in flux (95th percentile) induced by coal seam gas development as reported in the preliminary report from the GAB flux project (Sreekanth et al., 2017a). This approach is further being used for testing hypothesis about potential maximum flux losses from the GAB aquifer induced by CSG development and that will be reported in the companion report hypotheses testing on GAB fluxes. The parsimoniously parameterized approach was initially developed in the Bioregional Assessments Programme for modelling cumulative drawdown impacts from CSG and coal mine development (Sreekanth et al., 2017a) and was found to be most useful for quantifying potential impacts with high confidence (95th percentile). The parsimoniously parameterized approach used 81 parameters and identified a sub-set of 37 parameters to be varied in sensitivity and uncertainty analysis. These parameters related to hydraulic properties of 15 model layers, boundary conditions including diffuse recharge, flood recharge, irrigation recharge, river exchanges, and model stresses including coal seam gas development and coal mining. Observed depth dependence of hydraulic properties was accounted using a depth-dependent-decay function for these parameters. This approach helped reduce the number of hydraulic property parameters in the model while still accounting for the spatial heterogeneity. A detailed description of this approach is reported in the groundwater modelling report for the Namoi sub-region produced as part of the Bioregional Assessments Programme.

The second approach tested in this study employed 1672 parameters for parameterizing the groundwater model. Many of these parameters are used for characterizing the heterogeneity and spatial variability in the hydraulic properties in multiple layers of the model. Also, parameters were used to represent flood and irrigation recharge and river conductance. The spatial variability adopted in this parameterization approach is based on the data analysis and upscaling of hydraulic properties undertaken by Turnadge et al. (2017) as part of the Faults and Aquitards project for the Gunnedah Basin.

Turnadge et al (2017) collected core samples of key aquitard sequences within the Gunnedah and Surat basin formations in the region and correlated measured permeability with available geophysical borehole data. Porosity-permeability relationships were derived based on this and where applied to downhole porosity logs obtained from 97 exploration wells located across the

Gunnedah Basin (Turnadge et al., 2017). These were then used to estimate vertical distributions of permeability at various locations across the Gunnedah Basin. They applied upscaling approaches to upscale these core scale aquitard vertical hydraulic conductivity (Kv) values for inclusion in a regional-scale numerical groundwater model (CDM Smith, 2016). Further, they used this upscaled data to fit geostatistical variogram models to characterize the spatial co-variance of hydraulic properties in the aquitard sequences. A spherical variogram with a Sill of 0.764, Nugget of 0.327 and Range of 129 km was used to define the co-variance structure of hydraulic properties horizontally in different model layers. The highly parameterized modelling approach used in the current study is underpinned by this spatial covariance structure for characterising prior uncertainty in hydraulic properties.

Realizations of Kv for one model layer (layer 8) corresponding to the upper aquitard sequence generated using the above described approach is shown in Figure 2. The realizations of Kv for the same layer obtained using the parsimoniously parameterized approach are shown in Figure 3.
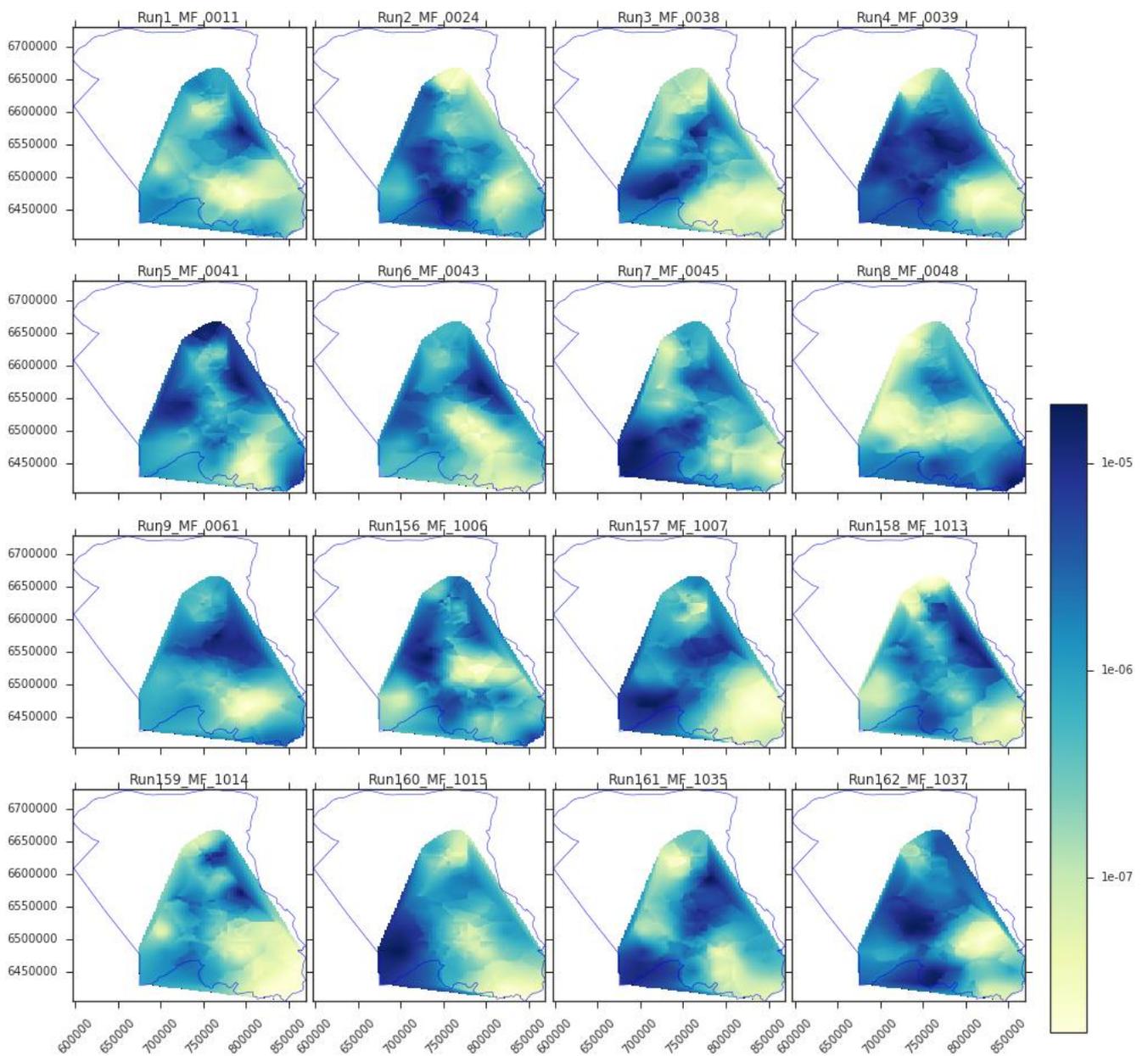
Figure 2: Sixteen out of 500 realizations of the vertical hydraulic conductivity of model layer 8 representative of the upper aquitard sequence generated in the highly parameterized model using a spherical variogram

Each of these two parameterization approaches have their own specific advantages. While the parsimoniously parameterized modelling approach was useful for providing conservative estimate of extreme changes (95th percentile) in drawdown and flux changes, the highly parameterized approach, being underpinned by data and geostatistical models on spatial variability of hydraulic properties, is likely to provide more realistic estimates of the median values of drawdown. The highly parameterized approach also provides better flexibility in computing parameter contributions to prediction uncertainty and analysis of data-worth. Given that these two are important focus areas of this study higher emphasis is provided in describing the details of the latter approach in this report. However, comparison of the 5th, 50th and 95th percentiles of drawdown predictions provided by the two approaches are provided in section 2.4. Additional modelling analysis and results relevant for data-worth analysis and optimal monitoring design including particle tracking analysis for water quality impact screening are described in the subsequent sections of this report.

2.4 Model calibration and uncertainty analysis

Calibration of a groundwater model using groundwater level/pressure and other observations is most valuable when the information available from observations are useful in constraining the model to make the predictions of interest. While groundwater observation from many bores are available for the Namoi region, most of these bores are in the Namoi alluvium and only a small proportion of the bores screen deeper formations in the Surat and Gunnedah basins. Thus, these observations are not very useful for constraining the model when the predictions of interest is propagation of CSG-induced drawdown in GAB and other formations. Hence calibration of the model using water level observations from surficial aquifers is of lesser significance in this study. Nonetheless, the model has been used together with the parameter estimation software suite (PEST) that enabled linear uncertainty and data-worth analyses that are based on First Order Second Moment (FOSM) sensitivities.

While model calibration was not in the immediate scope of the predictive analysis undertaken in this study, a notional calibration error analysis was undertaken using PEST as such an analysis provides useful insights about errors in the model structure and parameters. Model calibration to improve the matches between observed and predicted water levels were not warranted for this study and it is noteworthy that such a calibration exercise would also require appropriate updating of the boundary conditions including the well pumping rates represented in the model. The distribution of errors is also used to inform the observation error variance that is required for the linear uncertainty analysis. The errors were analysed for a combination of parameters that resulted in median predicted flux losses from the Pilliga Sandstone formation reported in Sreekanth et al. (2017a). Scatter plots of observed versus modelled heads is shown in Figure 4. The distribution of errors is shown in Figure 5. The error distribution showed that nearly 75% of the errors are in the range -10 m to 10 m. The errors in this range are more likely to be occurring because of uncertainties in the model parameters and boundary conditions. For example, in the current model set up average pumping rates are assigned to wells in respective aquifers based on total annual allocation under a conservative assumption that allocations are fully extracted every year. This is a conservative assumption used in the BA and GISERA studies and not considering potential impacts from resource development under an extreme scenario. Similarly, the parameterization that is evaluated for error analysis assumes homogenous values of hydraulic properties in the alluvial aquifers. Analysis of the time series of modelled and observed heads at several observations wells showed that the model simulations capturing the seasonal trends in water level changes in alluvial aquifers. This is reflective of the representation of the diffuse, irrigation and flood recharge underpinned by catchment hydrological models. Smaller errors in calibrating to these observations could be improved by accounting for the heterogeneity and spatial variability of hydraulic parameters in the alluvial formations. The match can also be improved using a more realistic representation of the alluvial groundwater pumping.
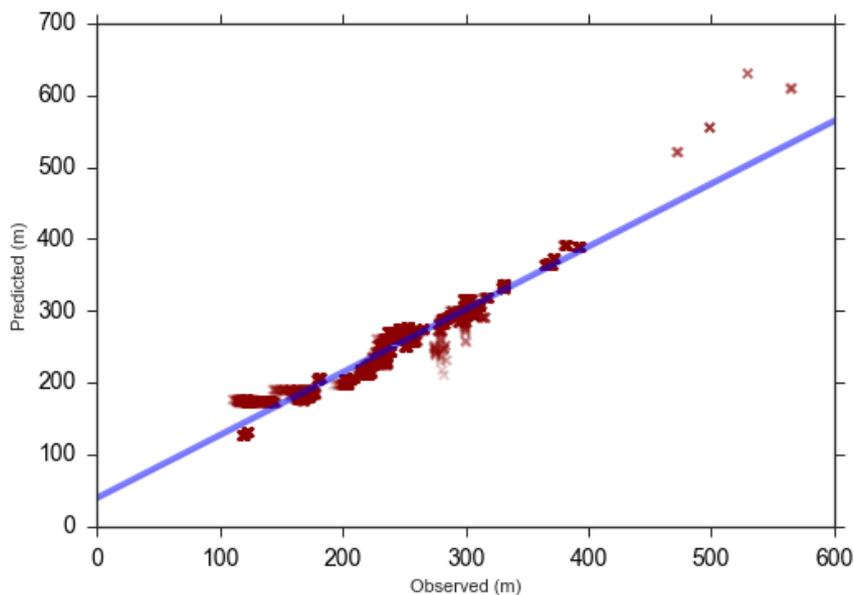
Figure 4: Scatter plot of observed versus simulated heads
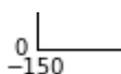
Error (m)

Figure 5: Distribution of prediction errors from limited calibration analysis

A small proportion of large errors indicate the possibility of some model structural errors and errors in aquifer assignments. The companion report on hydrogeochemical analysis (Raiber and Suckow, 2018) identifies the need for improving the assignment of bores in the region to distinct aquifers by matching information in bore completion reports to appropriate geological models.

As discussed earlier, the focus of the modelling undertaken in the current study is to investigate spatial trends in the groundwater drawdown in different model layers resulting from CSG development. The most important parameters that affect the drawdown propagation and flux across distinct model layers are the hydraulic characteristics of intervening formations. Sensitivity analysis reported in the Bioregional Assessments' modelling report for Namoi subregion (Sreekanth et al., 2017a) indicates that horizontal and vertical hydraulic conductivity values of the

model layers and spatial variability play the most important role in the prediction of drawdown and flux changes. Hence, a major focus of the study was placed on improved representation of hydraulic characteristics of the aquifer and aquitard sequences and quantifying the remaining predictive uncertainty and trends in drawdown for informing monitoring decisions.

Two different approaches of uncertainty analysis were applied. These are the non-linear uncertainty analysis based on Monte Carlo simulations and linear analysis based on propagation of variances. These are described in the following sub-sections and Section 4.

2.5   Non-linear analysis using Monte Carlo simulations

The non-linear analysis applied the Monte Carlo simulation approach similar to that applied in the Bioregional Assessment Programme for the Namoi model except that the highly parameterized model is used for exploring prediction uncertainty resulting from parameter uncertainty accounting for heterogeneity in the spatial characteristics of hydraulic properties.

The impact of the uncertainty of model inputs and parameters on the prediction of CSG-induced drawdown was analysed by doing an ensemble of predictive simulations consisting of many model runs. Uncertainty in the model inputs including recharge, SW-GW interactions and evapotranspiration were explicitly included in the uncertainty analysis by defining parameters that are relevant to these inputs. Uncertainty in horizontal and vertical hydraulic conductivity and specific storage were included by characterising the spatially varying parameter field using pilot points. As discussed earlier this parameterization scheme resulted in a total of 1672 parameters in the model.

The drawdown was predicted as the difference between the baseline and CSG development cases. This is accomplished by taking the difference between model predictions of two possible states of GAB groundwater resource – one corresponding to no depressurisation and the other corresponding to depressurisation based on a generic case of gas development. This accounted for the large scale uncertainty in CSG water production. The non-linear analysis resulted in simulation of CSG water extraction in the range 4.4 GL to 107.1 GL. This approach assumes that all other uses of groundwater remain unchanged over this period and quantifies only the changes in flux and water balance resulting from depressurisation only.

A total of 500 parameter combinations were evaluated for their predictive responses of CSG-induced drawdown. Latin Hypercube Sampling was undertaken to accomplish a representative sampling of these parameters from a 1672-dimensional parameter space. Characterization of the spatially varying hydraulic property fields from the pilot points was underpinned by observed spatial trends based on the data provided by Turnadge et al (2017). This enabled us to constrain the prior distribution of parameters underpinned by observed co-variance structure compared to the uniform distribution of parameters adopted in previous studies. This provided a significant improvement for the predictive analysis by directly helping to reduce the prediction uncertainty of CSG-induced drawdown impacts.

The 5th, 50th and 95th percentile of predicted drawdown in model layer 6, corresponding to the GAB aquifer, are show in Figure 6. The results show that the expected value of drawdown in the GAB aquifer is very small. The 95th percentile drawdown plot provides an estimate of probable

maximum drawdown that may be incurred in the GAB aquifer because of CSG development. The results indicate that CSG-induced drawdown in the GAB aquifer is unlikely to exceed 1.3 m.
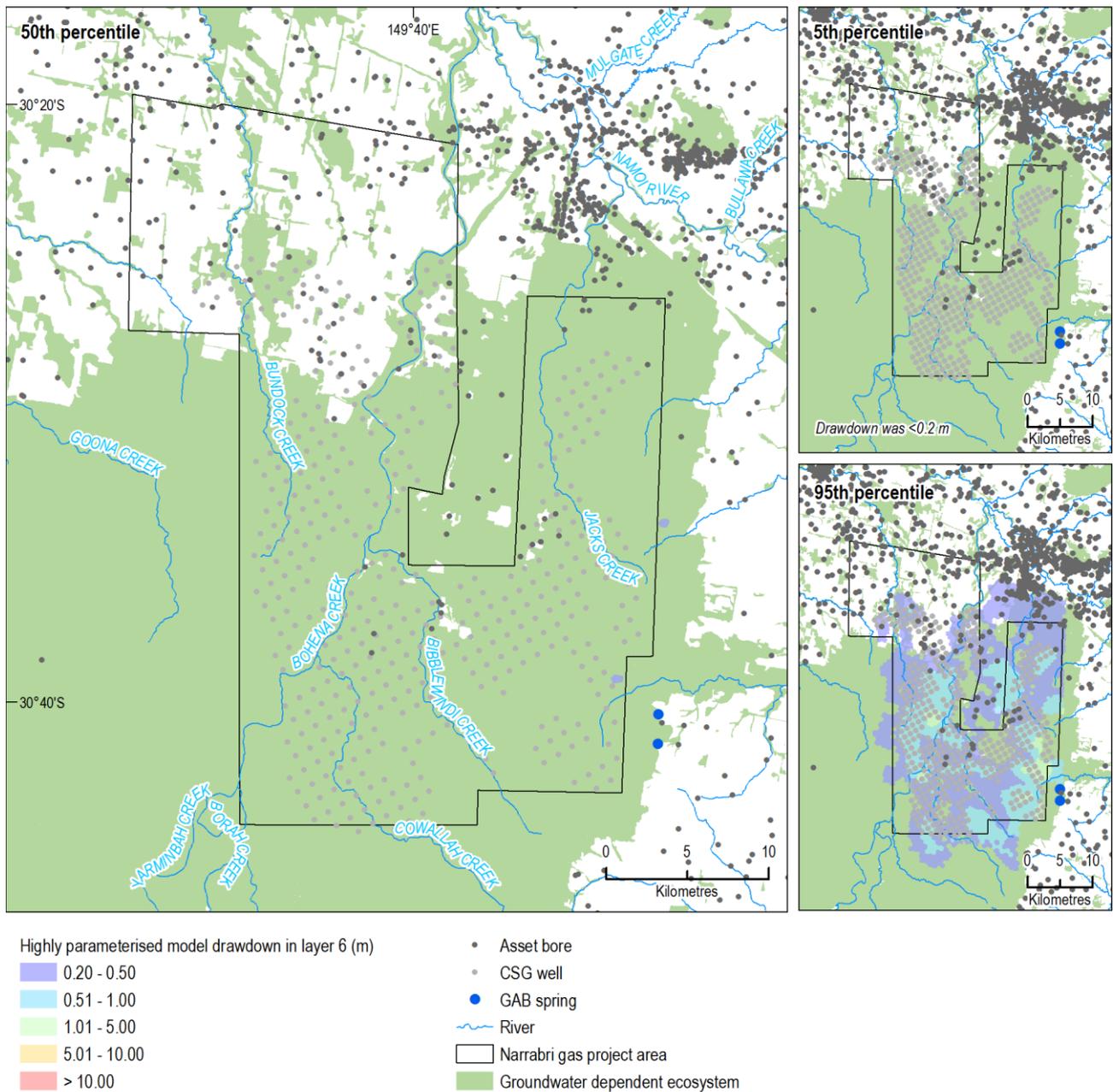
Figure 6: Percentiles of predicted drawdown in model layer 6 corresponding to the GAB aquifer Pilliga Sandstone obtained using the highly parameterized modelling approach

2.5.1 Comparison with parsimoniously parameterized approach

The Monte Carlo simulation was repeated using another set of 500 parameter combinations generated using the parsimoniously parameterized approach. Realizations of the 37-parameter combinations were generated using Latin Hypercube Sampling to ensure uniform sampling across the 37-dimensional parameter space. This approach resulted in 5th, 50th and 95th percentile of predicted drawdown in model layer 6 corresponding to the GAB aquifer as show in Figure 7. It may be observed that a larger magnitude and spatial extent of predicted drawdown has resulted from using this approach. This is because the parsimoniously parameterized modelling approach used

more conservative assumptions in characterizing the hydraulic characteristics in model layers corresponding to the aquitard and aquifer formations. Comparison of histograms of predicted drawdown at four selected receptor locations are also shown in Figure 8. Histograms of predicted drawdown from 40 other receptor locations are given in Appendix A.
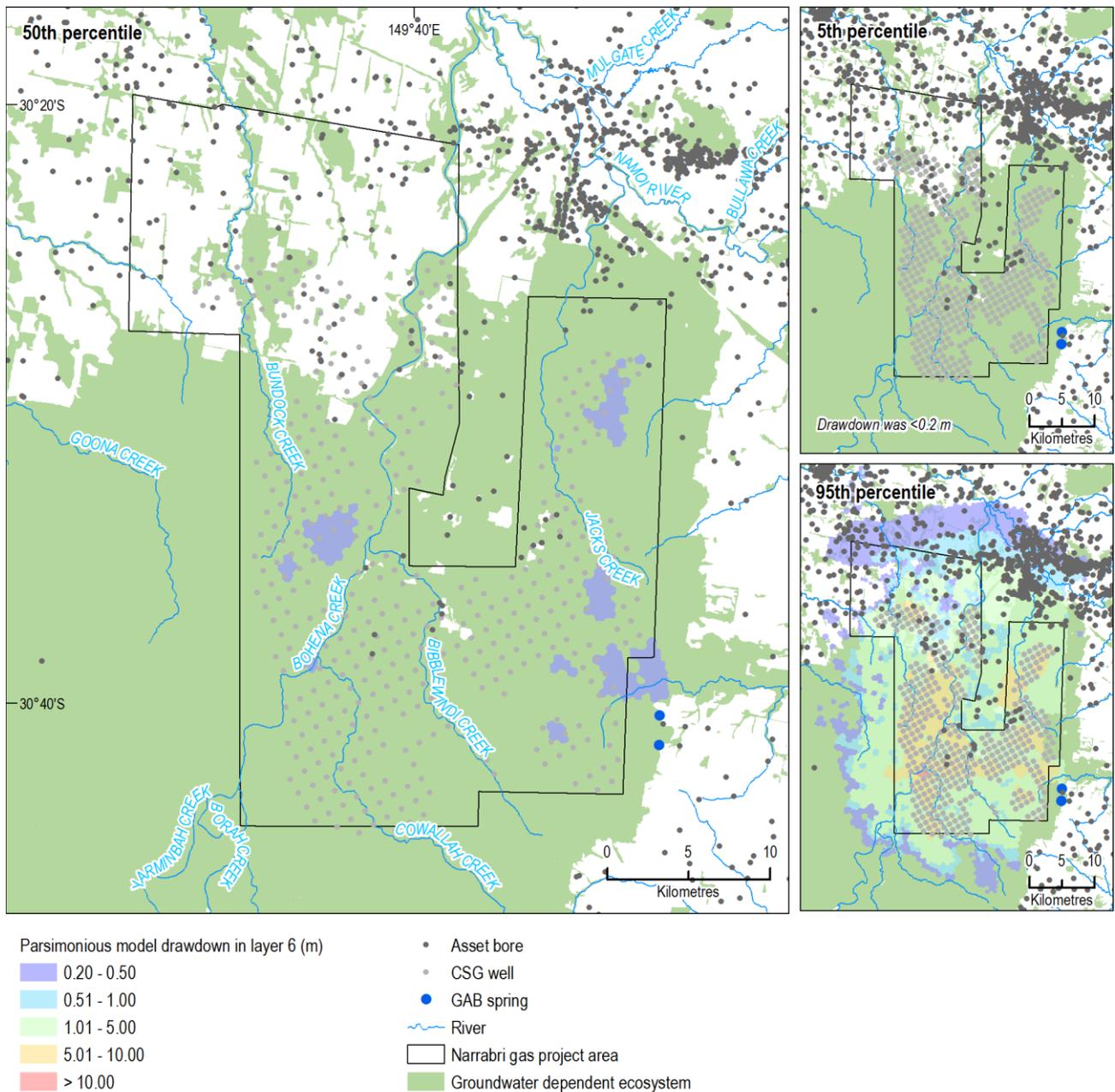
Figure 7: Percentiles of predicted drawdown in model layer 6 corresponding to the GAB aquifer Pilliga Sandstone obtained using the parsimoniously parameterized modelling approach

It was observed from the histograms that the distribution of predicted drawdown at many nodes are similar between the two approaches. But, in general the parsimoniously parameterized approach resulted in distributions with longer tail towards the maximum drawdown as compared to the highly parameterized approach. The lower number of parameters used in the former enables the sampling of extreme combinations of parameters and thus is more likely to result in prediction of largest drawdown impacts. However, for a relatively smaller number of nodes, the

highly parameterized modelling approach resulted in larger values of predicted maximum (95th percentile) drawdown than the parsimoniously parameterized approach.
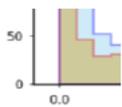
Figure 8: Distribution of predicted drawdown at four selected model nodes corresponding to receptor locations.

The standard deviation of predicted drawdown in model layer 6 obtained from the highly parameterized approach is shown in Figure 9. This contour plot is indicative of the predictive uncertainty in drawdown across the region. The standard deviation of predicted drawdown for selected locations obtained from the non-linear analysis is also compared with the estimates from linear analysis in Section 5.
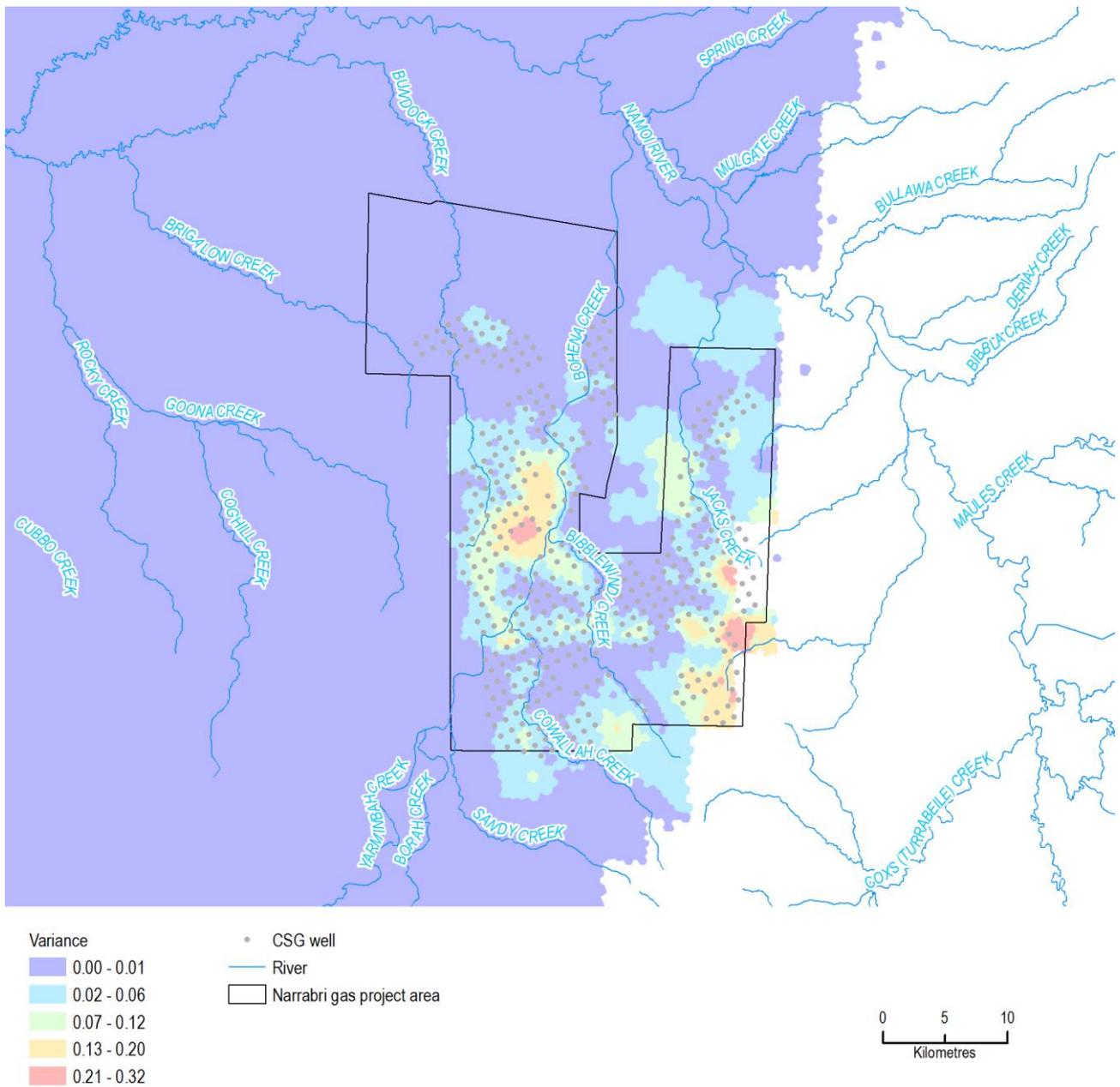
Figure 9: Standard deviation in drawdown predictions for model layer 6 from the highly parameterised model.

3 Probabilistic particle tracking as screening analysis for water quality impacts

Particle tracking analysis was undertaken for a screening analysis of water quality risks posed by the CSG wells. The contamination pathway explored in this analysis specifically related to the CSG well becoming a transport pathway for migration of contaminants or poorer quality water from deeper formations to migrate to the closest GAB aquifer: the Pilliga Sandstone. Specifically, the particle tracking analysis was undertaken to do a screening analysis to quantify the likelihood of such contaminant particles flowing along the groundwater flow direction and travelling to different risk receptors including GAB bores, springs or GDEs in the CSG development areas. Should such contamination pathways pose a high risk to these receptors, the monitoring network within the gas development areas may need to be designed with the objective of sentinel monitoring so that the monitoring bores are able to detect water quality changes well in advance of the contaminants travelling to sensitive risk receptors.

If contaminant particles reach the aquifer through CSG well bore de-bonding, the contamination risk for any receptor in the region is informed by the proximity of the receptor from the CSG well and the travel time between the CSG well and the receptor. The most important factor that influences the travel time between the source and the receptor is the groundwater flow velocity. Particle tracking analysis uses estimates of groundwater flow velocity from the groundwater flow model to calculate the distance travelled by particles and the corresponding travel time. Particles could be tracked forward in time whereby particles released from the sources could be tracked forward in time to calculate the distance travelled during specified time intervals until they reach a discharge point of a receptor location of interest. Alternatively, particles could be tracked backward in time from the receptor location to the source point. Both these approaches were implemented in this study. The particle tracking was undertaken using the recently developed MOD-PATH3DU software (Papadopoulos, 2016) that is compatible with MODFLOW-USG model developed using unstructured grids (Sreekanth et al., 2017a). The details of these analyses are presented in the following sub-sections.

3.1 Forward particle tracking analysis

In the forward particle tracking analysis, we simulated the travel path lines of particles released from 409 CSG well locations within the Narrabri Gas Project area. The analysis was undertaken for a conservative scenario assuming the well bore de-bonding results in the leakage of contaminant particles into the Pilliga Sandstone formation of the GAB aquifer. While well bore de-bonding occurring for all CSG wells at the same time is very unlikely, this scenario was constructed to investigate the travel paths from all CSG wells, assuming that accidental damage is equally likely for any of the wells. The simulation of particle travel paths was undertaken considering water extraction for CSG development.

While the focus of the monitoring network design presented in this study is for a simulation period of 100 years, one deterministic particle tracking for a much longer period of 3000 years was

undertaken to evaluate the travel path lines over a very long period. This simulation considered the model parameter combinations that resulted in the median predicted drawdown. The highly parameterized approach was used for solving the flow model for this simulation. The resulting path lines of travel over the 3000 year period is shown in Figure 10.
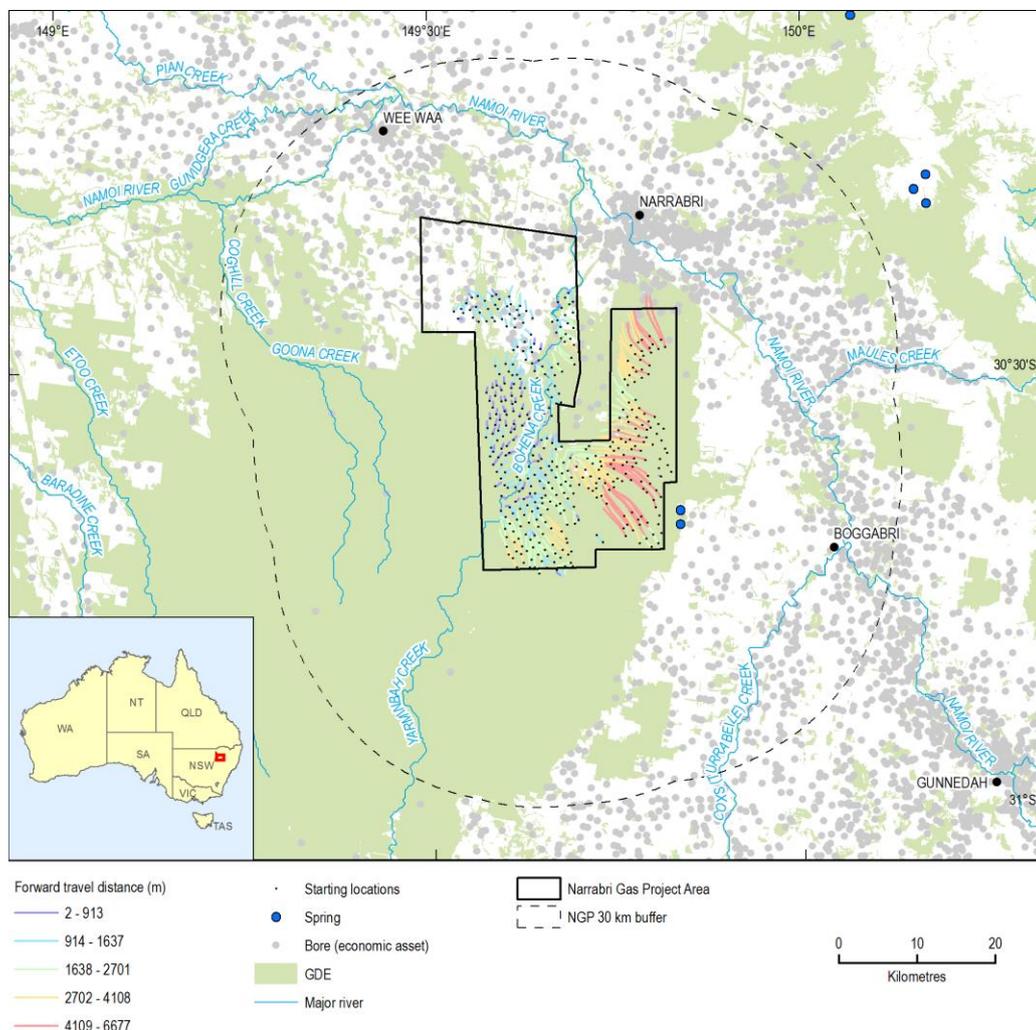

Figure 10: Forward particle tracking analysis over 3000 years from CSG wells to risk receptors

It may be observed from the map in Figure 10 that the travel path lines don't travel very far from the CSG wells in 3000 years. The maximum distances travelled by particles released from any CSG well location in the GAB aquifer is 6.7 km. If we consider this in perspective with the demarcated area for the Narrabri Gas Project, the particles don't exit the project area boundary except in the north eastern corner where the particles exit the boundary in about 3000 years. This indicated that contamination risks resulting from CSG well bore de-bonding to farmers' bores and other risk receptors beyond the project area can be ruled out with high confidence.

Subsequently probabilistic forward particle tracking analysis was undertaken for a simulation period of 100 years. Five hundred flow fields simulated using the highly parameterized flow model provided the head distributions required for undertaking the Monte Carlo particle tracking runs. The porosity characteristics in conjunction with the flow field for the aquifer (and aquitard formations) play important roles in defining the particle path travel times and path lines.

Randomly sampled porosity values sampled from the prior distribution were paired with different flow fields obtained from the Monte Carlo simulation of the flow model for undertaking the particle tracking model runs.

The particle travel path lines obtained from the Monte Carlo simulation for a period of 100 years between 2012 and 2102 is shown in Figure 11. These simulations assumed that a connection between the CSG target formations and the GAB aquifer were activated during the drilling of CSG bores resulting in release of contaminant particles. The results show that distances up to 1.1 km are travelled by particles within the shorter simulation period. It was noticed that particles travelled relatively larger distances towards the east of the Narrabri Gas Project area. This is because the GAB aquifer is outcropping in this area and relatively higher velocities of flow may be observed in this region than in deeper parts of the aquifer.

A total of 44 bores and 1036 km^2 of GDEs were within the maximum travel distance of 1127 m in any direction from CSG wells (from the multiple 100 year forward tracking runs). A total of 134 bores, 2325 km^2 of GDEs, and 2 GAB springs were within the maximum travel distance of 6677 m in any direction from CSG wells (from the single 3000 year forward tracking run). The results from the Monte Carlo simulations for the 100-year period showed that none of the receptor bores would be intersected by any of the particle tracks within the 100-year period.
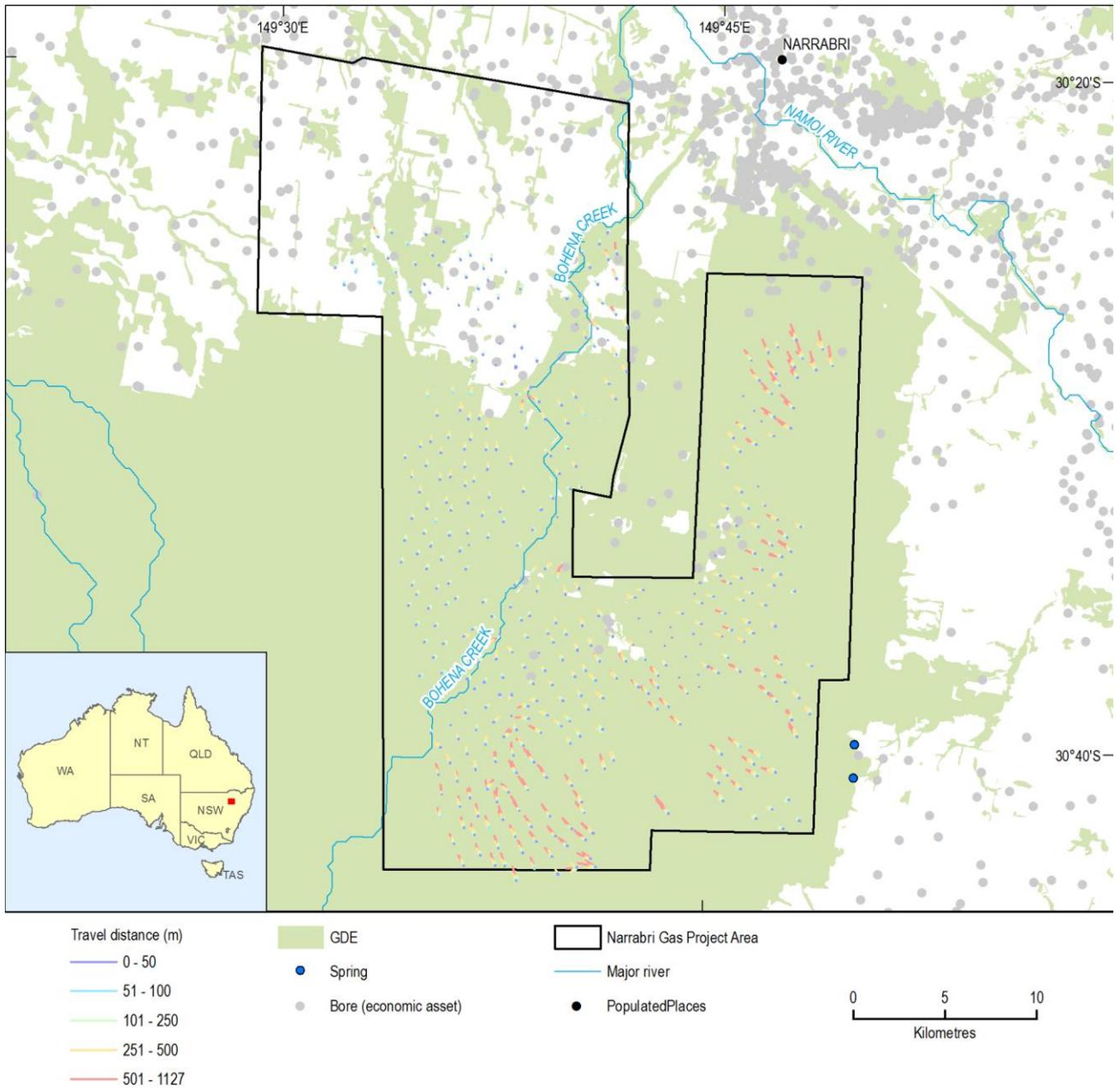
Figure 11: Forward particle tracks from multiple simulations over 100 years

The average velocity of particles tracked in the forward direction from multiple runs over 100 years was 0.39 m/yr (± 0.64 m), and the median velocity was 0.17 m/yr. Figure 12 presents a histogram of particle tracking velocities from this analysis.
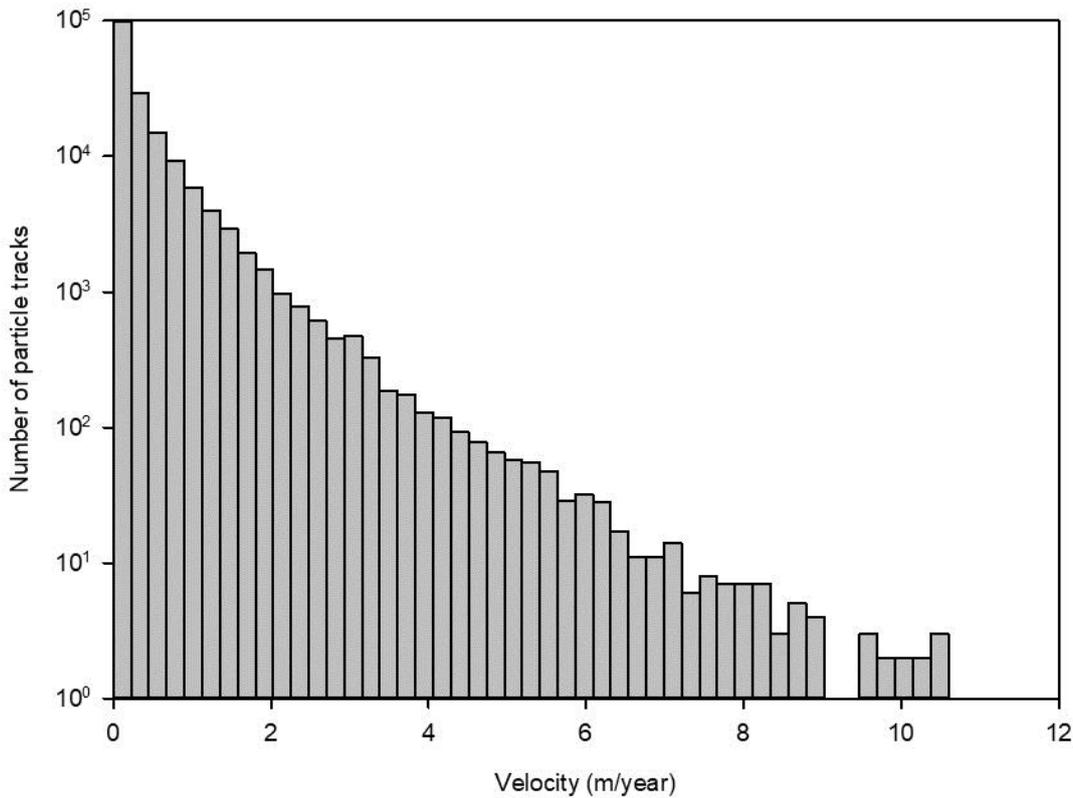
Figure 12 Frequency distribution of particle velocities from multiple forward tracking runs over 100 years.

Receptor locations and spatial clustering in relation to CSG wells and forward particle tracking paths are presented in Figure 13. Economic asset bore density was highest in north-east of the project area. Path line density (from multiple forward tracking runs over 100 years) was generally higher on the eastern side of the project area where particles tended to move greater distances. Asset bore distances from CSG wells were shortest along the northern edge of the well field in the general particle travel direction. There were 15 particle tracks whose trajectories were within 5 degrees of asset bores however none of the path lines reached the bores over the 100 year run time of this model.

The analysis indicated that there is relatively higher number of receptor bores in the north west of the gas project area. Particle track path line density is higher along the eastern edge of the gas project area and the receptor proximity to gas wells is highest towards the north west. Although the simulated forward particle tracks did not intersect with any of the receptor bores in the northwest of the gas project area, the areas where receptor bores are located within 1 km of CSG wells would be of more importance for monitoring water quality changes. Given that no receptor bores were intersected by probabilistically simulated particle tracks water quality monitoring objective was not included in the optimization based design of monitoring networks.
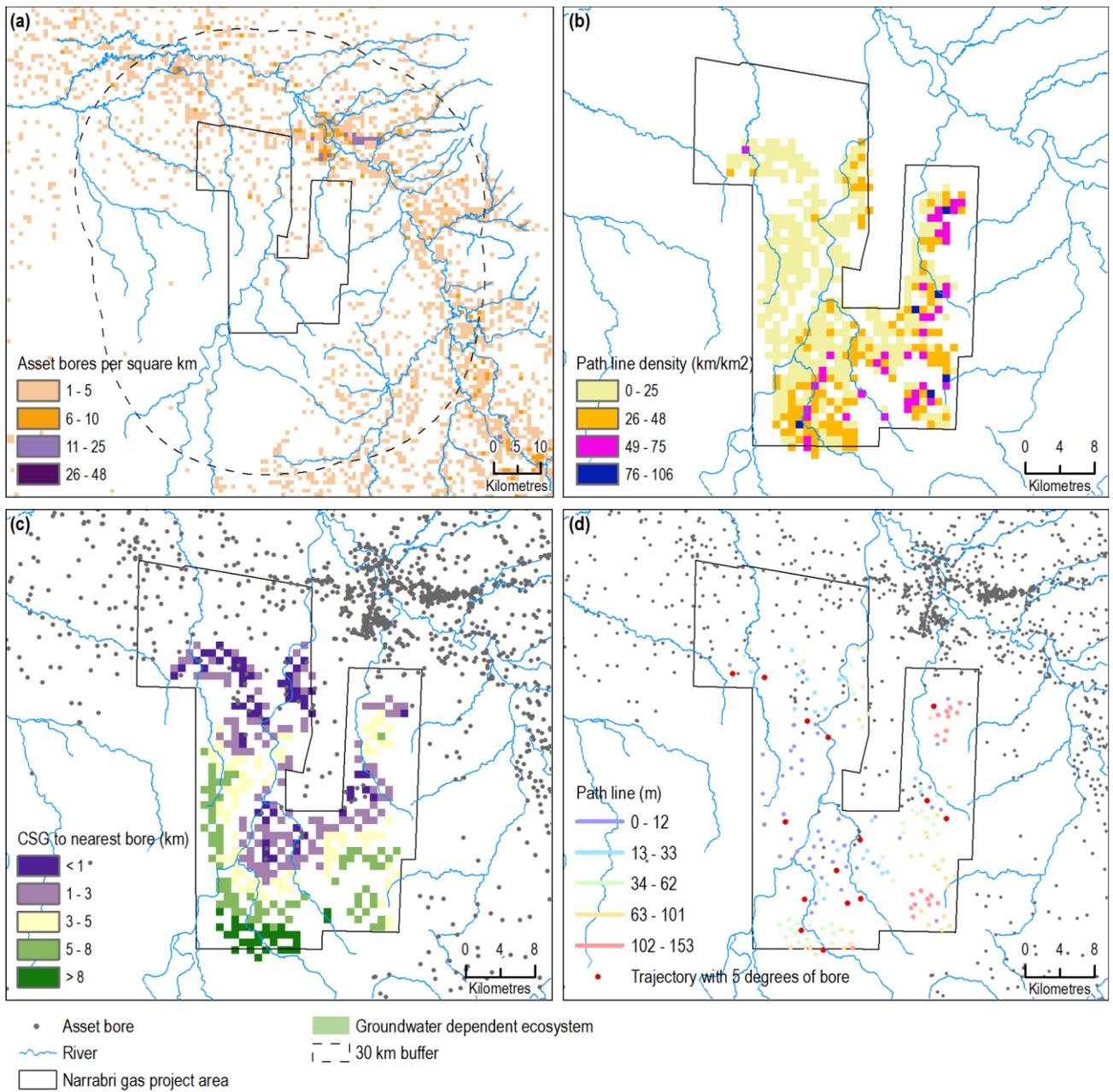
Figure 13: Forward tracking path line densities and trajectories in relation to receptor locations, proximities and densities

3.2   Reverse particle tracking analysis

Reverse particle tracking was undertaken to investigate the likelihood of the groundwater flow arriving at selected receptor locations that are at different distances within the 30-km buffer of the Narrabri Gas Project area within a 100 year simulation period may carry particles that were released from the CSG well bores. Similar to the forward particle tracking a single run of the particle tracking model was initially undertaken to investigate reverse particle tracks over a long period of time. The highly parameterized model parameter combinations that resulted in median predicted drawdown was used for this simulation. The particle tracks that resulted from this 3000-yr simulation is shown in Figure 14.

Similar to the forward particle tracking analysis, it was found that the longest distance of the reverse particle tracks within a 3000-year simulation period was 8.7 km. It was also observed that reverse particle tracks were closer to the outcrop of the Pilliga Sandstone formation along the eastern margin of the Narrabri Gas Project area (Figure 14).
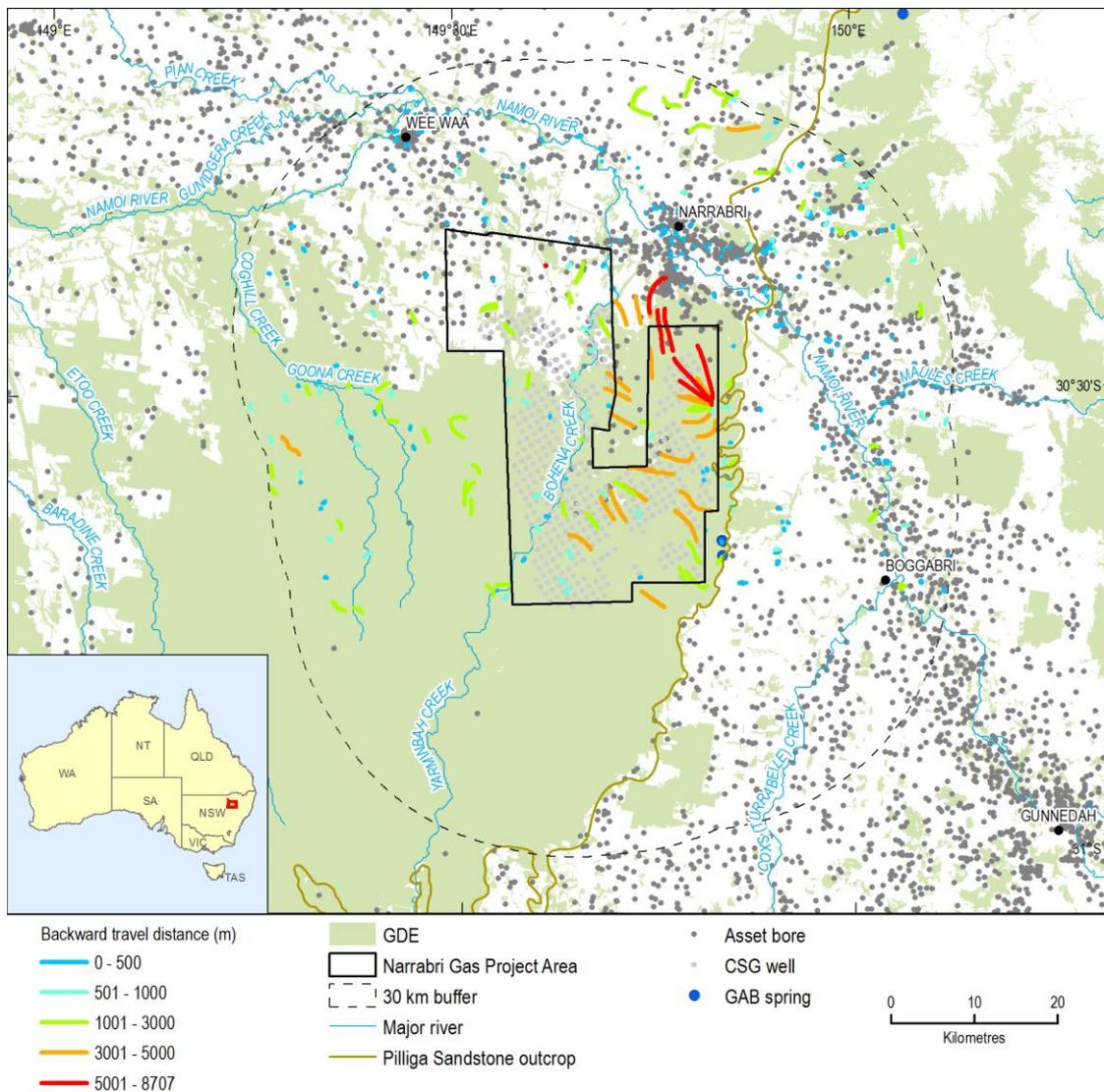

Figure 14: Backward particle tracking over 3000 years from receptors to the CSG wells

The reverse particle travel path times from one simulation for a period of 100 years is shown in Figure 15. The results showed that distances up to 0.5 km were travelled by particles within the shorter simulation period. Similar to forward particle tracks, particles travelled relatively larger distances towards the east of the Narrabri Gas Project area. The forward particle tracking analysis was further used in conjunction with spatial proximity and receptor density analyses as reported in the following section.
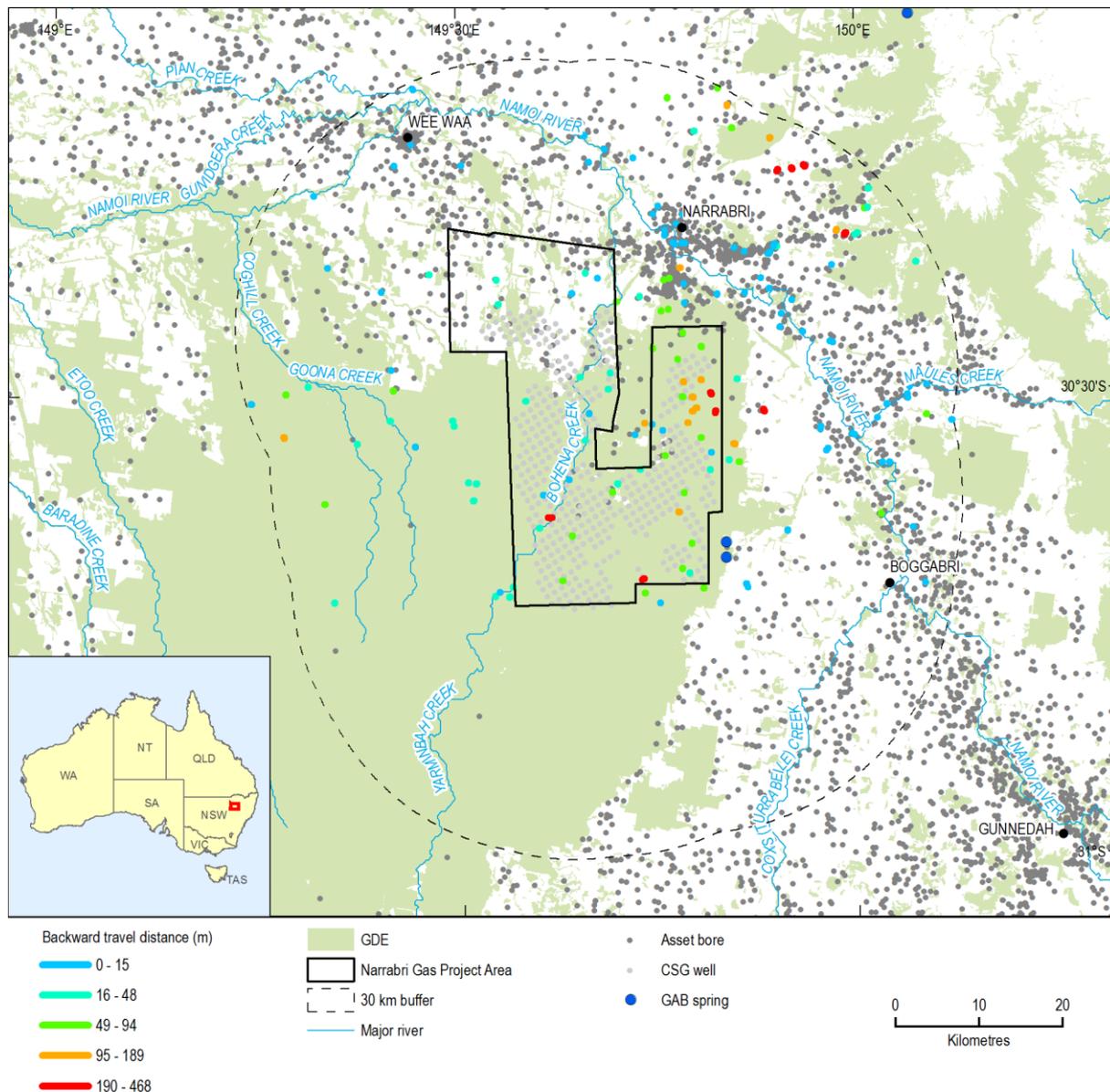

Figure 15: Reverse travel distances over 100 years simulated for selected risk receptors within the 30 km-buffer from the Narrabri Gas Project area

The analyses of travel times and distances both in forward and backward direction illustrated that groundwater flows very slowly in the GAB aquifer in this region. Independent analysis of groundwater flow velocities and recharge rates from environmental tracer analysis reported in the companion GISERA report (Raiber and Suckow, 2018) also estimated similarly small groundwater flow velocities with a maximum flow velocity of 3 m/year estimated from the spatial distribution of C-14 in the region. These results indicate that contamination risks for receptors located beyond the boundaries of the Narrabri Gas Project area are relatively low. Spatial analysis of the proximity

of receptors to the particle tracks within the Narrabri Gas Project area was undertaken to quantify the risks for these receptors. This analysis is reported in Section 5 of this report.

The particle tracking analysis undertaken in this study is restricted to investigating the particle travel times and distances from CSG wells and is used as a screening model for analysing the risks and to inform monitoring decisions. A recent study (Mallants et al., 2017) has used particle tracking modelling results together with dilution and attenuation characteristics of different CSG chemicals to perform contamination risk analysis with respect to individual chemicals. Such an analysis was not warranted in this study as the primary focus here was on data-worth analysis and spatial design of monitoring network. However, methods reported in Mallants et al. (2017) could be applied to the probabilistic particle tracking results from our study to undertake probabilistic quantification of contamination risks in this region accounting for the dilution and attenuation characteristics of different chemicals involved.

It is also noteworthy that the particle tracking resulted in a large number of travel paths and distances that are smaller than the cell size of the groundwater model used in this study. This indicates that groundwater flow velocities are small in this region. Hence it is suggested that more detailed studies on contamination transport, if warranted, should be done at a well-scale in order to better understand local scale contamination risks for well-defined scenarios. The approach we have used provided a quick and easy method for regional scale screening analysis of contamination risks from well integrity loss. Contamination risk quantification at local scales from events such as accidental spillage, leakage from lined ponds, oil tanks etc. are not dealt with in this study and are not within the scope of regional scale data-worth analysis and monitoring network design. A spatial analysis was undertaken in our study based on CSG and receptor bore density, proximity and travel path line characteristics were used to identify the necessity of water quality monitoring in a regional scale from a contamination risk/compliance monitoring perspective. The results of this analysis are presented in Section 5 of this report.

4   Spatial analysis of drawdown

4.1   Receptors within drawdown areas

The number or extent of risk receptors within areas where modelled drawdown in model layer 6 (corresponding to the GAB aquifer Pilliga Sandstone) exceeded 0.2m at 5th, 50th and 95th percentile levels are summarised in Table 2. No asset bores (production bores) or GAB springs are within drawdown areas at or below the median level of drawdown and there are no potential impacts at the 5th percentile level. Considerable areas of GDEs, mainly comprising stands of Eucalypt forest and similar vegetation are predicted to be within median and 95th percentile drawdown areas.

Table 2 Summary of risk receptors within areas of predicted drawdown in layer 6 exceeding 0.2m at 5th, 50th and 95th percentile levels for highly parameterised and parsimonious models.

	Highly parameterised model drawdown ≥0.2m		
Receptors	**5th**	**50th**	**95th**
No. asset bores	0	0	106
No. GAB springs	0	0	0
GDE area (km²)*	0	2	1437
	Parsimonious model drawdown ≥0.2m		
Receptors	**5th**	**50th**	**95th**
No. asset bores	0	0	650
No. GAB springs	0	0	0
GDE area (km²)*	0	71	2261

*NB mainly comprised of vegetation (Eucalypt forest), GDE features overlap each other

The intersection of drawdown areas with receptors is shown in Figure 16 for the highly parameterised model predictions.
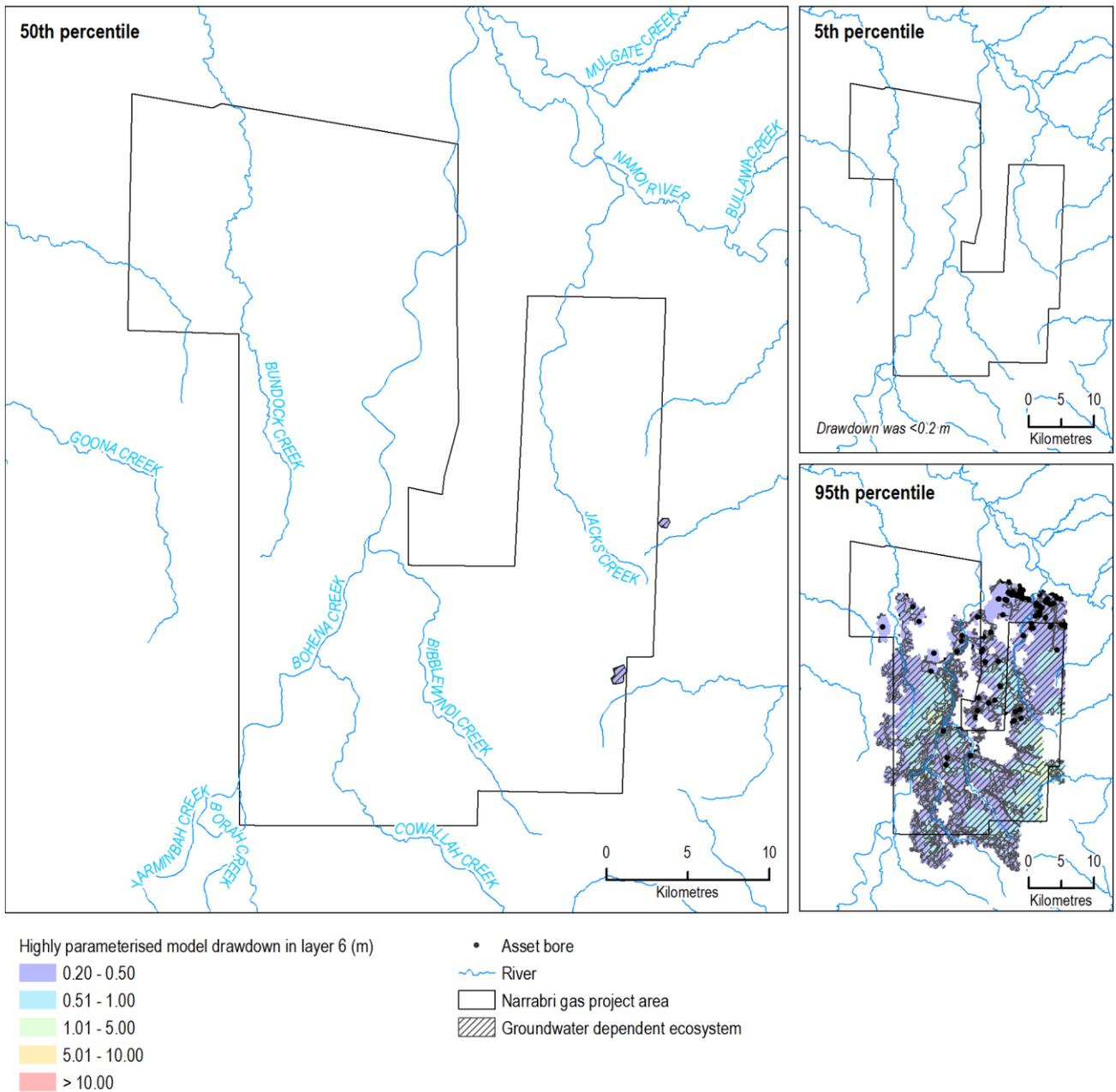
Figure 16 Receptors intersecting predicted drawdown areas ≥0.2 m in layer 6 from the highly parameterised model

4.2   Monitoring bore densities

The total area of predicted groundwater drawdown ≥0.2 m from the highly parameterised runs of 100 years was 811 km^2 at the 95th percentile level and 1 km^2 at the median level. There was no drawdown ≥0.2 m predicted at the 5th percentile. Current GAB monitoring bores within the 95th percentile drawdown area totalled 34 giving an average bore density of 4.2 bores per 100 km^2. Adding 10 new monitoring bores, as demonstrated in the following section on monitoring network design, in this area would increase density to 5.4 bores per 100 km^2. Current GAB monitoring bores are shown in Figure 17 with respect to GDEs and economic bore assets and the 95th percentile drawdown area.
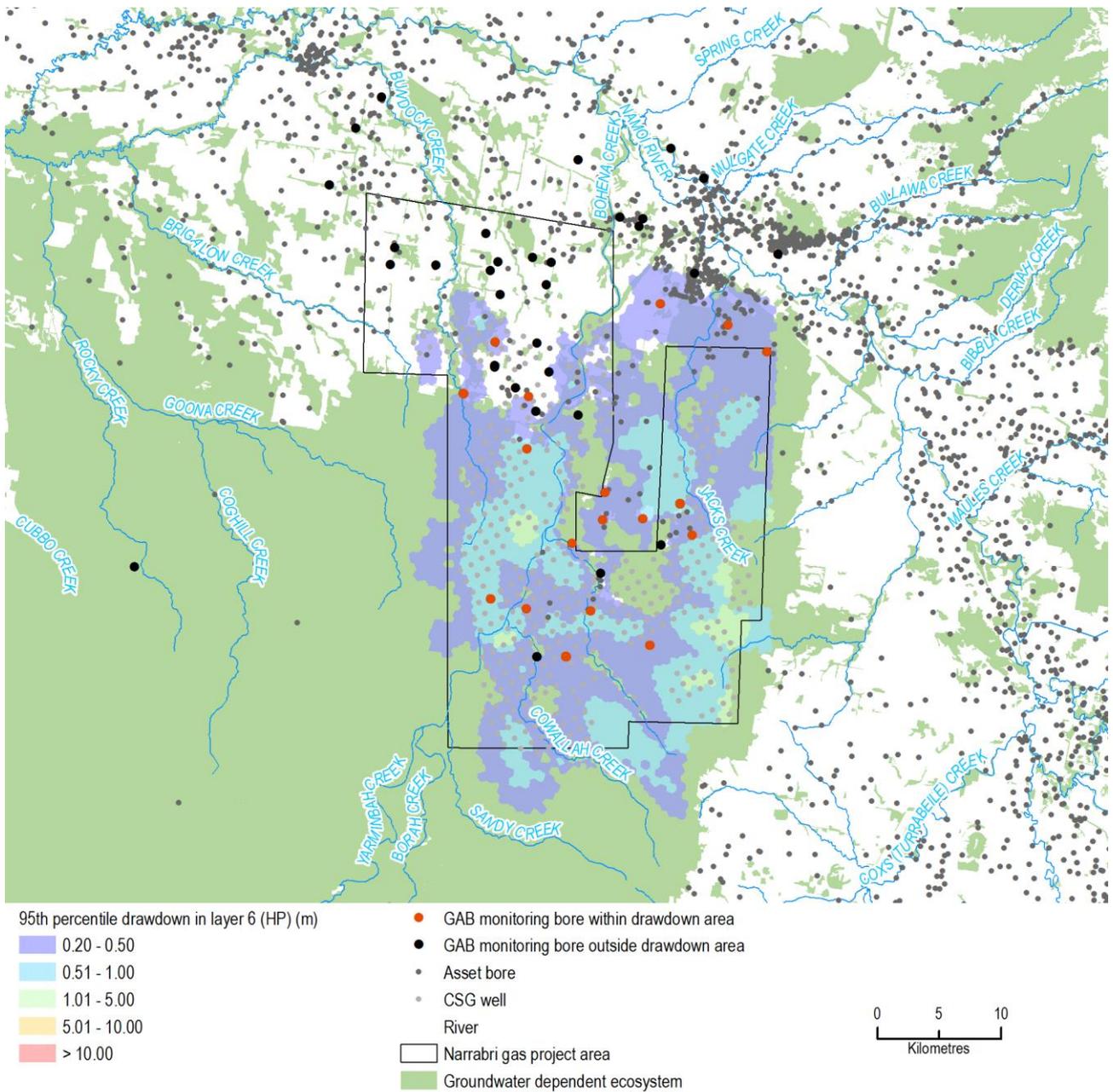
Figure 17: Current GAB monitoring bores within the 95th percentile drawdown area (≥0.2 m) from the highly parameterised model

5 Data-worth analysis

Methods based on linear uncertainty analysis can be used for evaluating the value of information used in calibration process. Because of the linearity assumption, First Order Second Moment (FOSM)-based methods can be employed with computational efficiency for purposes like obtaining increased prior knowledge about parameters, evaluating the worth of potential new observations or evaluating the worth of existing observations and parameter contributions to uncertainty. The theoretical basis of these approaches have been discussed in Fienen et al. (2010), Dausman et al. (2010), Tonkin et al. (2007) and White et al. (2016). Two different methods, but both based on similar linear analysis, were tested in this study to evaluate parameter estimability using existing and new data sets and relative worth of existing and new measurements.

Data-worth analysis was conducted to analyse the worth of currently existing and future monitoring data to inform prediction of CSG-induced drawdown at selected receptor locations. Fifty-seven receptors, hydrogeologically connected to different aquifer formations (and equivalent groundwater model layers) located within the 95th percentile drawdown extent (≥0.2 m), were used for the analysis. Estimability of model parameters using existing calibration data sets informs whether or not, and the extent to which, the existing data sets are able to constrain uncertainty in the CSG-induced drawdown prediction. The methods used for this analysis do not require the actual values for parameters or observations. Instead, only the sensitivities of predictions and observations with respect to parameters are used. This implies that the relative worth of future observations in informing the predictions of interest can also be calculated using these approaches. In this study we calculated the relative worth of drawdown measurements at 100 monitoring points across 5 model layers in informing the drawdown predictions.

5.1 Selected list of predictions/receptors for data-worth analysis

A total of 57 model nodes were selected for data-worth analysis. These nodes corresponded to locations, within which, risk receptors were identified and fell within 95th percentile predicted drawdown in the near surface aquifer (from all model layers). These nodes are shown in Figure 18.
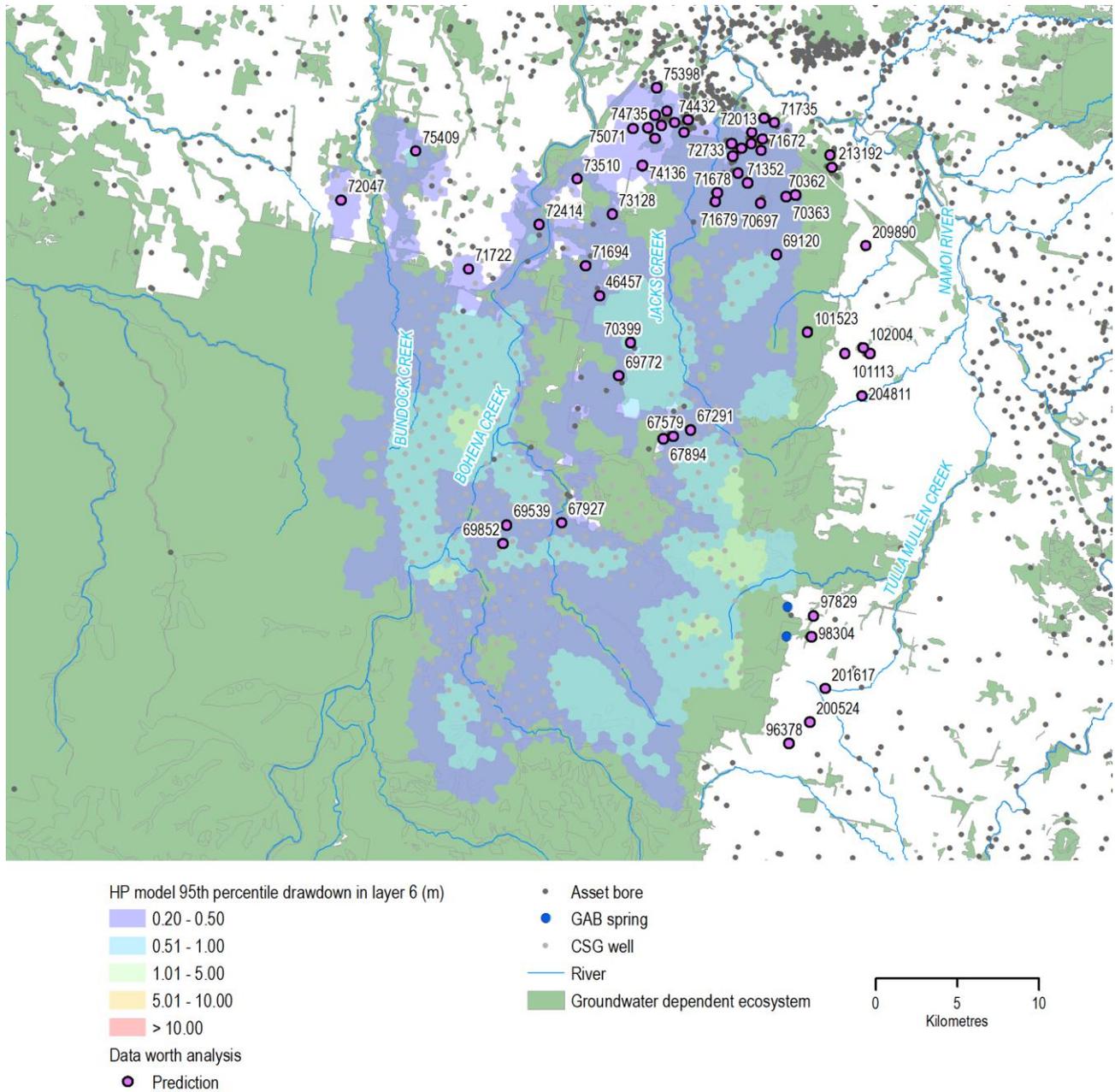

Figure 18: Selected receptor and potential monitoring bore locations with respect to predicted drawdown ≥0.2 m in model layer 6

5.2   Methods for linear analysis

The first method is based on Schur's complement. Schur's complement for linear uncertainty analysis can be viewed as a form of Bayes equation under the assumption of a linear model and a multivariate Gaussian distributions to describe the distribution of parameters, forecasts and observation noise (White et al., 2016, Tarantola, 2005, Fienen et al., 2010, Doherty, 2015). The Schur's complement approach, implemented in the software utility pyEMU (White et al., 2016), uses the following equation to estimate the posterior parameter covariance matrix $\overline{\Sigma}_\theta$ as:

$$\overline{\Sigma}_\theta = \Sigma_\theta - \Sigma_\theta J^T (J\Sigma_\theta J^T + \Sigma_\epsilon)^{-1} J\Sigma_\theta \tag{1}$$

where Σ_θ is the prior parameter covariance matrix, Σ_ϵ is the epistemic observation noise covariance matrix, and J is the Jacobian matrix of partial first derivatives of observations with respect to parameters. This equation highlights the behaviour of the inversion process (White et al., 2016). The first term (i.e. Σ_θ) represents the parameter uncertainty prior to inversion, and the second term (i.e. $\Sigma_\theta J^T (J\Sigma_\theta J^T + \Sigma_\epsilon)^{-1} J\Sigma_\theta$) encapsulates the inversion process, through the Jacobian matrix and both parameter and observation covariance, as mapping of information from observations to parameters.

Prior and posterior uncertainty estimates for forecast s, defined as $\sigma^2{}_s$ and $\bar{\sigma}_s^2$ respectively, can be easily calculated by projecting the requisite parameter covariance matrix to the forecast output space using a forecast sensitivity vector:

$$\sigma^2{}_s = y^T \Sigma_\theta y \tag{2}$$

and,

$$\bar{\sigma}^2{}_s = y^T \overline{\Sigma}_\theta y \tag{3}$$

where y is the vector of prediction sensitivity to each parameter. The details of the implementation of this method in pyEMU is described in White et al. (2017).

The linear analysis conceives the model as a matrix acting on a set of parameters to simulate the model outputs. These outputs can be of different types, with some corresponding to existing observations and others being future predictions of the variables of interest. The Bayesian linear uncertainty analysis, employed in this approach is based on the sensitivity of the model outputs with respect to the inputs. Thus, the method is independent on the absolute value of the outputs. The model sensitivity to parameters is encapsulated in the Jacobian matrix which is the matrix of the derivatives of the outputs with respect to the model parameters. This uncertainty analysis captures the uncertainty (variance) in the parameter sets and propagates it through the model operator to estimate uncertainty in model predictions. The estimate of prediction uncertainty is approximate because of the linearity assumption. The theory behind the linear uncertainty analysis methodology implemented in the PEST suite is briefly described in the following. Details regarding this methodology have been reported elsewhere (Christensen and Doherty, 2008, Dausman et al., 2010, Engelhardt et al., 2013).

Let p be an m length vector comprised of model parameters that follow a multivariate Gaussian distribution with covariance $C(p)$. Further, define an operator matrix Z that translates model parameters p to vector h, where h represents the observations of the system state comprising the model calibration set contaminated by measurement noise ϵ. Then, it follows that

$$h = Zp + \epsilon. \tag{4}$$

Let s be a scalar that represents a prediction made by the model, and let the vector **y** represent the sensitivities of this prediction to model parameters. Then s can be calculated using the relationship:

$$s = y^T p. \tag{5}$$

Combining (4) with (5) we get:

$$\begin{bmatrix} s \\ h \end{bmatrix} = \begin{bmatrix} y^T & 0^T \\ Z & I \end{bmatrix} \begin{bmatrix} p \\ \epsilon \end{bmatrix}. \tag{6}$$

where 0 is the zero-vector and I the identity matrix. We therefore can get a covariance matrix $C((s, h)^T)$ for the parameter vector of Equation (6) by using the propagation of error formula:

$$C \begin{pmatrix} s \\ h \end{pmatrix} = \begin{bmatrix} y^T & 0^T \\ Z & I \end{bmatrix} \begin{bmatrix} C(p) & 0 \\ 0 & C(\epsilon) \end{bmatrix} \begin{bmatrix} y & Z^T \\ 0 & I \end{bmatrix} = \begin{bmatrix} y^T C(p)y & y^T C(p)Z^T \\ ZC(p)y & ZC(p)Z^T + C(\epsilon) \end{bmatrix}, \tag{7}$$

where $C(p)$ is the covariance of innate parameter variability as defined before. Christensen and Doherty (2008) combined equations (4) and (7) to give a Bayesian formulation of predictive uncertainty as encapsulated by Equation (8):

$$\bar{\sigma}^2{}_s = y^T C(p)y - y^T C(p)Z^T (C(p)Z^T + C(\epsilon)^{-1} ZC(p)y. \tag{8}$$

The estimation of prior uncertainty for the prediction is identical across these two approaches. We used this formulation to evaluate prediction uncertainty in CSG-induced drawdown impacts at chosen risk receptors in this study. For this purpose the $C(p)$ matrix was obtained from the observed parameter covariance structure underpinned by the variogram. Similarly the $C(\epsilon)$ matrix representing the observation error variance was obtained from the limited calibration analysis. Equation (8) also forms the basis for computing the worth of groundwater monitoring observations in reducing the prediction uncertainties.

The results of data-worth analysis are reported in this section. Observations corresponding to each bore location were sequentially added to an initially empty calibration data set and Equation (8) was evaluated each time to compute the decrease in the prediction uncertainty achieved by the addition of each measurement. In a similar exercise, observations were sequentially removed from an initially complete calibration data set and the increase in the prediction uncertainty was computed by evaluating Equation (8). The increase and decrease in the uncertainty estimates

upon exclusion and inclusion, respectively, of observations into the calibration data set helped to identify the relative worth of each observation in informing the prediction.

The approach implemented in the PYEMU (White et al., 2016) was used to quantify the predictive uncertainty at 30 receptor nodes in model layer 6. The predictive uncertainty estimates for the drawdown at these nodes obtained from the linear analysis in comparison to standard deviation obtained from non-linear analysis is reported in the following sub-section. The comparison shows that the prediction uncertainty standard deviation estimated from the non-linear and linear analyses are comparable, and shows similar spatial trends. This reiterated the sufficiency of the analysis based on a linear approximation for data-worth analysis.

5.3   Comparison of linear and non-linear prediction uncertainty estimates

The methods described above underpin the data worth analysis and assume a linear response of the model prediction with respect to the parameters. To ensure that this underpinning assumption does not limit the applicability of the method for this case study we compared the prediction uncertainty at selected monitoring locations obtained using Equation (2) to the prediction variance obtained from the non-linear MC simulation reported in Section 3. The comparison is shown in Figure 19. The comparison shows that the uncertainty variance obtained from the two approaches are similar, warranting the use of a linear approach for data-worth analysis. It is also noteworthy from the figure that there is significant uncertainty in the prediction of drawdown in deeper model layers closer to the CSG targets (Layer 10 and Layer 14). However, uncertainty in the model layers associated with risk receptors (e.g. Layer 6 corresponding to the GAB aquifer) show smaller prediction uncertainty and enable meaningful conclusions to be drawn about drawdown risks from CSG development.
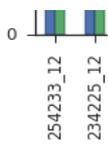
Figure 19: Comparison of uncertainty standard deviation in the prediction of drawdown at 30 model nodes obtained using linear and non-linear methods

5.4 Parameter estimability using existing calibration data set

The linear uncertainty analysis using the Schur's complement was implemented to quantify the percentage reduction in parameter uncertainty that can be achieved using the existing calibration data set. Forty parameters for which highest reduction in uncertainty can be achieved are shown in Figure 20. These parameters comprise of the multipliers for recharge, horizontal conductivity and specific storage pertaining mostly to aquifers close to the surface. The reason for this is fairly intuitive in that the observations used in the calibration data set comprises water level records from the alluvial and other water table aquifers. Similar results were produced by the analysis of parameter sensitivity to water levels in the Bioregional Assessments study using a parsimonious parameterization scheme for the same model. That sensitivity analysis showed that the observed water levels were most sensitive to parameters pertaining to river, diffuse recharge and hydraulic properties of the shallow aquifers.
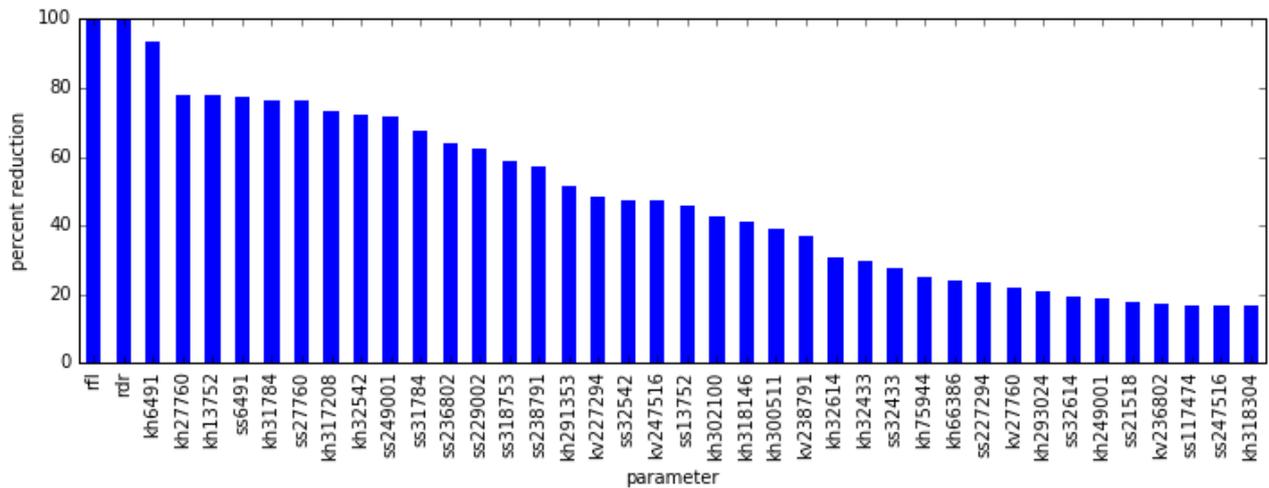
Figure 20: Parameter uncertainty reduction achieved using the existing calibration data set

The prior and posterior parameter uncertainty for the top 10 parameters are shown in Figure 21. These results imply that uncertainty in these parameters can be significantly reduced by having the observations present in the existing calibration data set.

Figure 21: Prior and posterior uncertainty for the 10 most estimable parameters

Further analysis was undertaken to calculate the parameter contributions to prediction uncertainty. The relative contribution of 5 parameter groups (kh – horizontal conductivity, kv – vertical conductivity, ss – specific storage, r- parameters pertaining to recharge and z – parameters pertaining to SW-GW interaction) to uncertainty in prediction of drawdown at the 57 receptors/monitoring bores of interest. The relative contribution of different parameter groups to prediction uncertainty at different receptor nodes is shown in Figure 22.
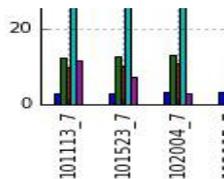
Figure 22: Relative contribution (unitless) of parameter groups to prediction uncertainty at 57 receptors/monitoring bores

Important information that is readily gleaned from these plots is that the prediction uncertainty of drawdown has distinct contributions from different parameter groups. The specific storage groups is the predominant parameter group for many receptors. Relatively higher contributions from the ss, z and kh groups are observed compared to the kv group. Uncertainty contributions of the z group are enhanced when the receptor is closer to the river network. For example, for the receptor nodes 69539, 69852 and 67927 are in outcrop areas of layer 6 and have a small 95th percentile predicted drawdown as shown in Figure 23. It may be observed from Figure 22 that the

parameter contribution of the river (z) group is highest for these receptors. This implies that the drawdown incurred at these bores depends on how fast/slow the CSG-induced flux losses from this area is replenished by additional inflows from (or reduction in base flows to) the river.

These results also indicate a more general implication that at this scale (noting that maximum drawdown is close to zero for the 50th percentile), drawdown incurred is more sensitive to how fast/slow the small flux losses from the aquifer are compensated by release of water from storage, river or recharge, than to the propagation of declining pressures in deeper formations to these aquifers.

This is because the aquitard formations that are present between the coal formations and aquifers, dampen the propagation of depressurization, and pressure changes are time lagged before they reach the GAB and water table aquifers. It is important, however, to note that vertical hydraulic conductivity is still an important parameter that governs the propagation of CSG-induced drawdown in to the upper layers. While this is intuitive, it could also be quantitatively evaluated by analysing the parameter contributions to prediction uncertainty of drawdown in model layers between the GAB aquifer and the coal bearing formations. The data-worth analysis for a selected number of potential monitoring bores is described in the following section.

5.5 Parameter and prediction uncertainty with respect to potential new monitoring data

A set of 20 potential monitoring bore locations were considered for this analysis. While these locations could virtually be anywhere within the model domain, we used the information available from the monitoring network design analysis reported in Section 6 to efficiently choose 20 preferred locations to do data-worth analysis. Data-worth analysis was undertaken to study the worth of data from multilevel piezometers at these locations. It is assumed that the multi-level piezometers will be installed to measure drawdown information at five different depths at these locations. The 20 locations chosen are shown in Figure 23.
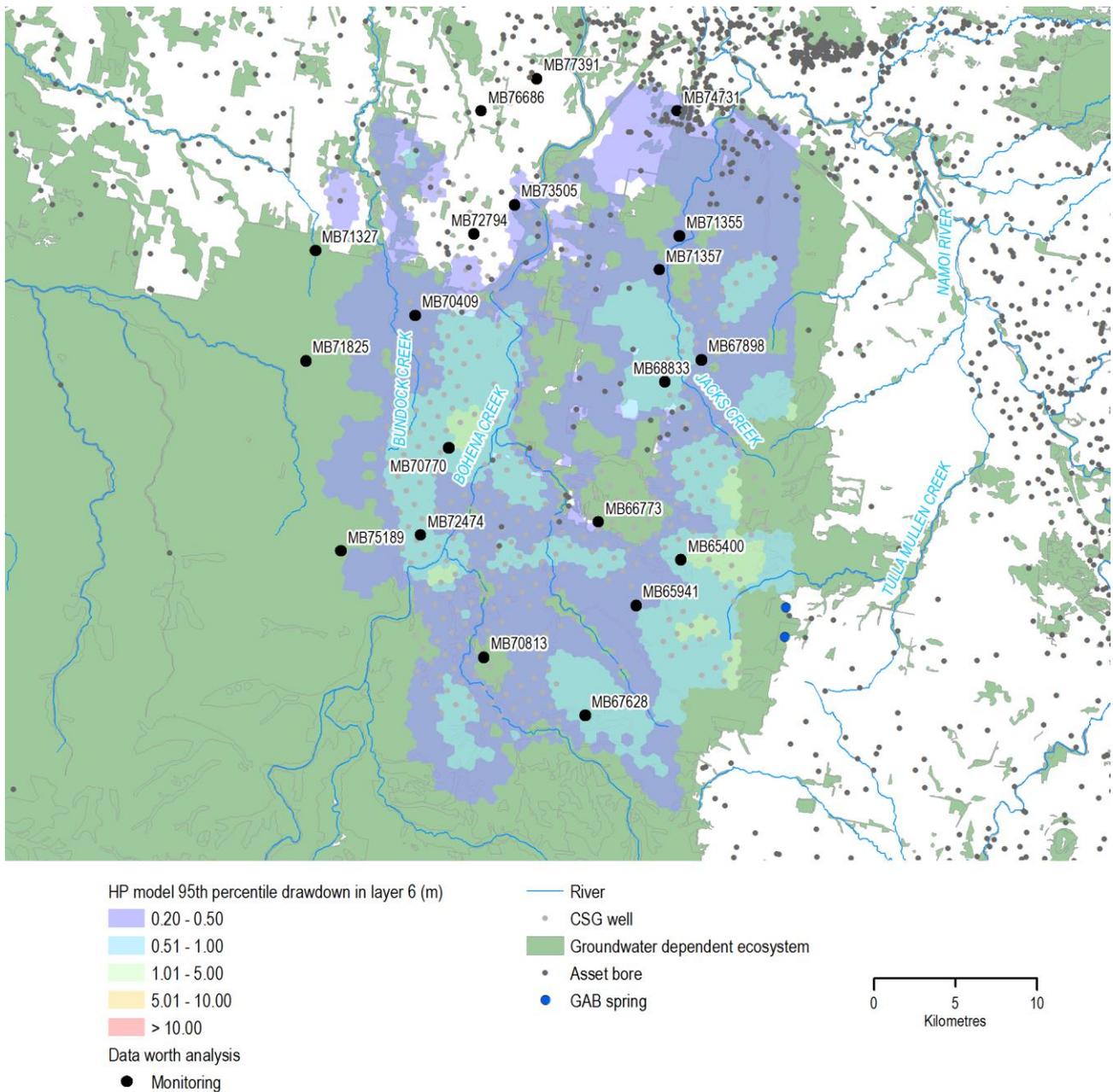
Figure 23: Locations of potential monitoring bores chosen for data-worth analysis with respect to 95th percentile drawdown ≥0.2 m in model layer 6

The data-worth and parameter contributions to uncertainty were evaluated layer-wise for these potential monitoring bores. The layer-wise analysis helps to understand which parameters contribute to prediction uncertainty of drawdown in that layer and, in turn, which monitoring data helps to reduce these parameter uncertainties. It is important to note that some of these model layers do not specifically follow any geological formations, except for the aquifer and coal bearing formations but roughly correspond to different inter-burden formations between the Hoskissons coal and Maules Creek formations and between Hoskissons coal and GAB aquifer formations. Thus, more generally the data-worth analysis for monitoring bores in different model layers inform the worth of drawdown measurements at different depths from the surface.

Model layer 8 corresponded to the inter-burden formation between the GAB aquifer (layer 6) and Hoskissons Coal formation (layer 10). Model layer 12 corresponded to the inter-burden between

the Hoskissons Coal and Maules Creek formations. Model layer 14 corresponded to the Maules Creek formation in which coal seams are present. The top 20 parameters for which parameter uncertainty reduction can be achieved by having monitoring bores in these layer are shown in Figure 24.
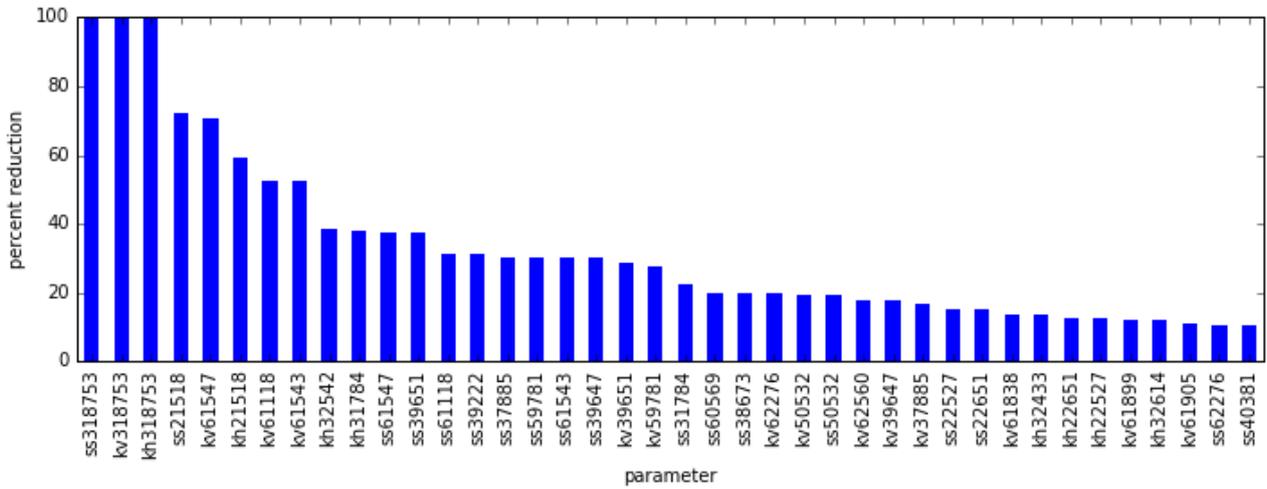
Figure 24: Parameter uncertainty reduction that can be achieved by using drawdown measurements from deeper model layers
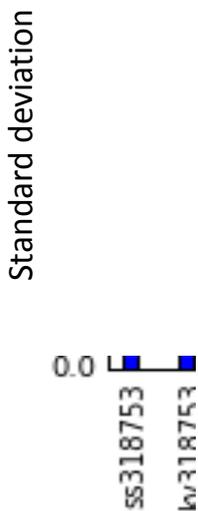
Figure 25: Prior and posterior parameter uncertainties for the top 20 parameters informed by additional monitoring data

Further insights can be inferred from Figure 25. Maximum percent reduction in uncertainty is achieved mostly for vertical hydraulic conductivity and specific storage values. Unlike the existing calibration data set which informs parameters pertaining to the horizontal hydraulic conductivity, storage, recharge and SW-GW interaction in shallower formations (Figure 21), and drawdown

measurements from deeper layers would help constrain predominantly parameters that influence the vertical propagation of depressurization. This implied that drawdown data obtained from these deeper layers together can inform the parameters that influence the propagation of drawdown upwards. The first three parameters for which maximum percent reduction in uncertainty can be achieved are the specific storage, horizontal and vertical conductivity at pilot point locations close to the CSG wells in the Maules Creek formation. However, other parameters for which uncertainty reduction can be achieved include specific storages, vertical and horizontal hydraulic conductivity in the layers above including the layers close to the surface. This indicates that drawdown measurements from the network can be used to systematically inform propagation of drawdown from the CSG target all the way to the surface. The prior and posterior uncertainty reduction in top 20 parameters using the data from the new monitoring data is shown in Figure 25.

Further, the parameter group contributions to drawdown prediction at these monitoring bore locations in deeper model layers were analysed. The parameter group contributions for drawdown predictions in multiple locations in model layers 8, 12 and 14 are shown in Figure 26. It is noteworthy that unlike the receptors/ bores used in the calibration analysis, vertical hydraulic conductivity is a key parameter informing the predictive uncertainty of drawdown at these new locations in deeper layers. This implies that new measurements of drawdown at these locations would inform the vertical hydraulic conductivity (and other hydraulic characteristics of deeper formations) which in-turn governs the vertical propagation of depressurization into upper layers where most of the risk receptors are present.

Relative data-worth of monitoring data from different depths in informing drawdown predictions can be calculated as a relative increase in prediction uncertainty owing to the removal of dataset from the calibration data set. Relative data worth was calculated for five different observation groups (i.e., monitoring data from model layers 7, 8, 12 and 14 and the observation group) corresponding to existing calibration data from shallower formation. Figure 27a illustrates the relative worth of monitoring observations in model layers 7, 8, 12 and 14 and existing observations from shallower wells in informing drawdown prediction at 20 locations within model layer 6 corresponding to the Pilliga Sandstone. It may be observed that the relatively highest data-worth for this purpose arise from observations in model layer 7, that lies immediately below model layer 6. This implies that measuring groundwater head in the Purlawaugh formation immediately below the Pilliga Sandstone (and evaluating it in comparison with pre-CSG heads) would be of relatively highest worth in informing drawdown predictions in the Pilliga Sandstone formation. Similar results were obtained from data-worth analysis for drawdown predictions in model layers 8 and 12 as shown in Figure 28. As readily observed from Figure 27 and Figure 28 when the prediction of interest is in a deep formation, monitoring information from deeper and adjacent layers provide relatively higher data worth. Figure 29 shows the relative data worth of these observation groups in predicting drawdown predictions in a number of identified receptors that are located within the 95th percentile drawdown extent. The results show that some of the existing calibration data provides the most data worth for some of these predictions while drawdown measurements from deeper formations also help to reduce prediction uncertainty. This result aligns with our previous interpretation that the parameters corresponding to shallower aquifers (hydraulic properties, recharge and SW-GW interactions) help to inform the drawdown at these receptor locations.

Percent contributions (unitless)

Monitoring bore

Figure 26: The hydraulic property parameter group percentage contributions to drawdown predictions at potential monitoring locations

a

b

Observation group

Figure 27: Relative data-worth of observation groups in informing a) drawdown prediction at monitoring locations in model layer 6 and b) drawdown prediction in model layer 7

a)

b)

Figure 28: Relative data worth of observation groups in informing a) drawdown prediction at monitoring locations in model layer 8 and b) drawdown prediction in model layer 12

a)

b)

0

Figure 29: Relative data worth of observation groups in informing drawdown prediction at receptor locations that are hydrogeologically connected to formations corresponding to model layers 6, 7 and 11

It is important to note that whilst separate observation groups have relatively distinct information content in constraining drawdown prediction in different model layers, monitoring bores also contribute to the prediction uncertainty at different locations, and exhibit spatial variability in their information content. This can be illustrated using spatial plots of relative data worth pertaining to drawdown predictions at individual locations. Spatial plots of relative data worth pertaining to two receptors are shown in Figure 30 illustrating two distinct characteristics of data-worth of a monitoring network.
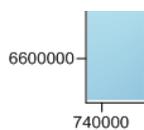
Figure 30: Spatial plots of relative data worth in predicting drawdown at two receptor locations in model layer 6 a) Node number 70399 and b) Node number 71722. The black dots indicate 20 potential monitoring bores from which the data worth was interpolated.

Figure 30a indicates that the close-by (potential) monitoring location MB68833 provides most data worth for drawdown prediction at node 70399. However, for the node 71722, relative data worth is spread across multiple bores in the network. Interestingly the second plot also indicates the redundancy in information when two monitoring bores are located close to each other (by low value of relative data worth in one of the two bores). It is noteworthy that the data worth analysis for potential monitoring bores considered two sets of 10 bores obtained from the monitoring network design analysis reported in the following section. Thus, one set contains redundant information and this is reflected in the data worth analysis.

6 Geostatistical analysis for baseline water quality and optimization

Another key aspect of understanding groundwater quality in the regional GAB aquifer formations, is an assessment of baseline water quality indicators. However, predictions of groundwater quality indicators are difficult, and often result in high levels of uncertainty, particularly in the face of the limited data commonly encountered in groundwater applications. Critically, interest is in determining locations for future boreholes which will reduce uncertainty in predictions of baseline water quality indicators within the GAB aquifer layers. We therefore outline a procedure in this section that will aide in determining optimal locations for monitoring baseline water quality values. The focus of this analysis is the hydrostratigraphic unit classified as Cadna-owie Hooray Sandstone and Pilliga Sandstone in the conceptual hydrogeological model developed as part of the Bioregional Assessments for the Namoi subregion (Aryal et al., 2017). While this is a composite unit comprising a few different hydrostratigraphic units of the GAB sequence, finer resolution of water quality data was not available for this study and hence we undertook the geostatistical analysis for the composite unit.

Specifically, we draw upon geostatistical methods to model baseline water quality variables of available data, namely kriging models. We use automated procedures to choose the appropriate variogram model and appropriate predictors in order to interpolate the baseline water quality variables across the outlined aquifer region and obtain a current measure of uncertainty. We then utilize a differential optimization algorithm in order to determine supplementary monitoring locations that would offer the greatest reduction in uncertainty across all baseline water quality variables of interest over the entire aquifer region of interest. Details and results of this analysis are provided in the remainder of this section.

6.1 Water quality parameters at monitoring locations

A sufficient amount of observations are needed in order to adequately estimate kriging models for baseline water quality variables. In what follows, we limit ourselves to baseline water quality variables in the study region for which at least five unique water quality observations were available. This resulted in 41 different baseline water quality variables for model fitting. The associated variables can be found in Table 3. These variables primarily include concentrations of various ions and other water quality parameters like pH and total dissolved solids. This resulted in 60 different observation locations with at least one baseline water quality variable of interest. Importantly, not all locations have each of the 41 baseline water quality variables chosen. The 60 locations can be found in Figure 32 as blue dots.

Additionally, for some of these variables additional predictors may be useful in explaining the spatial patterns observed. We therefore included the associated depth (corresponding to the mid-point of the hydrostratigraphic unit) of the Aquifer (herein labelled "depth") as a potential covariate in our analyses, as many water quality variables tend to vary with depth. Figure 31

shows associated depth values across the region of interest. For each of the 60 observation locations, we choose the associated depth value with location that the observation is closest to.

Table 3: Baseline water quality variables use for Kriging analysis

Variable	Unit	Variable	Unit
pH (lab)		Magnesium	mg/L
Alkalinity (Bicarbonate as CaCO3)	mg/L	Manganese	mg/L
Alkalinity (Bicarbonate)	mg/L	Nitrate (as NO3-)	mg/L
Alkalinity (total)	mg/L	Phosphate	µg/L
Alkalinity (total) as CaCO3	mg/L	Potassium	mg/L
Aluminium	mg/L	Silica	µg/L
Barium	mg/L	Sodium	mg/L
Boron	mg/L	Strontium	mg/L
Bromide	µg/L	Sulphate	mg/L
Calcium	mg/L	Uranium	
Chloride	mg/L	Zinc	µg/L
Copper	mg/L	Hardness	mg/L
Dissolved Oxygen	mg/L	d13C pptPDB	
Electrical conductivity (lab)	mg/L	Chromium	
Fluoride	mg/L	Total Organic Carbon	
Hydrogen	mg/L	Cobalt	
Helium	mg/L	Dissolved Organic Carbon	
Iodide	mg/L	Heterotrophic Plate Count (22°C)	
Iron	mg/L	Nickel	
Lead	mg/L	Total Dissolved Solids at 180°C	
Lithium	mg/L		
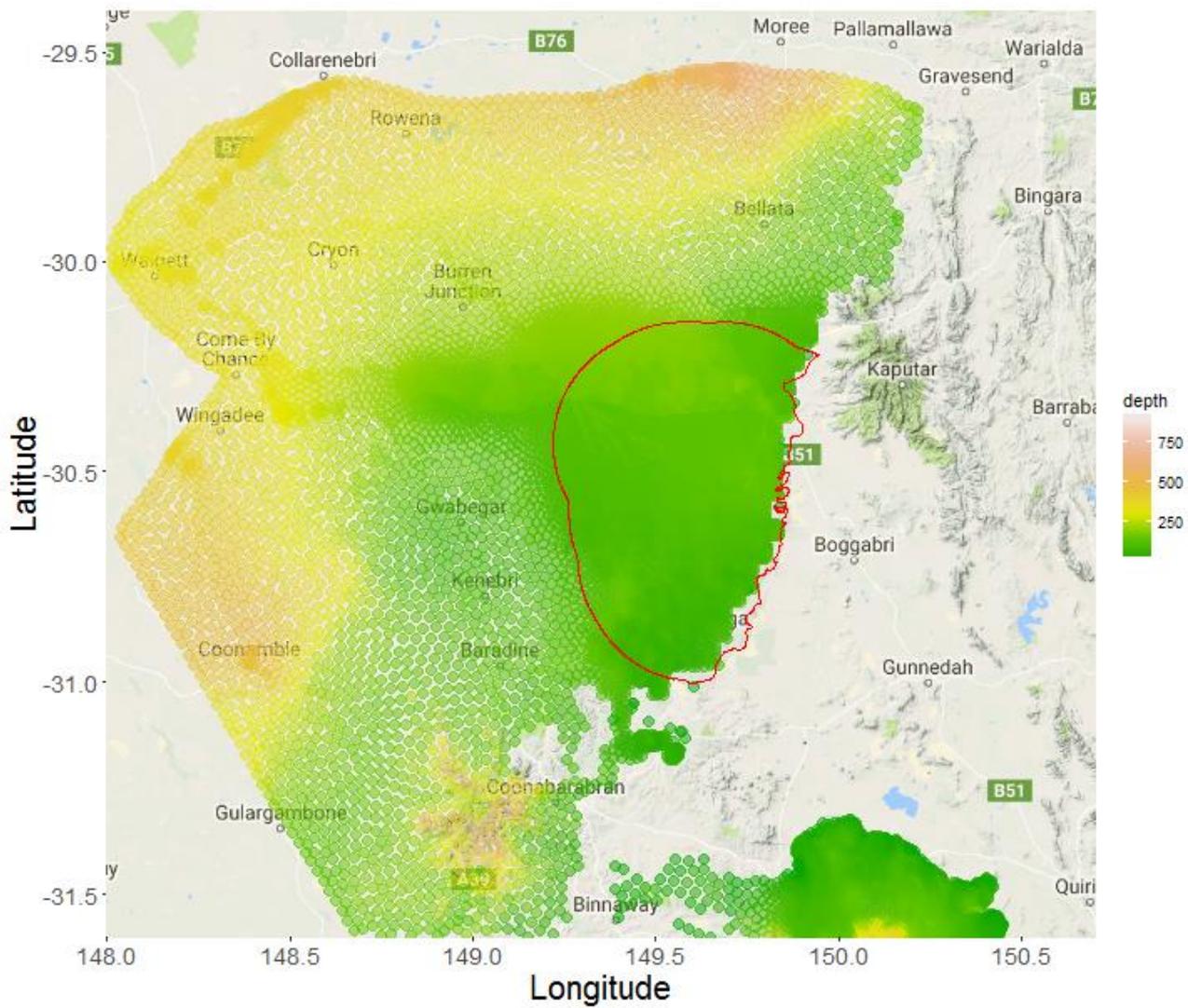
Figure 31: Depth to the mid-point of the hydrostratigraphic unit used as potential covariates in kriging models. The red outline shows the GAB aquifer region of interest

6.2   Geostatistical analysis

In order to predict the surfaces of the 41 baseline water quality variables across the Pilliga Sandstone region, we developed individual kriging models for each variable. Kriging models are Gaussian Processes where means and spatial covariances are used to describe spatial variability and may or may not include spatial explanatory variables. For simplicity, and due to the sparse nature of the data, we assume stationarity and continuity across the domain of interest. That is, we assume invariance with regard to translations in space. Mathematically, these models can be written as:

$$y(s) = \mu(s) + \epsilon(s) \tag{9}$$

where $y(s)$ is the spatially referenced observed baseline water quality variable, $\mu(s)$ is some spatially referenced unknown mean function and $\epsilon(s)$ the associated spatial error, with s the spatial location in the given domain. When the mean function is not inclusive of any additional covariates, Equation (9) is a specification of Kriging known as Ordinary Kriging (OK). One may wish to include additional predictors (as in our case, "depth"), in which $\mu(s) = x(s)^T \beta$. This is known as Universal Kriging (UK). In both cases, for interpolation at unknown location s_0 to obtain $y(s_0)$, we must estimate a covariance function $c(s, s_0)$. Critically, we assume covariance function depends only on the distance between locations s and s_0.

If we define c_0 as the vector of all covariances $c(s, s_0)$ between all observed values (denoted by vector y) and unknown location s_0, C as the covariance of all observed locations (estimated from the given covariance function), $c(s_0)$ as the covariance of the unknown location, and μ and $\mu(s_0)$ the mean vector and value of the observed and predicted locations, respectively, then it follows that a prediction of a baseline water quality variable at location s_0, denoted $y(s_0)$ can be obtained in a similar way of Equations (7) and (8) by

$$y(s_0)| \, y \sim N(\mu(s_0) + \, c_0^T C^{-1}(y - \mu), c(s_0) - c_0^T C^{-1} c_0) \, . \tag{10}$$

Equation (10) indicates that in order to obtain an estimate of the baseline water quality variables we must first estimate the covariance function $c()$ and a general form for the mean function $\mu()$. Therefore, the interpolated values for a given baseline water quality value at any location in the buffer region can be given as

$$\mu(s_0) + \, c_0^T C^{-1}(y - \mu). \tag{11}$$

and the estimated uncertainty as

$$c(s_0) - c_0^T C^{-1} c_0 \, . \tag{12}$$

As is common in geostatistical applications, we first obtain an empirical variogram and estimate a reasonable covariance function that will fit appropriately to the observed data. We restrict ourselves to covariance functions that are common amongst spatial data, those being the exponential, spherical, Gaussian, or Matérn covariance functions. Then, for both the OK and UK specifications, we fit all possible variograms using these covariance functions, and find the appropriate one that gives the minimum sum of squared residuals between the fitted and

empirical variograms. Ultimately, this yields 41 individual kriging models, one for each baseline water quality variable, and thus an estimate of the associated variability for each variable in the Pilliga region. For our purposes, we interpolate 20,000 different spatial pixels, giving a 487 m pixel resolution across the study region outlined in red in Figure 31.

6.3   Optimal monitoring design

We use the 41 fitted geostatistical Kriging models to determine optimal locations within the project region. Following the optimization work of Sreekanth et al. (2017b), we use the Differential Evolution (DE) algorithm (Storn and Price, 1997) taking the interpolated values of each baseline water quality variable from the kriging models in Section 6.2 as potential observations at sets of "proposed" monitoring locations. Using these additional observations, we then refitted the estimated Kriging models to see how the associated uncertainty across the project region would be reduced. We optimise the placement of these new monitoring locations using an objective function that targets the overall reduction in relative uncertainty across the gas project region. Specifically, we proceed as follows. First, we fit 41 kriging models corresponding to the 41 baseline water quality variables discussed in Section 6.2. Then, we interpolate at 20,000 spatial locations in the gas project region and obtain an initial estimate of uncertainty. Next, we implement the DE optimization algorithm using the DEoptim package in the R statistical programming language (Mullen et al., 2011; R Core Team, 2018).

In order to optimize over the gas project region using Differential Evolution, we choose an objective function that accounts for the total uncertainty reduced in the region. Specifically, we set our objective function to the mean relative variance reduction if we observed all 41 baseline water quality variables used for development of the Kriging models. That is, we first obtain the variance at each interpolated spatial location in the project region, for each water quality variable. We then choose a set of 10 random spatial locations in the gas project region. Next, we take the interpolated values for all baseline water quality variables at those 10 chosen spatial locations, treating those values as potentially new observed values. With those values, we refit the same kriging models and interpolate at the remaining spatial locations in the gas project region. From there, we obtain a new estimated variance at each spatial location. Finally, we find the mean reduction in variance across the region for all water quality variables, normalized by the initial fitted variance. We then choose a new set of spatial locations according to the Differential Evolution algorithm, and repeat until we find the 10 optimal locations with the maximum reduction in uncertainty in the region for all baseline water quality variables. For complete details on the Differential Evolution algorithm, we refer the reader to Mullen et al. (2011) and Sreekanth et al. (2017b).

Implementation of this algorithm results in 10 optimal spatial locations that will reduce the uncertainty in prediction across the project region, should those locations have new bore holes. Figure 32 shows the optimal locations of those boreholes, with Table 4 showing the latitude and longitude values of those 10 locations.

The reduction in the uncertainty for predictions of all baseline water quality variables is also important to inspect. While we are able to obtain maps of uncertainty for all 41, we present three examples of water quality variables and the reduction in predictive uncertainty from new

monitoring wells. The prediction uncertainty before and after the addition of the 10 new wells for 'pH field' are shown in, for 'electrical conductivity' in Figure 34, and for 'alkalinity bicarbonate' in Figure 35. In each of these three figures, the red dots show the optimization locations while the blue dots show locations where the water quality variable was observed in the gas project region.

We note that reduction in uncertainty is more uniform when there were no previous observations in the region (for example, in the case of 'pH field') than when some observations were already in the project region (for example, 'electrical conductivity'). Further additional predictor covariates such as 'depth' may also contribute to reduction in uncertainty, where this resulted in a better kriging model, such as the case for 'alkalinity bicarbonate'. This method does showcase the flexibility and value of additional information in either predictors (such as 'depth') or more water quality variables, if the data is available.
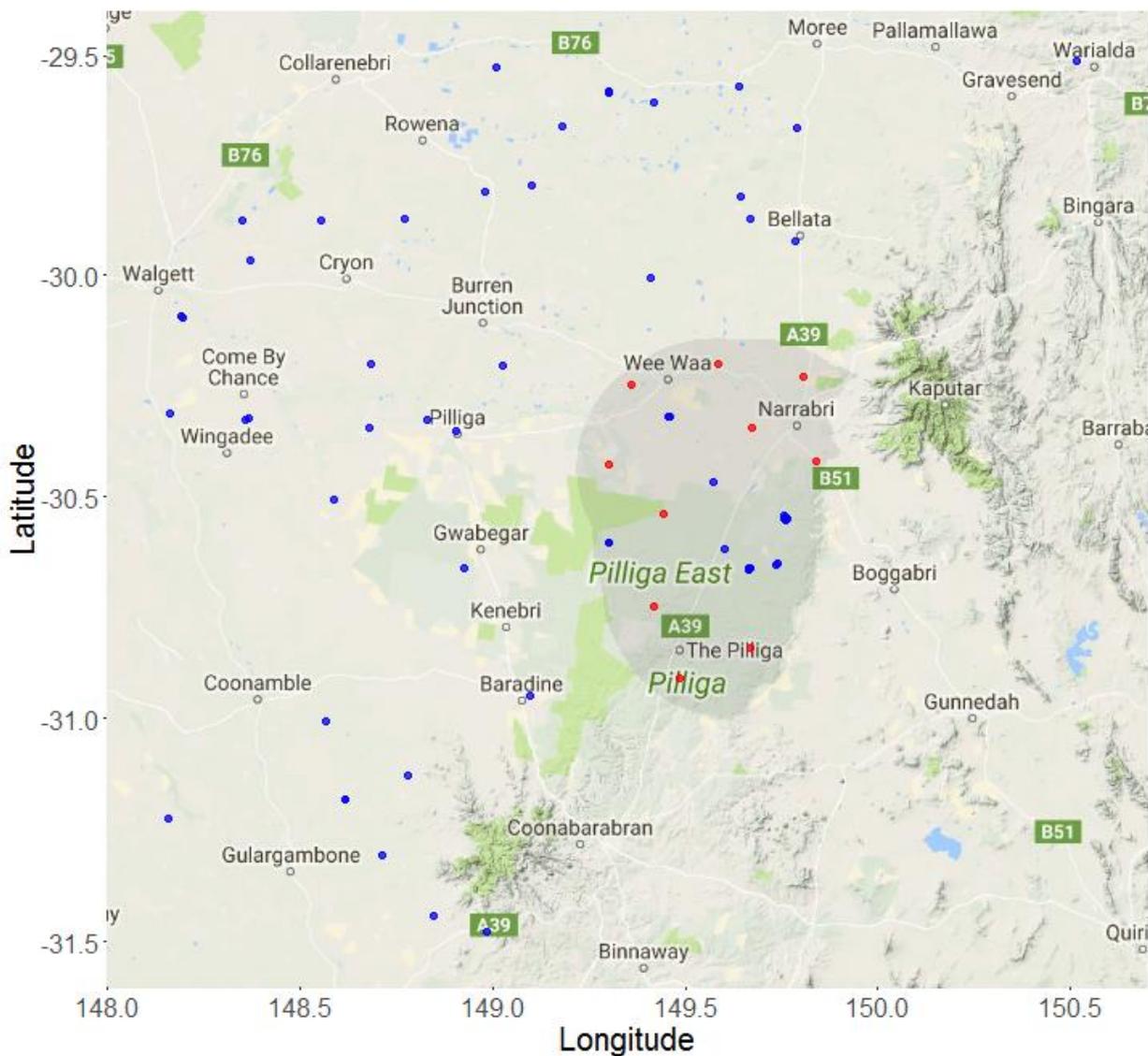
Figure 32: Potential new locations (red) for monitoring baseline water quality in the Cadna-Owie Hooray Sandstone unit and locations with at least one water quality observation (blue).

Table 4: Optimal locations for 10 wells using the Geostatistical analysis method of baseline water quality variables

Location	Longitude	Latitude
1	149.8067	-30.22994
2	149.4865	-30.9098
3	149.5831	-30.20046
4	149.6746	-30.34372
5	149.4197	-30.74785
6	149.8405	-30.42217
7	149.3006	-30.42883
8	149.4434	-30.53962
9	149.3602	-30.24888
10	149.6706	-30.84207
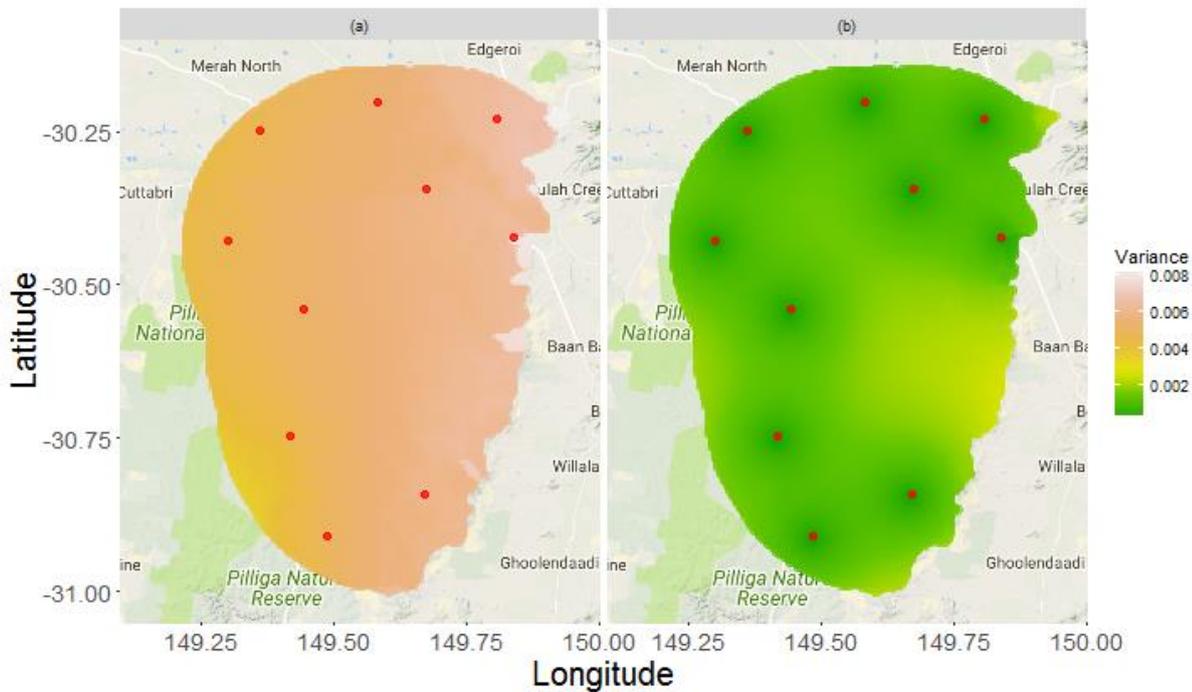
Figure 33: Estimated variance (a) without monitoring locations and (b) after optimization and supplementation with the new monitoring locations for "pH field" water quality variable. Red dots indicate potential new monitoring locations.
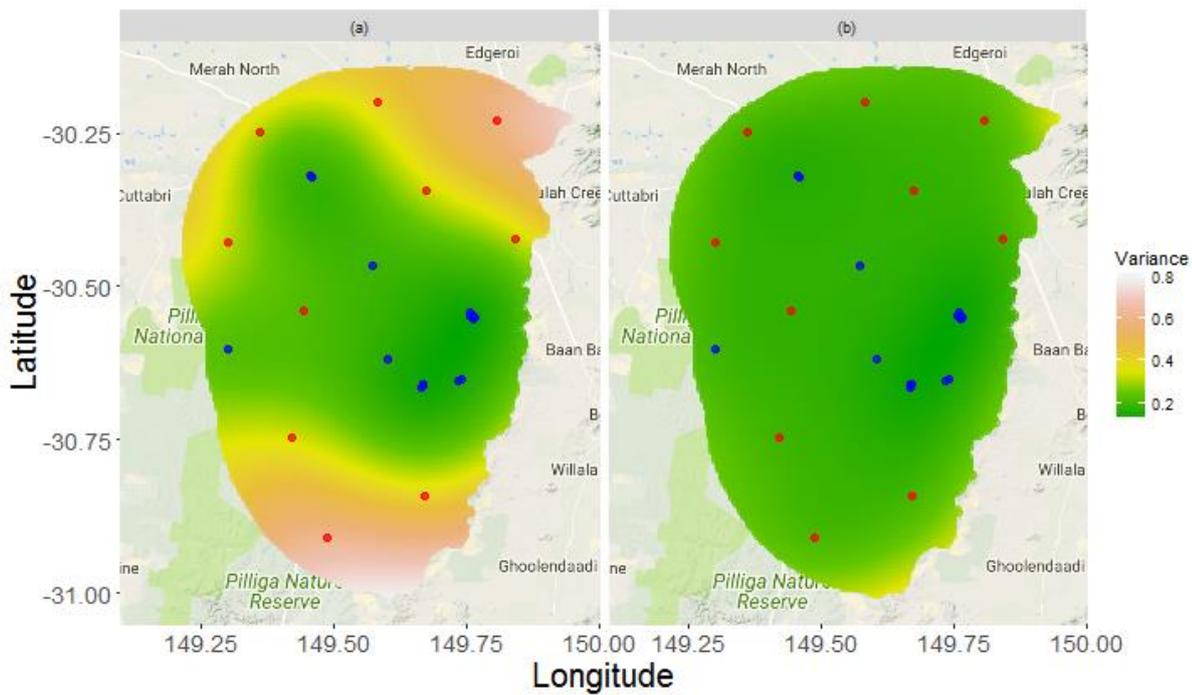
Figure 34: Estimated variance (a) without monitoring locations and (b) after optimization and supplementation with the new monitoring locations for "Electrical Conductivity" water quality variable. Red dots indicate potential new monitoring locations.
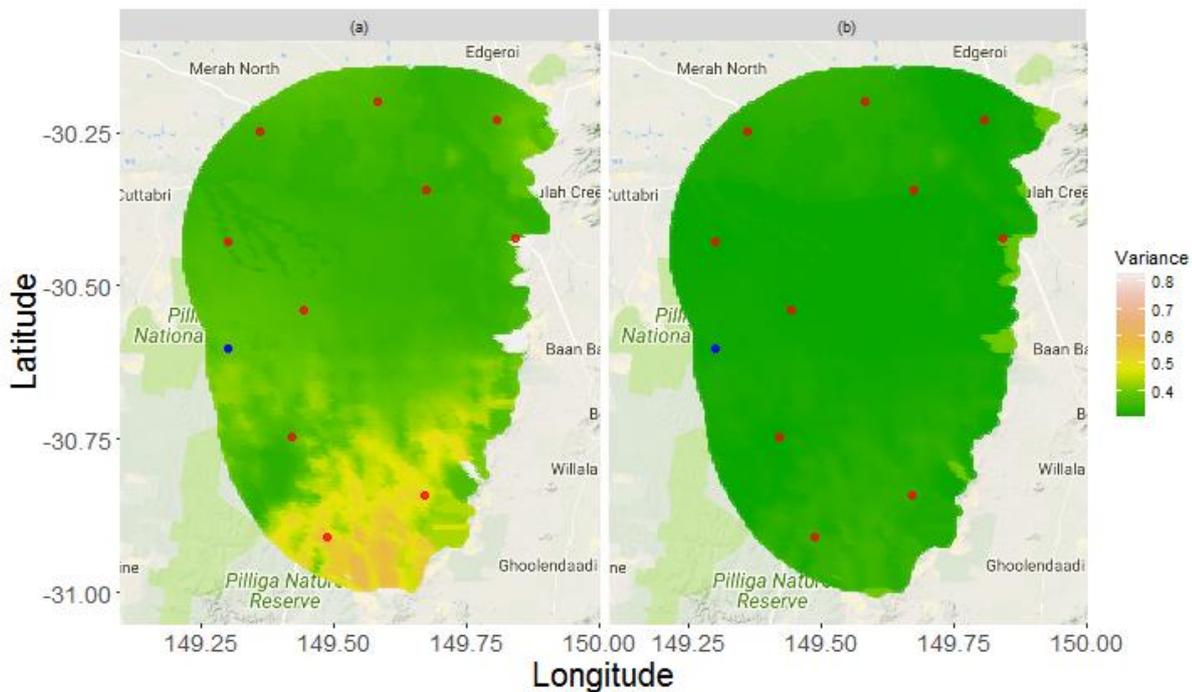
Figure 35: Estimated variance (a) without monitoring locations and (b) after optimization and supplementation with the new monitoring locations for "Alkalinity Bicarbonate" water quality variable. Red dots indicate potential new monitoring locations.

7 Optimal network design for monitoring drawdown

Optimization methods can also be used to find optimal borehole locations using simulations from a stochastic simulation model of groundwater flow specifically modelling maximum drawdown (e.g. CSG-induced pressure/water level changes), herein labelled D-max, and time to observe the maximum drawdown (herein labelled T-max). Critically, several model simulations are undertaken in order to provide a robust prediction of maximum drawdown and time. These simulations are then projected onto a lower dimensional field through basis functions that account for spatial correlation, correlation between layers, and correlation between D-max and T-max. We then use the resulting dimension reduction to determine optimal locations that will minimize prediction uncertainty regardless of what ultimately is the true D-max and T-max. A critical point is that such a design is robust to all possible simulations of both D-max and T-max in all considered layers.

7.1 Drawdown patterns in multiple model layers

Our methodology follows that of Sreekanth et al. (2017b). In order to generate optimal locations, we first must simulate several possible drawdown patterns across multiple model layers. For our purposes, we had 423 successful model runs of drawdown prediction for multiple layers in the gas project area. While the model used does generate multiple layers, we only focus on 6 of the layers for our example, those being layers 6, 7, 8, 10, 12, and 14. These correspond to the Pilliga Sandstone, the model layer immediately below the Pilliga Sandstone, upper aquitard sequence, Hoskissons Coal, lower aquitard sequence and the Maules Creek formations respectively. We further note that we use both D-max and T-max from all 6 layers, resulting in 12 different outputs per spatial location. It is worth noting that the chosen methodology is flexible to include additional layers, or only focus on a subset of layers.

The simulation model used does not output D-max and T-max at each spatial pixel in the Pilliga Sandstone aquifer region. Additionally, each layer was built on differing spatial resolutions. We therefore create a spatial grid in the project region and interpolate to each pixel each simulation. For this example, we use the same spatial grid used with the geostatistical optimization of baseline water quality variables discussed in Section 6. We use a linear triangulation (interpolation on a triangular network) to map the simulation model output onto the generated spatial grid. Additionally, we only use pixels that have modelled values for both D-max and T-max in all 6 layers for all 423 simulations. This resulted in use of 6418 spatial pixels at a resolution of 487m. Figure 36 - Figure 47 show examples of 4 model runs (simulation 1, 100, 200, and 400) for both D-max and T-max layers 6, 7, 8, 10, 12, and 14. These plots showcase the variability in the drawdown pattern simulations.
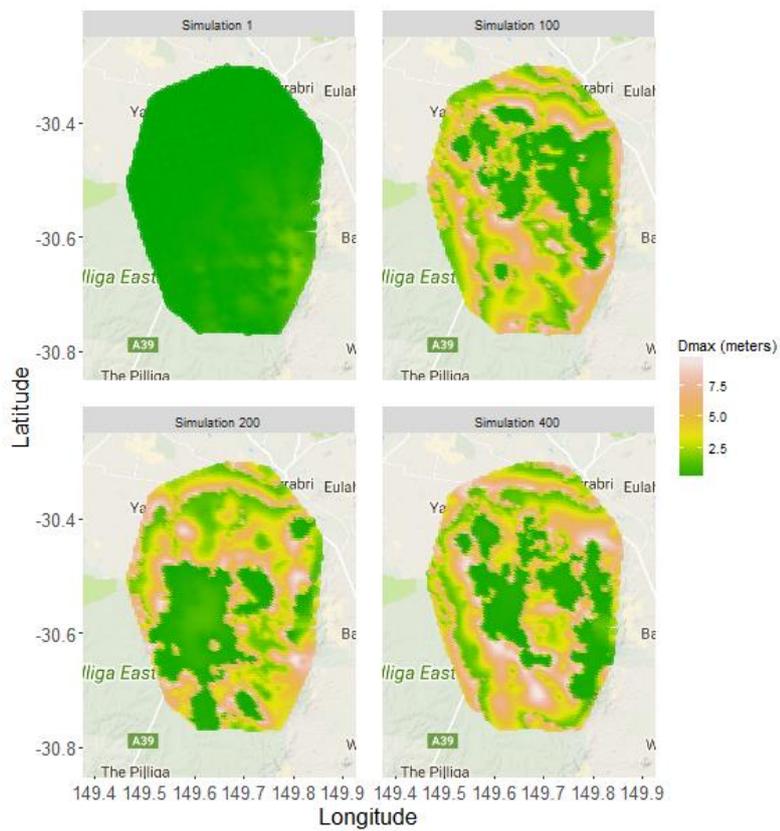
Figure 36: Four simulated examples of D-max at layer 6
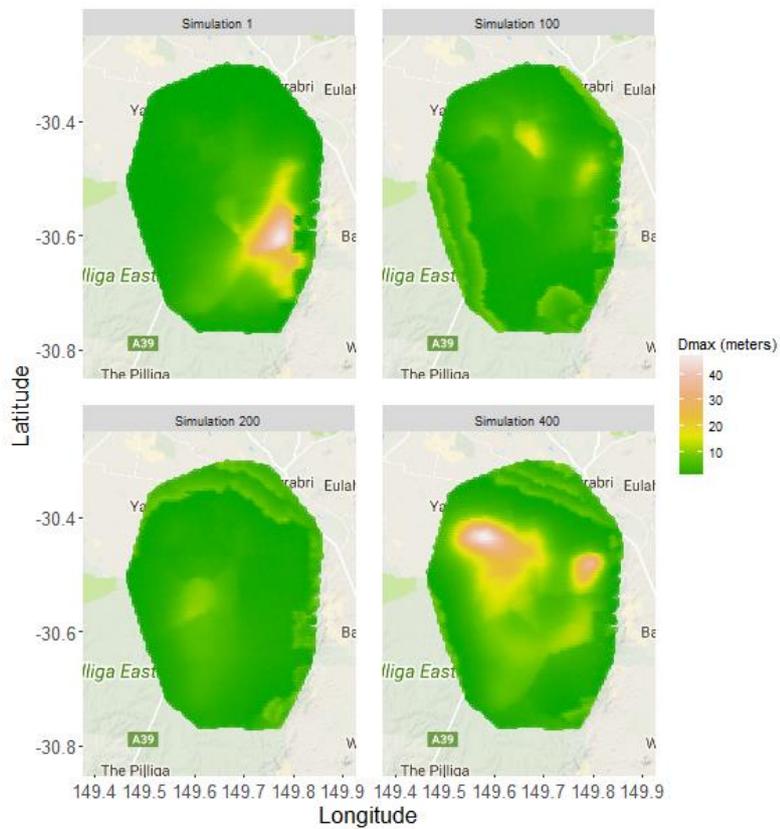
Figure 37: Four simulated examples of D-max at layer 7
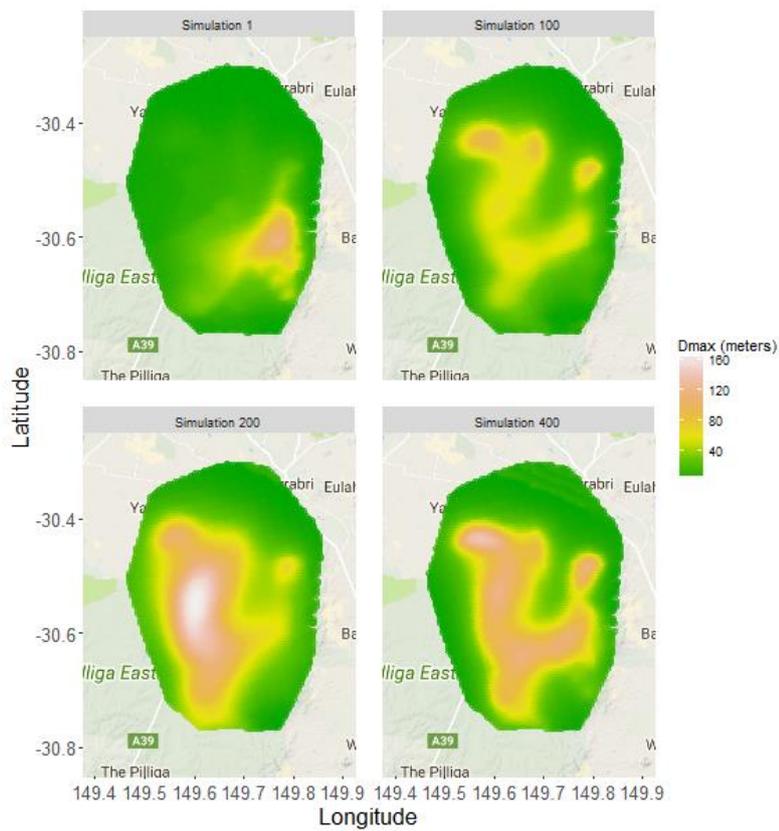
Figure 38: Four simulated examples of D-max at layer 8
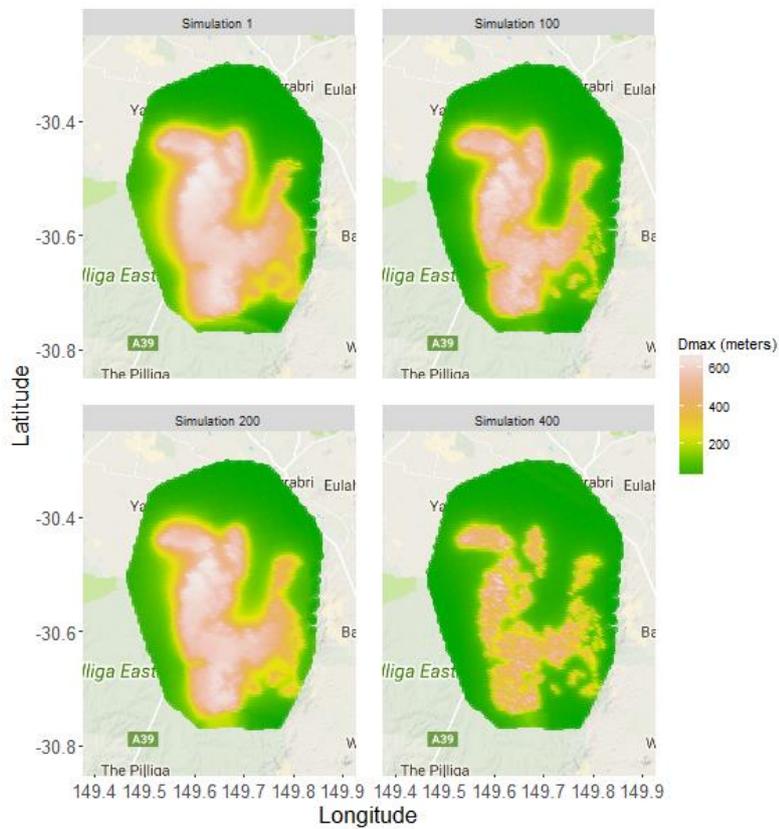
Figure 39: Four simulated examples of D-max at layer 10
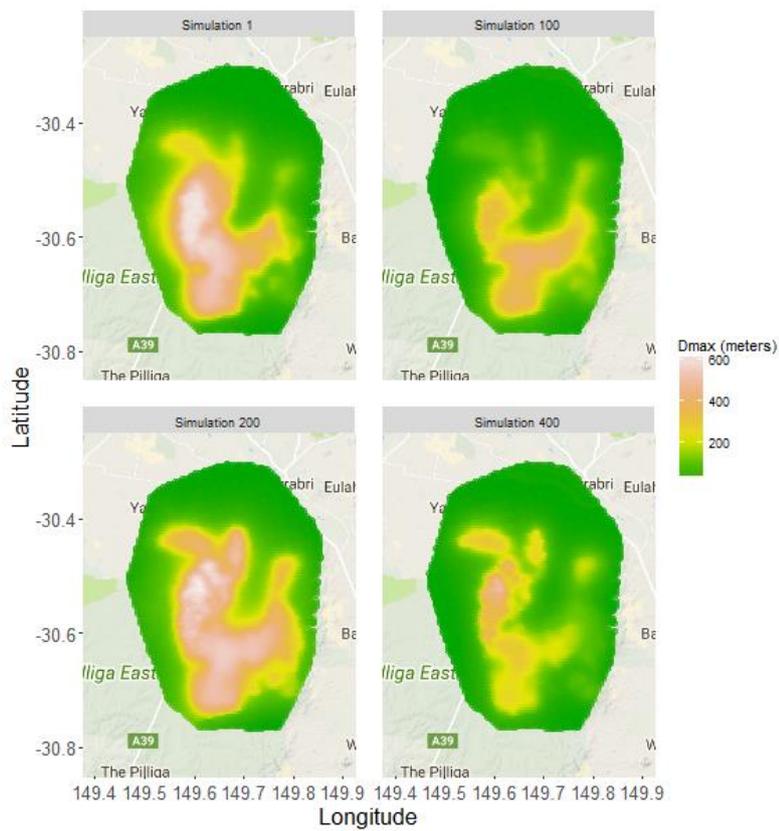
Figure 40: Four simulated examples of D-max at layer 12
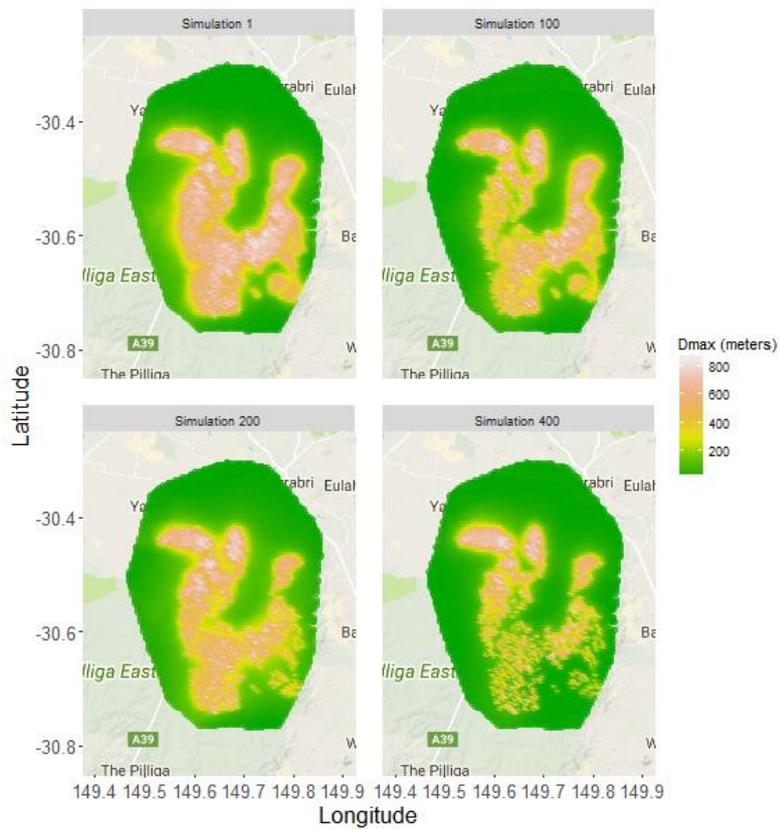
Figure 41: Four simulated examples of D-max at layer 14
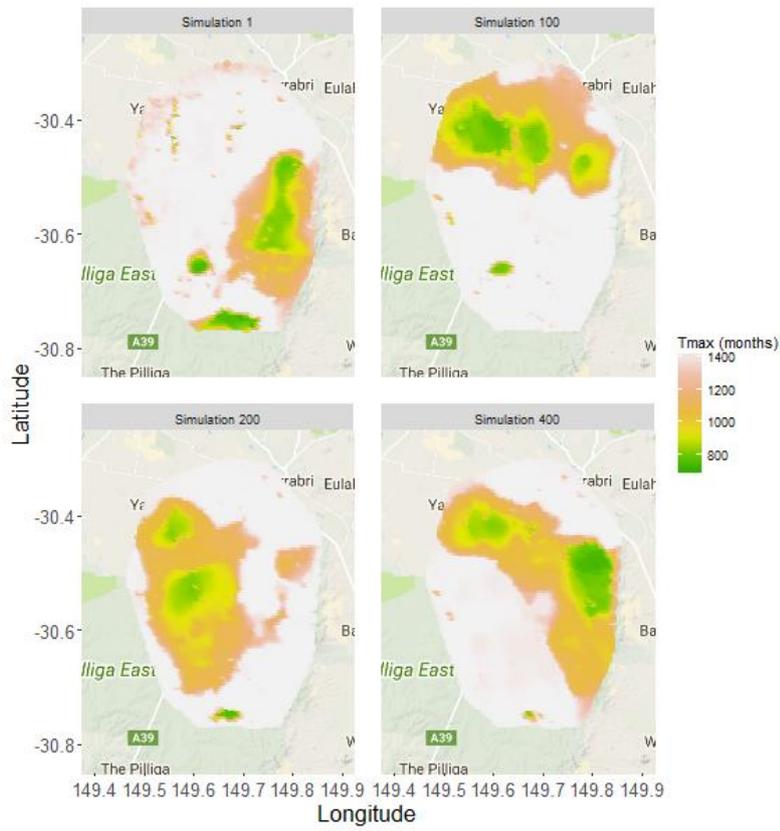
Figure 42: Four simulated examples of T-max at layer 6
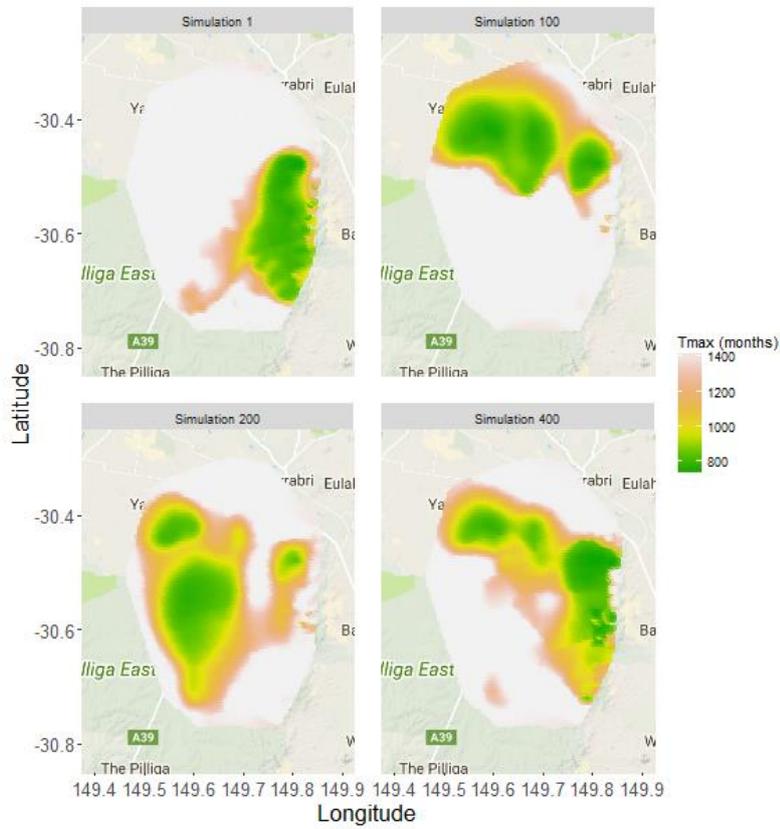
Figure 43: Four simulated examples of T-max at layer 7
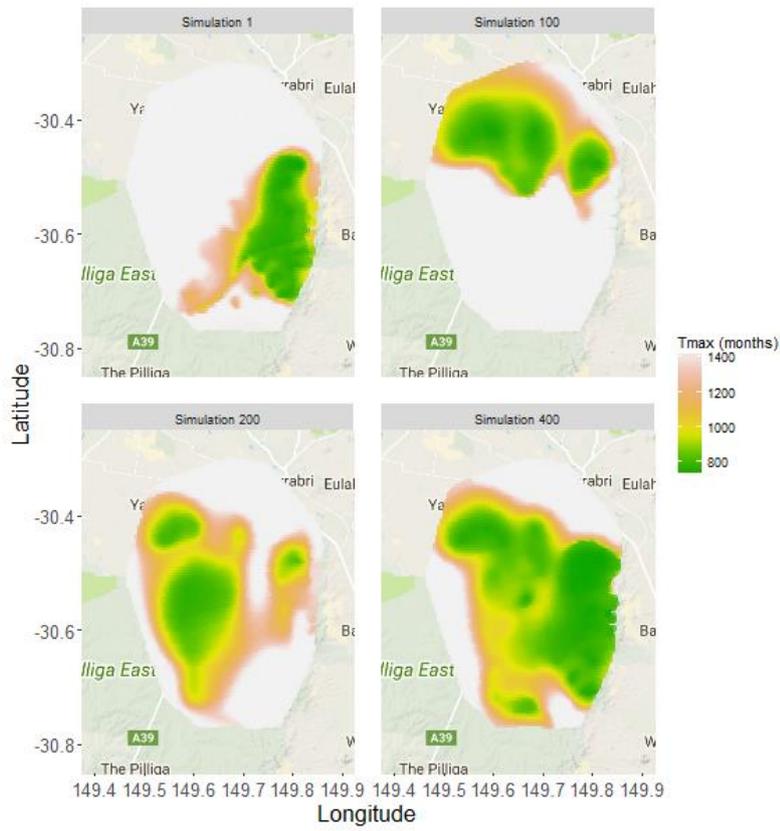
Figure 44: Four simulated examples of T-max at layer 8
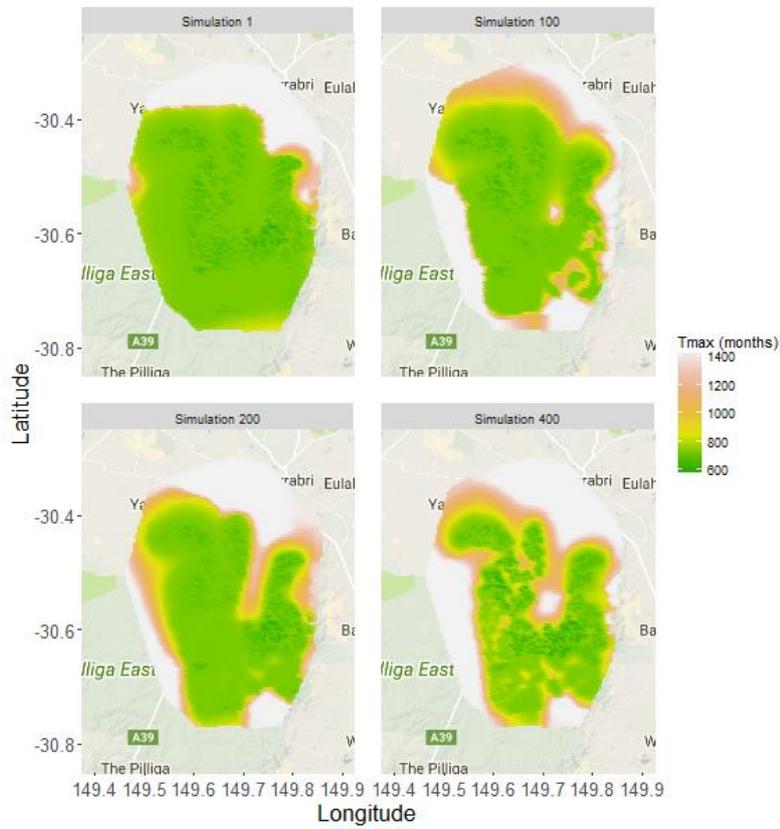
Figure 45: Four simulated examples of T-max at layer 10
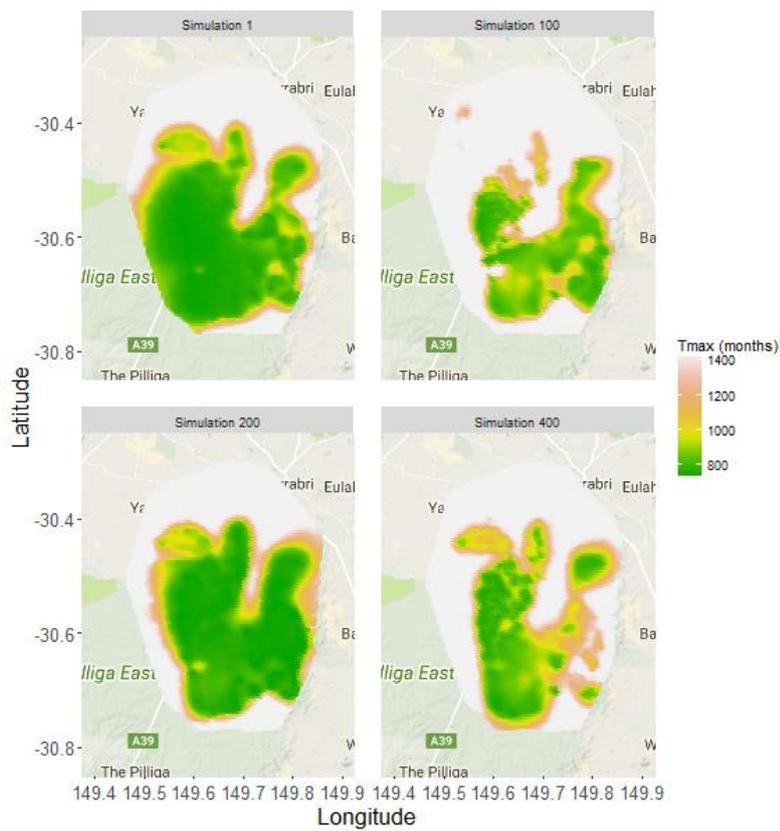
Figure 46: Four simulated examples of T-max at layer 12
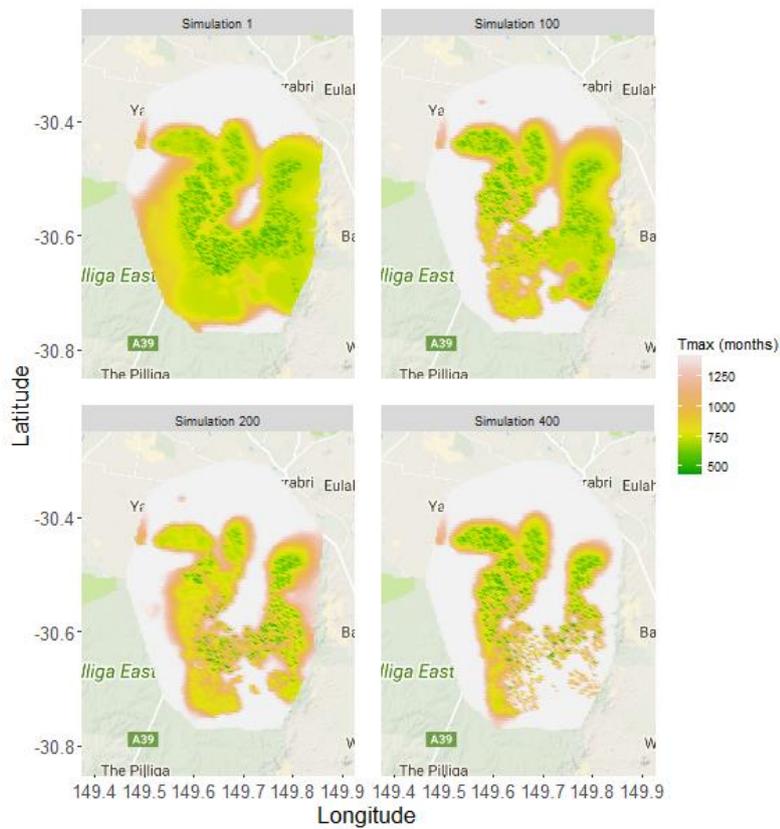
Figure 47: Four simulated examples of T-max at layer 14

7.2   Spatial Basis Functions

The methodology presented in Sreekanth et al. (2017b) is dependent upon an orthogonal basis function expansion. As noted in Cressie and Wikle (2011, Ch. 7), there are several choices for spatial basis functions, including empirical orthogonal functions (EOFs), Fourier basis functions, splines, and wavelets among others. Given their use in Sreekanth et al. (2017b) and their successful use in several model simulations in many disciplines (see Cressie and Wikle, 2011 for examples), we use EOFs to project the groundwater simulations into a lower dimensional space. Critically, these EOFs account for the spatial dependence as well as the interlayer and model output dependencies. Following Sreekanth et al. (2017b) we use the same number of EOFs as with the number of optimization wells, and in our case we use 10 EOFs. These 10 EOFs explain the majority of the variability in the groundwater simulation runs, and are calculated through a singular value decomposition of the model output. Importantly, the first EOF accounts for the most variability of all EOFs, followed by the second EOF, and so on. Figure 48 - Figure 59 show the spatial basis functions for all 6 layers of interest for both D-max and T-max. As described in Sreekanth et al. (2017b), the first EOF used is actually a vectorization of the first panel in these twelve figures combined, and likewise for the remaining EOFs. Critically, we are able to determine the spatial dependencies in the model output by examining the spatial structure of these EOFs.
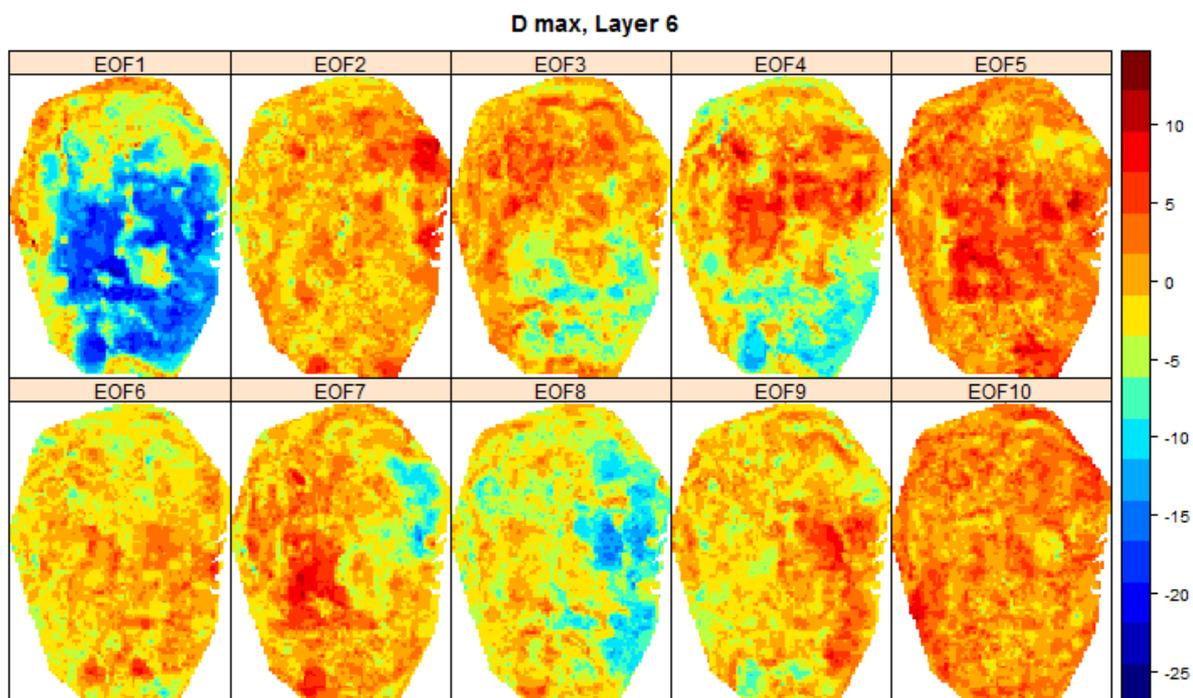
Figure 48: Spatial basis functions for maximum drawdown layer 6 (unitless). The subplots show the dominant ten empirical orthogonal functions (EOF) that explain the greatest proportion of variability in the ensemble flow model runs.
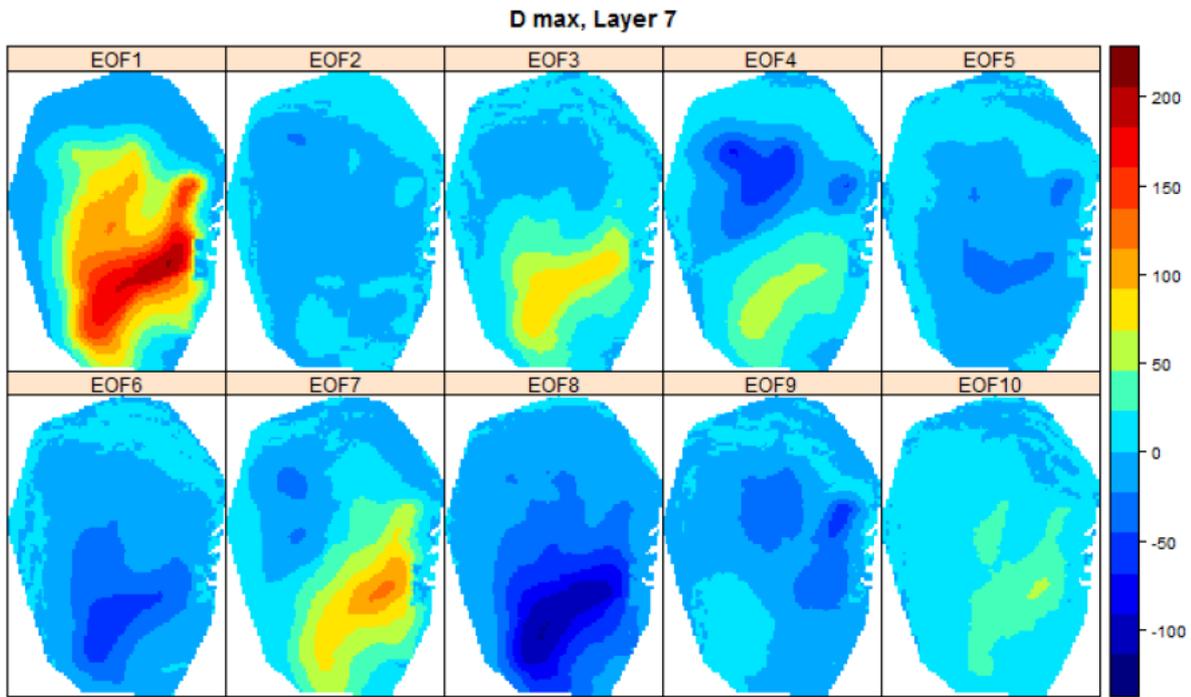
Figure 49: Spatial basis functions for maximum drawdown layer 7 (unitless). The subplots show the dominant ten empirical orthogonal functions (EOF) that explain the greatest proportion of variability in the ensemble model runs.
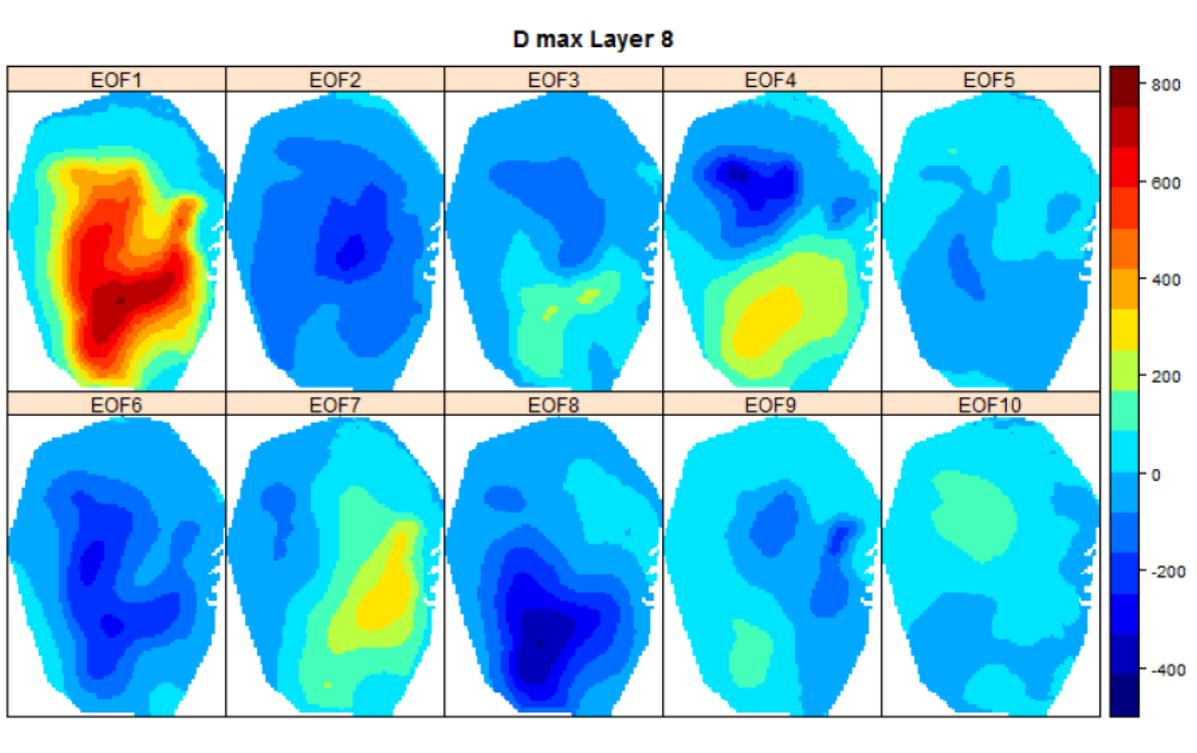
Figure 50: Spatial basis functions for maximum drawdown layer 8 (unitless). The subplots show the dominant ten empirical orthogonal functions (EOF) that explain the greatest proportion of variability in the ensemble model runs.
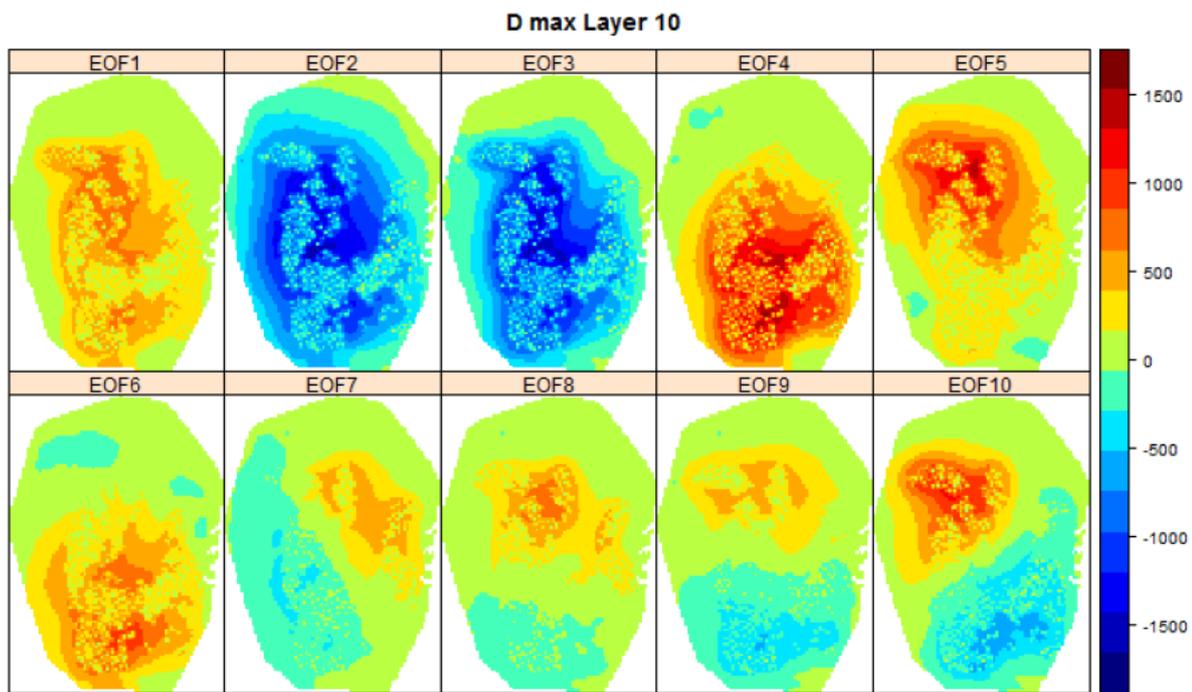
Figure 51: Spatial basis functions for maximum drawdown layer 10 (unitless). The subplots show the dominant ten empirical orthogonal functions (EOF) that explain the greatest proportion of variability in the ensemble flow model runs.
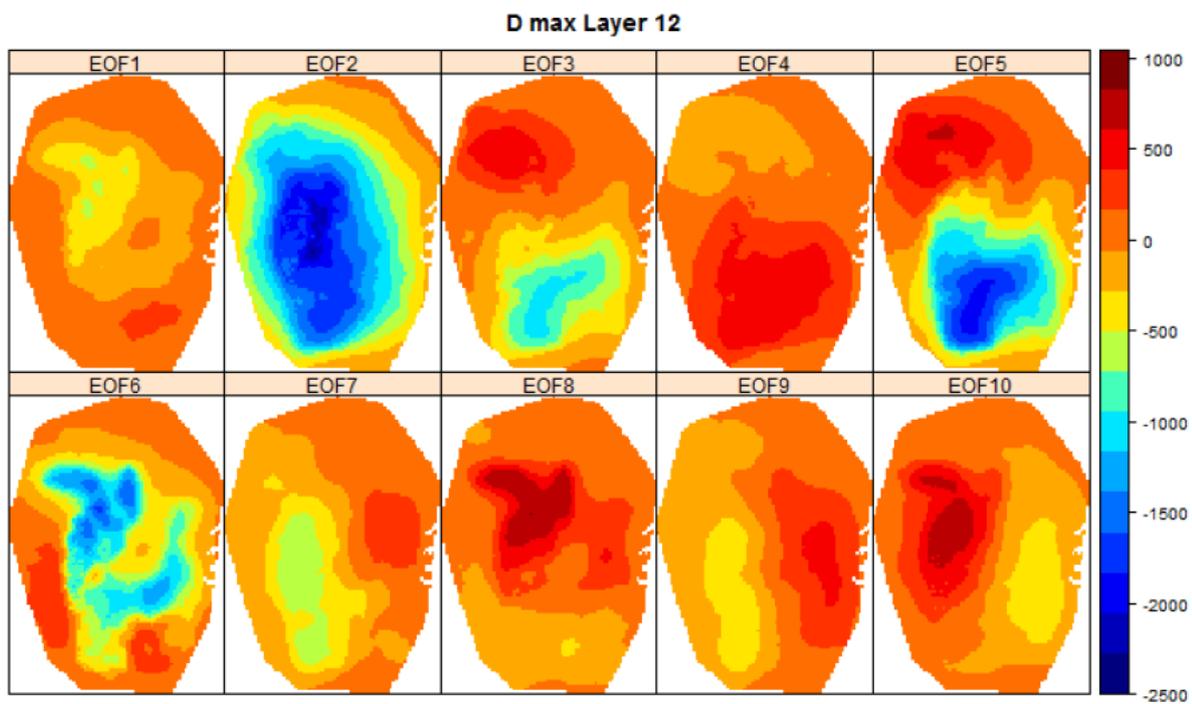
Figure 52: Spatial basis functions for maximum drawdown layer 12 (unitless). The subplots show the dominant ten empirical orthogonal functions (EOF) that explain the greatest proportion of variability in the ensemble flow model runs.
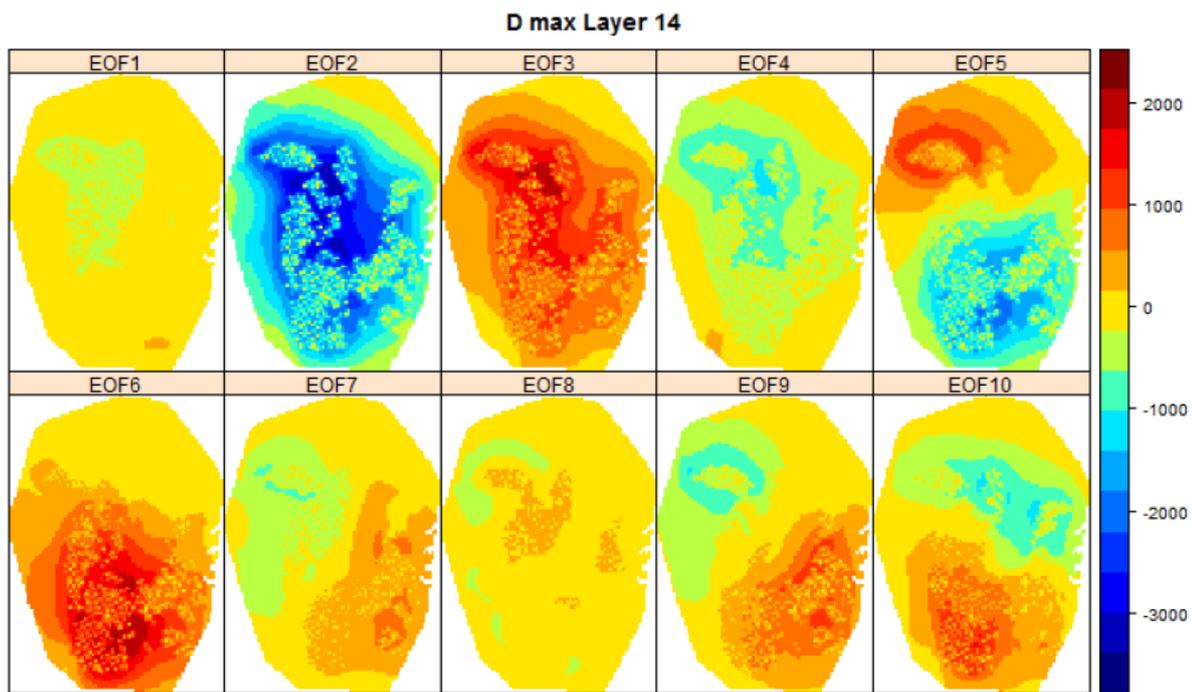
Figure 53: Spatial basis functions for maximum drawdown layer 14 (unitless). The subplots show the dominant ten empirical orthogonal functions (EOF) that explain the greatest proportion of variability in the ensemble flow model runs.
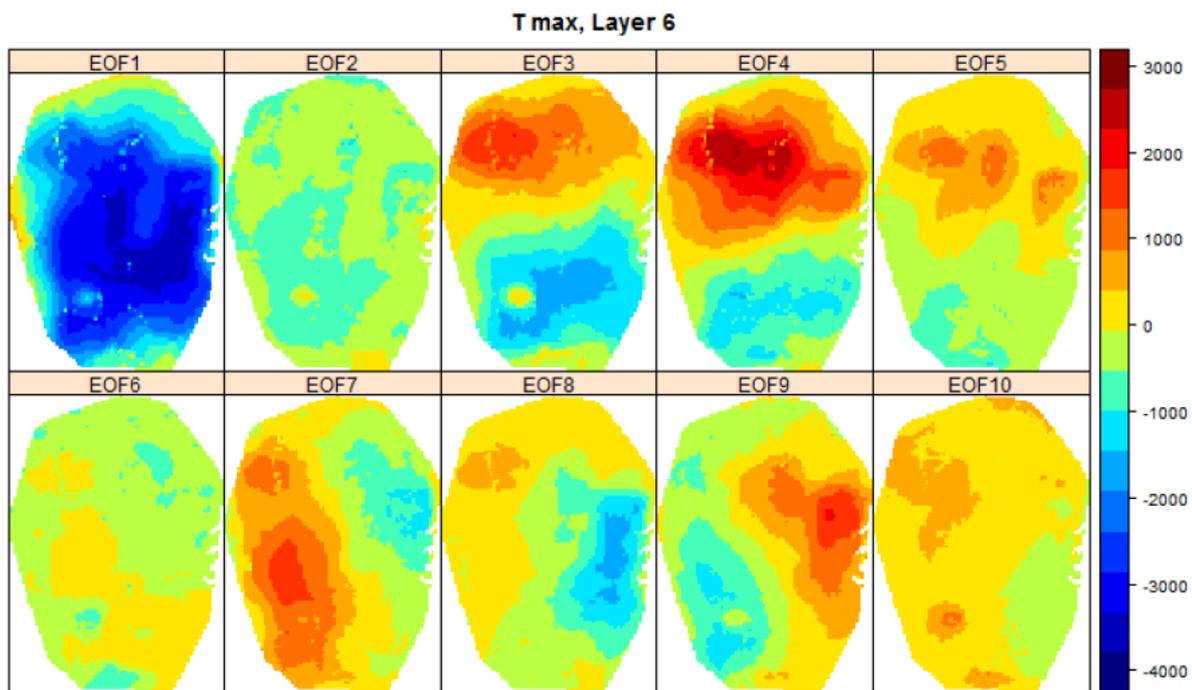
Figure 54: Spatial basis functions for time to maximum drawdown layer 6 (unitless). The subplots show the dominant ten empirical orthogonal functions (EOF) that explain the greatest proportion of variability in the ensemble flow model runs.
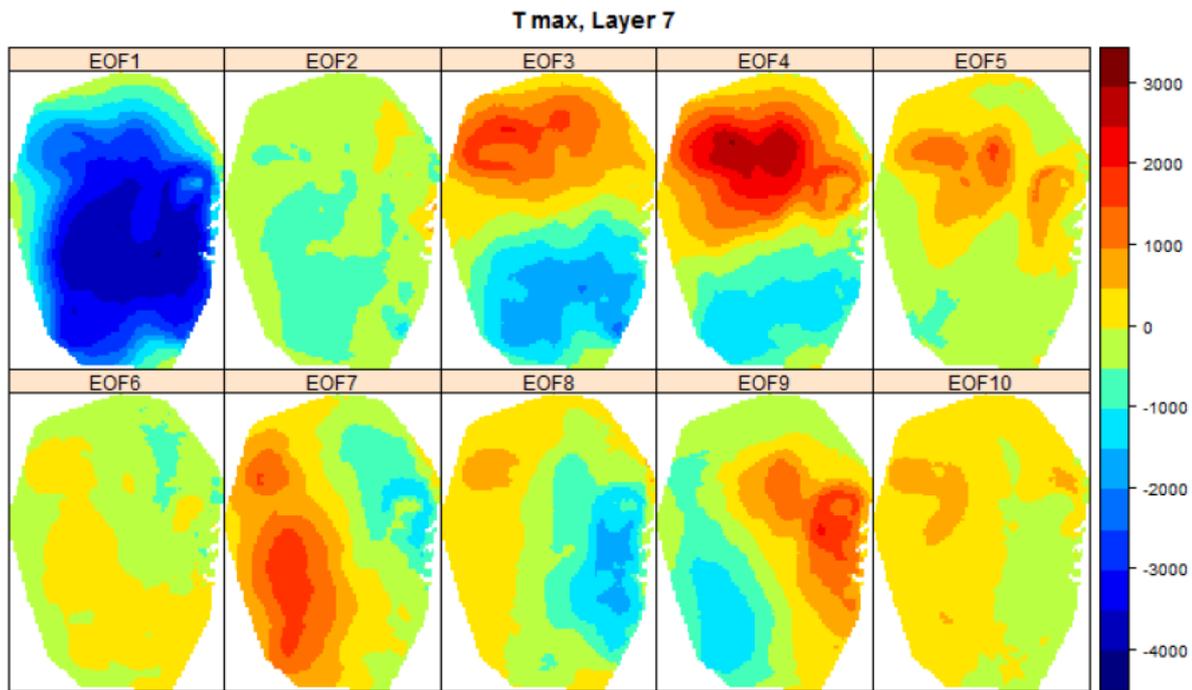
Figure 55: Spatial basis functions for time to maximum drawdown layer 7 (unitless). The subplots show the dominant ten empirical orthogonal functions (EOF) that explain the greatest proportion of variability in the ensemble flow model runs.
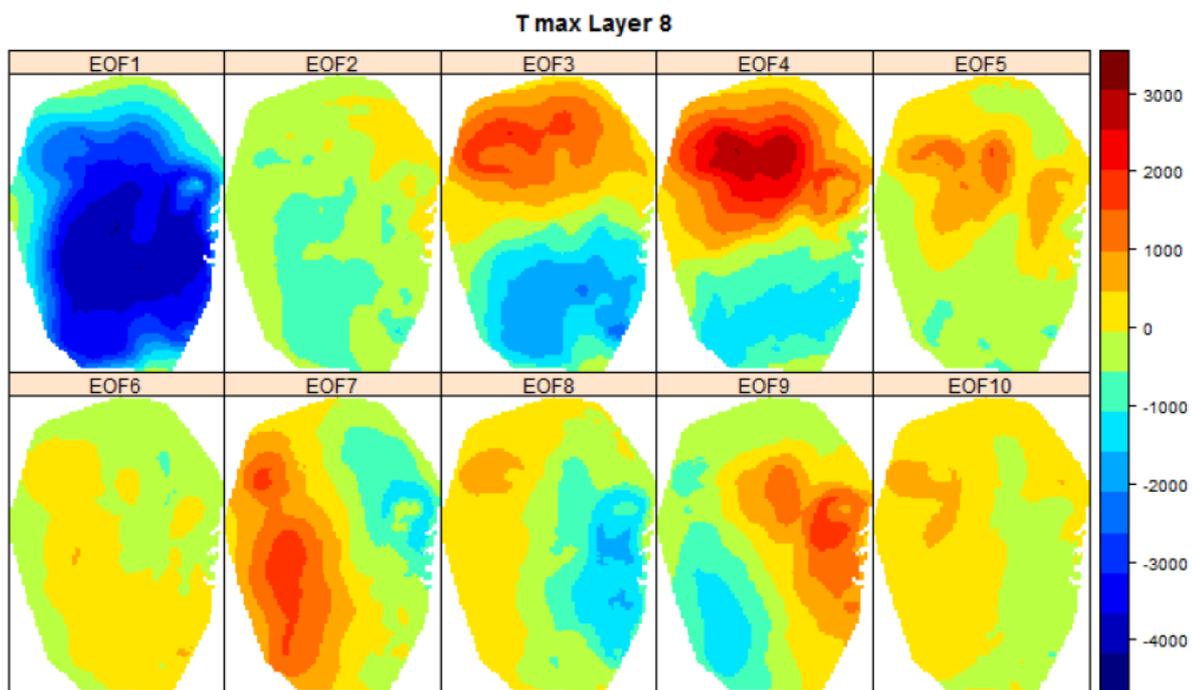
Figure 56: Spatial basis functions for time to maximum drawdown layer 8 (unitless). The subplots show the dominant ten empirical orthogonal functions (EOF) that explain the greatest proportion of variability in the ensemble flow model runs.
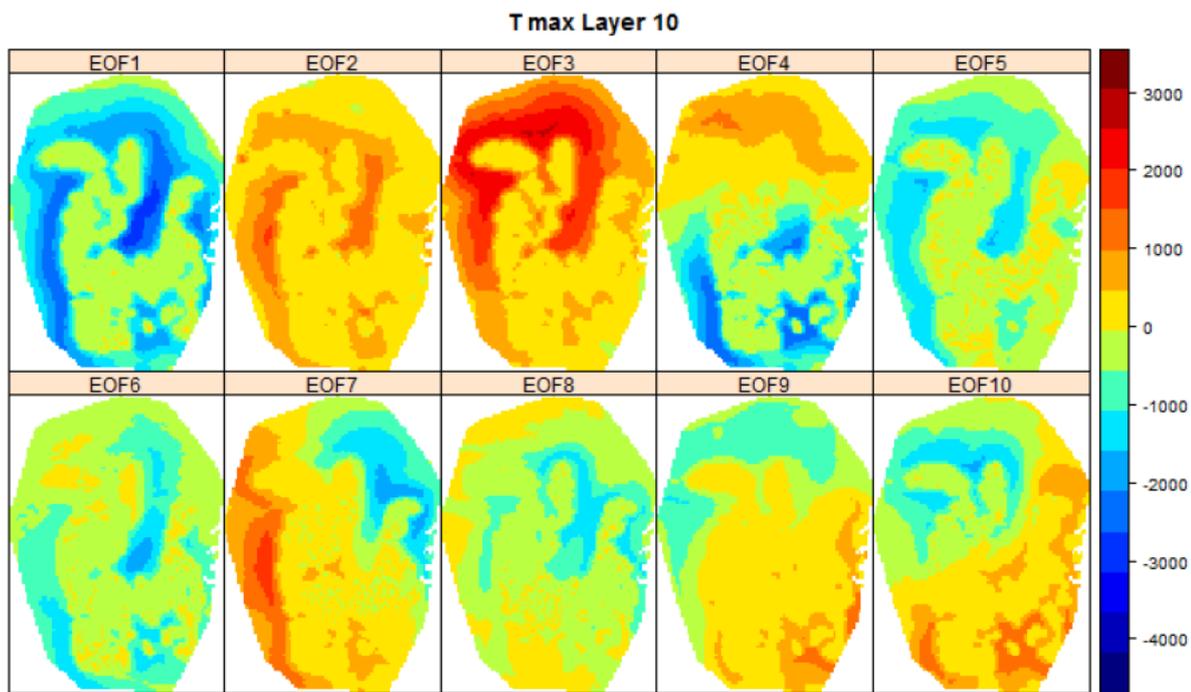
Figure 57: Spatial basis functions for time to maximum drawdown layer 10 (unitless). The subplots show the dominant ten empirical orthogonal functions (EOF) that explain the greatest proportion of variability in the ensemble flow model runs.
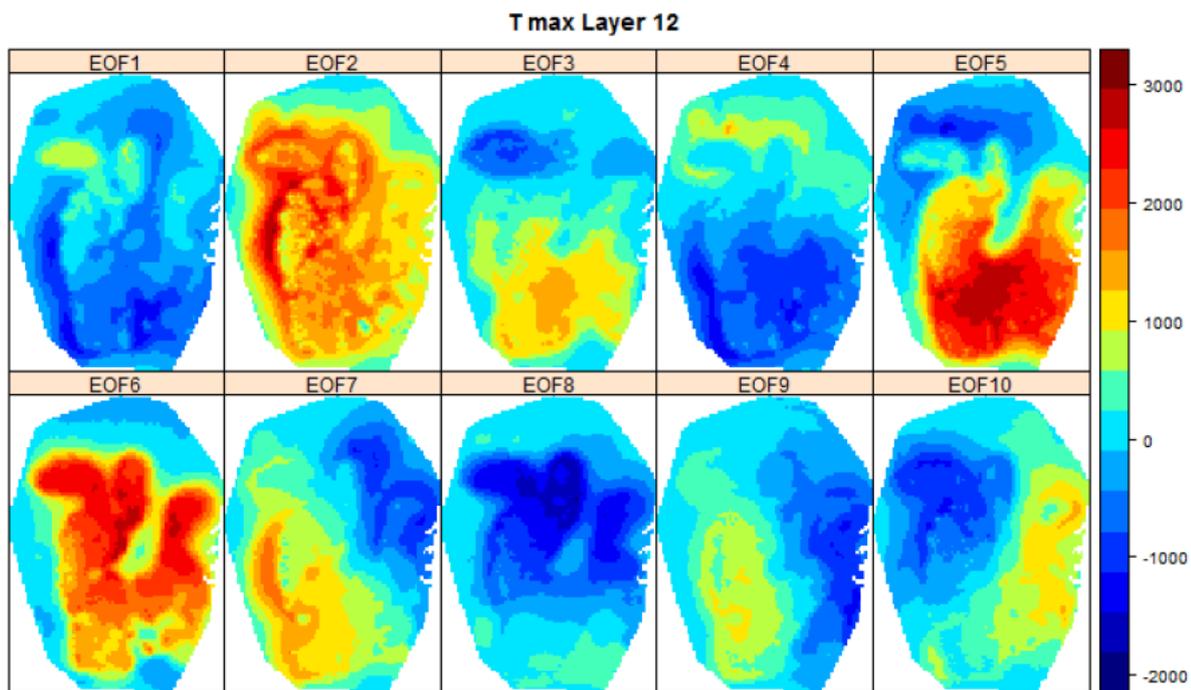
Figure 58: Spatial basis functions for time to maximum drawdown layer 12 (unitless). The subplots show the dominant ten empirical orthogonal functions (EOF) that explain the greatest proportion of variability in the ensemble flow model runs.

T max Layer 14
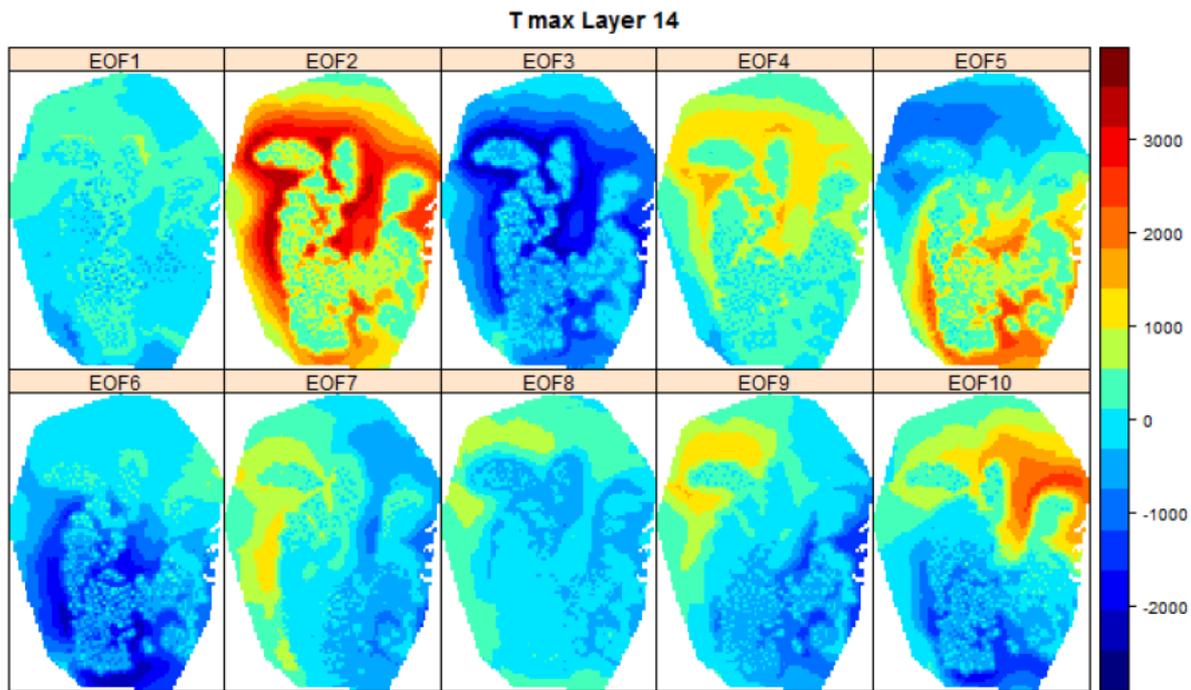
Figure 59: Spatial basis functions for time to maximum drawdown layer 14 (unitless). The subplots show the dominant ten empirical orthogonal functions (EOF) that explain the greatest proportion of variability in the ensemble flow model runs.

We also undertook an investigation into whether the number of model runs undertaken, sufficiently covered the range of outcomes possible from the model. To study this, we performed a cross-validation experiment where we used smaller numbers of randomly selected model runs to construct basis functions and examined how well the out-of-sample model runs could be recovered from all available EOFs from the in-sample model runs. Figure 60 shows that beyond around 200 model runs, the mean squared error in reconstructing the out-of-sample model runs decreases little and indicates that 200 model runs has adequately explored most of the patterns of variability in the model output. The 423 model runs generated were therefore deemed to be adequate for our purposes.
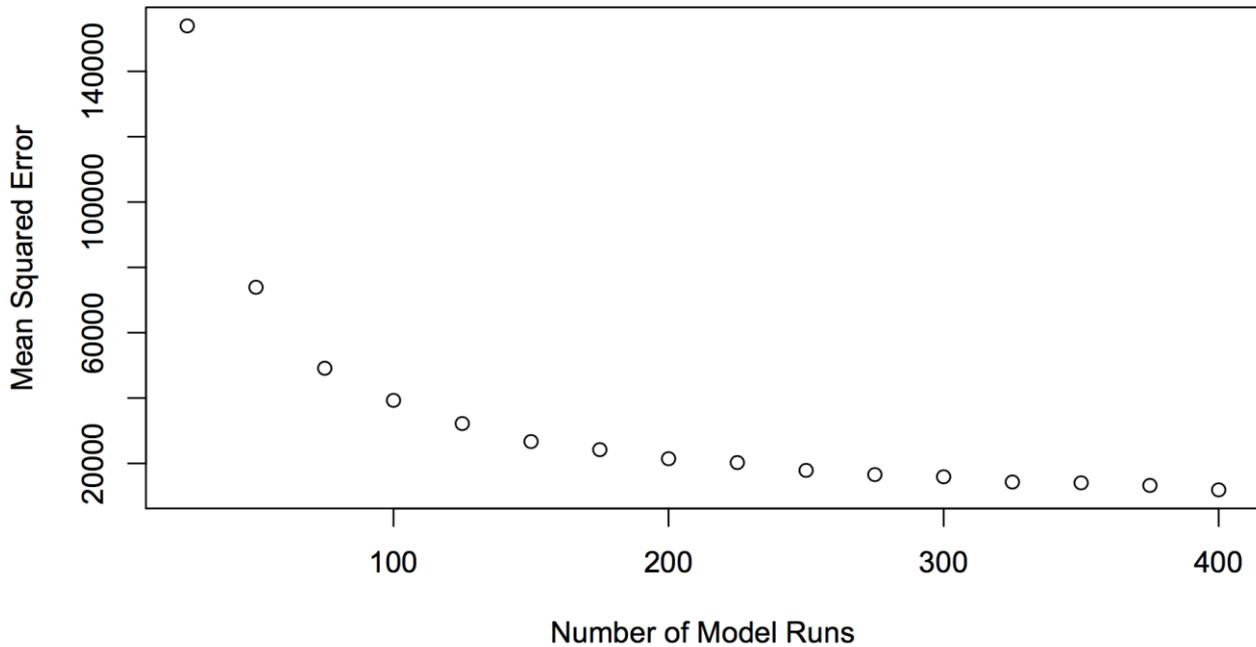
Figure 60: Reduction in out-of-sample mean squared error in a cross-validation experiment as the number of in-sample model runs is increased

7.3 Design for monitoring and reducing uncertainty in drawdown impacts

As with the geostatistical optimization from Section 6, we follow Sreekanth et al. (2017b) and use the Differential Evolution optimization algorithm. In this case, implementation of the DE algorithm in the reduced-rank space minimizes the uncertainty of the predicted drawdown pattern for the project region, regardless of what the true values of D-max and T-max across space are. As discussed before, we present 10 optimal borehole wells in the gas project region, restricted to all spatial locations with values associated with D-max and T-max for all 6 model layers of interest. The optimal locations can be found in Figure 61. For reference, this figure also highlights the entire project region (light grey) and the imposed restricted area from our groundwater simulation model (dark grey). The latitude and longitude locations for the 10 optimal locations can be found in Table 5.

We note that there is relatively uniform spatial coverage in the region that we use for our optimization, indicating that our method will be robust if future need arises to adjust interested layers or include additional spatial locations. The data-worth analysis reported in Section 4 also showed that multi-level piezometers in this region, in general, carries most worth in informing drawdown predictions in the GAB aquifer. The data-worth analysis used only the sensitivity of drawdown predictions to model parameters as opposed to the use of spatial trends in stochastic drawdown predictions used in the monitoring network design. Hence, the data-worth provides an independent verification of the utility of the monitoring network design.
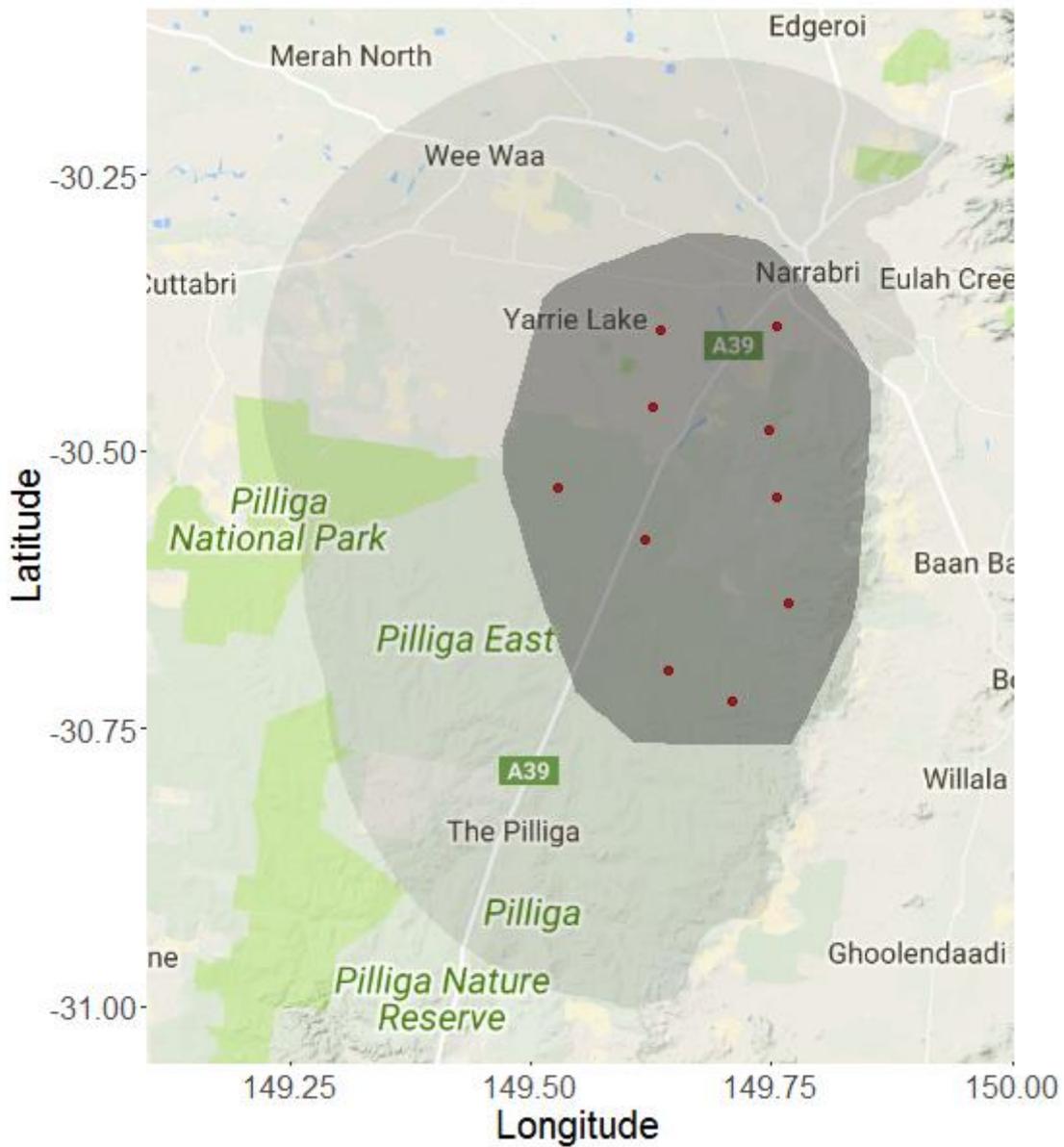
Figure 61: Optimisation results from the reduced-rank spatial prediction method. Red dots indicate potential new locations. Light grey region indicates the buffer region of interest. Dark grey indicates region with all 6 layers of interest for both maximum drawdown and distance travelled.

Table 5: Optimal locations for 10 optimal wells using the reduced rank simulation method.

Monitoring Well Number	Longitude	Latitude
1	149.6188	-30.5792
2	149.6423	-30.697
3	149.7673	-30.6375
4	149.709	-30.7263
5	149.634	-30.3906
6	149.6258	-30.4608
7	149.7477	-30.4802
8	149.7552	-30.3881
9	149.5265	-30.5329
10	149.7544	-30.5414

As discussed earlier, an addition of 10 monitoring bores results in an increase in the monitoring bore density from 4.2 bores / 100 km^2 to 5.4 bores / km^2 within the 95th percentile predicted drawdown area. In Queensland, a targeted monitoring network design for the Surat Cumulative Management area adopted a density of 0.3 bores/ 100 km^2. The design we considered in this study used a network of 10 bores with the intent of demonstrating the utility of the design and does not conform to any required density stipulated by regulators. Such decisions are made by the regulator considering broader aspects of water resource management in the region.

8   Conclusions

This study demonstrated an iterative approach for risk assessment and monitoring for adaptive management of impacts on onshore gas development on groundwater resources. The risks to groundwater quantity and quality are initially quantified with a groundwater model that is underpinned with limited available datasets. The models are then used to inform the relative worth of existing and potential future datasets, their types and locations that can be most useful in informing and reducing uncertainty in the predictions of interest. The predictive analysis is then used for designing optimal monitoring strategies.

Predictive analysis of drawdown indicates that the predicted median of CSG-induced drawdown for the important GAB aquifer, Pilliga Sandstone, is less than 0.2 m for most of the areas within and beyond the project area. The model simulations and uncertainty analysis indicated that it is highly likely that CSG-induced drawdown in the GAB aquifer, Pilliga Sandstone, would be less than 1.2 m if CSG development similar to the generic scenario considered in this study were to go ahead.

We also undertook probabilistic analysis of travel times and distances over which water quality changes would be expected to occur, should any contaminant particles migrate to the GAB aquifer because of well integrity loss or accidental damage. Probabilistic simulation of travel times using forward particle tracking indicated that the velocity of flow of groundwater in the GAB aquifer is very slow and particles are likely to travel a maximum of only 100s of metres within a design period of 100 years. The maximum travel distance obtained from the probabilistic particle tracking analysis was 468 m. Spatial analysis considering the proximity of CSG bores to risk receptor bores and particle travel distances and directions revealed that there were zero bores that may intercept the water flowing from CSG bore locations which would pose a contamination risk. Hence, water quality monitoring was not considered as one of the objectives in the optimization-based monitoring network design. Instead, the spatial proximity analysis was employed to identify areas where receptor bore density, CSG well proximity and travel path densities were high within the gas project area.

The spatial patterns of predicted drawdown and particle travel distances were used in a mathematical model reduction framework that enabled the knowledge to be used within an optimization routine to identify optimal locations for monitoring. The value of the monitoring bores as a network for minimising prediction uncertainty across the region is demonstrated by designing a 10-bore monitoring network that collects drawdown data from multiple depths at each location. Geostatistical analysis of baseline water quality data collected from the GAB aquifer, the Pilliga Sandstone, was undertaken to identify areas where estimates of water quality are uncertain. This knowledge was used together with a global optimization algorithm Differential Evolution to identify optimal locations for monitoring baseline water quality.

The relative worth of monitoring data collected at potential monitoring locations was further evaluated using the data-worth analysis techniques. Data worth analysis was undertaken to analyse the parameter contributions to predictive uncertainty and relative worth of monitoring data in reducing the prediction uncertainty of CSG-induced drawdown at these locations. Since the

predicted level of drawdown is small for the considered generic scenario of coal seam gas development, the drawdown simulated in the GAB aquifers are found to be more dependent on how the CSG-induced flux losses are compensated by release of water from storage and/or river or recharge processes.

The analysis indicated that hydraulic characteristics of deeper formations in the Surat and Gunnedah basins govern the likelihood of propagation of CSG-induced drawdown into the GAB and other aquifers closer to the surface. For example, it was found that the data worth for informing drawdown predictions in the Pilliga Sandstone was for monitoring data collected from the model layer immediately below the Pilliga Sandstone (i.e. corresponding to the Purlawaugh formation). While the monitoring network design presented in our study used the spatial patterns of predicted drawdown in multiple model layers, the data worth analysis was independent of the absolute values of predictions and used only the sensitivity of predictions to observations and parameters. Thus, data worth analysis provided an independent method for evaluating the veracity of the design. The predictive analysis, monitoring network design and data worth analysis, in general, indicated that having multi-level piezometers in and around the 95th percentile drawdown extent can help reduce uncertainty in CSG-induced drawdown predictions. The probabilistic particle tracking and CSG well-to-receptor proximity analysis helped to identify areas of relatively higher importance for monitoring water quality impacts, while the analysis indicated that water quality risks pertaining to the regional scale that was considered in this study is very low.

Appendix A Histograms of drawdown prediction at 40 receptor locations

The following showcases the histograms of drawdown prediction at 40 different receptor locations. Note that lp refers to parsimoniously parameterized while hp refers to highly parameterized models.

(a)

(b)

(c)

4_46307: dmax (m)

(a)

(b)

(c)

5_56354: dmax (m)

(a)

(b)

(c)

7_111314: dmax (m)

(a)

(b)

(c)

7_112037: dmax (m)

8_130599: dmax (m)

8_131843: dmax (m)

10_194545: dmax (m)

11_204022: dmax (m)

12_240815: dmax (m)

13_274779: dmax (m)

(a)

(b)

(c)

13_276710: dmax (m)

(a)

(b)

(c)

13_281631: dmax (m)

14_308559: dmax (m)

14_308883: dmax (m)

14_316760: dmax (m)

15_323781: dmax (m)

15_326464: dmax (m)

15_340263: dmax (m)
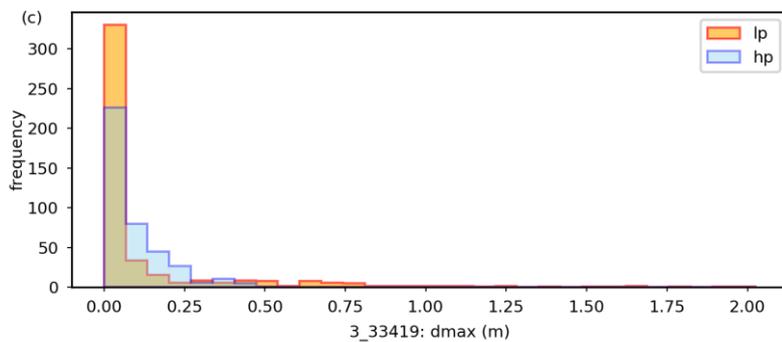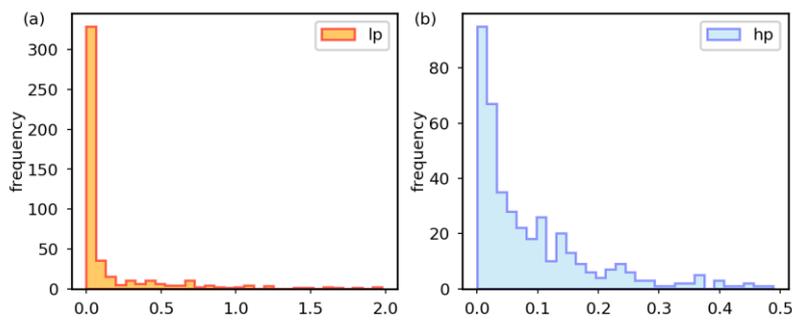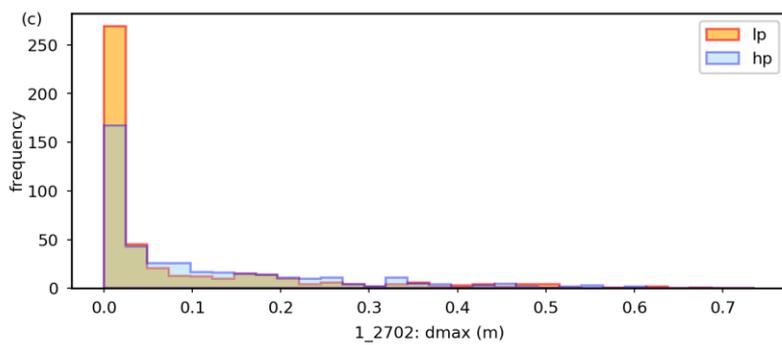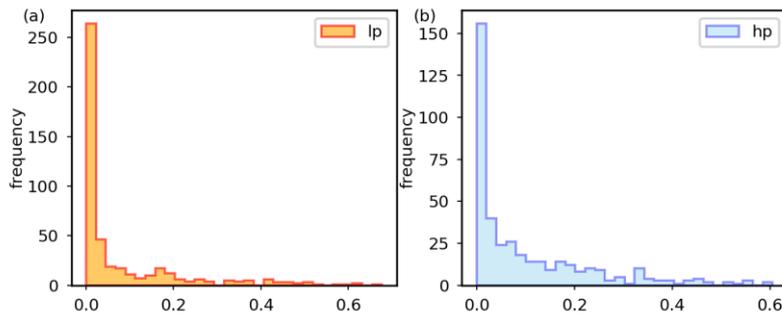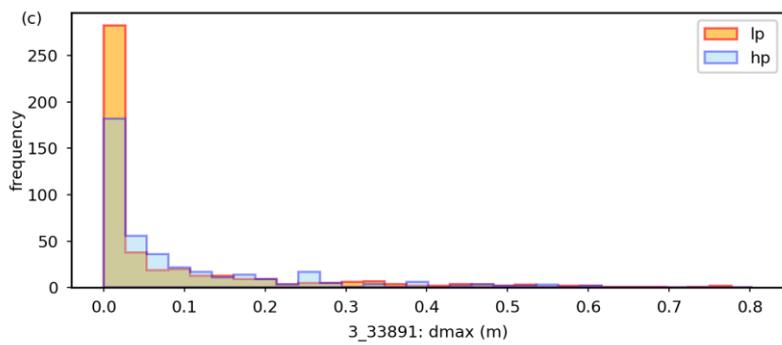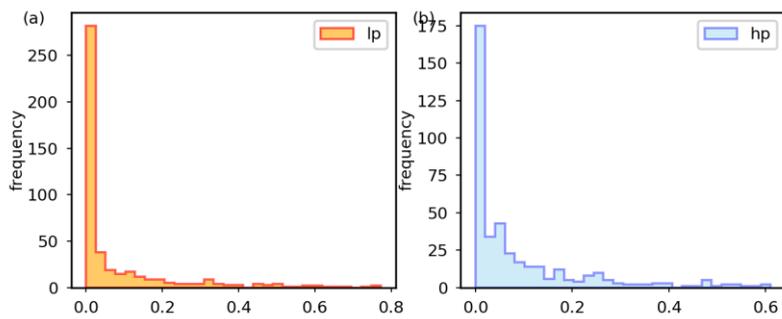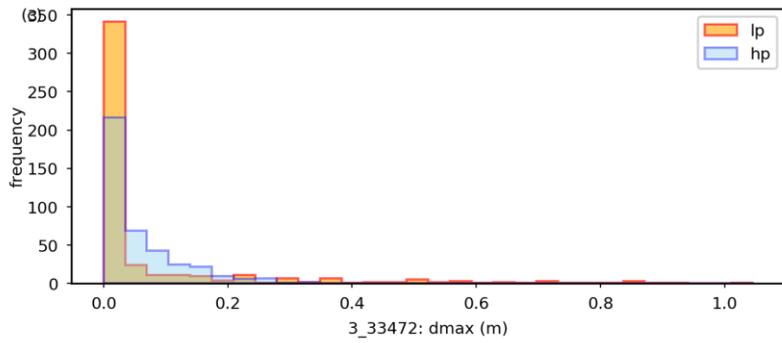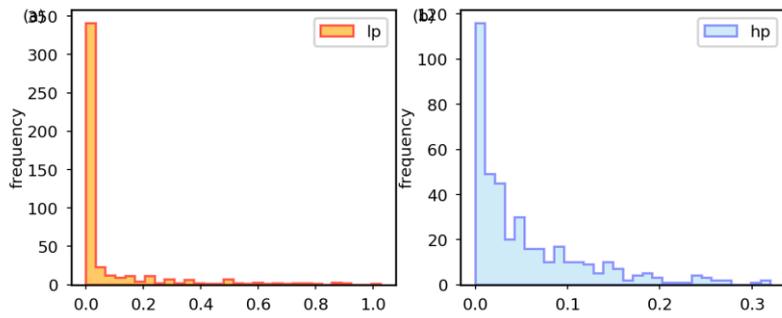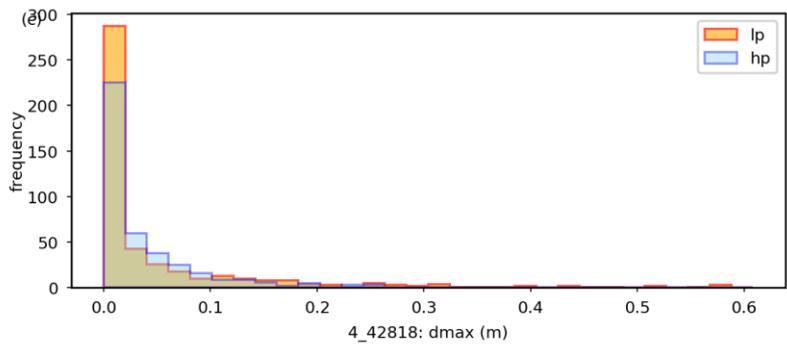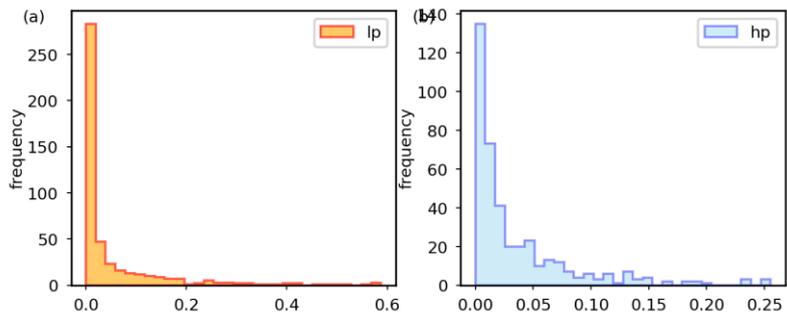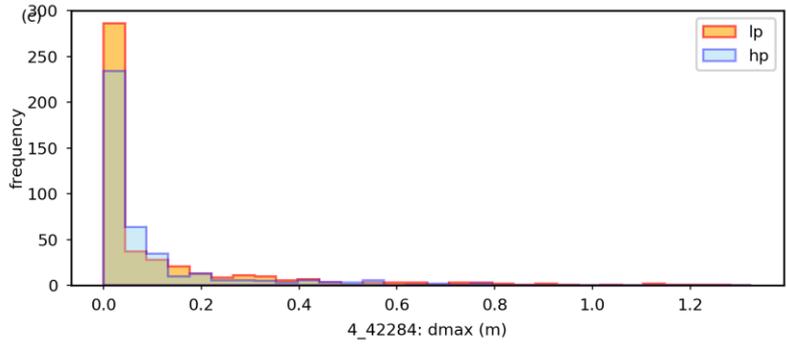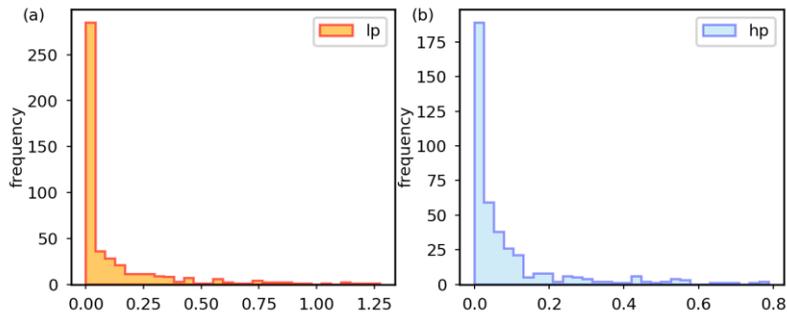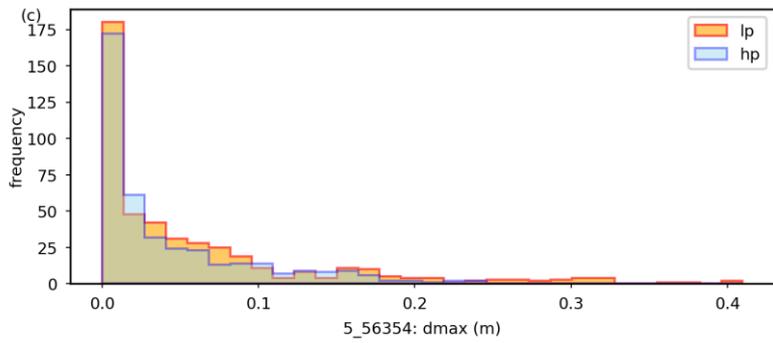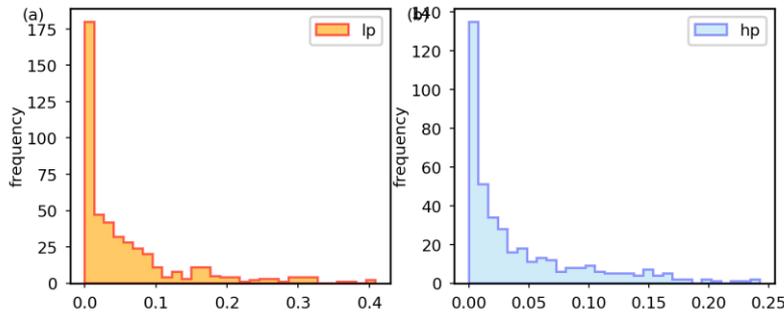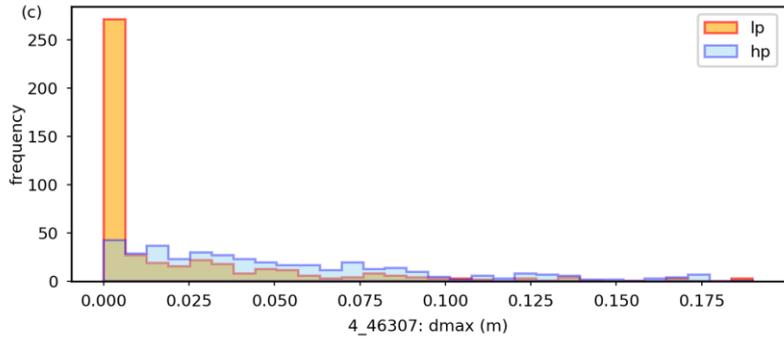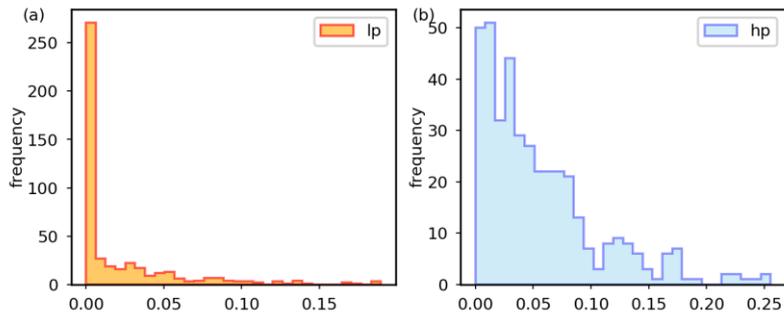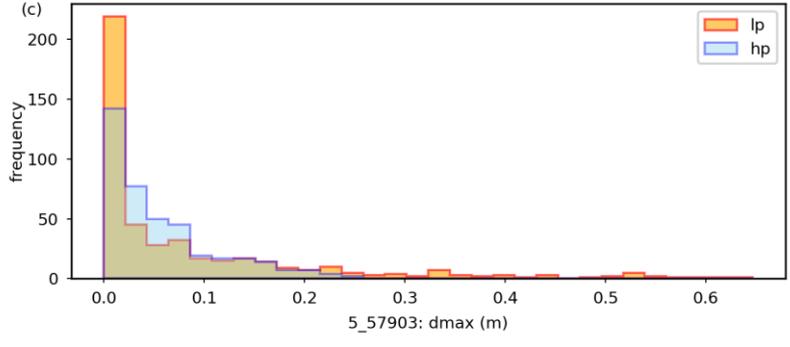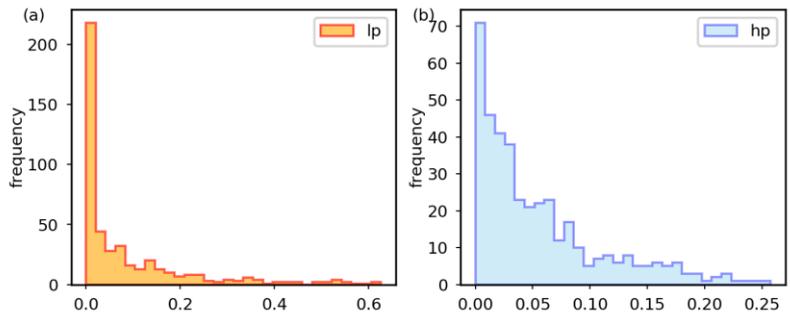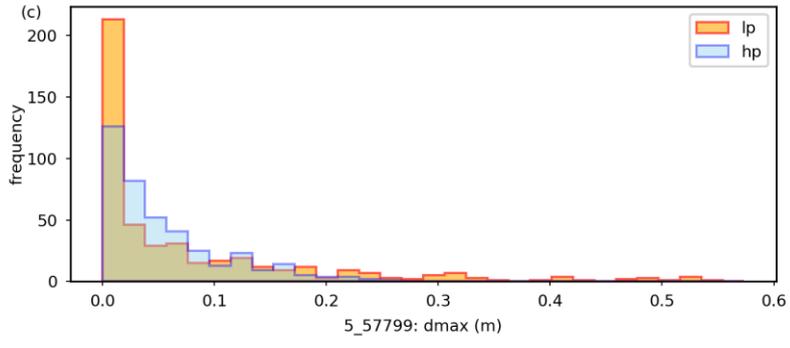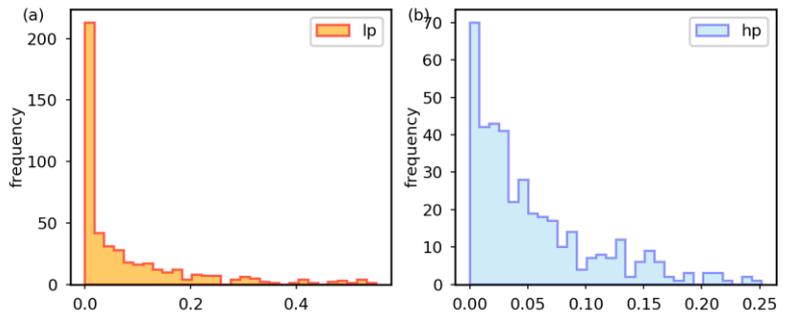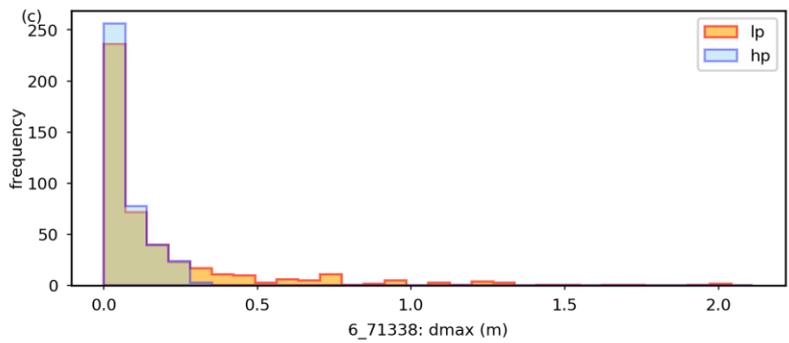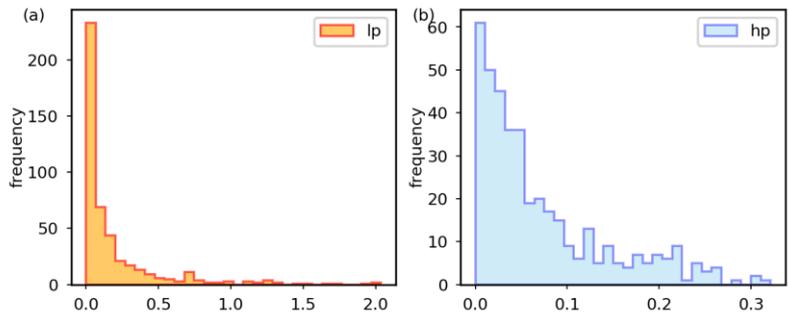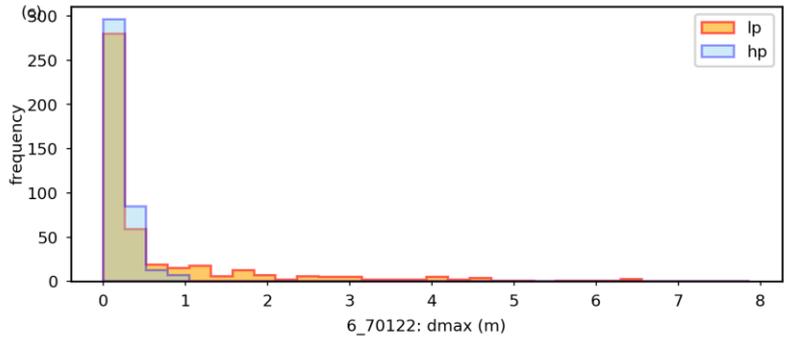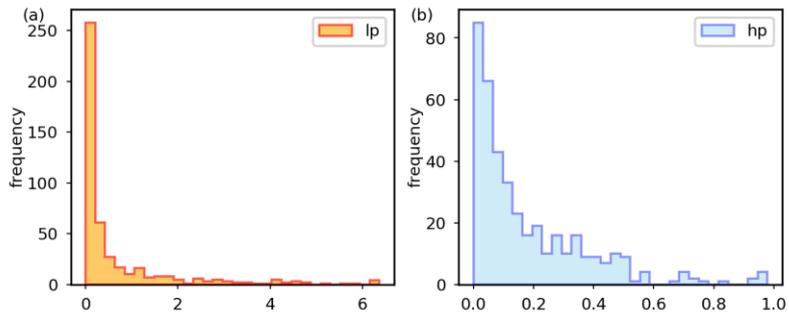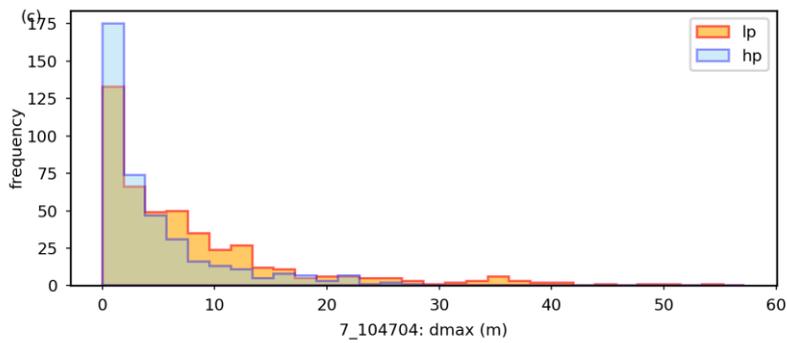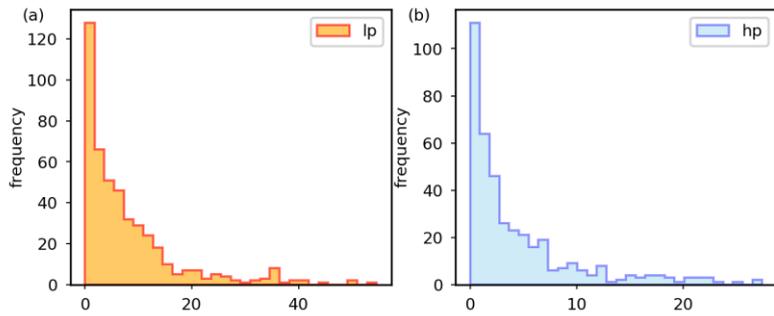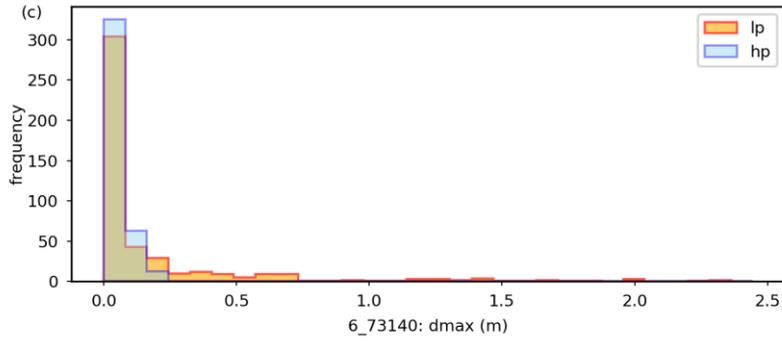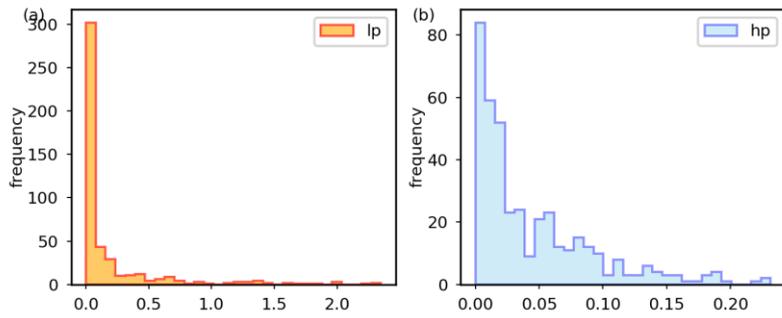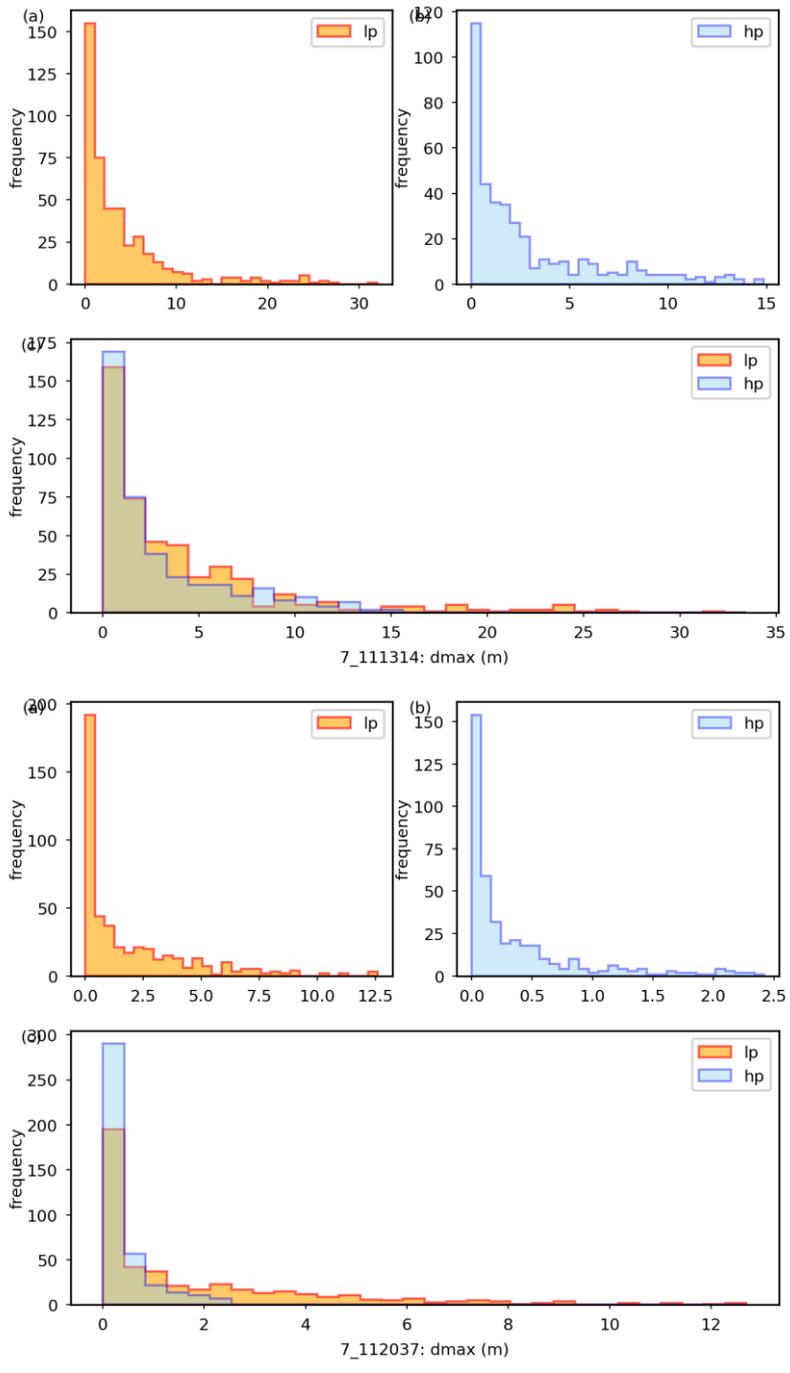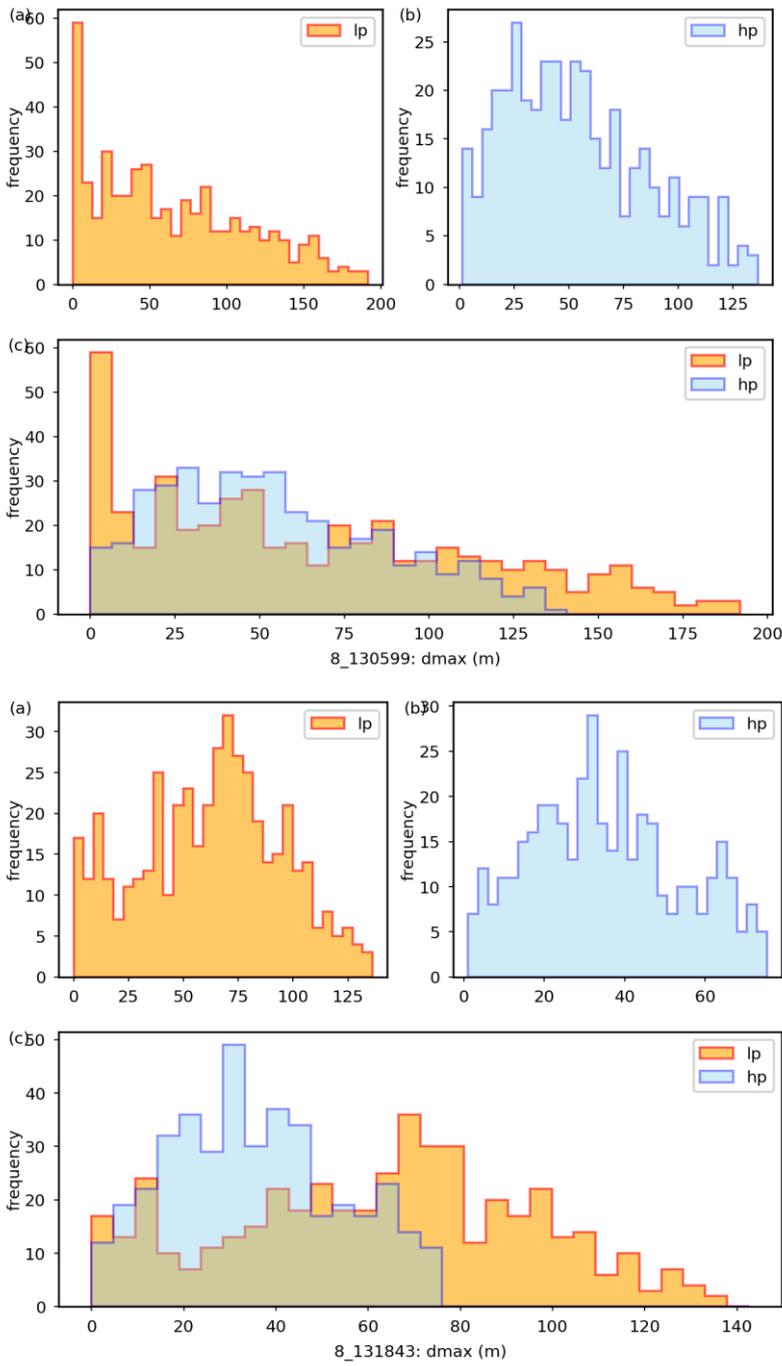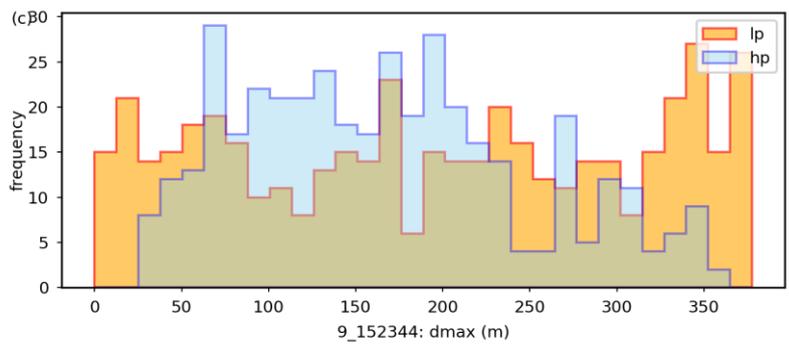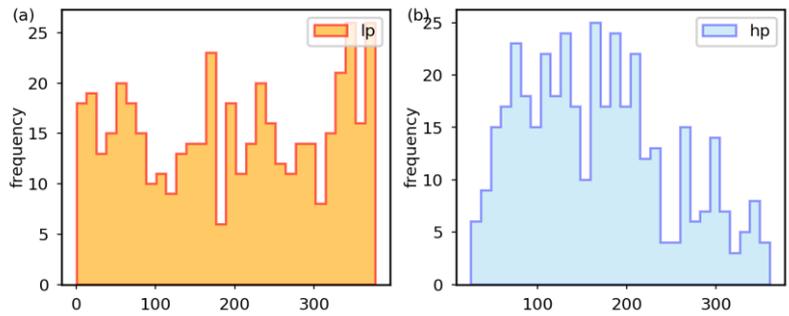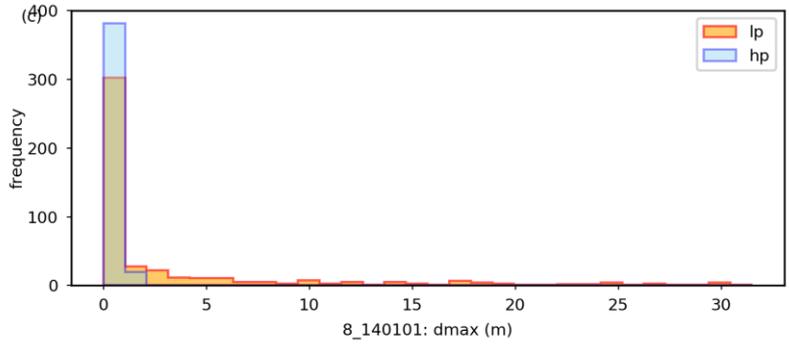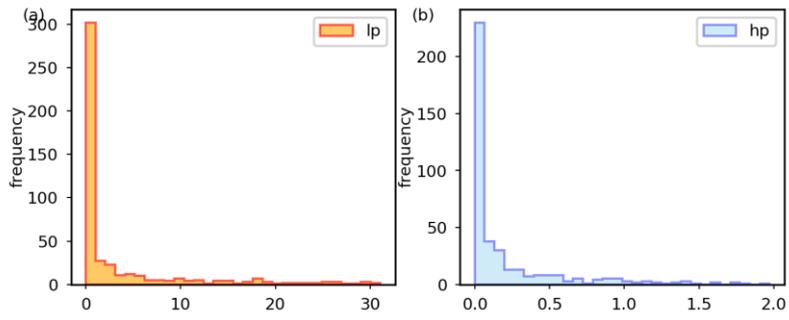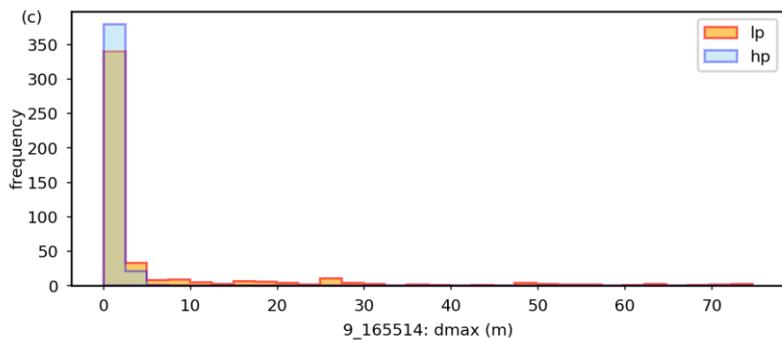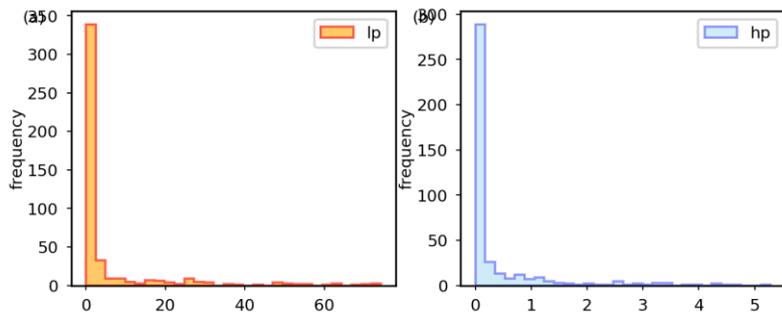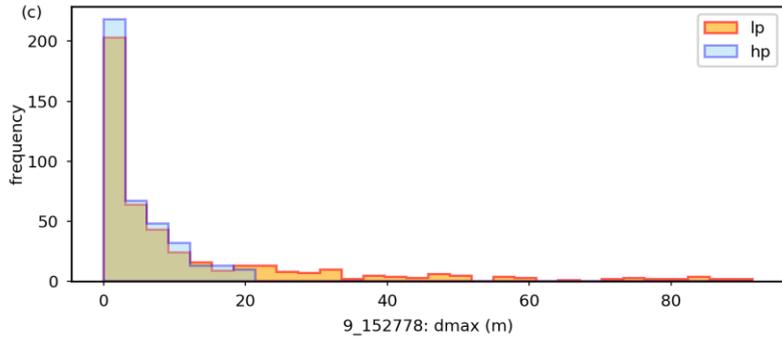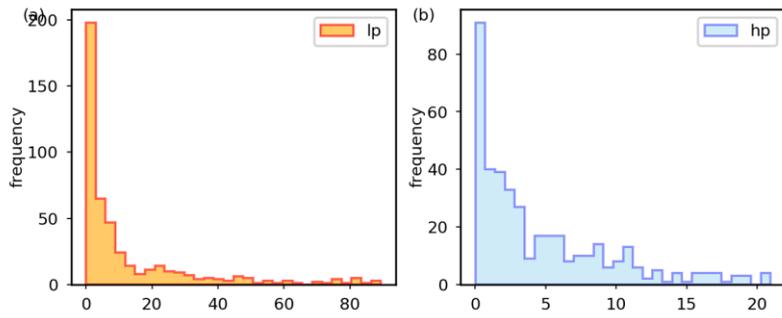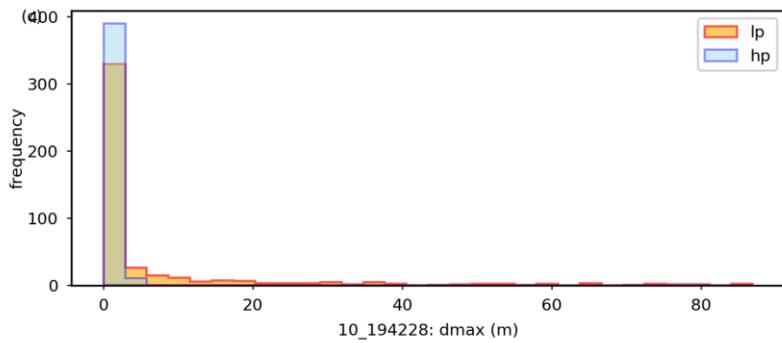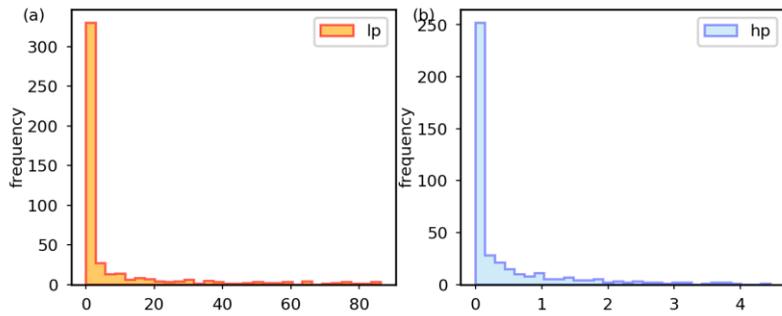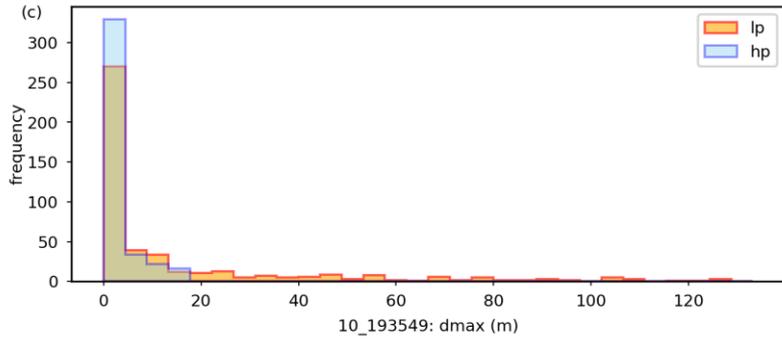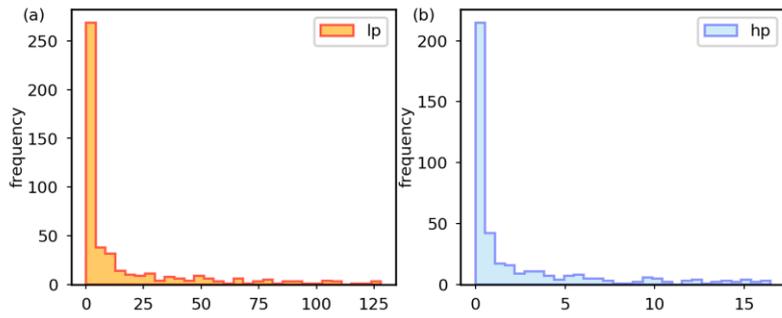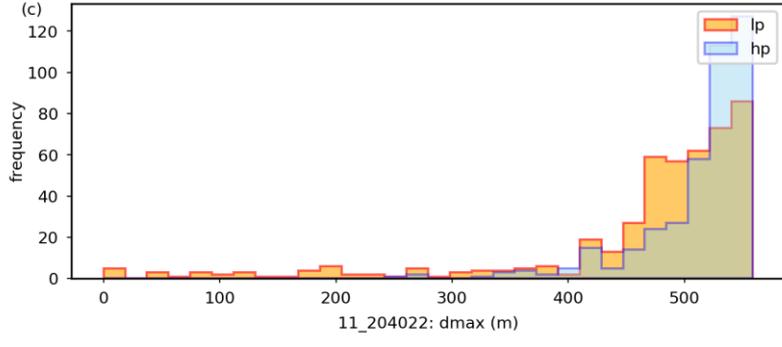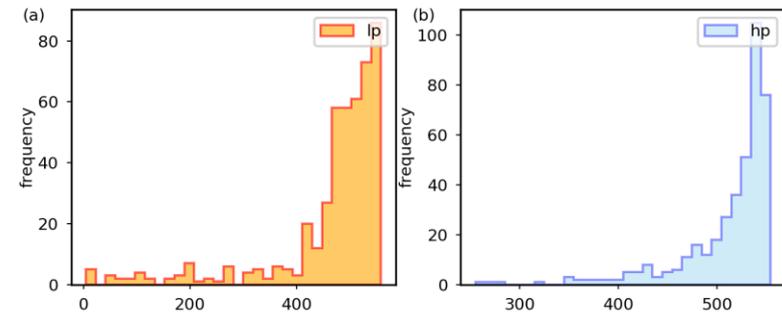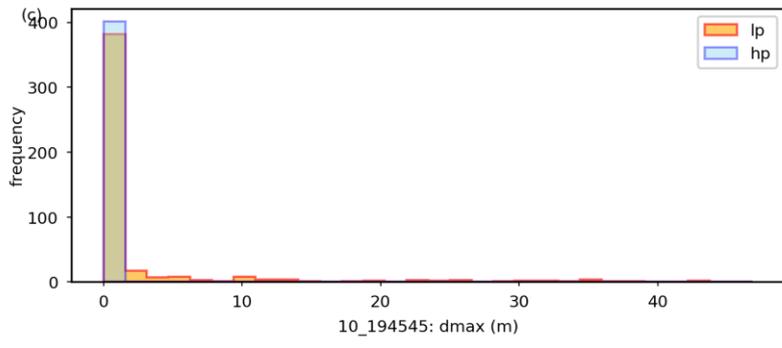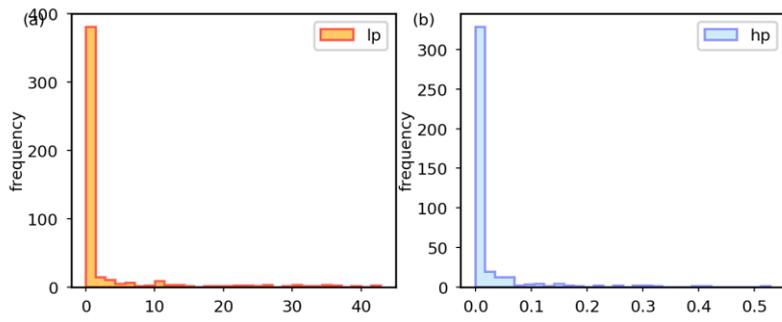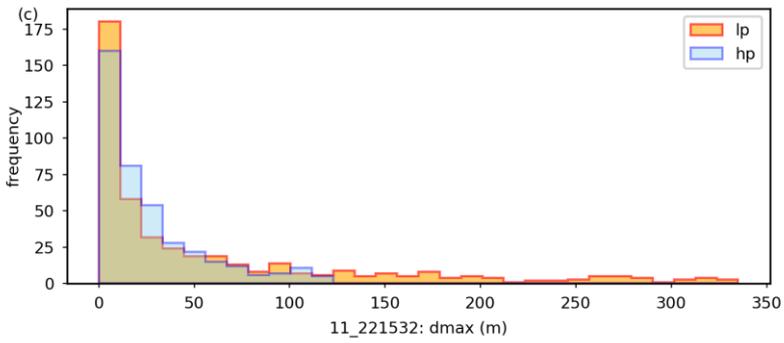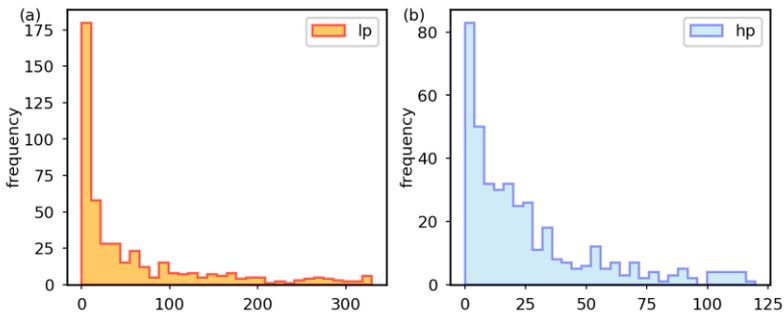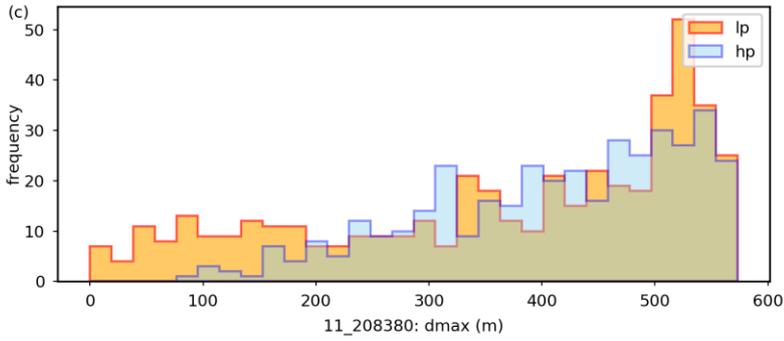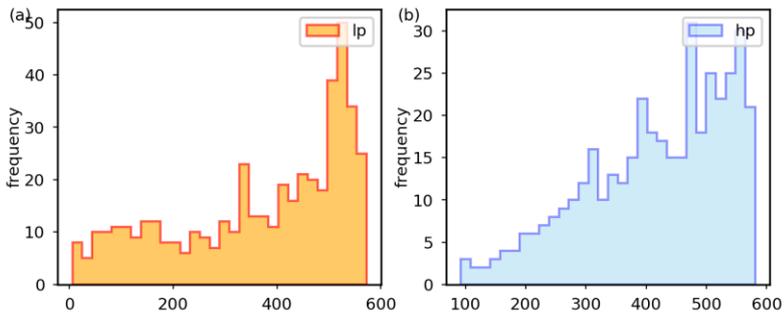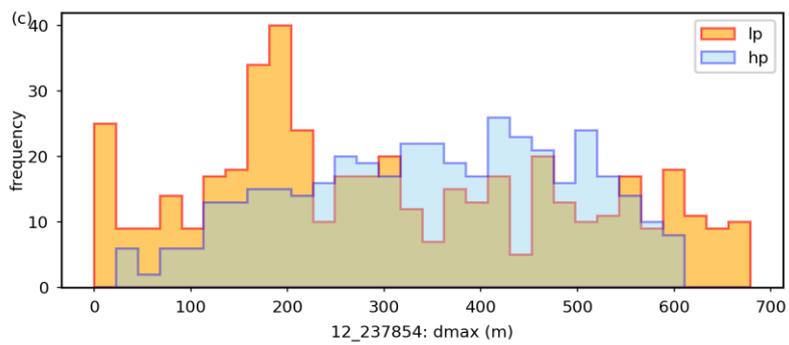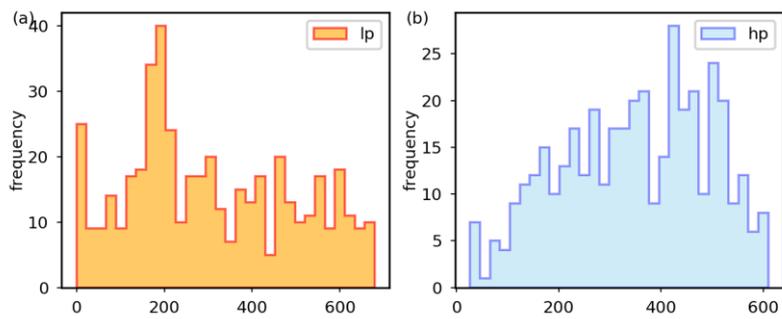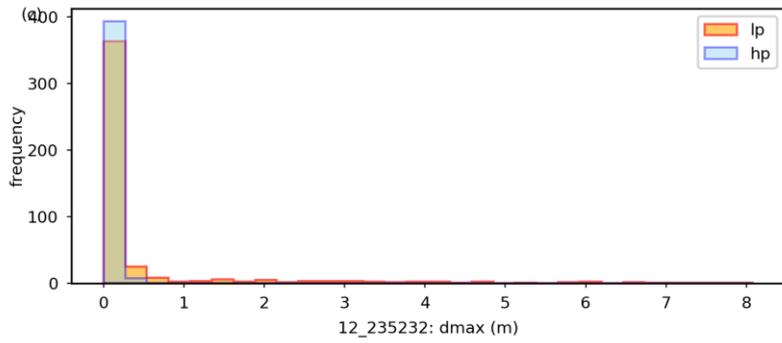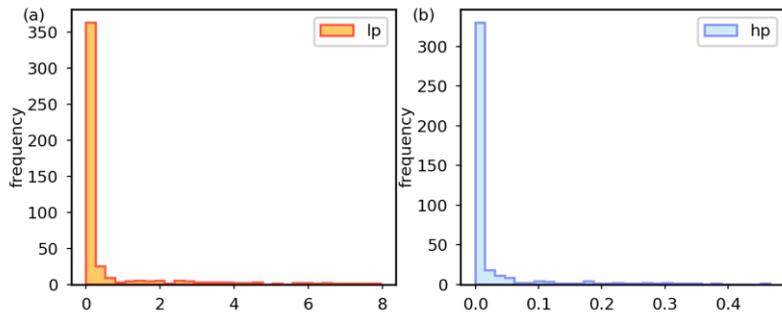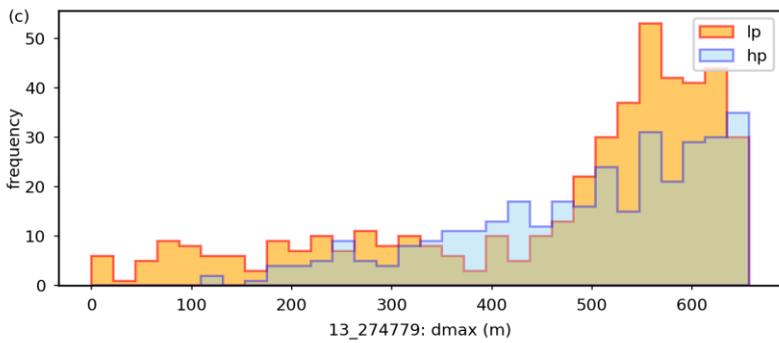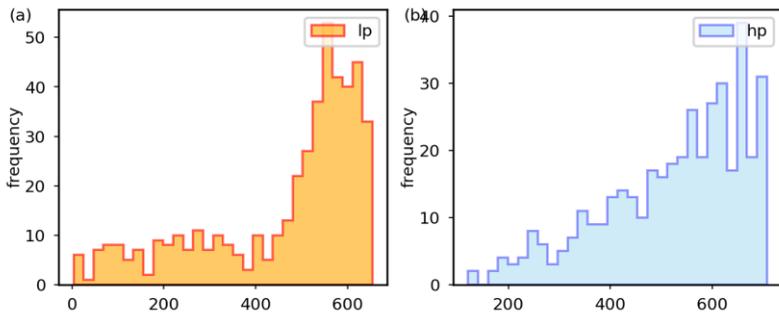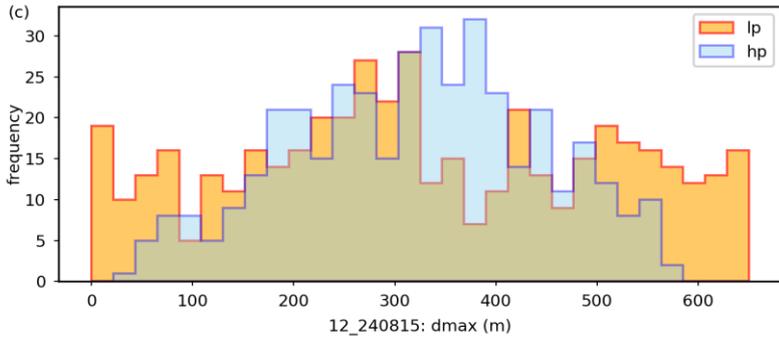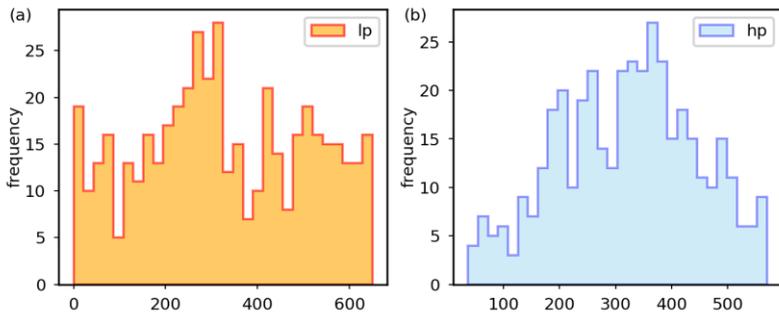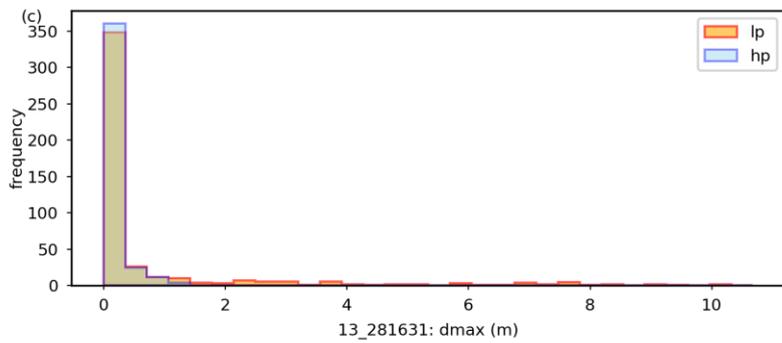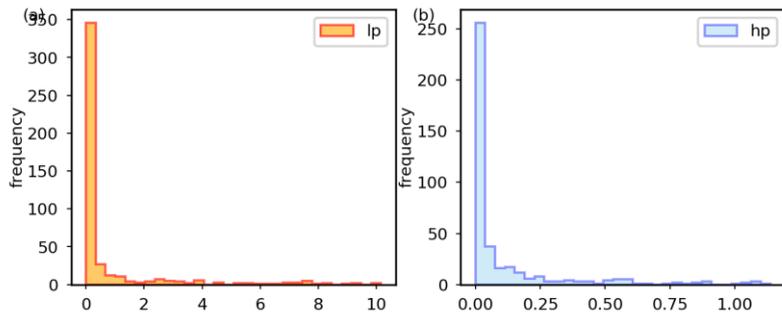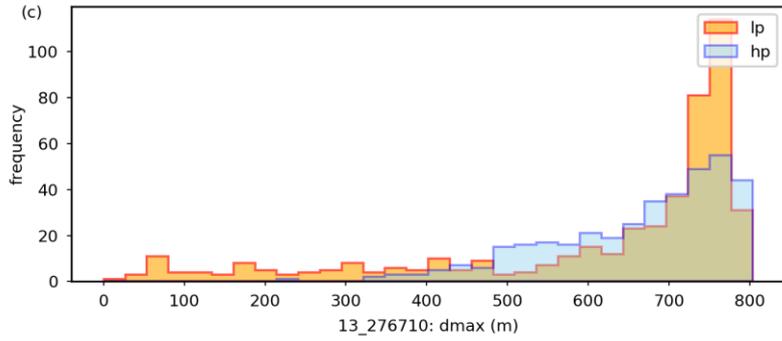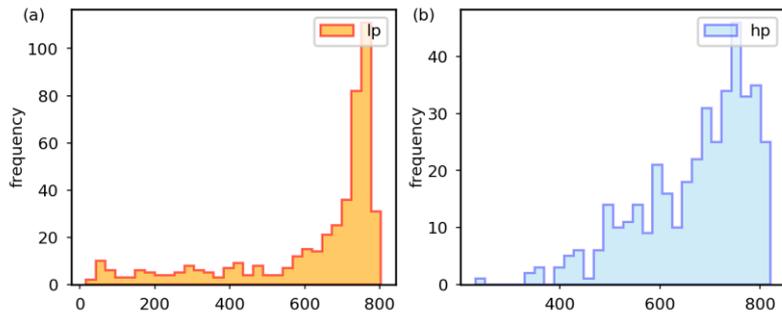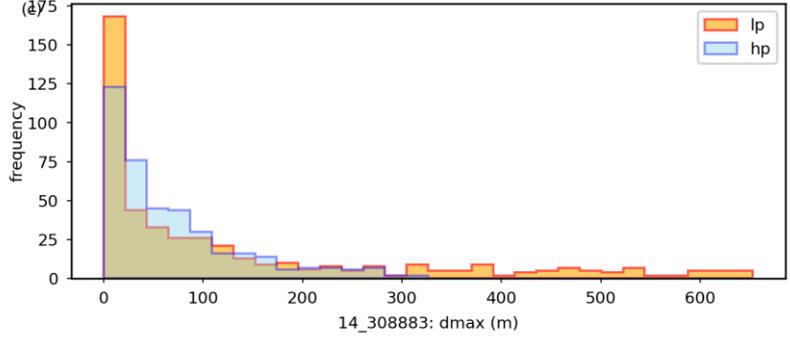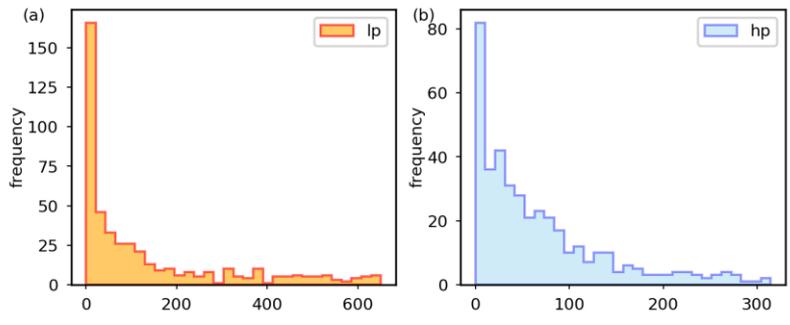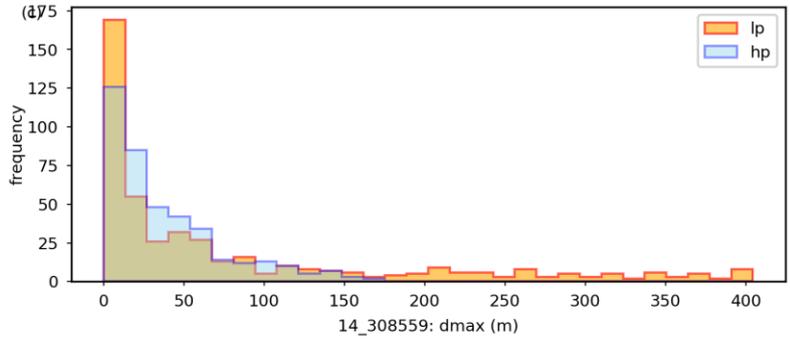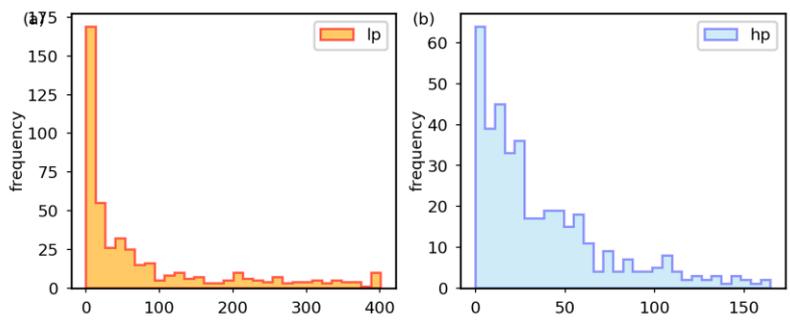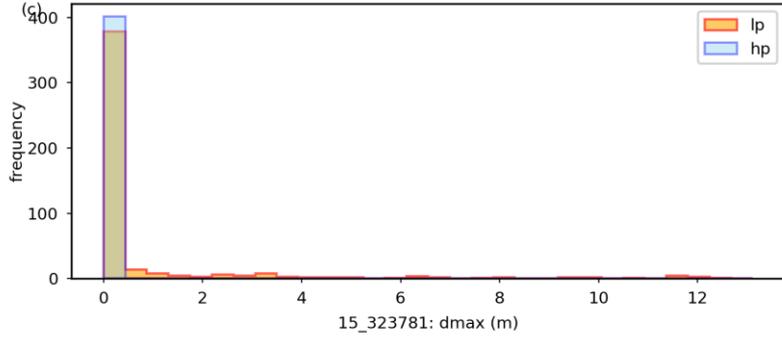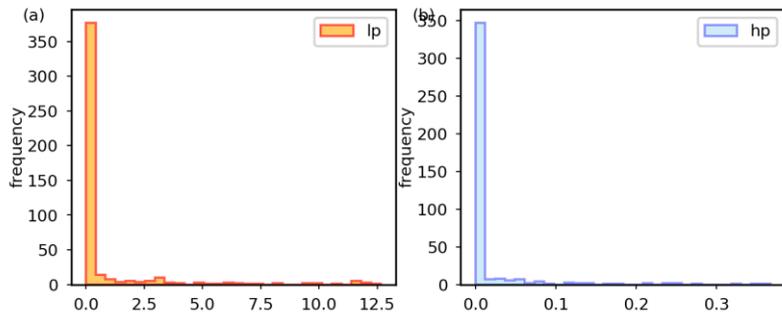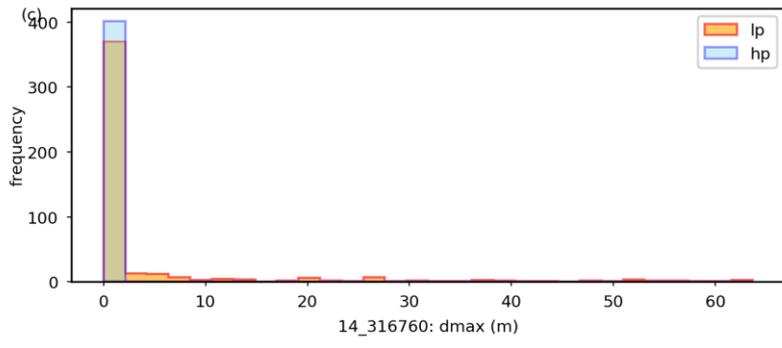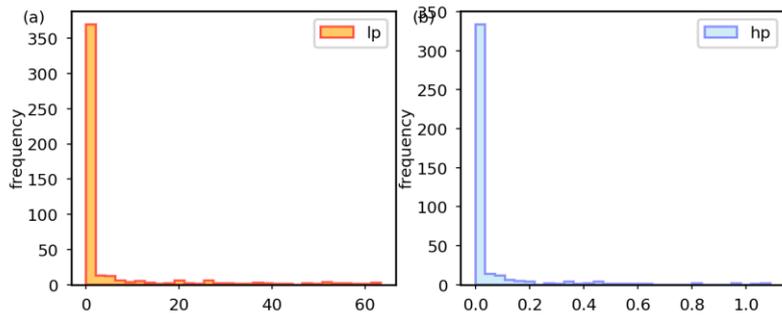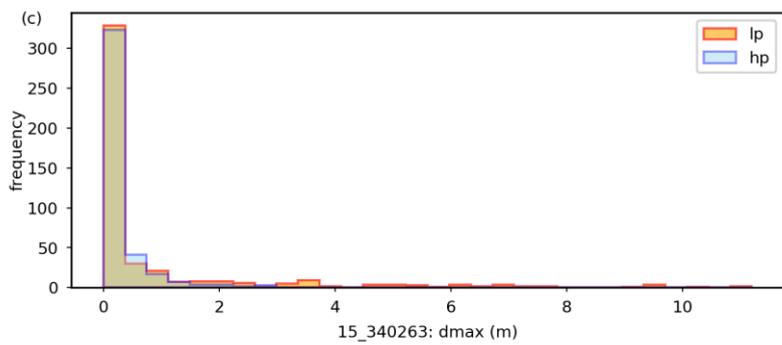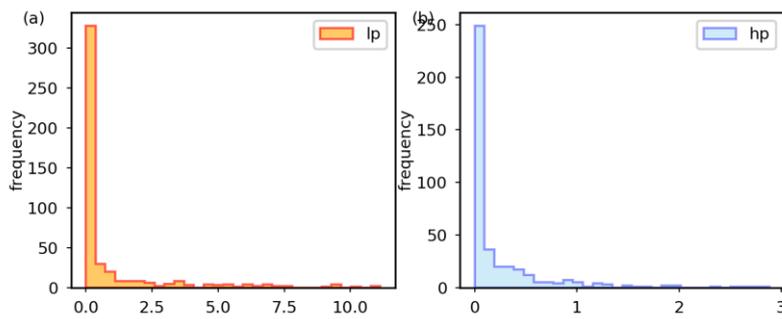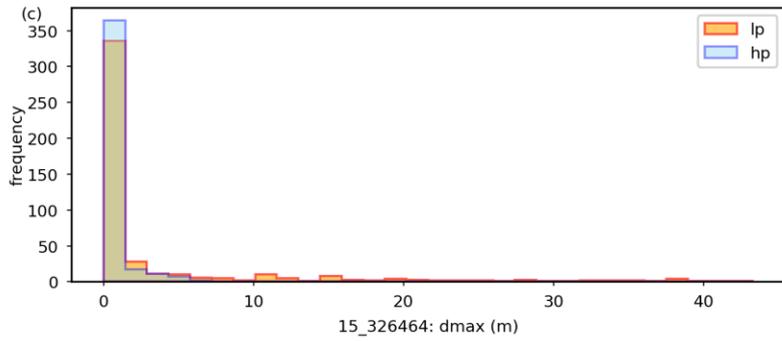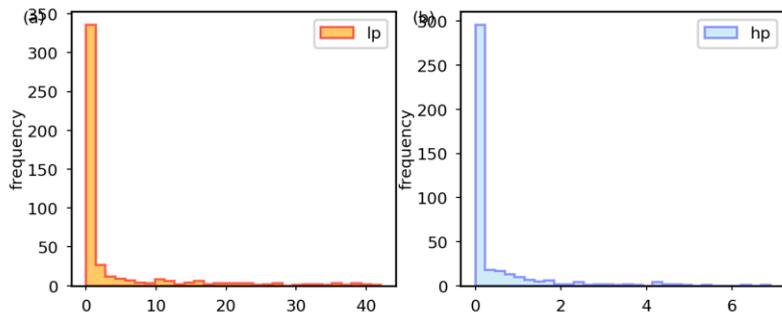
References

CDM Smith (2016) Narrabri Gas Project Groundwater Impact Assessment Report, Santos Limited.

Christensen, S., & Doherty, J. (2008). Predictive error dependencies when using pilot points and singular value decomposition in groundwater model calibration. *Advances in Water Resources*, *31*(4), 674-700.

Cressie, N. and Wikle, C.K. (2011) *Statistics for Spatio-Temporal Data*. Wiley & Sons.

Crosbie, R., Peeters, L. and Carey, H. (2016) Groundwater modelling. Submethodology M07 from the Bioregional Assessment Technical Programme. Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia, Australia. http://data.bioregionalassessments.gov.au/submethodology/M07.

Dausman, A.M., Doherty, J., Langevin, C.D. and Sukop, M.C. (2010) Quantifying data worth toward reducing predictive uncertainty. *Groundwater* 48(5), 729–740.

Doherty, J. (2015) *PEST and its Utility Support Software, Theory*. Watermark Numerical Publishing.

Engelhardt, I., Prommer, H., Moore, C., Schulz, M., Schüth, C., & Ternes, T. A. (2013). Suitability of temperature, hydraulic heads, and acesulfame to quantify wastewater-related fluxes in the hyporheic and riparian zone. *Water Resources Research*, *49*(1), 426-440.

Fienen, M.N., Doherty, J.E., Hunt, R.J. and Reeves, H.W. (2010) *Using prediction uncertainty analysis to design hydrologic monitoring networks: Example applications from the great lakes water availability pilot project* (No. 2010-5159). US Geological Survey.

Dirk Mallants, Simon Apte, James Kear, Chris Turnadge, Sreekanth Janardhanan, Dennis Gonzalez, Mike Williams, Zuorong Chen, Rai Kookana, Andrew Taylor, Matthias Raiber, Merrin Adams, Jody Bruce, Henning Prommer (2017). Deeper groundwater hazard screening research, prepared by the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Canberra.

Mullen, K.M., Ardia, D., Gil, D., Windover, D. and Cline, J. (2011) DEoptim: An R package for global optimization by differential evolution *J. Stat. Software* 40(6), 1–26.

Papadopoulos, S and Associates (2016) mod-PATH3DU, A groundwater path and travel time simulator, http://www.sspa.com/software/mod-path3du

Raiber, M. and Suckow, A. (2018) Hydrochemistry of the Pilliga Sandstone aquifer in NSW – data availability, preliminary assessment of spatial patterns and conceptual model uncertainties. CSIRO, Australia

R Core Team (2018) *R: A Language and Environment for Statistical Computing, R Found. for Stat. Comput.*, Vienna. [Available at http://www.R-project.org/.]

Sreekanth, J., Crosbie, R., Pickett, T., Cui, T., Peeters, L., Slatter, E., Northey, J., Merrin, L.E., Davies, P., Miotlinski, K., Schmid, W. and Herr, A. (2017a) Groundwater numerical modelling for the Namoi subregion. Product 2.6.2 for the Namoi subregion from the Northern Inland Catchments Bioregional Assessment. Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia, Australia.

Sreekanth, J., Lau, H. and Pagendam, D. (2017b) Design of optimal groundwater monitoring well network using stochastic modeling and reduced-rank spatial prediction. *Water Resources Research* 53(8), 6821–6840.

Sreekanth, J., Cui, T., Pickett, T., & Barrett, D. (2017c). Uncertainty analysis of CSG-induced GAB flux and water balance changes in the Narrabri Gas Project area.

Storn, R. and Price, K. (1997) Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. *J. Global Optim.* 11(4), 341–359. doi:10.1023/A:1008202821328.

Tarantola, A. (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation*.

Society for Industrial and Applied mathematics, Philadelphia.

Tonkin, M., Doherty, J. and Moore, C. (2007) Efficient nonlinear predictive error variance for highly parameterized models. *Water Resources Research 43*(7). doi:10.1029/2006WR005348

Tonkin, M. and Doherty, J. (2009) Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques. *Water Resources Research* 45, W00B10, doi:10.1029/2007WR006678.

Turnadge, C., Mallants, D. and Peeters, L (2017) *Sensitivity and uncertainty analysis of a regional-scale groundwater flow model featuring coal seam gas extraction*. CSIRO, Australia.

White, J.T., Fienen, M.N. and Doherty, J.E. (2016) A python framework for environmental model uncertainty analysis. *Environmental Modelling & Software 85*, 217–228.

CONTACT US
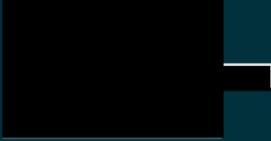
AT CSIRO, WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today – for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australia's catalyst for innovation.

CSIRO. WE IMAGINE. WE COLLABORATE. WE INNOVATE.

FOR FURTHER INFORMATION

CSIRO Land and Water