Where is All the Water?

Final Report – December 2021

Citation

Please cite as: Andersen, M.S.[5], Close, J.[1], Hu, T.[4], Legge, S.[1], McCallum, D.[4], Mukhopadhyay. S.[2], Runcie, P.[4], Rutlidge, H.[5], Tamsitt, L.[3], Shearan, B.[2], Tregoning, P.[1], Vervoort, R.W.[6] (2021) Where is All the Water? NSW Government report prepared by the NSW Smart Sensing Network.

1. Australian National University
2. Macquarie University
3. NSW Department of Planning, Industry and Environment
4. NSW Smart Sensing Network
5. UNSW Sydney
6. The University of Sydney

W: nssn.org.au

E: admin@nssn.org.au

Twitter: @nswsensing

Foreword

The NSW Office of the Chief Scientist and Engineer is proud to have contributed financial assistance, together with the NSW Department of Planning, Industry and Environment (DPIE), and the NSW Smart Sensing Network (NSSN) to the *Where is All the Water?* project.

The idea for an intricate sensing project to aid understanding of catchments in NSW arose from a number of preceding reviews and projects originating from my own office and that of then NSW Minister for Water, Property and Housing, Hon. Melinda Pavey, and Dr Jim Bentley, Deputy Secretary (CEO of the NSW Water Sector), Department of Planning, Industry and Environment.

The early stages of the *Where Is All the Water?* project coincided with of one of the worst droughts on record, culminating in the catastrophic bushfires of 2020; the preparation and delivery of the project was during the worst health crisis in living memory; while the final stages and reporting have been impacted by extreme flooding in 2021-22.

In July 2020 my office released *its Review of water related data collections, data infrastructure and capabilities.* Also requested by Minister Pavey, the report is prompted by the question: do we have the right data, of sufficient quality, and in useable form to make well-informed decisions? While that report is primary concerned with water data governance and management, it does show there is a lack of simple, open, telemetered data across NSW.

The *Where is All the Water?* project has shown for the first time that low-cost sensor networks can respond to the Australian infrastructure problem of great distances and low population. It proves that local gravity measurements – soon likely to be augmented with quantum sensing capability – are a technology that can help us map the underground; and that satellites have an important role to play in the management of resources. It also reveals that we still have a lot to learn about aquifer recharge. Wrapped around this is the power of data, particularly the analysis of uncertainty, to reveal problems and solutions in our physical world.

Given the Ministerial impetus to innovate, the NSW Government agencies responsible in the management of water are encouraged to read, understand, follow-on and where practical, adopt the recommendations of this report.

Prof Hugh Durant-Whyte

NSW Chief Scientist and Engineer

Executive Summary

The *Where is All the Water?* exploratory research project helps NSW government agencies and other organisations who want to improve aspects of their water management by overcoming gaps and discrepancies in the data of water assets, and provides a research platform for integrating different types of sensors and the data analytics used to aid modelling, predictions and decision making. This is made possible by being able to bring together innovative sensors such as satellite and ground level gravity sensors, to low-cost sensor modules that can accept a wide range of physical sensors (temperature, moisture, water levels, and so on), Bayesian analytical tools and an expert understanding of the state's hydrology.

The project helps us better understand *Where is All the Water?*, improving our broad measurement capabilities, and reducing the extent of "unaccounted differences" (which can be greater than 25%), due to our inability to measure and understand variable natural processes. Agencies seek the latest research, scientific and technical tools to manage the capture, storage and release of water: The right amount of water, in the right place, at the right time. Furthermore, they seek to understand how better sensing can provide data to improve knowledge of ground water resources, what groundwater is influenced by and influences on; and thus aid more robust, timely, scale appropriate decision making.

This collaborative program brought together university and government partners with the aim of developing a probabilistic modelling framework and roadmap to improve integrated and evidence-based management of water resources in NSW.

The project partners included a number of university partners. The **Australian National University** (ANU), led two sub-projects; Local gravity sensing, and an investigation into satellite gravity measurements of Australian water. **Macquarie University** brought expertise in low-cost monitoring, demonstrating the utility of high spatial resolution sensing using low-cost sensors. Probabilistic modelling was done by the **University of Sydney**'s ARC Training Centre in Data Analytics for Resources and Environments (DARE). The **University of NSW** (UNSW) brought a wealth of background knowledge on all aspects of hydrology, and specifically led an investigation into recharge mechanisms of aquifers.

The NSW Department of Planning and Environment (DPE)[1] were the key audience for the work and seen as the main problem owner. The research outcomes of this project are intended to provide tools that benefit all of NSW, but particularly those functions of water management overseen by DPE. With responsibility for monitoring, compliance and education around water laws, NRAR are an important audience and provider of the problem statement. With detailed knowledge of the NSW water catchments and holder of important data sets, WaterNSW are an important audience and provider of the problem statements.

The NSW Smart Sensing Network coalesced the partners and coordinated the project and assisted with funding. They also bring resourcing from the NSW Office of the Chief Scientist and Engineer.

[1] The agency was named DPIE over the life of the project. The name became the Department of Planning and Environment (DPE) after project activities ceased.

Key findings

This project has demonstrated that through multi-disciplinary collaboration, better understanding and information can be collected than each method alone. The synergy of physical sensing and the use of data provides a clear example of this; through data analysis we can provide stronger statistical understanding of what types of physical monitoring is required and where, how sensor fusion can be beneficial, and an understanding of where remote methods provide the most value.

Gaps in water data necessitate the need for more physical sensors to improve our understanding. A quick look at the NSW state map of water sensor distribution shows large areas not currently sensed or telemetered. Low-cost sensors, by which we mean sensors not bearing the high costs of weather stations, borehole drilling, wired communication, manual servicing, data collection and so on, are an option to explore as they can disengage agencies from commitments to expensive infrastructure. Low-cost sensors placed at strategic locations can provide data that can reduce data uncertainties that arise from modelling assumptions. The sensors can also be deployed as reference points for large-scale sensing techniques such as emerging gravity sensing methods.

A gravity signal cannot be shielded or blocked by anything. This means information about the underground can be ascertained by a sensor above ground, without the need for digging, tunnelling, insertion of probes, chemical analysis or other contact. ANU previously demonstrated that a leaking water pipe underground in an urban environment can be detected with a gravimeter, a lawnmower-sized device above ground. Simulations of gravitational signals from realistic models of groundwater systems show that quantum gravity sensors can inform catchment water accounting through remote sensing of the total water mass changes, bring new information about the underground to existing data sets.

Satellites that measure changes in the gravitational pull of water on Earth (the NASA Gravity Recovery and Climate Experiment (GRACE) missions) can be used to quantify changes in total water storage at global, continental and basin scales. Combined with other data sources (e.g., soil moisture and ground-based gravity observations and models), the spatial resolution can be improved and estimates of groundwater changes can be made. Space gravity data provide an effective means of observing volumetric changes in water resources state-wide in places currently lacking in situ measurements and goes beyond remote imaging methods which rely on processing of two-dimensional data. NASA have used these satellites to measure the amount of water diminished because of the Californian drought, to give early clues to flood danger in the Mississippi delta, and such data have helped find water deposits in the central Asian deserts.

There are significant unknown losses and gains through natural processes in the water balance of NSW river systems. The University of Sydney (Data Analytics for Resources & Environments Centre) has developed a prototype formal framework to quantify these losses and the associated uncertainties. The preliminary data analysis has highlighted clear uncertainties in gauging data, groundwater space and time observations and understanding river width. In prioritising the list of uncertainty of physical variables this work guides investment in data collection and placement of new sensors, and more generally reduces uncertainty in water balance terms and reduce risk for decision makers.

UNSW showed that focussed recharge (e.g. inflows from surface water) is the main groundwater recharge mechanism in areas adjacent to intermittent streams, with on average one to two recharge events per year. In contrast, diffuse recharge of groundwater (e.g. rain percolating down through the soil) is more episodic, with only one to two significant events, depending on location, over a ten-year period. This process of focussed groundwater recharge illustrates a mechanism to explain surface water transmission losses of flow releases in dry streams. This process also potentially means that direct rainfall on fields may not be a determining factor for groundwater recharge, especially if flow is generated by rainfall further upstream.

Cumulatively this work has demonstrated the importance of a broad collaboration of researchers working in close harmony with government agencies responsible for water management. When

water administration decisions are at their most critical, such as in times of drought (low flow use cases), tools being developed by researchers in universities can be translated for adoption in water management.

Established hydrological experts (e.g. DPE, DARE, UNSW, WaterNSW) are able to continue to develop their understanding of water knowing that plentiful data, from both a cheaper proliferation of sensors of established physical parameters (Macquarie) and new techniques such as those around gravity (ANU). The inverse is also true, with those developing sensors, best able to do so when working side by side with those closest to the water issues.

The input of multiple types of new sensor empowered with Bayesian inference gives real world credence to emerging paradigms like *sensor fusion* and *predictive intelligence*, and provides an excellent platform to show off advances in *quantum science*.

Recommendations

Continue engagement with closer collaboration of agency technical and scientific staff with researchers. This would enable better guidance of researchers, a better understanding of the tools in use, how they can be improved how advanced knowledge is used in the agency setting.

A greater understanding of how information about water is used in agency decision making will help researchers direct their efforts more efficiently. This current work has been somewhat *researcher push*, this needs to be balanced in future work with *industry pull.*

The data techniques, in particular the uncertainty analysis of DARE, should be applied as soon as possible specifically towards a real agency (NRAR and WaterNSW) problem, such as an area with particularly high uncertainty (20% to 50% unaccounted differences), over a stretch of river (Namoi, Lachlan, Macquarie and Borders Rivers have all been highlighted as examples). DARE has commenced this work, and given a detailed proposal, broken down into the various parameters of uncertainty, to complete it.

Benefits can be realised in further campaign style (e.g. 1 to 3 year) deployment of low-cost sensor sets to address specific gaps, capture events and tune models. Coastal areas currently lacking data may be of interest. Low-cost sensors can be used as calibration points as various remote sensing (satellite, aircraft, drone and sentinel devices) methods are developed. Further developments in the uncertainty models which will help determine the most efficient deployment of sensors.

Tasks such as water asset management, irrigation, surface water releases and storage decisions may be informed by the findings on recharge mechanisms, their timing and amounts. Also models, predictive and otherwise, about water behaviour should consider these findings, particularly any models that make assumptions about how aquifer recharge is related to rainfall and other hydrologic processes such as evapotranspiration and soil moisture storage.

Gravimetry field trails should be sponsored. Ideally this would require the purchase of a classical gravimeter (Scintrex CG-6). Several experiments are proposed: measuring the gravity signal during the daily displacement of water in the Snowy Mountains dams, deploying the gravimeter to survey a square kilometre array in the Maules Creek study area for a year, observing the gravity signal a series of saturated and unsaturated, layered gravel and sand pits. Site deployment of a quantum gravimeter would also bring exciting results. All these experiments work towards achieving confident day-to-day measurement of underground water using gravimetry.

Further interrogation of satellite gravity data could aid modellers as the large-scale remote sensing of the total water storage is not provided by any other means. An assessment can be made if high flow event can be tracked and time intervals can be shortened to five days. Environmental sensing is one of the most often touted non-military justifications for the current expansion of Australia's space capability. Stakeholders in space and water sectors should continue to liaise and long-term aspirations should be held for satellite assisted water monitoring.

Water as an Asset: Considering key water bodies as assets, each with many attributes (size, shape and other various qualities changing over time), can simplify the way in which input from multiple sensors can update asset attributes tables and models, and different analysis methods can control and predict the water assets into the future. This draws on established asset management and prediction tools.

A thorough translation of the research knowledge from universities to the agencies, could be undertaken over three and half years, in the vicinity of less than a million dollars per year. This figure is scalable; there are many smaller pieces of research and consultancy work that the universities can provide for considerably smaller sums and shorter timeframes.

Table of Contents

How this report is structured

Part 1 of this report describes the Outcomes. This part first gives an overview of what the combined effort of the collaborative project, then describes what has been achieved in each of the five sub-projects.

Part 2 is a Technology Roadmap to help explain to agencies and other stakeholders on how better sensing might be used across the state.

Part 3 describes the Research Project in terms of traditional scientific study: Background, objectives, approach, results, and so on. Project requirements such as Milestones and Case Studies are described here, as is the governance, budget and project management.

Part 4 provides each of the Sub-project Reports, largely unchanged from how they were submitted by the sub-project, authors with small edits to ensure alignment of reporting style. All these reports can be provided as a standalone report.

References, glossary and appendices follow.

1 Part One: Outputs

1.1 Combining new sensing and data tools to improve water knowledge

The *Where is All the Water?* project has demonstrated that the latest university research into different methods of advanced sensing and data analytics can be brought together under a single umbrella to bring awareness of new tools that have real potential in helping NSW government agencies and other stakeholders overcome challenges in water management and decision making, and develop a more complete knowledge of water in NSW (see Figure 1.1).

The project successfully deployed (despite the ravages of COVID-19) low-cost sensors in the Namoi catchment which provide telemetered data on soil moisture, rainfall and temperature, and other physical parameters can readily be added. This demonstrates that deploying new sensors to provide additional data can be done as part of a relatively small and simple project.

The principle of gravity sensing of water has been shown to be possible using a traditional gravity meter, and that this will be improved in orders of magnitude as quantum sensors are developed. This frees up sensing of water from digging, boreholes, probes and other laborious methods.

Gravity sensing from NASA satellites (GRACE) has shown water movement at a continental to state level. This way of remote sensing shows water flows at scale and coverage not possible by other means.

Recharge mechanisms of NSW aquifers have been found to be predominantly diffuse recharge (e.g. via rivers), rather than focussed recharge (e.g. rain percolating through the soil), a result which may explain downstream water losses and can inform assumptions in water models.

All these sensing methods can inform and be informed by data methods. A probabilistic modelling framework was developed to explain and quantify unaccounted differences in the NSW General Purpose Water Accounting Reports (GPWARs) for major rivers up to catchments.

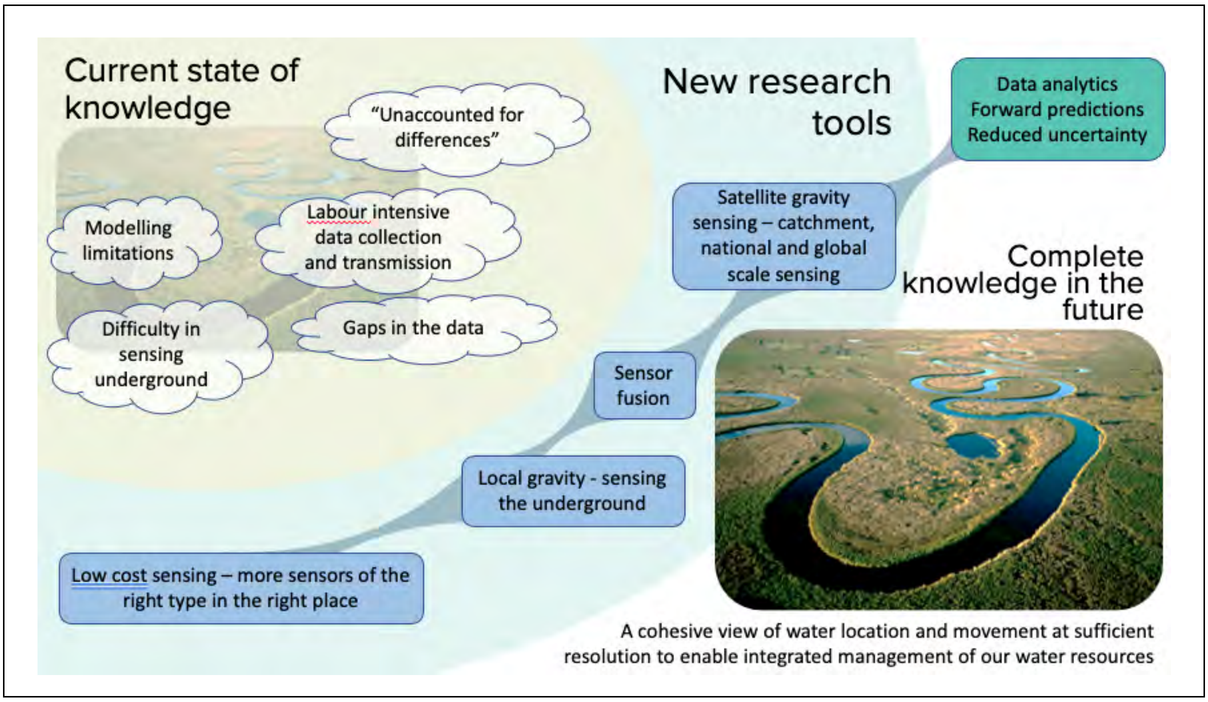


Figure 1.1: New research tools are intended to drive to a more complete state of knowledge.

1.2 New sensing tools and data methodologies

This section describes five new ways of understanding water in our catchments.

1.2.1 Low-cost sensing

Macquarie University has installed low-cost sensors in the Namoi catchment. They generate data at a fraction of the cost of weather stations and other infrastructure heavy sensing systems, and utilise regular transmission technology (LoraWan) outside areas of domestic internet and mobile phone coverage. The efficiency with which they were installed indicates that increased sensing and telemetry of water data can be readily facilitated.

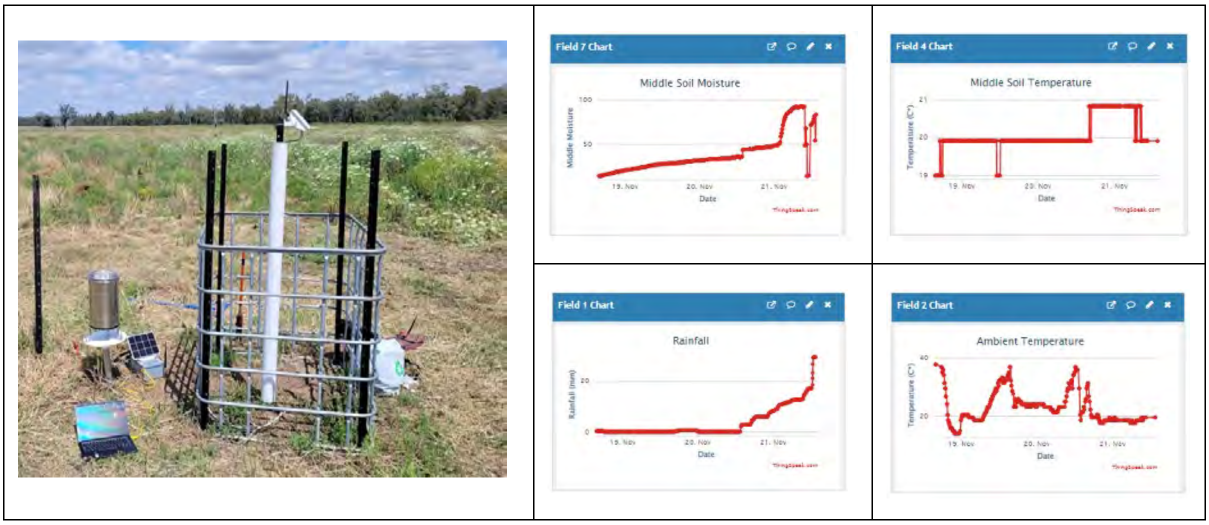


Figure 1.2: Left: a sensor node installed in Narrabri. Right: data generated for soil moisture, rainfall, soil temperature and ambient temperature.

Gaps in the current NSW water monitoring network necessitates the need for the deployment of more physical, telemetry-based sensors. The deployment of additional low-cost sensors in strategic locations throughout NSW will aid in closing the gaps in water data and further reduce data uncertainties.

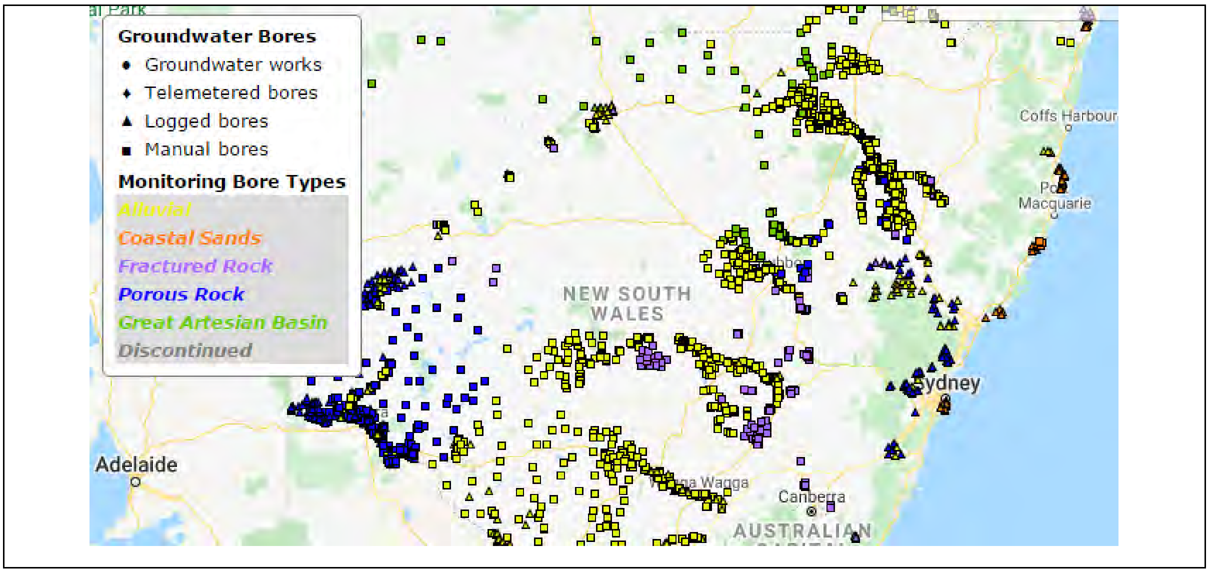


Figure 1.3: WaterNSW data clearly showing the lack of monitoring in many regions, in particular telemetered groundwater monitoring. https://realtimedata.waternsw.com.au/).

Macquarie University addressed the problem statements through developing prototype sensor nodes to monitor soil moisture content at differing depths, rainfall and ambient environmental conditions. Deploying a large number of these low-cost nodes would allow for a verification and greater understanding of rainfall variability and identification of events that lead up to and cause diffuse recharge to occur. These low-cost sensors will provide the means to fill in the gaps within current water data available. Current and emerging technologies have limitations in their scale and resolution, and this creates the gaps that cost-effective traditional sensors can satisfy.

This research showcases the design, development and deployment of a low-cost telemetry-based sensing system for monitoring various environmental parameters including soil moisture and temperature at an array of depths, rainfall and ambient conditions from any remote location throughout NSW. The Macquarie University workstream aims to aid in decreasing data uncertainties through developing and deploying the proposed low-cost systems to provide an increase in the spatial and temporal resolution of data that is currently available. The collected sensor data is transferred and stored in the cloud periodically through the Long-Range Wide Area Network (LoRaWAN) communication protocol. Adaptation and data collection from the deployment of many proposed low-cost sensor nodes in targeted locations throughout NSW will provide a major breakthrough in addressing current gaps in the knowledge of water location and movement throughout the state.

About Macquarie University Engineering

Macquarie's University's engineering researchers are dedicated to creating technological solutions to problems relevant to society's health and environment – solutions that expand the capability of people to achieve their goals. They have research strengths in electromagnetic and antenna design, energy conversion and management, integrated wireless communication systems, nonlinear electronics, guided-wave optics and photonics, very-large-scale integration, and wireless communications and networking.

1.2.2 Local gravity sensing

The ANU simulated a real-world test case to investigate the gravitational signals from near surface groundwater. This feasibility study showed that the gravitational signals produced by recharge to the groundwater systems will be measurable by the next generation of portable quantum sensors. Measurement campaigns utilising these sensors will be able to inform the location of water and in conjunction with exiting sensors and boreholes, improve knowledge on both soil moisture, evapotranspiration rates and groundwater location.

ANU also provided larger test worlds based on slope data for satellite gravity (GRACE) modelling. This helped to explore the feasibility of space based gravimetry surveys as a water monitoring technique.

They focused specifically on the ground water aspect using gravity signals. The amount of water that is leached from a stream or the surface into the ground was built as a toy world and the expected gravitational signals were simulated. Measurements of gravity over time will help inform water release agencies the expected loss and gain of water resources to and from the ground and through evapotranspiration.

Uniquely, since gravity is a mass-based signal, with sufficient knowledge of the underground system and the appropriate sensor, gravity has the potential to be a catch all measurement for water, able to cumulatively see the water table, soil moisture and surface water.

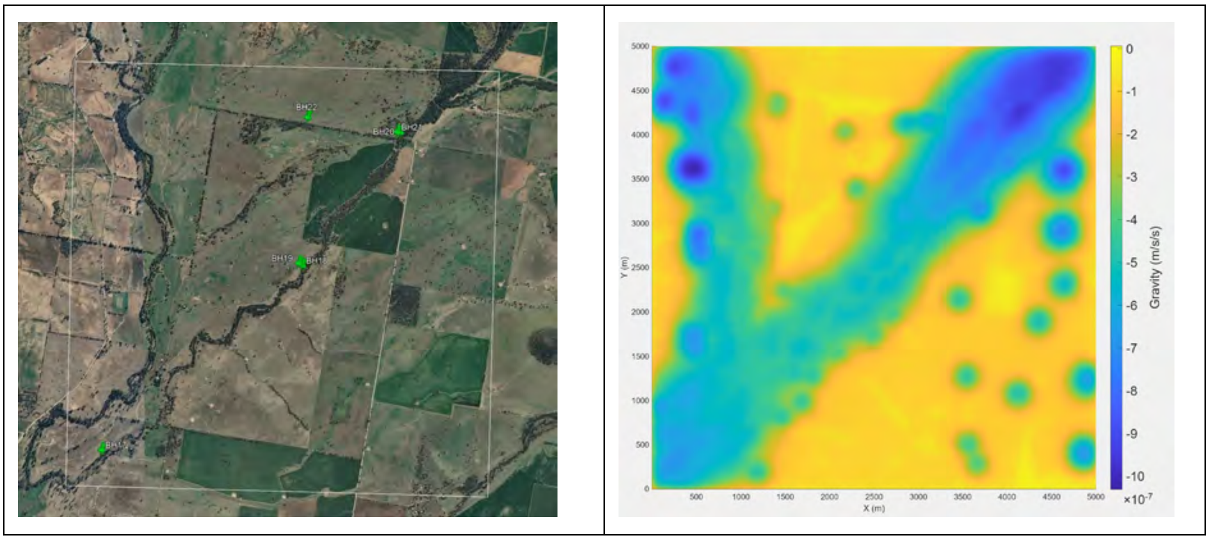


Figure 1.4: Left: arial map of the Y-shaped study area of Middle Creek between Gunnedah and Narrabri, NSW. Right: Plot of simulated differential gravity signal (g_z) with a large amount of surface, dam and groundwater in the system. Simulated survey done at 50m above ground level (drone).

About ANU Department of Quantum Science and Technology

The Department of Quantum Science and Technology pursues basic scientific experimental and theoretical research in the quantum mechanics of many body systems and translates basic research to devices, applications and commercialisation. At present, the department is focussed on the quantum mechanics of photons, cooled atoms, NV centres, quasi-particles and solid state rare earth systems to investigate the fundamental physics of measurement, squeezing, entanglement, and quantum control among other many-body phenomena and the translation of those ideas and developments to applications in communication, computing and sensing.

1.2.3 Satellite gravity measurements of Australian water

Satellite gravity measurements clearly show the movement of total water, that is ground water, surface water and soil moisture, through New South Wales

A pair of NASA satellites orbiting Earth are sensitive to changes in the strength of the gravity field, indicating changes in mass of water, ice, earthquake deformation and mantle convection. In Australia, changes in mass are dominated by hydrological processes. NASA have used these satellites to measure the amount of water diminished because of the Californian drought, give early clues to flood danger in the Mississippi delta, and helped find water deposits in the central Asian deserts (https://www.nasa.gov/mission_pages/Grace/).

In the *Where is All the Water?* project, assessments were made of the accuracy with which a realistic signal of changes in water volumes could be recovered from space gravity data when estimated at a ~200x200 km scale. Maps were produced of changes in total water storage during heavy rain events in 2019 as displayed in Figure 1.5. (See Figure 4.26 for larger plots and other years).

Interactions with DPIE led to the investigations of whether flood waters in southern Queensland can be tracked into NSW rivers.

Space gravity data provides over-arching constraints on how much water resources have changed in the landscape, integrated over groundwater, soil moisture and surface water stores. These constraints can be coupled with in situ measurements and/or water models to improve the accuracy of the knowledge of how much water is present in any location and, therefore, what storage/releases are appropriate. This can help inform agencies to manage to capture, store and release water. (i.e the WaterNSW problem statement given in section 3.2).

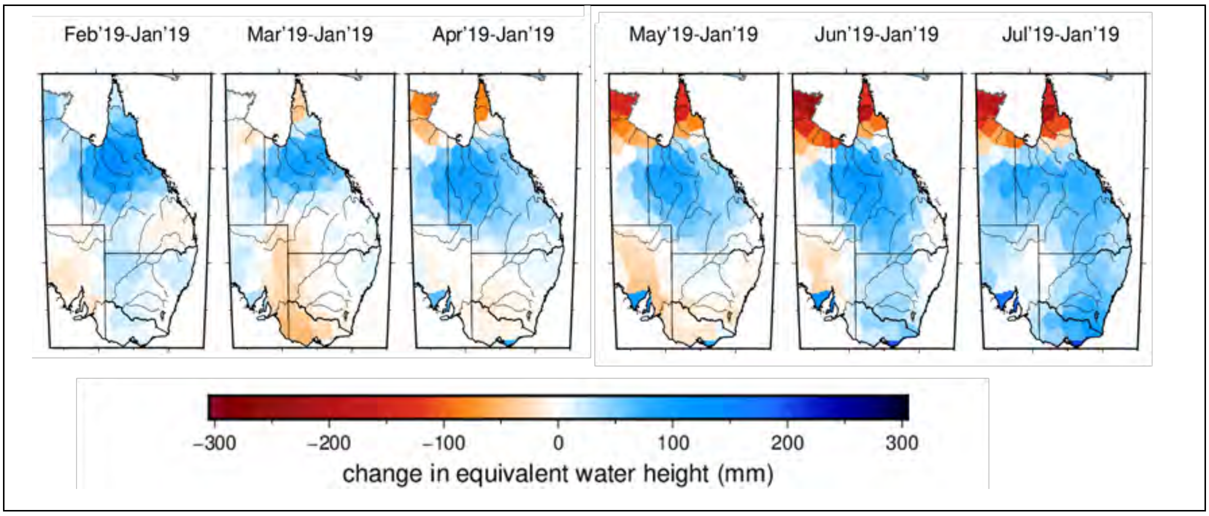


Figure 1.5: Total water storage across eastern Australia February to July 2019 with January 2019 as the baseline. Hydraulic activity such as the dry season in the Top End (darker red indicating a reduction in water volume), or the flow of water from monsoonal storms in Queensland (darker blue), down the Channel Country to NSW, can be observed.

Space gravity observations may provide a means to determine whether the unaccounted water has been either removed from the river system(s), has evaporated, or has replenished groundwater stores as it flowed downstream. Evaporation would cause a loss of mass, removal would relocate the mass beside the river while replenishment would leave the mass behind as the flow moved downstream. The difficulty of using space gravity observations for this purpose is one of spatial resolution: mass estimates are simply not of sufficiently high spatial resolution to resolve these signals. The possibility of using actual signals in the measurements (rather than estimates over 200x200 km) is investigated in this report. This can help inform agencies to explain unaccounted for differences. (i.e. the NRAR problem statement given in section 3.2).

Changes in groundwater can be estimated through either a simple subtraction [total water storage minus (soil moisture + surface water)] or derived more accurately by assimilating into hydrology models the total water storage estimates (and shallow soil moisture) from remote sensing data. This improves knowledge of ground water and aid groundwater decision making. (i.e the DPIE problem statement given in section 3.2).

The significance of the work undertaken in this project is that satellite gravity data can provide information at a broad scale, which is otherwise typically not available.

About ANU Institute for Water Futures

The Institute for Water Futures (IWF) leads research to quantify and understand change and enable action in Australia and beyond. Though our work we grow capabilities across the water sector to inform decisions that anticipate our increasingly complex uncertain water futures.

1.2.4 Data uncertainty analysis

The data analytics sub-project produced a probabilistic modelling framework for the water balance, useful up to catchment scales, for the purpose of making water management decisions at the state level in NSW under uncertain conditions.

A pilot Bayesian framework was developed that quantifies uncertainties in the components that contribute to the water balance in stretches of the river and the interaction between groundwater and surface water. This framework was tested to identify processes that have the highest uncertainties in both space and time, which can assist in decisions on placement of new sensors and identifies gaps in current data and knowledge.

Different inversion options for gravity data were tested and it was identified that this problem is not well constrained and needs more research and data analysis to draft models of the underground.

The framework developed addresses in particular the need for agencies to have better tools to explain unaccounted for differences. (i.e. the NRAR problem statement given in section 3.2), by identifying parts of the water balance with the highest uncertainty. Better understanding of uncertainty feeds into better decision making in all parts of the river and groundwater system and can guide placement of novel sensors.

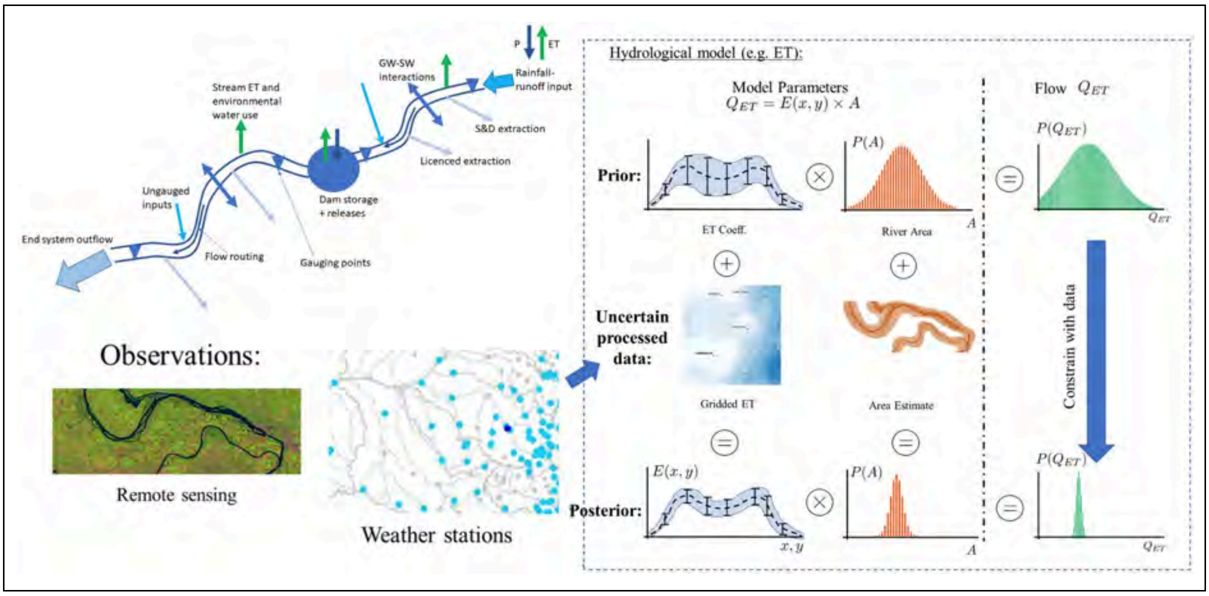


Figure 1.6: Top left: Stretch of river with components of uncertainty. Bottom left: Observed data may come from a range of sources, examples include remote sensing and weather stations. Right: The model described how a component of the water balance, in this case evapotranspiration (ET), can be analysed considering several parameters. The techniques lead to a reduction in the uncertainty as represented by the narrowing of the green distributions to the right of the figure.

About DARE

The University of Sydney's ARC Training Centre in Data Analytics for Resources and Environments (DARE) will provide PhD candidates in world leading data science through an innovative, collaborative program between industry, government and academia. All students will undertake a cohort-based training program in data science prior to selecting a specific data science project and domain. Data science research projects will be applied against real world challenges through an industry placement program with one of our Partner Organisations.

1.2.5 Hydrological recharge mechanisms

The project was supported by vast experience in measurement of NSW water resources. There was considerable background knowledge of hydrogeological conceptual models and data. The project had access to data from the NCRIS Groundwater (GW) Infrastructure field sites as well as information from NSW-DPIE and NRAR and broad general groundwater knowledge. This provided the 'ground-truthing', guidance and benchmarks for the new sensing methods.

Specifically, this sub-project showed that focussed recharge is the predominate groundwater recharge mechanism in areas adjacent to intermittent streams rather that diffuse recharge,

The existing NCRIS Groundwater data sets provided timeseries of over 10 years of groundwater levels from various bores at two locations Wellington and Maules Creek. From this UNSW was able to determine that significant diffuse recharge events have happened about twice in the last decade, whereas focussed recharge from streams happens once or twice a year. While bores can give us insights into groundwater processes in the saturated zone, a knowledge gap exists around water movement through the unsaturated zone. Collaboration between UNSW and ANU has framed early thinking about use of gravity sensors and soil moisture sensors to trace water movement in the subsurface.

The processes below the intermittent Middle Creek illustrate the mechanism of surface water transmission losses (e.g., focused groundwater recharge) for flow releases in dry streams. Conceptually the location of the groundwater table at the time of the release partly controls the amount of loss. E.g., the lower the groundwater table (the greater the unsaturated zone), the greater the surface water losses. This can help inform agencies to make decisions around the capture, store and release water. (i.e the WaterNSW problem statement given in section 3.2).

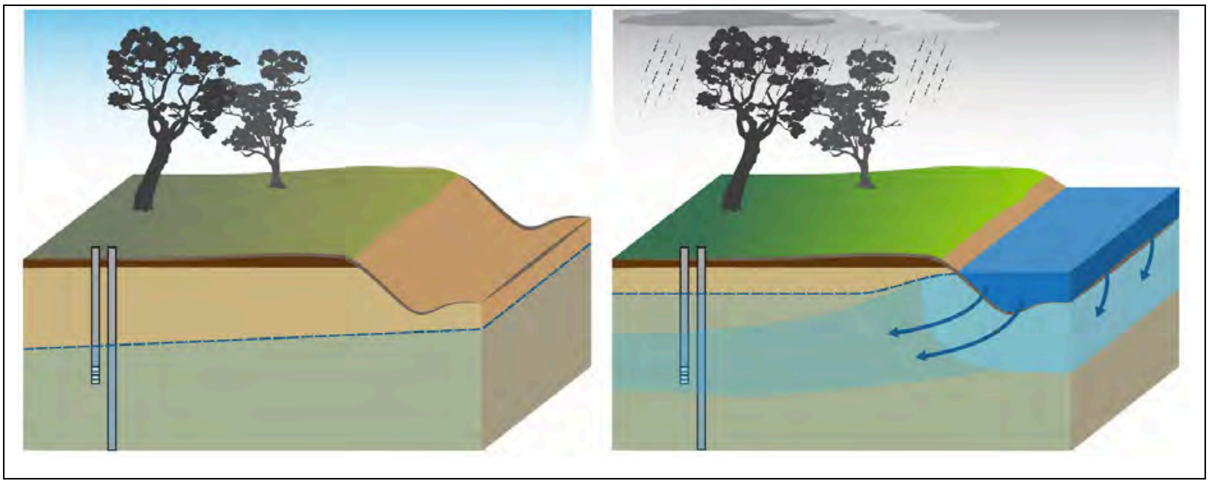


Figure 1.7: Representation of focussed recharge. Left: the aquifer in a region is represented below the dotted blue line. Right: Recharge of the aquifer has been shown to be more frequently attributable to focussed recharge from adjacent streams, rather than from diffuse recharge of rainfall over the terrain above the aquifer.

This project has quantified aspects of groundwater recharge and determined the relative importance of the different mechanisms of recharge at the NCRIS Wellington and Maules Creek catchments. These insights, which in principle are applicable at other similar semi-arid/arid sites, are important for sustainable groundwater management. This provides knowledge of ground water resources, what groundwater is influenced by and influences on; and thus aid more robust, timely, scale appropriate decision making. (i.e. the DPIE problem statement given in section 3.2).

About the UNSW Water Research Centre

The Water Research Centre (WRC) is an international leading university centre that provides multidisciplinary research in water resources, engineering, management and the development of tools for environmental management and sustainability for improving aquatic and atmospheric environments. With its two research locations; WRC at the Kensington campus and the Water Research Laboratory (WRL) located at Manly Vale, operates as an externally funded University of New South Wales (UNSW) research centre within the School of Civil and Environmental Engineering.

2 Part Two: A Technology Roadmap for better understanding of water assets

The project has come about from a number of starting points; ministerial decree, the establishment of an Australian Research Council training centre (DARE), the release of the report from the Office of the Chief Scientist and Engineer, *Review of water related data collections, data infrastructure and capabilities* (2020), and the emergence of the NSW networks, the NSSN being one, as valuable respondents to government agency challenges.

This project was proposed as an exploratory project. In the parlance of Technology Readiness Levels (TRLs) (see section 0) we were observing basic principles and formulating concepts (TRL 1 and 2). This project has taken some aspects along to proof of concept (TRL3) right through to demonstration in relevant environments (TRL6, in the case of low-cost sensing).

The next step in the roadmap is to work closely with agencies to develop a prototype asset management tool that best suits their needs. This allows us to guide further development efforts towards validation and demonstration in an operational environment (TRL5, 6 and 7).

2.1 Before the project: gaps, limitations, and unknowns

The management of water in Australia is complex with myriad stakeholders across jurisdictions. In NSW. In NSW three of the agencies with responsibility for water management are the WaterNSW, Natural Resources Access Regulator and the Department of Environment and Planning (formerly DPIE). While three agencies are proud to be world leading in their processes and outcomes, the harsh Australian environment, climate change and population growth mean there is always a need to innovate, overcome challenges and optimise use resources.

In a series of workshops early in the collaboration in 2021, the agencies gave presentations to demonstrate where their needs were. These were summarised thus:

- WaterNSW: How does this research help them to capture, store and release water: The right amount of water, in the right place, at the right time.
- NRAR: How can the research address the issue of the “25%” unaccounted for water in some of the water balances.
- DPIE: How can this provide data to improve knowledge of ground water resources, what GW is influenced by and influences on; and thus aid more robust, timely, scale appropriate decision making.

The potential sensing modalities are potentially infinite and listing every type of water sensor in use in NSW is likely to be futile. However an overall impression can be gained of what new technologies currently under research may contribute. The NSW Office of Chief Scientist and Engineer undertook a review of water data in NSW, see Figure 2.1. From this, we can understand the types of situations where water data is being retrieved from.

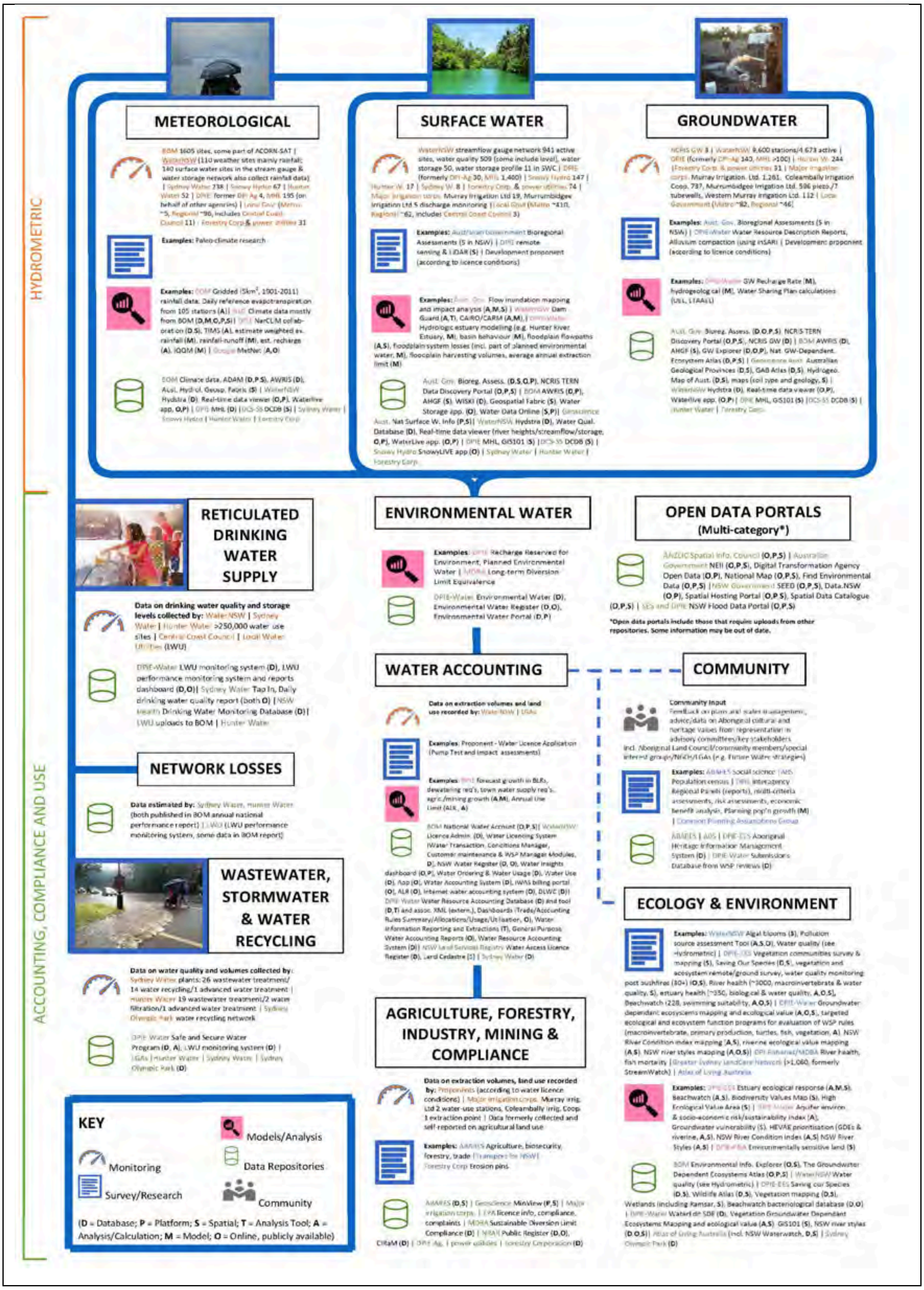


Figure 2.1: Water data in NSW. Review of water-related data collections, data infrastructure and capabilities, July 2020, OCSE [1].

2.1.1 Water as an asset

Water entities can be considered as assets, with a range of attributes, including, quite simply, volume. Attributes of the asset can be updated as new data comes in. The asset database can be built up and developed into a dynamic database with temporal information. Different and new sensing modalities can easily be included as they bring updated information about the asset attributes.

Sydney Water successfully used such an approach for reliable identification and prediction of at-risk pipes [2] [3]. A similar approach could be used to identify at-risk (e.g. depleted or overflown) water assets. This could be further augmented with uncertainty analysis to provide the probability of risk.

An asset management approach also permits consideration to sensor fusion, whereby the understanding of assets comes from disparate sensors, using the data analytics such as Bayesian techniques that can help reduce uncertainty. Such approaches are being developed in defence and medical technology applications.

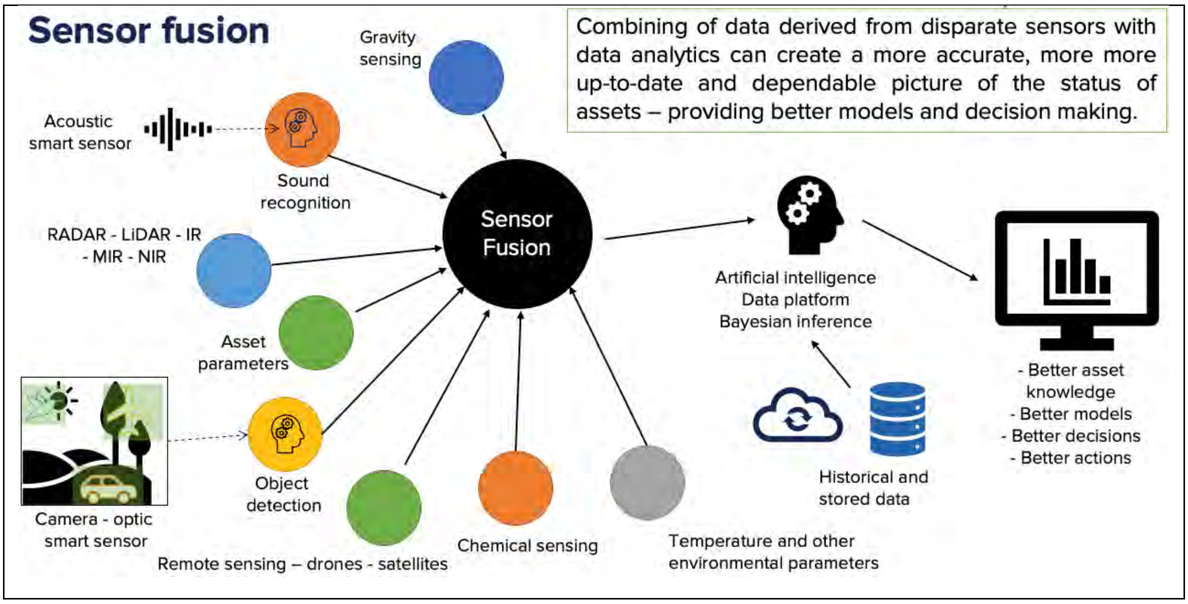


Figure 2.2: Sensor fusion and data analytics to improve understanding of assets.

2.2 Advancements of WIATW sub-projects

The progress of the sub-projects can be visualised using the Bayesian framework developed by DARE. This initially uses existing Data to develop the Model. It is strongly linked to the process conceptualisation in the model (including discussions with UNSW) for both the river system framework and the gravity inversion work. For the gravity inversion, DARE has completed testing using synthetic data, which has led to identifying gaps in understanding of the conceptualisation of underground structures. A proposal to refine the conceptualisation of underground structures is an ongoing discussion among research partners including ANU and UNSW.

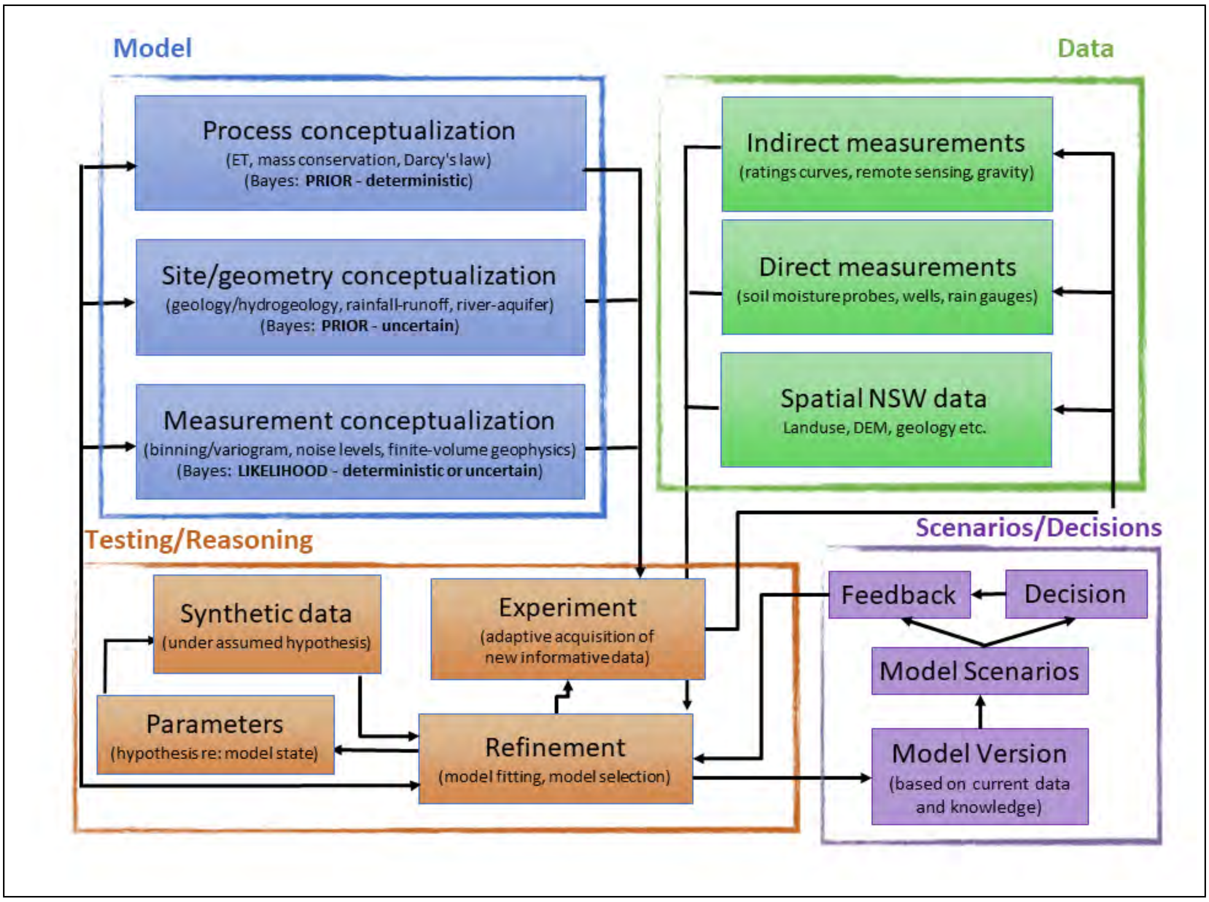


Figure 2.3: Where is All the Water? project Bayesian framework. This considered the project in terms of key aspects of Model, Data, Testing/ Reasoning and Scenarios and Decisions.

Low-cost sensing

The Macquarie University sub-project plays a strong role within the project model. As the modelling, simulations and resulting decision making are all dependent upon the initial data utilised, it is important to have the right amount and type of data. Macquarie aims to obtain and contrast relevant data that is of a high temporal and spatial resolution which can then be utilised for modelling and decision making.

The Macquarie sub-project fits into the direct measurement role, providing ground-truthing real-time data that can be compared with or utilizes as a calibration point for other types of remote sensing. Based on the output and decision making from the modelling, changes to the position and/or temporal resolution of data being collected from the developed systems can be adjusted as required.

Local gravity sensing

Have simulated indirect measurements of gravity based on a real system provides a new source of data.

Satellite gravity sensing

Indirect measurement: the space gravity data provides indirect estimates of changes in total water storage, averaged over 200x200km regions. This quantity cannot be measured by any other technique at these large scales; therefore, this is a unique measurement. It provides information at a broad scale, permitting the regional/basin-scale patterns of water movement to be interpreted.

Experiment/simulation/defining parameters: The space gravity estimates of total water storage can feed into the testing/reasoning components of the Roadmap. New parameters – or, at least, new constraints on existing parameters – need to be added into model state calculations to ensure that the integrated changes in model estimates of water quantities match the broad-scale average changes obtained from the space gravity data.

Data uncertainty

The pilot applications of the Bayesian framework from DARE also fit into the Testing/ Reasoning. The results of the pilot will inform updates to the model, and as indicated, the results highlight gaps in the data collection. Once the models have been refined, these can be used for decision making (conditional on the data and the specific model realisation). However, the Bayesian approach is strongly linked to the idea that a model can only be as good as the data and the current understanding. Therefore, feedback from the stakeholders using the model will generate further improvements in models, data and experimentation. Agile management approaches will allow this continual optimisation.

Hydrological recharge mechanisms

A key contribution of UNSW to this project is groundwater data of high spatial and temporal resolution from two NCRIS sites at Wellington and Maules Creek that provide direct measurements of groundwater levels over a 10-year period and hence provided ground-truthing for the wider project. This data has formed the basis of process and site conceptualisations that have informed our conclusions about groundwater recharge. This data has also informed the conceptual model and physical parameters for the model testing for local gravity sensing (Sub-project 2).

2.3 Operational use

The following paragraphs discuss how the techniques can be used by the relevant agencies.

Low-cost sensing

Utilizing the LoRaWAN Network – which stands for long range Wide area network, this communication protocol allows for these low-power sensor nodes to communicate small amounts of data at regular periods at very low power consumption. Using LoRaWAN therefore means we can deploy these systems anywhere within NSW and maintain their power via the use of small solar panels. Further information on their specific use cases is highlighted in the report.

Local gravity measurements

The sensors needed to deploy this technology terrestrially and on small scales are still in development.

Once developed gravitational surveys could be used to monitor interesting or problem areas. For example, if a particular catchment consistently has less water in it than expected, a number of gravitational surveys could be done and the mass balances calculated and compared to existing models for that catchment. The extra data source can be used to identify where water is been lost to ground water recharge or to evapotranspiration.

Satellite gravity measurements

For large scale use cases the data from satellite gravity missions can provide high quality estimates of changes in total water storage at large scale. At spatial resolution of ~350-400 km these estimates will be uncorrelated, whereas as smaller spatial resolution of ~200 km some smearing/leakage of signal occurs spatially.

In low flow use cases the data from satellite gravity analysis has a sensitivity of ~20 mm water thickness across the spatial region of the estimate. Provided the changes in total water storage exceed this magnitude, the changes would be detectable.

For high flow use cases there is no upper bound on the magnitude of signals that can be detected and quantified from satellite gravity missions.

Data analytics

The key difference between existing technologies and the approach developed by DARE is that the Bayesian framework through Bayes formula explicitly includes the quantification of uncertainty of all sources of uncertainty in the model. The framework applies to all the possible use cases, although the different elements of the framework and the choices of priors will differ by use case. In all cases, there is a need for investment (as outlined in the DARE milestone 1 report) to create a workflow from data to a model that can be used for operational applications.

Recharge Mechanisms

In the high flow use-case, especially after a dry period without surface flow, in intermittent streams the groundwater table will be lower than the stream bed (i.e., disconnected). When stream flow is generated upstream (i.e., a dam-release or rainfall), focussed recharge will occur with the surface water infiltrating down to the water table and laterally away from the stream. This work has shown that this is the predominate recharge mechanism for intermittent creeks in semi-arid/arid regions. This potentially means that direct rainfall at such sites may not be a determining factor for groundwater recharge, especially if stream flow is generated by rainfall further upstream, as is the case with the Middle Creek site.

It is important to realise that this does not happen to the same extent for perennial streams where the surface water and groundwater levels are in a dynamic equilibrium. For perennial streams a flow release may lead to bank storage effects with the stored water returning to the stream after the flow-release (unless large groundwater abstraction is happening at the same time). This same behaviour can be observed in intermittent streams during periods of sustained surface water flow, as the groundwater table is then connected to the stream.

2.3.1 Technology review

Agencies have requested a comparison of the techniques for better water management in comparison with those currently in use. Discussion of the various sensing techniques has been provided in the previous Milestone reports [4] and is discussed as part of the final report from Macquarie (see section 4.1.3).

The selection process of the technologies to investigate had several strands as described in the introductory remarks from the NSW Chief Scientist and Engineer (see Foreword, pg. 2). There was a Ministerial directive leading to the major NSW Water Industry Innovation Workshop in October 2019 attended by over 100 delegates. From this a longlist of areas of investigation was presented to the Water CEO, Jim Bentley. A shortlist of four priorities (projects about Avoiding Fish Kills, Flood Plain Harvesting, and Data Management were the other three) was considered before a consensus was reached that the *Where is All the Water?* project would be the most informative.

The NSW Smart Sensing Network (NSSN) has been working at the vanguard of translating sensor technology from research to the assistance of industry and government since 2016. The NSSN follows the latest research and technology worldwide in physical sensing, such as chemical and optical techniques. The inclusion of the low-cost sensing sub-project in this project allows the versatility for any novel developments in physical sensing to be adapted where feasible. NSW is also embarking on a Space agenda and there is a state-wide drive into Quantum science. The inclusion of the satellite gravity and low-cost sensing sub-projects address these domains respectively. No conversation about sensing is complete without a consideration of how the data is processed. While many analytical tools, the reduction of data uncertainly by Bayesian inference is one of the most exciting.

In terms of readiness the data techniques are ready to be applied to problem right now. Likewise, the timing of deployment of low-cost sensors will be dependent on financing, planning and logistics, with the actual sensors themselves being ready now. Further analysis of satellite derived water data and calculations of the utility of quantum gravimetry can commence as soon as possible.

Deployment of special Australian satellites for environmental sensing is due within 10 years, while commercially available quantum devices suitable to this problem are seen in five-year horizon.

The low-cost sensing proposed by Macquarie university is not beholden to any particular design, manufacturer or style of sensor. The nodes can adapt the most cost-effective transducer, and use telemetry provided by the most suitable transmission protocols to the local context.

Satellite derived data will of course be related the cost of satellites. Sending a dedicated sensor into space is an exercise in the order of tens of millions of dollars. Once installed vast amounts of data, with global coverage, including detailed knowledge of NSW and the MDB is forthcoming. Given Australia's push into space, agencies should not shy away from the idea of commissioning special sensors.

Cubesats[2] have given indication payloads might be in the order of $100,000s to deploy, suggesting dedicated satellite sensing of the NSW water environment may be within agency budgets. The SmartSat CRC[3], CSIRO Aquawatch [5] and the Waratah Seed Mission[4] should also be followed.

Classical gravity sensors, gravimeters, are now routinely available from a number of suppliers. They are used for many scientific and industrial uses, particularly mining. The preferred gravimeter is in the order of $150,000 (Scintrex-CG6).

Data analytics does not require the installation of hardware, though data engineers are a skill set trade attracting premium salary.

Currently the most common method for monitoring ground water is through the drilling of bore holes. This method is expensive; the global rule of thumb is that a borehole with pump (no sensing) costs approximately $US10,000. The Australian experience is that they are typically between $A5,000 to $A20,000. The cost of putting a sensor in an existing borehole is relatively cheap, though telemetered sensors are rare. Labour and travel costs of retrieving data from remote sites adds up. It is also highly discreet; to do an area survey a grid of boreholes would need to be dug and monitored, additionally this method is not typically sensitive to water retained in the soil and it invasively disturbs the soil structure, altering porosity and other parameters around the bore hole itself. Gravity surveys comparatively are non-invasive and can be performed anywhere, been able to remotely sense changes in water mass, regardless of conditions such as salinity, ground hardness, etc. This makes these sensors relatively unique in the field of hydrology and ideal for measuring water volumes.

2.3.1.1 Advantages of new sensors and data tools

Sensors are operating at very different area and length scales (as mentioned in the Milestone reports) and the only way to get the full picture is to find a way to combine all of these. Low-cost in-situ sensors are the most crowded space (lots of companies and the likes of CSIRO's SENSEI playing here), but these have nothing like the scalability and ability to cover wide areas offered by gravity sensors.

Figure 2.4 is a 2-axis comparison of sensing technology that considers scalability and resolution, while table Table 2.1 provides a summary of different sensor types.

---

[2] https://www.cubesat.org
[3] https://smartsatcrc.com/
[4] https://www.waratahseed.space/

This exercise can only fully be completed when agencies work closely with researchers to elicit the scope, parameters and criteria they currently use for sensing selection.

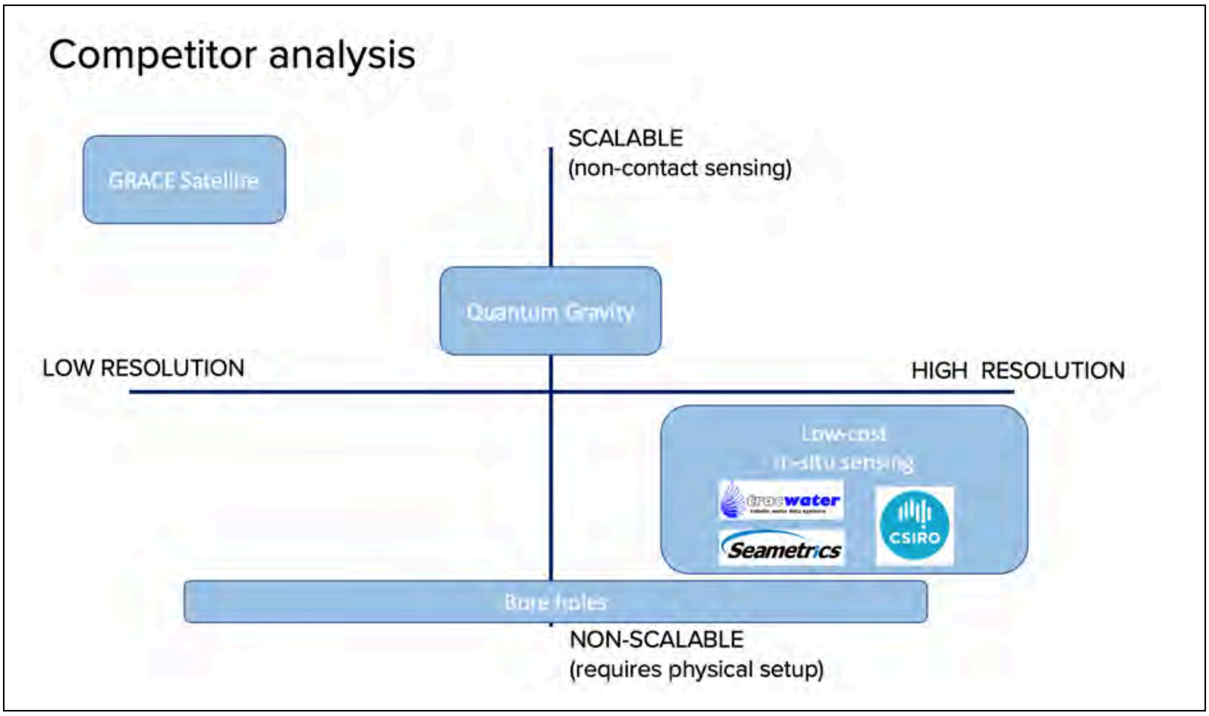


Figure 2.4: Competitor analysis of scalability and resolution for various sensing modes.

Table 2.1: Summary of sensors types. ✓ indicates a positive or enabling attribute, ✗ indicates a negative or constraining attribute, and ⊖ a neutral attribute.

Technology	Accuracy	Resolution	Scalability	Cost	Frequency	Maturity
Boreholes	✓	✗	✗	✗	✓	✓
In-situ sensing (traditional)	✓	✓	✗	⊖	✓	✓
Remote sensing (drones, robots)	⊖	✓	⊖	✗	✗	⊖
In-situ sensing (low-cost)	✓	✓	✗	✓	✓	⊖
Quantum/gravimetry	✓	⊖	✓	⊖	⊖	✗
Satellite gravimetry	⊖	✗	✓	✗	⊖	✗
Data analytics	⊖	⊖	✓	⊖	✓	⊖

A note on cost

For each of the sensing modes researchers will need to compare the utility of the data produced with the cost of acquiring that data. This requires more detailed information from the agencies of the types of sensors they are using, how they are being used and the various up-front and ongoing costs of accessing and analysing the data.

Agencies must work with researchers and provide the criteria including scope, parameters and other details of how they cost sensing presently for fair comparison to be made.

2.4 Future phases

2.4.1 Low-cost sensing

In the immediate term, the next step is to develop many more of the prototype nodes and deploy them at a site of interest. This could include the land offered by The University of Sydney Narrabri L'lara farm, or another site in collaboration with the deployment of gravity sensors with ANU. This would provide an outcome of achieving 20 – 30 deployable sensor nodes monitoring soil moisture and rainfall to collect data and reduce uncertainties.

In the short to medium term, further research and analysis regarding the inclusion of additional sensors would be the next goal. As we are currently monitoring rainfall and soil moisture quantity, the inclusion of quality sensors would be appropriate. For example, measuring rainfall quality (hardness, pH etc.) and/or soil quality which could then be analysed to potentially provide a relationship to the moisture content.

At Macquarie university we specialise in developing low-cost sensors, and the inclusion of these could provide a significant increase in the amount of information each of these nodes provide.

Once many sensor nodes have been deployed and are confidently operating and collecting data in an accurate and useful measure. Deploying these nodes at strategic locations within a large-scale testing area and using them as calibration points for large-scale remote sensing is the long-term outcome. These nodes could be easily deployed as calibration points for sensing parameters such as rainfall, temperature and soil moisture (e.g. using mobile cell-tower signals to determine the amount of rainfall in an area of interest).

2.4.2 Local gravity measurement and modelling

2.4.2.1 Gravity field tests

The next stage is to go into the field and test the model developed in this project. By preforming iterative surveys and long time based single point surveys of gravity the signals produced can be compared to that of the simulation to further inform the models of groundwater and evapotranspiration. Surveys will ideally be taken over a period of months to years, incorporating both low flow and high flow events to show significant changes in groundwater.

These surveys can be initially preformed with existing classical gravimeters but would also benefit from using some of the prototype quantum sensors in development, especially to recalibrate the classical sensors and reduce drift.

Further deployment of this technology depends on both the results of field trials and the ongoing development of these quantum sensors. Initial field trials will be ground based with larger sensors moved and placed manually. With the development of smaller, lighter quantum sensors the feasibility of a drone based aerial survey becomes an interesting prospect. Sensors deployed on drones will have the ability to dwell and measure gravity close to the ground and make measurement surveys across a grid automatically.

2.4.2.2 Gravity signal inversion experiment

Much could be garnered from an experiment to measure the gravity signal of some highly controlled source. In the first four experiments of Figure 2.5, there is absolute confidence that the two states will give a different signal. Experimental variables such as volume, mass, ratios could be varied to understand the gravity signal, eventually leading to more complex situations like experiment 5.

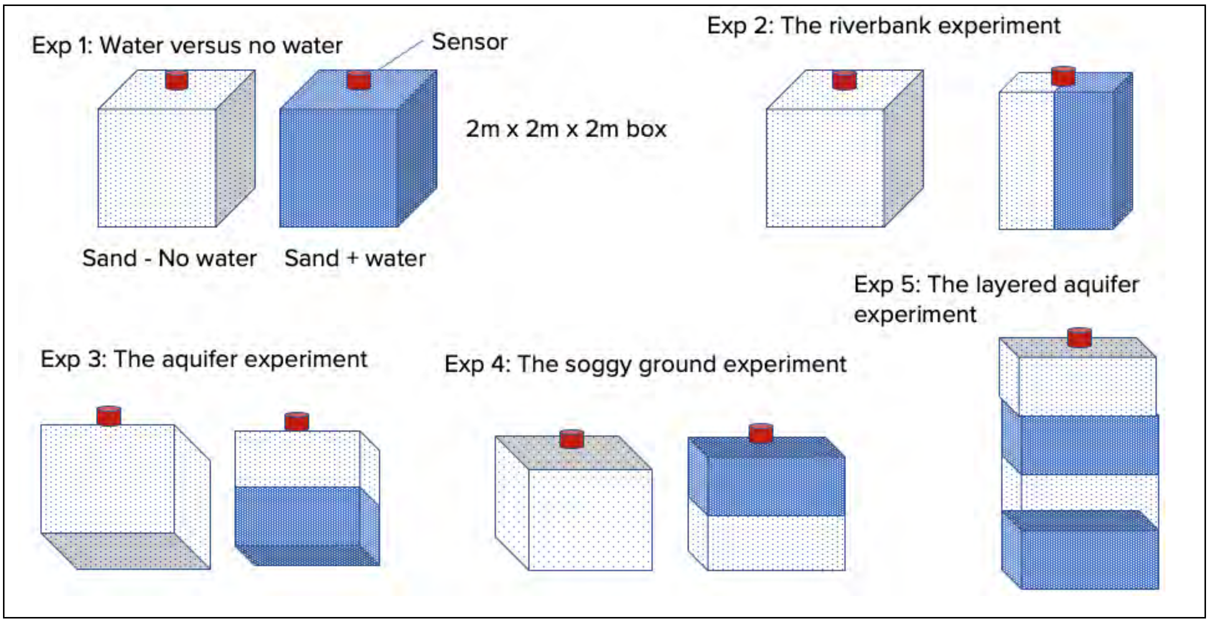


Figure 2.5: A series of experiments around the inversion of the gravity signal. Measurements from the gravity sensor (red disc) would be taken from an array of positions on top of the box. The water and sand + water conditions would be varied, with increasing complexity, including more realistic riverbanks.

2.4.2.3 Gravity signal of water bodies experiment

Another interesting experiment to try will take gravimeters to the sites of regular massive changes in water volume, such as the Snowy Mountains hydro-electricity reservoir. Learnings from this more controlled situation, could then be taken to a more complex environment, such as a dam release.

Figure 2.6: Measuring gravity changes in Snowy mountains dams, which have regular known changes of considerable mass, will help to understand the gravity change of water in unknown situations. (Photo: https://www.snowyhydro.com.au/).

2.4.3 Satellite gravity measurements

Stakeholders can benefit from this work immediately through the analysis and interpretation of the quantification of changes in water at the 200-400 km scale. The maps provided here from the analysis of GRACE Follow-On data show clearly the propagation of large rainfall events across Queensland and NSW in 2019 and 2020. Significant differences are evident between the two years.

Future work could increase the spatial resolution to enable interpretations at the catchment or sub-catchment scale through:

- The assimilation of changes in total water storage (and satellite-based measurements of soil moisture) into land surface models. This permits the space gravity data to be down-sampled – yielding higher spatial resolution – and also to separate the changes in water into different layers (groundwater, soil moisture, surface)
- The assessment of whether the inter-satellite observations along the ground track of the satellites can permit the tracking of high flow data through river systems,
- The creation of estimates at time intervals shorter than monthly. The ground tracks repeat every ~5 days; therefore, more frequent temporal estimates could be generated, although a more coarse spatial resolution would likely become necessary.

2.4.4 Data uncertainty

The general approach developed by DARE: a probabilistic modelling framework for uncertainty quantification using Bayesian analysis can be applied to several different "accounting systems" and observation networks in natural resources, including groundwater monitoring, soil carbon monitoring, water quality monitoring. We demonstrated this for the gravity signal inversion problem, which quickly highlights the main limitation in the conceptualisation of the underground. Both pilot approaches developed in this project will need expanding to become operational

Specifically:

1. As indicated, the key development need for the gravity signal inversion is a better model of the underground before the work outlined in the DARE milestone 1 report can be undertaken. The work by Fuentes et al. (2021) can serve as a starting point to develop a state-wide model of the underground where DARE would collaborate with ANU and UNSW. This model of the underground would have applications across state governments in natural resources, as well applications based on the inversions of gravity signals at both local and large scale.
2. To further develop the probabilistic modelling framework to an operational a further 3 phases would be required. Phase 1 would be to expand the pilot to a full river-groundwater system in one catchment as outlined in the milestone 1 report. Phase 2 would involve testing, developing training and refinement of the framework with stakeholders. Finally, phase 3 would involve operationalising the framework so it can be applied as a workflow by stakeholders.
3. The outcomes of phase 1 from 2) (above) can be used to guide testing of cheap sensor placement (MQ) and data collection campaigns (ANU), as this will for the river system indicate the greatest uncertainties in time and space.

More generally there is an opportunity in two additional areas:

1. Improving the communication of uncertainty and risk as part of the operational activities at DPIE, WaterNSW and NRAR. It is clear that there is a need to improve the “common language” around uncertainty between scientists at DARE and scientists and managers at the stakeholder organisations and development of communication to the public.
2. Generalised auditing of the current data collection networks for all stakeholders using an uncertainty based approach. This is initially purely a historical data analysis, but can be extended to include simulation or prediction of future data collection network improvements.

2.4.5 Groundwater recharge

While we have been able to draw conclusions about groundwater recharge timing and thresholds and the relative importance of the different mechanisms through this work, quantifying recharge amounts (mass) and determining the spatial variability of recharge remains a challenge. Measuring soil moisture in a reliable way at various spatial locations at one site will provide further insights into the spatial and temporal variability of groundwater recharge and ultimately inform sustainable management of groundwater and provide insights about how the generation of this resource may change with climate change. This can potentially be achieved using low-cost sensors (as with Sub-project 1) to create a network of monitoring points at a site and/or by using gravity sensors (Sub-project 2). The use of gravity sensors for measuring soil moisture does have the advantage of not disturbing the subsurface potentially leading to more reliable measurements of soil moisture. But most importantly, timelapse measurements of gravity directly measures the change in mass in the soil profile without the additional uncertainties of estimating porosity, specific yield and other soil parameter typically needed for traditional recharge estimates.

A ‘stretch goal’ would be to increase the capacity of and reduce costs of radio-tracer analysis (Carbon-14, tritium, etc.) so that they become routine measurements. An alternative would be to develop sensor methods to measure these continuously and in-situ (very sci-fi at this stage).

2.4.6 Summary of future work

Table 2.2: Summary of suggested future activities.

Activity name	Description	Lead partner/s	Resources and timing
0.1 Technology review	Comparison of current sensing technologies in use across the state by various agencies, with a range of off the shelf and possible future technologies still in the research stage	DPIE, MQ	$50,000 Short term
0.2 Water research capability review	Review of research capability across all NSW/ACT research organisations	NSSN	$50,000 Short term
1.1 Proliferation of low-cost sensors	Develop many more of the prototype nodes (e.g. 20 to 30) and deploy them at a site of interest. Combine with gravity sensors	DPIE, MQ, ANU	$150,000 Short term
1.2 Additional sensing parameters	Add additional parameters including hardness, pH, turbidity, oxygen etc.) and soil quality	WaterNSW, MQ	$150,000 Mid term
1.3 Wide area sensor deployment	Deploying low-cost sensor nodes at strategic locations over large-scale (>100s km) testing area and using them as calibration points for evolving remote sensing as being provided by drones, planes, satellites and sentinel devices.	MQ, ANU, DPIE, WaterNSW	$300,000 long term
2.1 Gravity meter purchase	Quantum sensing field experiments. Purchase/hire of gravity meter	ANU, UNSW	$370,000 Short term
2.1.1 Inversion experiment	Inversion experiment. Gravity sensor measuring various water soaked configurations of a 8m^3 volume	ANU, USYD, DPIE	$150,000 Short term
2.1.2 Snowy mountains experiment	Snowy mountains experiment. Install gravity meter at a hydroelectricity reservoir with mass daily changes in water.	ANU, DPIE (SMEC)	$150,000 Short term
2.2 Gravity groundwater survey	Iterative array surveys of a groundwater site, such as Maules Creek for several years	ANU, UNSW	$450,000 Medium term
2.3 Quantum sensors for water	Further development of quantum sensors. A contribution could direct research towards water and environmental use.	ANU	$500,000 long term
3.1 Enhancing models with satellite data	Assimilation of total water storage from satellite data into water models. Develop higher spatial resolution.	ANU	$150,000 Short term
3.2 High flow data	Assessment of whether the inter-satellite observations along the ground track of the satellites can permit the tracking of high flow data through river systems	ANU	$150,000 Medium term

3.3 NSW space and satellite development	Follow the general development of space and satellite progress in NSW, including Cube Sat, Smart Sat (CSIRO), SmartSat CRC, and Waratah Seed particularly with regards offerings for water and environmental sensing.	DPIE, ANU,	$N/A Long term
4.0 Communicating uncertainty	Improving the communication of uncertainty and risk between practitioners.	USYD, DPIE, WaterNSW and NRAR.	$N/A Short term
4.1 Modelling and sensing of the underground	Improve knowledge of the underground, in particular porosity, density at a scale and resolution to better interpret the signals of gravity sensors.	ANU, MQ, USYD, UNSW	$100,000 Short – mid term
4.2.1 Probabilistic modelling framework: P1	Phase 1: Expand the current pilot to a full river-groundwater system in one catchment as outlined	USYD, NRAR, DPIE, WaterNSW	$100,000 - $200,000 Short term
4.2.2 Probabilistic modelling framework: P2	Phase 2: Testing, developing training and refinement of the framework with stakeholders.	USYD, NRAR, DPIE, WaterNSW	$100,000 - $200,000 Mid term
4.2.3 Probabilistic modelling framework: P3	Phase 3. Operationalising the framework so it can be applied as a workflow by stakeholders. Advise on low-cost, gravity and quantum sensor deployment.	USYD, NRAR, DPIE, WaterNSW, MQ, ANU	$100,000 - $200,000 Long term
5.1 Soil moisture for recharge	Measuring soil moisture to provide further insights into the spatial and temporal variability of groundwater recharge	UNSW, MQ	$150,000 Short term
5.2 Gravity and timelapse for recharge	Timelapse measurements of gravity to. measures the change in mass in the soil profile without the additional uncertainties of estimating porosity, specific yield and other soil parameters	UNSW, ANU	$150,000 Mid term
5.3 Radio tracer analysis	Increase the capacity of and reduce costs of radio-tracer analysis (Carbon-14, tritium, etc.) so that they become routine measurements	UNSW	$150,000 Long term

2.4.7 A roadmap for water research translation in NSW

A road map for translation of university water research

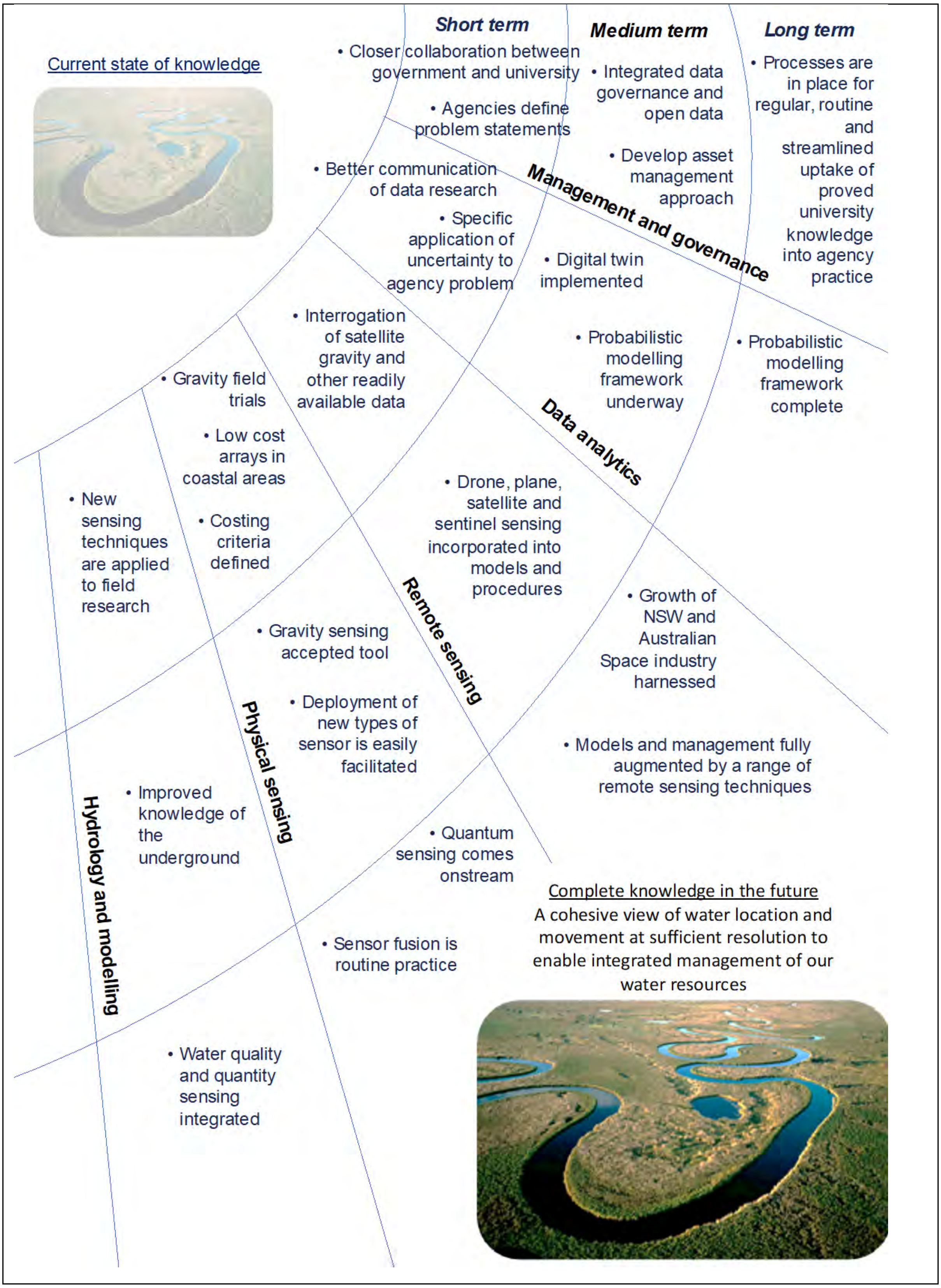


Figure 2.7: Project road-map to help solve the problem, Where is All the Water?

3 Part Three: The *Where is All the Water?* research project

3.1 Background

Water is a valuable and limited natural resource. Agriculture, various industries, the environment and metropolitan water all compete for a share of surface and ground water resources in NSW. These demands are dynamic; growing populations, changing land use and climate change are placing increasing pressure on our state's water resources.

Managing this resource in an integrated manner requires an accurate understanding of where water is and how it moves. Although we have a good understanding of many aspects of water, there are gaps in our knowledge. Agencies such as WaterNSW and NRAR regularly report data showing at least 25%, equivalent to hundreds of gigalitres, of water in catchments as "unaccounted differences".

The availability of new and emerging sensing technologies and data analysis techniques will enable us to increase certainty in the information we do have and obtain new information to fill the gaps in our knowledge.

3.2 Objectives

Three NSW agencies with key responsibility for water management; DPIE[5], NRAR[6] and WaterNSW[7] presented to the research group on the 25 February 2021. These talks gave further information on how they experience the unaccounted differences and suggestions on how better data and sensing can assist them. High level summaries were thus made:

- How can research help to manage to capture, store and release water: The right amount of water, in the right place, at the right time. (WaterNSW)
- How can the research address the issue of the "25%" unaccounted for water in some of the water balances. (NRAR)
- How can this provide data to improve knowledge of ground water resources, what groundwater is influenced by and influences on; and thus aid more robust, timely, scale appropriate decision making. (DPIE)

The two high-level deliverables the project worked towards were:

- A detailed approach that combines existing data sets with the latest developments in low-cost sensing, quantum gravity sensing, gravity data sets and data fusion techniques to address gaps in our current knowledge of water location and movement in NSW, thus addressing the issue of *Where is All the Water?*
- Collaborations between researchers and NSW Government agencies, that work towards a project design, roadmap, and prototype model to demonstrate how these agencies can use available and new data sources together in ways that increase certainty of water location and movement.

---

[5] Ground Water in NSW, Sue Hamilton, 2021, DPIE
[6] Where Is the Missing Water in NSW: The Known Unknowns, the Unknown Unknowns and the Role of NRAR, Ivars Reinfelds, 2021, NRAR
[7] Where Is All the Water, Dan Berry, WaterNSW 2021

In the early stage of the project an overall problem statement was summarised thus:

- How can we obtain a cohesive view of water location and movement at sufficient resolution to enable integrated management of our water resources.

Over the course of the project it was determined to further prove the utility of the work by applying the capability to a set of use cases. The definition and response to these use cases is provided in section 3.3.1.3.

3.3 Approach

The *Where Is All the Water?* project aims to combine existing datasets with new and advancing sensor technologies through data fusion techniques to address current gaps in the knowledge of water location and movement throughout NSW. Currently, there are significant gaps within the NSW water monitoring network, necessitating the need for the deployment of more physical, telemetry-based sensors.

3.3.1 Methodology

The vision of this project is to provide NSW agencies with better sensing and data tools, that reducing uncertainty around unaccounted for differences in the water balance, thus enabling better management of water resources.

3.3.1.1 Sensor fusion

Sensing technologies already in use include subsurface water level and pressure telemetry, water usage metering, water depth and flow in rivers and storage locations and satellite imagery to characterise land use and environmental health. This project sought to bring several new sensing technologies to provide complimentary data to existing data sources:

1. Satellite-based gravity measurement – enabling large area, relatively low spatial resolution measurement of subsurface water
2. Quantum and classical gravimetry – medium and high-resolution mapping of surface and subsurface water. Offers unique insight into large surface water to groundwater flows
3. Low-cost environmental sensors: enabling cost effective high spatial and temporal resolution in-situ sensing.
4. Use of a range of existing data sets from WaterNSW, BOM, NCRIs and others
5. A better understanding of the uncertainty of sensed data, particularly by the use of Bayesian inference.

The data fusion model plays an important role in understanding the degree to which new data sources can improve our understanding of water location and movement. The way in which this combines with physical sensing data is represented in Figure 3.1.

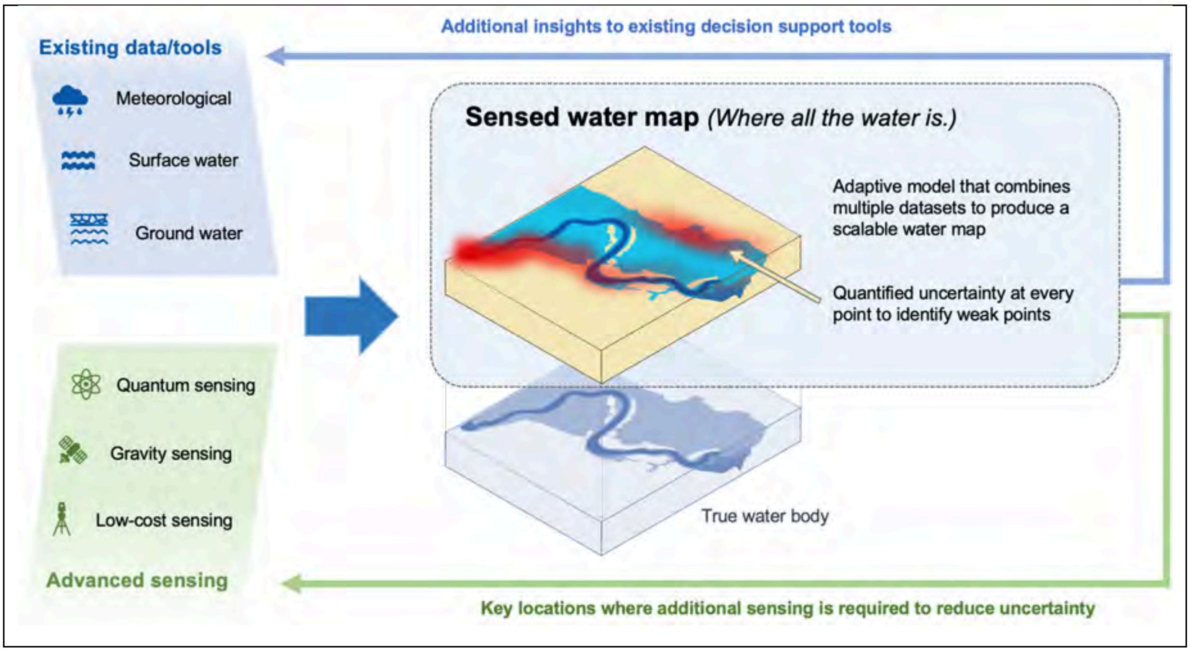


Figure 3.1: Improved understanding of the waterscape by combining data and advanced physical sensing.

3.3.1.2 Milestones

A project plan with a series of milestones was agreed to guide the project to the deliverable and addressing the problem statements. These were presented in three-month intervals. See earlier reports addressing Milestones 1 and 2. This, the final report considers the third set of milestones, with the primary contributor to the milestone in brackets.

1. Utility of high spatial resolution sensing using low-cost sensors. (Macquarie University)
2. Results of modelling gravitational fields and gradients of a simulated ground water system (ANU – Local gravity measurements)
3. Scoping of properties required of a next generation gravity sensor (ANU – Local gravity measurements)
4. Characterisation of measurements available from satellite gravimetry including uncertainties at various spatial scales (ANU – Satellite gravity measurements)
5. Methodology to provide indications of uncertainty quantified estimates of the parameters of the preliminary model. (DARE – The University of Sydney)
6. Hydrogeological conceptual models and data and reporting on groundwater recharge (UNSW)
7. Findings regarding to what extent water quality data can complement water quantity data. (All)
8. Technology review. (All)
9. Roadmap for translation of research to day to day use by DPIE and other organisations, based on use cases, proposals for further work (NSW Smart Sensing Network).

3.3.1.3 Use cases

As is normal for an intellectual study of this kind, original parameters of enquiry are informed along the way. Part way through the project it was decided to consider use cases, potential, real-world scenarios where the research finding might best show their utility. Practical use cases should have criteria that considered situations about which:

- The analysis will provide a useful outcomes, for researchers and agencies
- Take advantage of the multi-disciplinary approach in the project
- They are consistent with the last six months work
- They have suitable data and resources available
- Can be worked into proposals for further work

A large-scale use case: A large-scale use case was proposed that cover enough of the spatial domain that it can deliver interpretable outcomes using the satellite grace data (> 500 km). The large-scale use case gives an overview of the shifts in the water balance as a result of a large flow through the system: How much is retained in the system, how much evaporates and how much is delivered downstream.

Low flow use case: Low flow uses cases are critical as they are likely to be caused by drought and other stresses. This is when management of water resources is extremely important. It is also known that various agency models that work adequately in times of regular flow, do not always hold when there is limited flow.

High flow use case: In contrast to the Low-flow use case, a High-flow use case covers a flooding event or a very wet period (such as early 2021 or 2016). There are many questions here: can overall volume balances can be trusted, what is the accuracy of the gauging network, what fraction of the flood water is recharging aquifers either directly via the stream channel, from inundated floodplains or by rainfall across the landscape, and what remains in surface water and floodplains or evaporates? In particular the attention on the different recharge processes can build on the smaller scale case studies in this case study.

3.3.2 Management and governance

3.3.2.1 Project Budget

Table 3.1: Project Income

DPIE	$150,000
OCSE	$150,000
NSSN	$150,000
In-kind: NSSN, DPIE, Macquarie, ANU x 2, DARE x 2 and UNSW management and professorial support	$400,000
	$850,000

Table 3.2: Project Expenditure

Macquarie researcher costs	$90,000
ANU researcher costs	$180,000
UNSW researcher costs	$90,000
The University of Sydney (DARE) researcher costs	$90,000
In-kind expenditure	$400,000
	$850,000

3.3.2.2 Project management

An adapted Scrum methodology was used over the nine-month life of the project. This included online Sprint meetings every three weeks which was the ideal time for collaborators to meet, provide brief reports and identify tasks requiring collaboration.

Milestones and reports were delivered at three-month intervals.

The project was largely undertaken during periods of COVID-19 restrictions.

3.3.2.3 Steering Committee

Three Steering Committee meetings were held over the life of the project at three monthly intervals. In addition to the project researchers, representatives from NRAR, WaterNSW and several senior staff from DPIE were present. A brief Terms of Reference was installed at the start of the project. Minutes of each meeting are kept on file.

3.4 Results

3.4.1 Milestone Results

3.4.1.1 Utility of high spatial resolution sensing using low-cost sensors

Macquarie University (MQ) had the deliverable of developing prototype nodes for utilization of point-source measurements to increase the temporal and spatial resolution of data from stations that are currently available. MQ designed and developed initial prototype nodes and have commenced testing and deployment of them. This included first deciding on the parameters that need to be sensed, which was rainfall, ambient temperature and humidity and soil moisture and temperature at differing depths. Then an initial low-cost technology review was undertaken for the chosen parameters, this is included in-detail within the report. Off-the-shelf sensors were then purchased and an electronic sampling system was designed to utilize all the necessary sensors and transmit data wirelessly utilizing LoRaWAN communication (see section 4.1.3.4). The development of the prototype nodes was initially conducted at a residential home due to the COVID-19 restrictions that were in place. Once the restrictions had been lifted, development of the first two prototypes were completed on-campus and then one was deployed at the Narrabri L'lara farm for testing.

More nodes are being developed now to deploy alongside the initial prototypes to achieve an overall very high spatial and temporal resolution sensing system which can be deployed to any desired testing site.

MQ is currently researching the inclusion of water quality data within these sensing systems, for example utilizing low-cost nitrate sensors within the soil to both inform about soil conditions as well as potentially developing an algorithm to relate the soil moisture content to nitrate readings. There is room for inclusion of numerous water quality sensors and they could provide additional critical information.

3.4.1.2 Results of modelling gravitational fields and gradients of a simulated ground water system

A flexible gravitational simulation package was developed in MATLAB that can model gravitational and gravity gradient signals from arbitrary three-dimensional mass distributions defined via voxels. Following this, through collaboration with hydrology experts and existing borehole data, a voxel world was created that simulates the area around Middle Creek farm (see Figure 4.15 and Figure 4.17). This voxel world was then simulated and the gravitational signal in the vertical direction (g_z) is shown in Figure 4.20, the vertical gravitational gradient in the vertical direction (G_{zz}) is shown in Figure 4..

3.4.1.3 Scoping of properties required of a next generation gravity sensor

The results from the gravitational simulation give an indication of the expected signal magnitude between dry and wet conditions and for various survey heights above the ground. The largest signals in G_z are of the order of 10^{-6} ms^{-2} and the smallest simulated are of the order of 10^{-8} ms^{-2} This sets a minimum viable sensitivity in the range of 10^{-7} ms^{-2}. Such sensitivities have already been demonstrated to this level in laboratory-based devices.

Quantum sensing devices, based on falling cold atoms, offer increase in sensitivity of orders of magnitude.

3.4.1.4 Characterisation of measurements available from satellite gravimetry including uncertainties at various spatial scales

A simulation study was performed to assess the accuracy with which signals could be recovered on ~200x200 km entities in the Maules Creek region. (These entities are referred to as mascons, short for mass concentrations) For a 150 mm simulated signal – which is realistic for most situations in Australia - some smearing occurs in the recovered signal from where the signal should actually be located. Errors of ~10-15% occur in estimates up to 400 km away but are insignificant beyond that distance (Figure 4.24).

An alternate approach to estimating mass changes on mascon tiles is to look at the magnitude of the signal in the actual inter-satellite measurements themselves. This was assessed for both the simulation over Maules Creek and also to determine whether March 2020 flood waters could be tracked from southern Queensland into the northern rivers of NSW (Figure 4.29). While there is some indication that this might be possible, the interpretation is limited by the 5-day repeat time of the satellites flying over the same location. A careful interpretation of the available data is required, with cross-referencing of available stream flow data, to determine whether this is a viable method for obtaining higher spatial and temporal resolution.

3.4.1.5 Methodology to provide indications of uncertainty quantified estimates of the parameters of the preliminary model

As a first step DARE developed an in-depth analysis of the existing water accounting framework for the Namoi river in 2018/2019 from DPIE and quantified the uncertainty for the reported components based on readily available information. This has subsequently been extended to describe a formal uncertainty framework that can analyse the uncertainties in the river/groundwater system. This framework could not have been achieved without in-depth discussions with other sub-projects, in particular UNSW.

Collaborating with the ANU sub-projects, DARE has provided an analysis of the potential uncertainties in inversion of the gravitational measurements. This has highlighted that measurement errors are a negligible contribution to the uncertainty and the uncertainty is mostly related to the unknown structure of the underground.

In general, the framework is a fully specified Bayesian likelihood framework, which explicitly defines the posterior distributions of the water balance estimates in both time and space. This is mostly built on Gaussian Process modelling for the spatial and temporal integration.

For these two identified analyses, DARE has developed estimates of the time and effort required in developing this into a fully operational system for a river basin.

A pilot analysis of a low flow case study in the Namoi river has identified the following:

- An example which quantifies the uncertainty in stream flow gauging indicating that the main model estimation uncertainty is in the high flow estimates (due to a low number of observed values), but that measurement errors are a major component of the low flow estimates;
- An example which quantifies the uncertainty in evapotranspiration from open water and riparian vegetation highlighting that observations of the river width in time and space is a major uncertainty; and

- An example which quantifies the uncertainty in the groundwater surface water fluxes indicates the main uncertainty is in the sparseness in time and space of the groundwater observations combined with the interpolation of the surface water.

Uncertainties in water quality, which were not assessed, are likely to be even greater than for water quantity simply due to the scarcity of measurements. Most likely, interpolation of existing data and water quantity data would be used to extend data series, which introduces further uncertainty.

3.4.1.6 Hydrogeological conceptual models and data and reporting on groundwater recharge

Groundwater recharge is one of the components of the water balance that is notoriously hard to determine. There are two mechanisms of groundwater recharge, diffuse recharge where the groundwater table is directly recharged by rainfall through the general land-surface and focussed recharge where the groundwater table is recharged by flow in stream/river channels. This work has shown that for sites close to a stream or river the predominate recharge mechanism is focussed recharge, with on average one recharge event per year. This compares to three potential diffuse recharge events over 10 years at Middle Creek Farm, for example. At the Wellington NCRIS sites one significant diffuse recharge event was observed over the 10 years of monitoring, with other smaller events also occurring but some were not widespread across the site, suggestive of spatial variability in diffuse recharge.

This work also showed that antecedent conditions (i.e., soil moisture deficit) are an important determinant, along with rainfall amount when it comes to the likelihood of diffuse groundwater recharge occurring.

Based on the NCRIS data sets utilised in this project we have developed hydrological conceptual models for different use cases, based on the Maules Creek Catchment. These conceptual models have informed our work with the use cases (section 3.3.1.3), with further detail provided in the sub-project report (section 4.5).

3.4.1.7 Findings regarding to what extent water quality data can complement water quantity data.

The low-cost sensor nodes developed by Macquarie University facilitate the inclusion of sensors for water quality parameters. Alongside the development of the prototype system, alternative or developing techniques will be scoped out and analysed for future implementation. This includes the potential addition of low-cost water quality sensors to monitor hydro geochemistry, such as (pH, EC, NO^3, CO^3 and SO^4).

Neither local nor satellite gravity measurements offer a direct measurement of water quality. It can perhaps be said that a considerable number of water quality problems (e.g. deoxygenation) would be mitigated by regular flows of water in appropriate quantities. The CSIRO Aquawatch [5] program intends to investigate water quality parameters using satellite derived information.

Knowing the age of groundwater can be a very useful for estimating the groundwater residence time or renewal rate but integrated over the catchment and therefore compliment soil and vadose estimates of recharge. A ‘stretch goal’ would be to increase the capacity of and reduce costs of radio-tracer analysis (Carbon-14, tritium etc.) so that they become routine measurements. An alternative would be to develop sensor methods to measure these continuously and in-situ (very sci-fi at this stage).

Frequent measurements of a suite of dissolved ions and potential contaminants may also inform on water quality trends threatening quality of use and environmental receptors. Such changes may also inform on hydrologic changes due to management and climate change. Again, in-situ continuous measurements would be beneficial to capture ‘hot moments’ i.e., times of pulses or greater change. This can currently be achieved for the total load of dissolved ions by electrical conductivity probes, but they are prone to measurement drift.

3.4.2 Use case results

3.4.2.1 A large scale use case

A large-scale use case was proposed that cover enough of the spatial domain that it can deliver interpretable outcomes using the satellite (GRACE) data (> 500 km). The large-scale use case gives an overview of the shifts in the water balance as a result of a large flow through the system: How much is retained in the system, how much evaporates and how much is delivered downstream.

Comparing the results of this scale satellite gravity data monitoring to traditional water balance analysis (difference between gauge observations, groundwater bore responses) will identify anomalies and this will be integrated in hydrological modelling as a constraint.

In other words, while the satellite gravity data starkly shows information at a 200km^2 scale, this does provide a tool for finer use. if there was a significant inflow of water into region X, why is subregion X’ reporting a deficiency of water. If region Y is in overall deficiency of water, why is subregion y’ using excess water?

With refinement and further interrogation, satellite gravity data might well show some classic NSW river events such as the flow of water from the Northern Rivers or South-East Queensland to the Menindee lakes along the Darling.

Comparing the satellite gravity data analysis to operational forecast estimates will inform downstream impacts related to upstream management decisions. In addition, the volume estimates of water retained in the landscape compared to groundwater observations will confirm recharge volumes and groundwater recovery. Longer term satellite gravity observations will further inform how much of the stored water evaporates and how much is stored longer term in the landscape, for example, is the groundwater truly recovered after drought, or is most of this stored in the upper soil layer and evaporates?

The deployment of low-cost sensors to capture information of a release event over 1000km is an entirely feasible prospect.

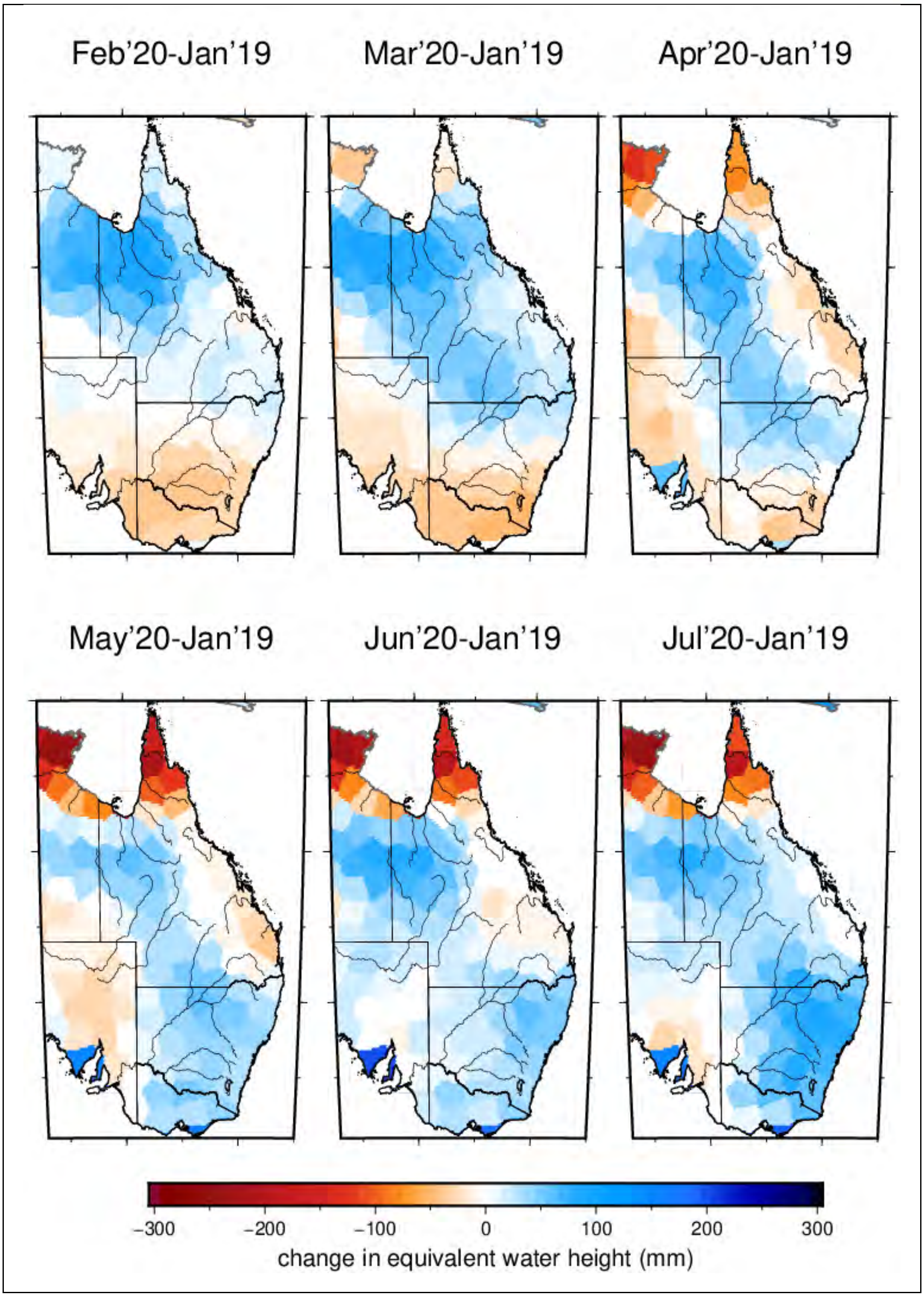


Figure 3.2: Large scale water movements as represented by the change in gravity over the first half of 2020.

3.4.2.2 Low flow use case

Tactical water releases from storage are made to assist a number of downstream clients that may include towns, environmental wetlands or agricultural irrigation areas. During very dry periods (such as late 2019) considerable transmission losses have been reported, indicating that water released (in response to a water order) has not being delivered to the client. The question is whether this water is lost to groundwater, evapotranspiration, is extracted by other users, or bypassed the client in some way. Increased transmission losses during dry periods are currently taken into account by WaterNSW, but it is unclear if these estimates are accurate. Given that all the data and algorithms to derive these estimates are uncertain, there are opportunities for reviewing improved observation methods, optimised locations for observations or improved estimates of variables.

As an initial case study DARE focussed on the Namoi River between Keepit dam and Gunnedah, including the inflows from the Peel and the Mooki, to derive uncertainties at different spatial locations for groundwater losses, evapotranspiration losses and surface water observations. More specifically, the outcome of this study would be able to pinpoint the spatial locations and times with the highest uncertainties and the variables (for example evapotranspiration estimates or streamflow estimates) with the highest uncertainties. The scope of the case study is a well-instrumented reach, but the methods can, with appropriate data, be extended to larger scales or longer reaches (such as the large-scale use case mentioned above).

The UNSW work has drawn on prior information to define reliable initial estimates from detailed studies such as the Maules Creek and Wellington case study. For example, this will inform initial estimates of fluxes to groundwater.

In addition, the results of this study will pinpoint locations and times where deployment of novel sensors (MQ Uni) or gravity measurements (ANU) are most valuable and can then be easily coupled to trial field studies.

NRAR should consider the data analytics techniques and improved understanding of GW/SW interaction.

3.4.2.3 High flow use case

A)

A larger scale high flow case study in the Namoi catchment concentrates on the river reach from Boggabri to Walgett. Starting at Boggabri has the advantage of including the past research at the small-scale long-term Maules Creek study site, but as pointed out, complications with inflows and contributions from different geological formations may increase complexity. Constraining to the lower Namoi west of Narrabri can provide a cleaner option (possibly with sparser data. Using the lower Namoi will be large enough to also use GRACE satellite data providing a further mass balance check as suggested in the large-scale study.

The outcomes of this case study specifically inform water sharing plans, 45rovideng large scale understanding of the volumes of water in the different parts of the water balance and provide better understanding of flood volume estimate uncertainties in large catchments. Following a recent flood event (for example in 202), local gravity measurement campaigns can provide more detailed estimates of changes in groundwater volumes, for example different recharge volumes close to the river channel and on the back plains, identifying differences between diffuse and direct recharge. In addition, cheap sensing of rainfall and soil moisture, can further inform water balance estimates by characterising spatial variability and providing verification or rainfall radar estimates.

B)

A companion high flow study will provide more information on the actual recharge processes and how streams and groundwater recover after drought. This study will focus on the historical work and investigations of this project at the Maules Creek site that ANU-Gram and UNSW have been concentrating on, as well as the investigations at the Wellington site. It will take advantage of the wealth of existing data, understanding and knowledge to provide a focussed and detailed investigation in the local hydrological processes. This can further inform the larger scale studies by providing understanding of the "how" to augment the "what" in the larger scale analysis. This work links into the testing and concept proof of low-cost sensors and the ANU-gram local gravity measurements. At this smaller scale, processes can be studied in more detail and particularly aim to unravel the balance between "diffuse" recharge and "direct" recharge. This provides important information about the resilience of the groundwater system and informs water sharing decision making. For initial testing

Low-cost sensors are currently installed at "Llara", a 1000 ha farm and part of The University of Sydney Plant Breeding Institute. (12656 Newell Hwy, Narrabri NSW 2390), alongside capacitance probe moisture sensors installed as part of a landscape rehydration project.

Given the scale of the more recent satellite products (for example 10 m pixels for Sentinel 2), satellite quantification of flood volumes will further inform the overall water balance estimates, linking to the high flow use case A.

3.5 Impact and Conclusion

The *Where is All the Water?* project has shown how low-cost sensing can fill gaps in NSW data; local gravity sensing of river catchments has been shown to be feasible, and the movement of water through the landscape can be measured using satellite derived gravity data. Bayesian data analytics techniques have shown they have the potential to reduce uncertainty in agency water data reports, while predominance of focussed, over diffuse rechange mechanisms has been demonstrated.

The key outputs of this body of research are reported in more detail in Part 1, while the overarching narrative is described in the Executive Summary at the start of this report.

This work has informed a road-map for the uptake of university research into water sensing and data analytics, see Part 2.

4 Part Four: Sub-project Reports

4.1 Low-cost sensing of water catchments

Subhas Mukhopadhyay, Brady Shearan
Macquarie University

4.1.1 Abstract

The *Where Is All the Water?* project aims to combine existing datasets with new and advancing sensor technologies through data fusion techniques to address current gaps in the knowledge of water location and movement throughout NSW. Currently, there are significant gaps within the NSW water monitoring network, necessitating the need for the deployment of more physical, telemetry-based sensors. This research showcases the design, development and deployment of a low-cost telemetry-based sensing system for monitoring various environmental parameters including soil moisture and temperature at an array of depths, rainfall and ambient conditions from any remote location throughout NSW. The Macquarie University workstream aims to aid in decreasing data uncertainties through developing and deploying the proposed low-cost systems to provide an increase in the spatial and temporal resolution of data that is currently available. The collected sensor data is transferred and stored in the cloud periodically through the Long-Range Wide Area Network (LoRaWAN) communication protocol. Adaptation and data collection from the deployment of a large number of the proposed low-cost sensor nodes in targeted locations throughout NSW will provide a major breakthrough in addressing current gaps in the knowledge of water location and movement throughout the state.

4.1.2 Introduction

The NSW Chief Scientist & Engineer's Review of water-related data collections, data infrastructure and capabilities released in July 2020 highlights the significant gaps in the water monitoring network, particularly the large variation in density of sensors between rural and metro areas [1]. The majority of current sites at which water data collection occurs, excluding water storage locations (dams) still utilize manual data collection methods (Figure 4.1). This means that most data are not automatically transferred to data repositories in real-time, and manual site visits are required to collect logged sensor data. It is also stated that the current network is insufficient for confidently measuring physical/chemical water parameters as well as understanding the surface and groundwater interactions that are occurring. Therefore, it is apparent that significant technological upgrades to the monitoring network need to be made, and through the research and development of low-cost telemetry-based sensor systems, both the data uncertainties and gaps in the monitoring networks can be greatly reduced.

The Macquarie University sub-project is developing low-cost sensor nodes to be deployed to monitor traditional environmental parameters to increase the temporal and spatial resolution of historical data that currently exists. Low-cost sensors will be deployed to provide ground-truthing real-time data that is optimised to maintain high spatial and temporal resolution, aiming to provide a better understanding of what is causing the significant uncertainties in the whereabouts of the 'missing' water that is largely unaccounted for throughout NSW.

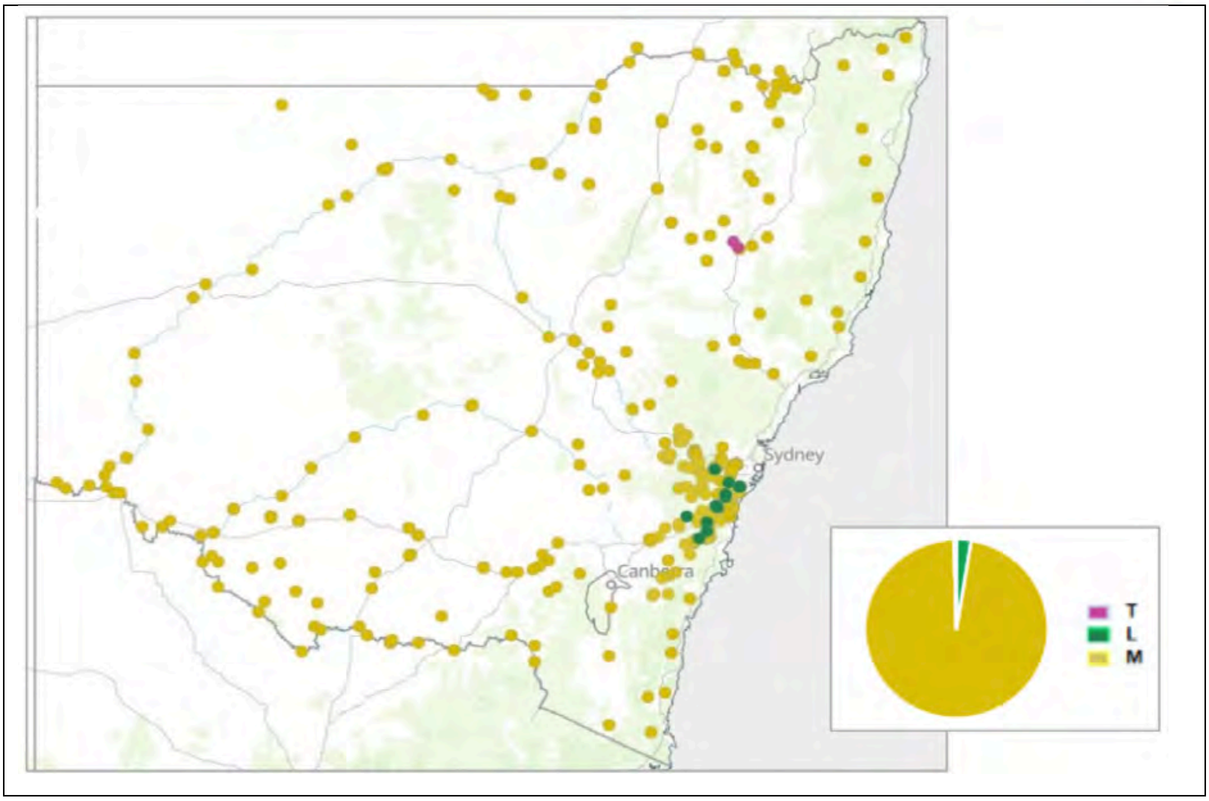


Figure 4.1: WaterNSW water quality monitoring network sites. T: Telemetered, L: Logger only, Manual.

Currently, there is significant amounts of water loses throughout waterways within NSW that are largely unaccounted for. These differences can be categorised into three main areas of interest, including surface evaporation, riparian vegetation transpiration, groundwater flux (loss to GW, GW to trees).

One of the main parameters of focus for the prototype system is groundwater recharge, which will be analysed through measurements of the change in soil temperature and moisture over time. This will aid the modelling in understanding the reactions occurring between surface and ground water and investigate losses corresponding with groundwater flux.

The emergence of Wireless Sensor Networks (WSNs) are driving the ever-developing Internet-of-Things area. They have received a great deal of attention across a diverse range of research areas with a particularly recent interest in environmental monitoring. Deploying WSNs over larger and larger areas has resulted in the development of communication protocols which are able to utilize sub-GHz frequency bands for communication over significant distances. These frequency bands provide superior propagation characteristics in comparison to higher frequency bands, though these typically can only maintain very low data rates, which is suited towards sensor applications.

The proposed system will employ low-cost sensing techniques to monitor traditional parameters whilst utilizing wireless connectivity to improve upon the spatial and temporal characteristics of data that is currently available. The system will measure rainfall, soil moisture and temperature using a sensor array, as well as ambient temperature and humidity. These parameters will be measured through low-cost methods to augment official gauging stations at unmeasured locations. The new sensors will be developed as Internet of Things (IoT) enabled sensor nodes such that the measured data will be uploaded to cloud in real-time without any manual intervention.

The developed low-cost sensing systems will not have any restrictions in terms of use at any specific location, apart from the individual sensors general placement constraints. The use of LoRaWAN communication was chosen due to the minimal deployment costs and constraints alongside significantly low power consumption and being deployable in both urban and remote testing locations.

4.1.3 Technology review of low-cost sensors

The following will explain an initial review of the low-cost technologies available for the decided sensor system parameters.

4.1.3.1 Rainfall Sensors

Historically, the most commonly employed technique for point-source rainfall measurements is the use of a tipping-bucket rain gauge. These utilize a mechanical fulcrum that tips when water is filtered into the system through funnelling (Figure 4.2). The volume of each side of the fulcrum can be equated to the amount of water that has fallen over time [6]. Though more recently, low-cost techniques have become more popular within research applications. These include capacitive sensors as well as ultrasonic sensors, which utilize low-cost, low-power techniques to determine the water level of a container after rainfall has occurred [7].

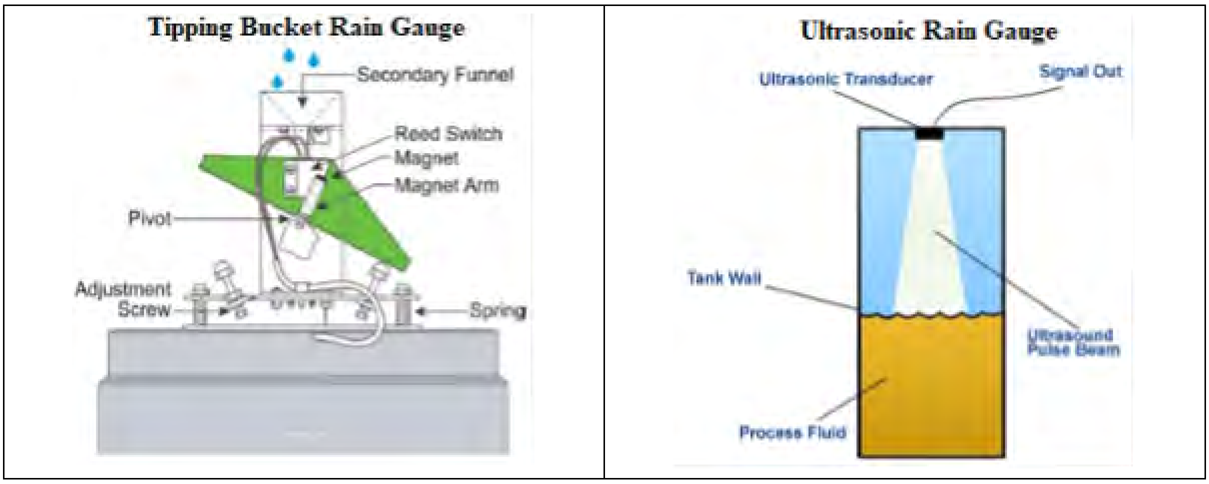


Figure 4.2: Tipping bucket and ultrasonic rain gauge. [8, 9]

An ultrasonic sensor was deployed within our designed system due to the process of water being stored and level-measured in a container before being emptied. This allows for the addition of water quality sensors to be easily implemented into the system at a future point, providing additional information on the quality of the rainwater.

In terms of larger scale applications, recently there have been research outputs investigating the use of remote sensing applications to monitor rainfall over large areas. One of the most recent techniques used is the estimate of rainfall through measuring the change in telecommunication signal strength between two or more radio towers. This technique is only very recently seeing research and development application within Australia [9]. Similarly, to the mobile towers, a similar remote sensing application could be conducted utilizing the LoRaWAN signals that are produced by each of the sensor nodes, acting as a 'portable cell tower". Though this type of technique is still in a stage of very early development and was not deemed appropriate for the timeline of the initial stages of this project.

An ideal outcome of this project in the long-term, would to be to utilize a similar type of remote sensing technique and deploy the designed sensor nodes as ground-truthing calibration points over a much larger scale.

4.1.3.2 Soil temperature and moisture

All soil moisture measurement methods have their own specific limitations and challenges, with ideal sensors being chosen based upon their size, accuracy, cost, response, ruggedness and other parameters. Often the actual sensors used within these applications are not expensive themselves, but rather the interfacing/sampling electronics obtaining data from the sensors are expensive.

Currently, there are a wide variety of techniques that are utilised to develop sensors to conduct soil moisture measurements. These include dielectric, gravimetric, tensiometer, neutron, gamma-ray project and remote sensing. A detailed review and comparison of each of these technologies can be seen in Table 4.1 with further explanation available in [10].

Table 4.1: Available Soil Moisture Sensor Technologies.

Sensor / Manufacturer	**Principle**	**Brief Description**	**Cost**
SEN0193 DFRobot (China)	Capacitive	The simplistic single probe capacitive sensor is coated with insulating paint to prevent corrosion and scratching for long-term deployment. Though additional waterproofing may be required.	$15
HD3910 Delta Ohm (USA)	Capacitive	Three-probe capacitive sensor which is suitable for measurements in small volumes, giving direct volumetric water content readings from 0-60%.	$370
10HS Decagon Devices (USA)	Dielectric principle	Distinct instrument sensitivity to soil type, thus indicating the necessity for specific individual soil calibration. More costly than similar devices.	$400 / $900 (With LoRa)
TDR-315N Acclima company (USA)	Time-domain reflectometry (TDR)	TDR based small 3-probe based soil moisture sensor. It achieves acceptable accuracies for managing irrigations at the site with low salinity and low clay content.	$650
ECH-GS3 Decagon Devices (USA)	Dielectric principle	These sensors offer research-grade accuracy at an economical price. Though often result in a more costly installation process.	$850
SoilVUE10 Campbell Scientific (USA)	TDR	This is a soil water content profile sensor, measuring soil moisture, electrical conductivity, and temperature profiles	>$1000
EnviroPro EnviroProSoilProbes (Australia)	Capacitive	Field of influence is significantly larger than capacitive sensors of a similar diameter. This allows for more meaningful soil measurements (Moisture, temperature and EC).	>$1000

The most cost-effective way of measuring soil moisture from a point-source perspective is through utilizing low-cost capacitive sensors. These commonly used sensors are an efficient way of monitoring the moisture content, though these sensors are prone to rapid corrosion and error over long periods of use. Therefore, when selective a sensor for use, it was important to select one that has been coated with an anti-corrosive, waterproof material. This type of coating also provides scratch-resistance which is of particular importance when installing the sensors into coarse soil.

An example of the types of capacitive soil moisture probs can be seen in Figure 4.3, with the sensor that was deployed in the final system being the SEN0193 soil moisture sensor (b), manufactured by DFRobot [11]. These were the most appropriate low-cost capacitive sensor for use within our initial prototype system, due to significantly low installation costs and ease of interfacing. Other more costly capacitive-based solutions operate using the same principle, though they often use an array of capacitive probes to provide continuous readings of the soil at desired depths. We are replicating this utilizing an array of low-cost SEN0193 sensors.

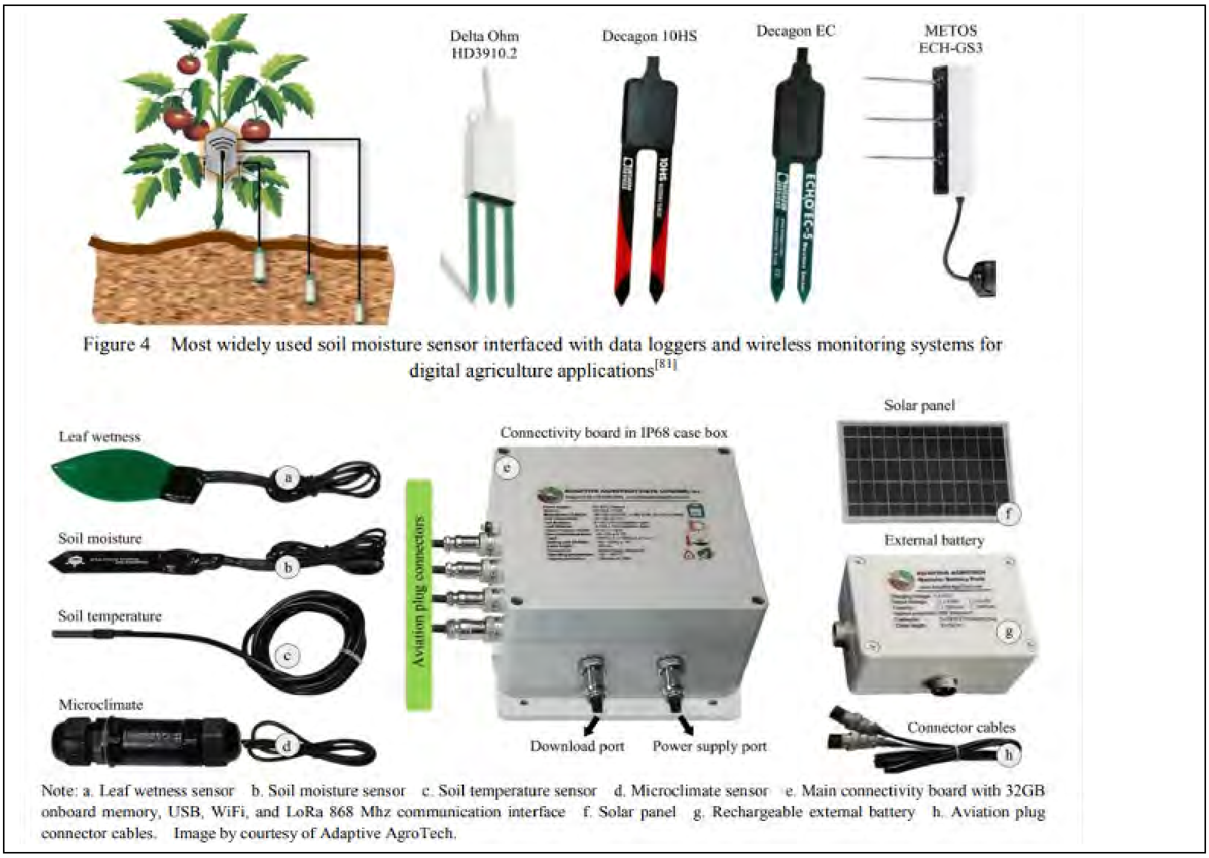


Figure 4.3: Common soil moisture sensors and data logger systems [5].

Soil temperatures measurements are generally limited by the cost of the sensors, with point-source measurements being the cheapest and non-invasive measurements being the more expensive route. These prototype nodes were designed to conduct point-source measurements; therefore, standardised miniature digital thermometers are the most appropriate.

Currently there is research surrounding the use of fibre-optic cabling as an invasive technique to measure the change in soil temperature over significant distances, though these techniques still currently accompany significant development and deployment costs [10]. Significant research is currently being undertaken to improve remote and proximal sensing techniques including Infrared and microwave signal systems such as RFID [11]. These currently are highly expensive and both hold limitations in their measurement depths and accuracy in small-scale applications, making them less than ideal for recharge monitoring.

For the temperature sensors, waterproofing and scratch resisting becomes of significant importance for deployment over long periods to maintain sensor accuracy and reliability. Therefore, it was decided that the SHT-30 would be deployed within the initial prototype nodes. These sensors are epoxy coated and placed inside a metal mesh encasing, ensuring they are completely waterproof and employable within corrosive environments. Other sensors such as the soil temperature sensor Figure 4.3 are deployable only in certain soil types that do not hold heavy metal corrosion rates. As the final deployable location was not finalised during the design, soil conditions were unknown and therefore the safest option was to deploy a rugged and shielded temperature sensor.

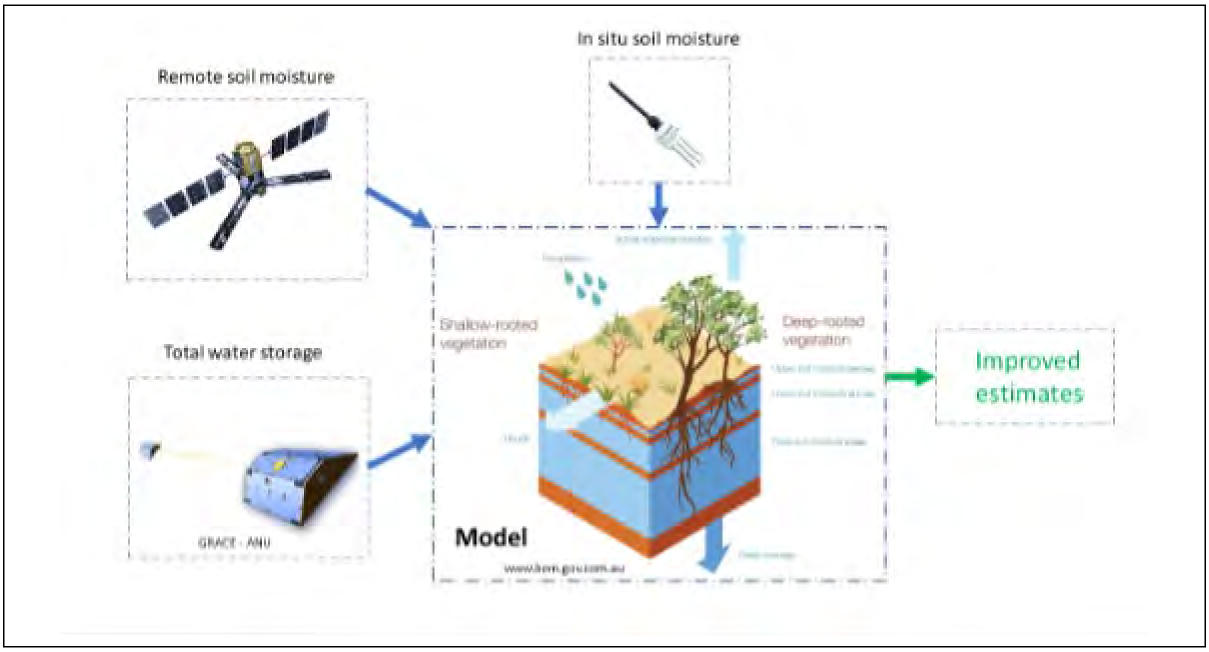


Figure 4.4: Combination of remote and in-situ sensing techniques for improved water balance estimates.

Again, an ideal outcome of this project in the long-term in terms of soil moisture and temperature, would to be to utilize a remote sensing technique such as infrared or RFID sensing and deploy the designed sensor nodes as ground-truthing calibration points over a much larger scale. These could then be utilised alongside gravity sensing to greatly improve the large-scale estimates of the water-balance equation, as exhibited in Figure 4.4.

4.1.3.3 Ambient Temperature & Humidity

Numerous types of low-cost temperature and humidity sensors are readily available for purchase, making the main deciding factor being the size and ruggedness. The low-cost sensor we plan to deploy in our system is one that has been utilised for numerous research projects within our laboratory at Macquarie University, the “CCS811/BME280 Environmental Combination Sensor”. This sensor provides ambient temperature and humidity readings, as well as the optional monitoring of multiple other environmental parameters including CO2, atmospheric pressure and total volatile organic compounds. Although these may not be the main parameters of interest, measuring these parameters accurately alongside our system may provide useful data that can be analysed and aid in modelling or finding trends within the data.

4.1.3.4 Telecommunication

The selection of an appropriate telecommunication protocol is the most critical component of a wireless sensor node. Many different protocols are available, providing a wide range of properties for different applications. For our application, we require a communication network that is very low-power and allows for regular transmissions. There are numerous protocols available for our application, as can be seen in Table 4.2.

Table 4.2: Communications protocol comparison.

Technology	Frequency	Data Rate	Range	Power Usage	Cost
3G / 4G	Cellular Bands	10 Mbps	>15km	High	High
Bluetooth / BLE	Sub-GHz, 2.4 Ghz	1-3 Mbps	< 100m	Low	Low
Wi-Fi	Sub-Ghz, 2.4 GHz, 5 Ghz	0.1-54 Mbps	< 100m	Medium	Low
Zigbee	2.4 GHz	250 Kbps	~ 200m	Low	Medium
NB-IoT	Sub-GHz	< 1 Kbps	> 10 km	Low	Medium
LoRa	*Sub-Ghz*	*< 50 Kbps*	*~ 200m*	*Low*	*Low*

LoRaWAN, which stands for 'long-range wide area network', was chosen as the communication protocol to use for the prototype nodes as this developing technology offers the lowest power consumption and the most cost-effective data transmission as the sensor data payload size is very small. The LoRaWAN protocol allows for data to be transmitted over significantly larger distances than a typical Wi-Fi network, though at much lower data rates [12].

The sensor nodes proposed within this research do not require significant quantities of data to be transmitted, though rather it requires regular small data transfers over significant periods of time. Therefore, as the nodes are ideally to be placed in the field for up to years at a time, the LoRaWAN network is the best option for the application of these nodes. Fortunately, both the testing site at Narrabri farm and Macquarie University have installed a LoRaWAN gateway, which allows for the sensor nodes to be easily connect to the network.

There has been development of the use of low-earth orbit satellites to act as gateways for LoRaWAN nodes, as depicted in Figure 4.5. This is particularly useful for remote sites of interest in which 3G and other communication options are not available. There are Australian companies working on commercialising this technology, including FleetSpace [13], which are aiming to offer satellite gateway connections for as a low as $2 USD per device annually. Moving forward, this would be an ideal solution to removing the requirement of a gateway node when deploying a large number of nodes over significant distances.

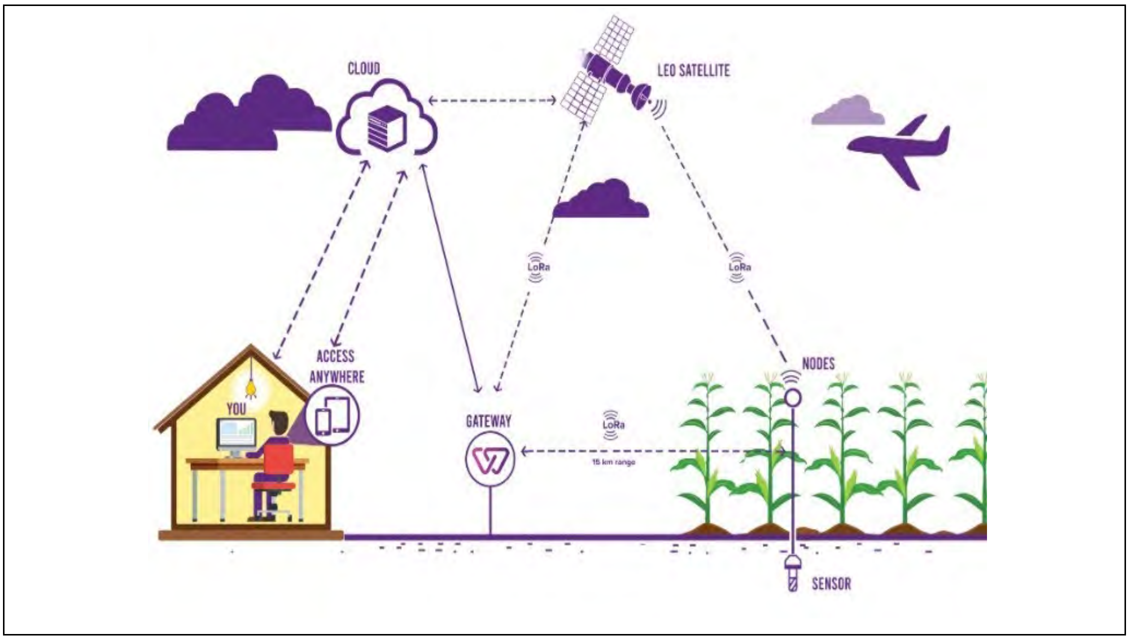


Figure 4.5: LoRaWAN LEO satellite connection example diagram.

4.1.4 Sensor Node Design

The prototype system utilizes an array of three soil moisture and three soil temperature sensors, an Arduino Uno microcontroller, an environmental combination sensor, a rain gauge, a solar panel, solar power conversion shield and a rechargeable battery to maintain energy autonomy. The soil moisture sensors are connected to the Arduino as analogue inputs and the soil temperature sensors are connected using the I^{2}C communication protocol, through an I^{2}C multiplexer as the temperature sensors all shared the same local address. The solar charger shield and the LoRa shield are also connected with the main microcontroller through Arduino compatible shields. The 6000 mAh rechargeable battery and 6 V solar panel are connected with the power management block. A summary of the types of sensors and other electronics are listed in Table 4.3, the circuit block diagram of the proposed system is shown in Figure 7.1 and a flow diagram for system operation in Figure 7.2.

Table 4.3: Electronics used in the prototype system.

Component Name	Description
Arduino Uno	Microcontroller used for interfacing sensors and components
DFRobot Soil Moisture Sensor	Capacitive soil moisture sensor with protective coating to prevent corrosion and scratching
SHT-30 Temperature Sensor	Weatherproof temperature sensor with metal encasing for soil deployment
CCS811/BME280 Environmental Combo	BME280 provides humidity (%), temperature (◦C), and barometric pressure (Pa)
I2C Multiplexer	I2C multiplexer board to combat device addressing issues.
Rain Gauge	Tipping Bucket / Ultrasonic Rain gauge
6V 6W Solar Panel	Solar panel for Voltaic
Seed Solar Panel Conversion Shield	Solar panel power conversion shield for Arduino
LoRa Shield for Arduino	915 MHz Long-range transceiver
Polymer Lithium-Ion Battery	3.7 V 6000 mAh Rechargeable battery
ANT-916-CW-HWR-SMA	External antenna

Battery calculation for developed sensor node

The sensor node system remains on for the entire period it is deployed. Sensors are set into an idle state when not in use for battery autonomy. The system was developed to run in 15-minute transmission cycles, taking only around one second to obtain data then transmit it. Once data is collected, the sensors that are able to be programmed go in to an idle state while the LoRa shield goes into an active mode for data transmission. Once the data is transmitted the shield goes into an idle state and the system goes into a low-power state. Table 4.4 shows the current consumption by the sensor node.

Table 4.4: Node current consumption calculation.

No.	Name of the Component	Run time	Current consumption (mA)	Current consumption in 1 cycle (mAsec)
1.	Arduino Uno MCU	15 min	20	= 20 × 60 × 15 = 18000
2.	3x Soil Moisture	15 min	15	= 15 × 15 × 60 = 13500
3.	3x Soil Temperature	15 min	1	= 1 ×15 × 60 = 900
4.	Env. combo (idle)	14 min 59 s	1	= 1 × (14 × 60 + 59) = 899
5.	Env. combo (active)	1 s	13	= 13 × 1 = 13
6.	LoRa shield(idle)	14min 59 sec 940 msec	1	= 1 × (14×60+59+940/1000) = 899.94
7.	LoRa shield (Active)	60 msec	120	= 120 × (60/1000) =7.2

The total current drawn by the system in if data is collected twice daily is found as,

$$I_n = \frac{18000 + 13500 + 900 + 899 + 13 + 899.94 + 7.2}{60 * 15} = \frac{34219.14}{900} = 38.021 \text{ mA}$$

Power consumption, P_s is calculated as

$$\Rightarrow P_s = \text{V} \times I_n = 3.7 \times 38.02 = 140.678 \text{ mVA}$$

If the power converter efficiency is 0.85%, the minimum mVA required from the battery,

$$P_{bs} = \frac{140.678}{0.85} = 165.50 \text{ mVA}$$

As the voltage of the battery is 3.7 V, the required discharge from the battery is,

$$I_{dis} = \frac{165.50}{3.7} = 44.73 \text{ mA}$$

Hence, the lifetime of 3.7 V 6000 mAh battery is found as:

$$B_{Life} = \frac{6000}{44.73} = 134.13 \text{ hrs}$$

If the battery is 100% discharged, the 3.7 V 6000 mAh battery can sustain up to 134 hours (Approx. 5 and a half days).

The tipping bucket rain gauge will only draw a very minimal amount of current on days that it is raining, so it was not necessary to include within calculations. The system will survive between 5 to 6 days maximum without solar power input. The solar panels utilized are 6V, 6W panels, providing enough power to sustain the system provided there is sufficient sunlight.

4.1.5 Developed Nodes

Two nodes were constructed for testing and deployment, as seen in Figure 4.6. These were developed as point-source monitoring nodes with initial deployment planned to be within a low-vegetation area at Macquarie University. Therefore, it was decided the node could be developed without the need for significant distance between the sensors and the electrical junction box which would be placed directly next to the soil sensors. This is not ideal for large-scale application as the LoRaWAN node should have a direct line-of-sight to the gateway, though there is not a significant distance between the gateway at Macquarie so this was not considered a critical component.

The nodes consisted of a soil sensor array attached to a hollow PVC piping, IP68 weatherproof Electrical junction enclosure, solar panel attached to the roof of the enclosure, an antenna for LoRa communication and each has their own rain gauge; one tipping bucket and one ultrasonic gauge. Cabling was protected with plastic conduit as well as all cabling and antenna connections to the junction box being sealed with silicon epoxy to prevent rainwater / moisture from entering and interfering with electronics.

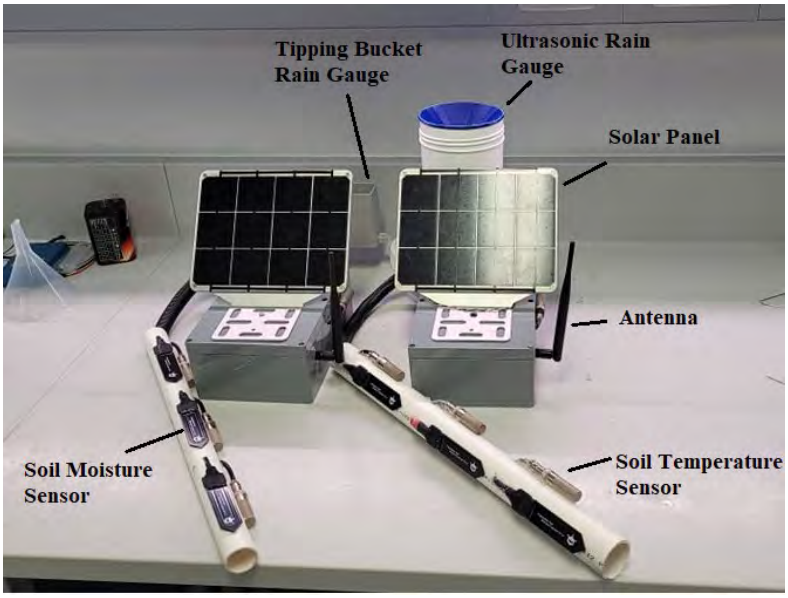


Figure 4.6: Developed prototype nodes.

4.1.6 Study Location

As part of the collaborative effort between the sub-projects, a testing site was offer to Macquarie for use by USYD for initial scoping of deployment requirements and future use for deployment of sensor nodes. The location was L'lara farm, which is part of a large commercial farm partially owned by USYD and utilised for agricultural research purposes.

Figure 4.7: Field testing site at L'lara Farm, Narrabri, NSW.

There are currently over 32 professionally installed soil moisture sensors installed on the property, which operate utilizing a public LoRaWAN network. These sensors required a surrounding caging to prevent livestock from damaging the sensors, as seen in Figure 4.7. These encaged locations alongside usage of the LoRaWAN network were offered to Macquarie for future deployments. This is ideal as it is as close as possible to a real deployment area as possible whilst already having a network to make the data transmission and storage a much simpler process.

One of the developed test nodes was deployed alongside previously implemented sensors. This particular location also held a large rain gauge and it was deemed appropriate to deploy a node alongside it to compare and contrast collected data. Therefore, the tipping-bucket rain gauge was installed inside the gating, as seen in Figure 4.8.

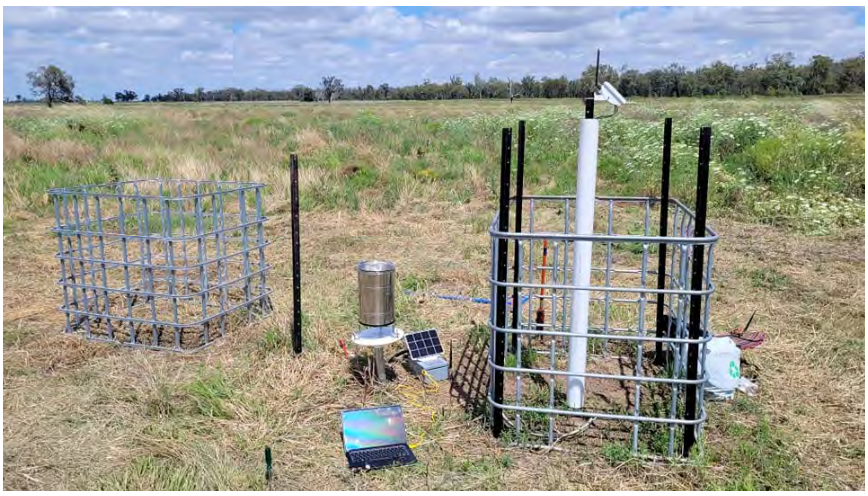


Figure 4.8: Installation site of node.

This location was the only spot with minimal surrounding vegetation, allowing for the node the connect to the LoRaWAN network without being deployed on a pole / at significant height. It was found that vegetation surrounding the node would have a significant impact upon the Signal-to-Noise Ratio (SNR). To combat this, the spreading factor was increase to compensate for the reduced line-of-site to the gateway. The spreading factor controls the 'chirp rate', which in term controls the speed of the data transmission. A lower spreading factor means faster chirps and therefore a higher data transmission rate, though it also means a reduction in the range of LoRa transmissions. Therefore, the spreading factor was increased to the maximum value on the day of installation resulting in larger battery consumption then the system was initially designed for.

The soil sensor array was installed utilizing a 50mm soil sampler probe to dig a test pit next to the pre-installed rain gauge. The arrangement of the sensor array results in a significant amount of space between the capacitive moisture sensors and the surrounding soil. A mixture of soil, sand and water was created within a container and poured to fill the gaps between the sensors and soil. It was then expected to take a period of time for the soil to settle and dry. Heavy rain was expected for the following day, which would result in the soil wetting, compacting and eventually drying around the sensors.

4.1.7 Initial results

Initial data was collected from the installed node at L'lara farm and stored on a ThingSpeak server. The stored data can be seen in Figure 4.9, with the initial data being collected over a 3-day period. The node stopped transmitting data approximately 3.5 days after being installed. As the period of fieldwork for installation was for one day, on the 18th of November, I was not able to stay around to trouble-shoot any issues and problems that would occur after my visit. This was the major downside of deploying the initial nodes at the L'lara location, particularly before being able to undertake significant testing at a closer location such as Macquarie.

From the data that was collected, it could be seen the temperature sensors functioned well, with top soil temperature showing larger fluxes as compared to the deeper soil sensors. The temperature data values were shortened from float (decimal points) to integer values to lower the transmission payload size in an attempt to compensate for the increased spreading factor required. The decimal accuracy was not a large concern for the initial node as the main objective was to deploy and assess the performance of the system.

The soil moisture data saw a gradual increase in the moisture readings, this can be assumed to be due to an adjustment period as the soil mixture poured into the pit was settling in place and contacting the sensor. The sensor array maintains soil depths of 15cm, 30cm, 45cm for top, middle and bottom respectively.

When the rainfall event began on 21/11/21, the middle and bottom moisture sensors were able to pick up and record the event showing an increase in moisture, whereas the top soil sensor did not. This could be due to the soil-sand mixture moving downwards as it wet and losing direct contact with the top capacitive sensor. A more detailed and confirmable explanation can be obtained when the node is visually inspected during the next visit.

The node stopped transmission late on 21/11/21, this most likely being due to a battery and lack of solar availability issue or potentially water damage from the rainfall event. The erratic readings from the soil moisture sensors shortly before transmission ending may be due to either moisture effecting the on-board sensor electronics or low-voltage complications from the discharged battery.

4.1.7.1 Discussion

Throughout this report, a low-cost sensing system for monitoring of soil moisture and temperature, rainfall and ambient conditions has been explained in detail. Macquarie university had the deliverable of developing prototype nodes for utilization of point-source measurements to increase the temporal and spatial resolution of data from weather stations that are currently available.

The development of the nodes was delayed due to the COVID-19 lockdown, though an initial test deployment and further scoping of system requirements was able to take place at Narrabri. The systems were developed to be initially tested at Macquarie, though one was deployed at Narrabri to achieve more realistic testing conditions and allow for our workstream to identify areas of concern with the node design and make any necessary improvements.

More time is required to validate the testing of the nodes, though with minimal adjustments and a few system improvements it is clear these nodes can be deployed at a location of interest to gather data that is of high spatial and temporal resolution.

The use of low-cost sensors which are developed at Macquarie University provides numerous advantages, including that prototyping can be readily and rapidly achieved to develop sensors nodes that are fit for purpose and meet the unique research and agency requirements.

The developed nodes can be easily adapted for use alongside different measurement platforms and made compatible for any required communications protocols to allow for the systems to be successfully deployed at any catchment or area of interest throughout NSW. Differing sensors of different shapes and sizes for different physical parameters can be added or swapped in and out of the nodes with ease. This allows for the research and developmental work to be decoupled and de-risked from reliance on a commercial partner.

Figure 4.9: Deployed sensor node data stored on a designated ThingSpeak channel.

4.1.8 Future Work

In terms of future work to be completed, this is broken down into three categories; Immediate term, medium-term and long-term goals.

In the immediate term, the next step is to develop many more of the prototype nodes and deploy them at a site of interest. This first involves making the necessary design improvements noted from the Narrabri deployment and then manufacturing many more prototype nodes for deployment. This includes further ruggedisation if necessary, such as further epoxy putting and replacing the plastic conduit with stainless steel conduit. This will come as a result of analysing the node once we are able to retrieve it from Narrabri. A sensor node deployment risk matrix (Appendix 1) was created for reference and

The testing site offered by USYD – Narrabri / L'lara farm would be the most likely location, though possibly another site in collaboration with the deployment of ANU's gravity sensors could see fit. This would provide an outcome of achieving 20 – 30 deployable sensor nodes monitoring soil moisture and rainfall to collect data and reduce uncertainties. Another area of interest is investigating water percolating through soil not only straight downwards in a vertical direction, but laterally through interpretation of soil moisture data between nodes. There are many analytical relationships that can be investigated from the data collected from the deployment of a swarm of closely deployed nodes.

In the short to medium term, further research and analysis regarding the inclusion of additional sensors would be the next goal. As we are currently monitoring rainfall and soil moisture quantity, the inclusion of quality sensors would be appropriate. For example, measuring rainfall quality (hardness and pH etc.) and/or soil quality which could then be analysed to potentially provide a relationship to the moisture content. At Macquarie university we specialise in developing low-cost sensors, and the inclusion of these could provide a significant increase in the amount of information each of these nodes provide.

Once many sensor nodes have been deployed and are confidently operating and collecting data in an accurate and useful measure. Deploying these nodes at strategic locations within a large-scale testing area and using them as calibration points for large-scale remote sensing is the long-term outcome. These nodes could be easily deployed as calibration points for sensing parameters such as rainfall, temperature and soil moisture (e.g. using mobile cell-tower signals to determine the amount of rainfall in an area of interest).

Risk Matrix for low-cost sensor deployment

Table 4.5: Risk Matrix for low-cost sensor deployment.

	Key Risks	**Mitigation Measures**	**Key questions to determine mitigation effectiveness**
Flood	Flood-induced high water levels, causing electrical equipment to be submerged for significant periods.	Ensure structure is tall enough for safe clearance under foreseeable flood levels. Ensuring the structure is stable and will not drift away during floods (e.g. Anchor Bolting Structure). Sensors should be waterproof and electrical junction box silicone/epoxy sealed to provide weatherproofing.	Where can the system be deployed within a flood-prone area to have sufficient clearance to safely accommodate a 1:200 year flood without submerging components?
Wind Storm	Vegetation/tree branches could fall and damage system.	Deploy structure in an area of clear vegetation. Ensure structure is stable and can withstand strong winds. (Particularly important for rainfall sensing).	Is the structure deployed close-by to tree branches that may cause damage? Can the system maintain stability during strong winds?
Fire	Bush fires can cause immediate damage if system catches on fire. Aftermath impacts including ash covering and damaging sensors.	Installation of system should occur in a location with minimal surrounding vegetation	Is the system deployed in a wild-fire possible area? If so, how close to vulnerable vegetation?

Condensation	Electrical enclosures exposed to moisture from condensation due to changing environmental temperatures.	Ensuring main electronic interfacing components are in waterproof enclosures and potentially the additional use of silicone/epoxy putting to ensure weatherproofing	Are the electronics / enclosure designed to prevent & protect from water/moisture damage?
Livestock	Livestock chewing through cabling / damaging system	Run cabling through PVC conduit. Potentially run cabling from sensors to junction box underground.	Are livestock able to chew through protective conduit? What do currently deployed systems use to minimize this risk?
Vandalism	Damage/theft of systems due to vandalism	Signage, lock electrical junction box, blend into vegetation	Is the system deployed in an area of known risk of theft? Where are current weather stations and sensors deployed and how do they prevent vandalism?

4.2 Local gravity sensing of creeks and rivers

Samuel Legge, John Close

ANU, Canberra

4.2.1 Gravity sensing overview

The concept of using gravity as a signal to detect water is simple. Changes in mass due to water movement will results in a change to the local gravitational field. By monitoring the local gravitational field, it is therefore possible to gain information about the general location and amount of water beneath the surface. This overview section will give a generalised background on gravity signals and sensing.

The primary goal of this investigation is to assess the feasibility of using current and next generation gravity sensors as a method of detecting and managing water resources. To achieve this the gravitational signal of groundwater in a real-world system was simulated. The sections following this one will detail the construction of a toy world based off real world data and the resulting gravitation signals from water variations within this toy world.

4.2.1.1 Density variations

When considering the expected gravitational signal due to water it is easiest to think about it in terms of the density variations water causes. Liquid water has a typical density of d_{water}=1000 $kg.m^{-3}$ and has a strong differential gravity signal to the typical atmosphere with a density of ~1 $kg.m^{-3}$, water vapor is neutrally buoyant and has no differential gravitational signal to the atmosphere.

Water that flows or rests above the ground displaces the atmosphere and fully fills that volume. Water under the ground exists within the voids between the porous earth. The ratio of liquid water to solid ground is variable depending on the ground composition and is known as the void ratio α. When liquid water flows into the ground its effective density is reduced by this ratio. For example, sandy soil that has a void ratio of α_{sand}=0.3 has 30% of its volume as space that can be saturated by water. When this occurs, the effective density of the water that is saturating the sand is given as $d_{eff} = d_{water} \times \alpha_{sand} = 300kg.m^{-3}$.

This is the maximum amount of water that can exist in this ground type. As water travels through the ground there is a typical retention ratio r that through surface tension will hold onto the water that travels through it. This mean that should the saturated ground water drain away from an area due to gravity, it will retain some moisture, for our sand example this value is typically r_{sand}=0.5 such that after draining it will still hold a water at a density given by $d_{ret} = d_{eff} \times r_{sand} = 150kg.m^{-3}$.

This retained water will not drain from the sand under the influence of gravity however it can still be removed, typically through evapotranspiration at the ground surface or from plant roots extracting the moisture left in the ground.

4.2.1.2 Differential measurements

Using gravity to measure water cannot easily be done in a single direct measurement as any such measurement will contain the strong background gravitational field of earth, as well as local scale gravitational signals such as dense mineral deposits and topographic variations.

By taking two gravitational measurements at differing times, the difference between the measurements can be taken to give the time differential signal. It can be assumed that the background field is typically constant and hence is removed from the differential measurement. The remaining signal will show a gravitational signal due to any density changes that have occurred between surveys. It can be assumed that this is primarily due to water movement as this typically occurs on a much faster timescale than geological variations. However, it should be noted that any faster geological events such as landslides, earthquakes, significant erosion, or human interference such as mining or quarries may generate measurable differential gravity signals. It is expected that such changes will be able to be accounted for or small compared to groundwater signals.

4.2.1.3 Gravity and gravity gradients

When preforming a gravity survey, the typical expectation is that the senser will measuring the acceleration due to gravity towards the centre of earth mass, there are however other potential measurements that may be of use such as the lateral gravitational acceleration or the gradient of the gravity signal. Gravity is a vector filed and as such has a component in the three cartesian directions shown in Figure 4.10 and represented as $g\_$ where the subscript indicates the direction. This leads to nine components of the gravity gradient field as there is a gradient to each gravity component in each cartesian direction. In Figure 4.10 the gradients are labelled as $G\_\_$ where the first subscript is the cartesian direction of the gradient and the second is the gravitational filed component of that gradient. For example, the G_{zx} gradient is the rate of change of the g_x gravity vector in the Z direction.

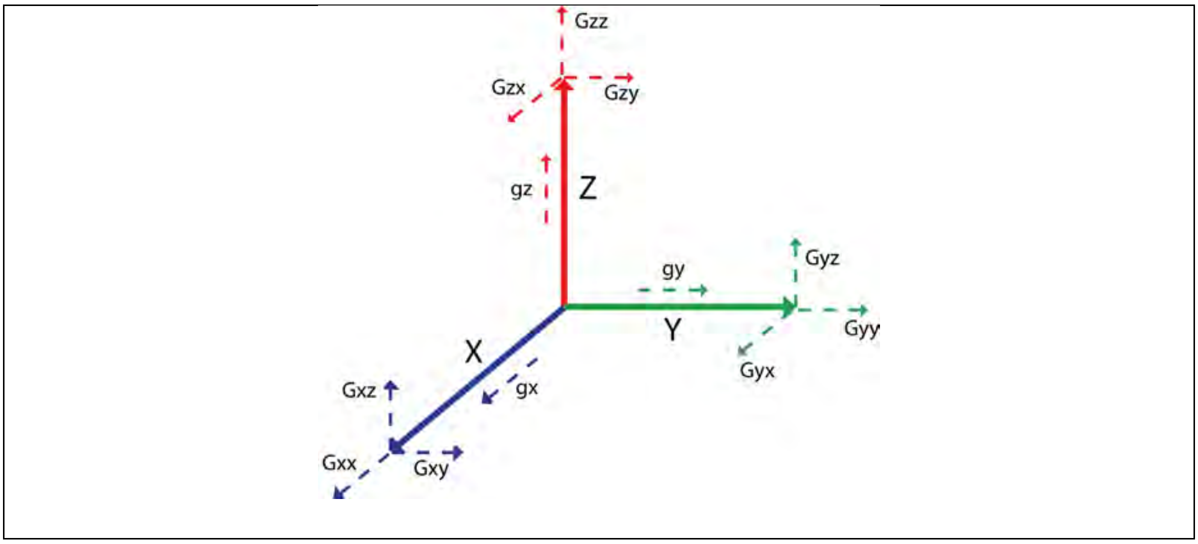


Figure 4.10: Diagram of the gravitational vector field in g_x g_y g_z and the corresponding nine gradient components. Reference https://en.wikipedia.org/wiki/Gravity_gradiometry

4.2.1.4 Typical gravity signals

To understand measured differential gravity signals it is important to consider what various shapes of density variations will look like. Here we will investigate the typical gravitational signals for a spherical mass source and a flat disk mass source, knowledge of these two basic structures will allow simple qualitative interpretation of the simulated gravity measurements.

Gausses law dictates that there is no difference in gravitational signal between a point source or a spherical source of mass, assuming that gravity is measured outside of the mass sphere. The gravitational signal from such sources can be easily modelled using newtons law of universal gravitation for both the gravitational g_z and gravity gradient G_{zz} signals. For all simulations in this section the mass is considered to be centred at the origin and the height is taken above the mass in the Z direction. The signals for a point mass are given by the following equations.

$$g_z = \frac{Gm}{r^2}$$

$$G_{zz} = \frac{2Gm}{r^3}$$

Where *G* is the gravitational constant, *m* is the mass and *r* is the distance from the centre of mass. When looking at the signals generated by such sources on a logarithmic scale as shown in Figure 4.11, we can see the clear inverse square relationship of the gravitational signal and the inverse cube relationship of the gradient.

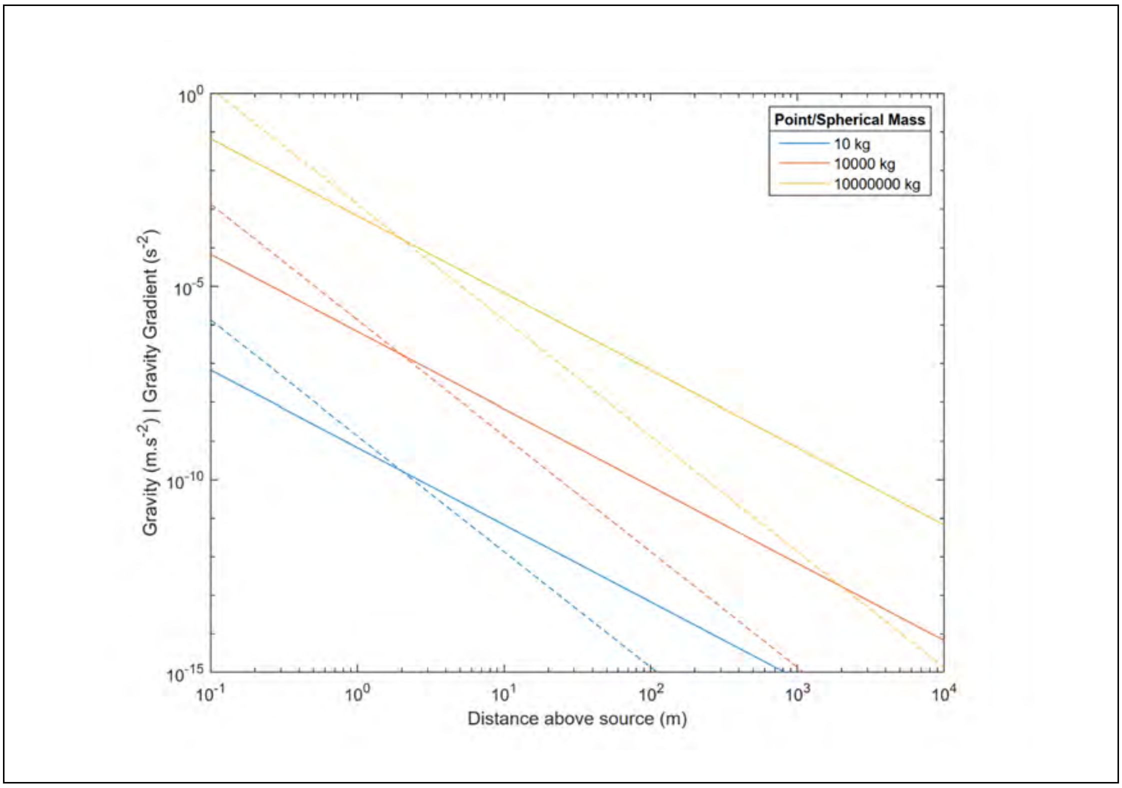


Figure 4.11: Gravitational signal (solid line) and gravitational gradient signal (dashed line) of three point/sphere mass sources.

In the real world we do not always expect approximately spherical sources, in fact, we know that water typically spreads out and fills large areas, resulting in sheets of mass that have a different gravitational signature. The gravitational signal given by an infinite plane of homogeneous density is constant and hence has a gravitational gradient of zero. The signals for an infinite plane are given by the following equations.

$$g_z = 2\pi G\rho t$$

$$G_{zz} = 0$$

Where ρ is the plane density and t is the thickness of the plane. By comparison the gravitational signal of a finite disk has two regimes. The first, when close to the surface relative to the disk radius, acts like an infinite plane and the second, where the distance increases to greater than the disk radius approximates a point mass source. The signals for a finite disk are given by the following equations.

$$g_z = 2\pi G\rho t\left(1 - \frac{r}{\sqrt{R^2 + r^2}}\right)$$

$$G_{zz} = \frac{2\pi G\rho t R^2}{(R^2 + r^2)^{\frac{3}{2}}}$$

Where R is the disk radius. Examples of this are given in Figure 4.12 and Figure 4.13.

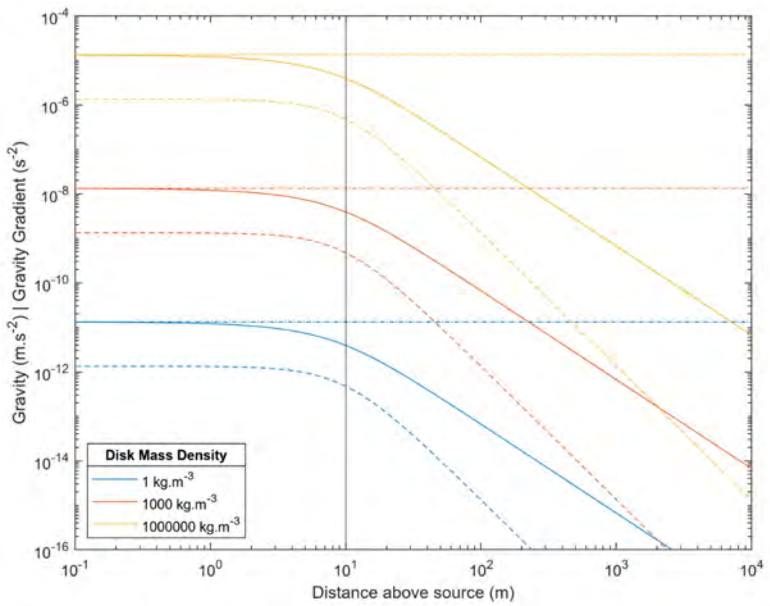


Figure 4.12: Gravitational signal (solid line) and gravitational gradient signal (dashed line) of three finite disk mass sources with varying density, a set thickness of 31.8 mm and a radius of 10 m. The flat line (dash dot) shows the equivalent gravitational signal for an infinite disk of this thickness and density, the infinite disk has no gravitational gradient signal.

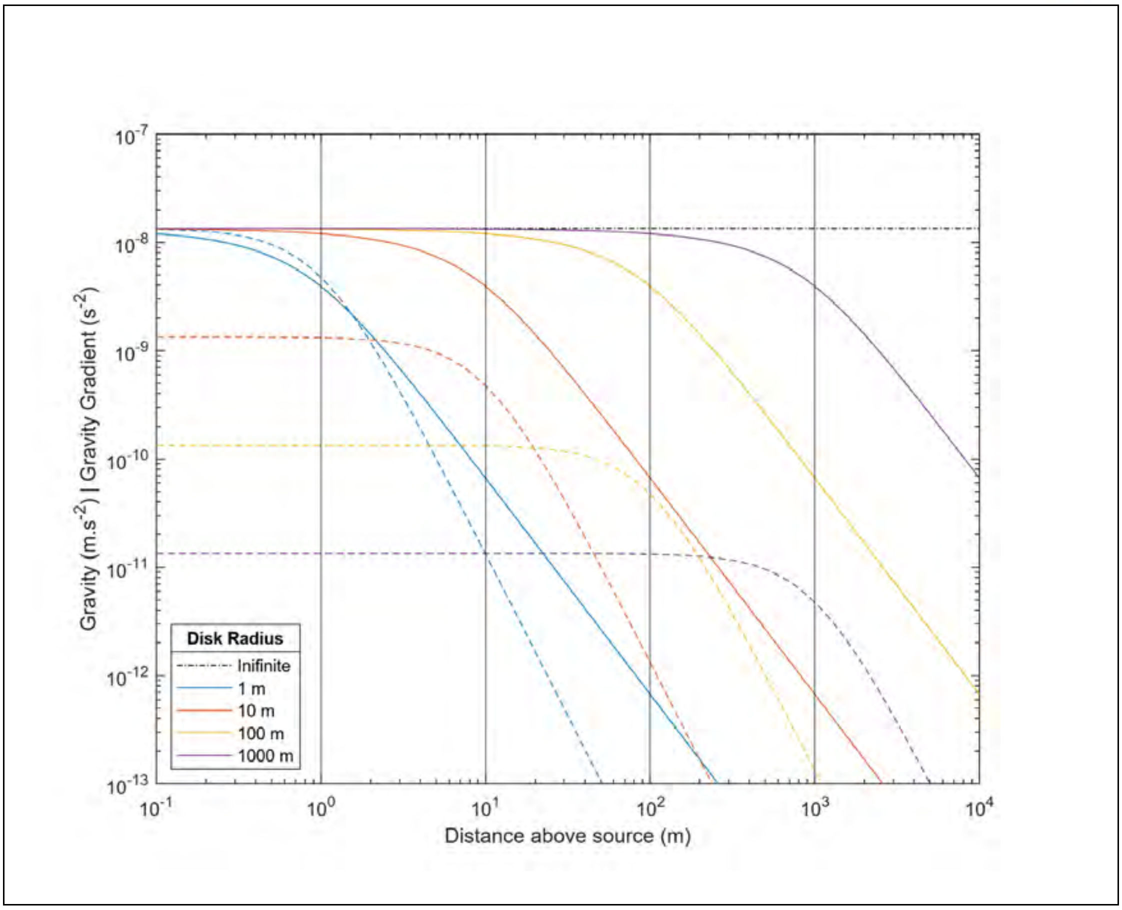


Figure 4.13: Gravitational signal (solid line) and gravitational gradient signal (dashed line) of four finite disk mass sources with varying radius, a set density of 1000 $kg.m^{-3}$ and a set thickness of 31.8 mm. The flat line (dash dot) shows the equivalent gravitational signal for an infinite disk of this thickness and density, the infinite disk has no gravitational gradient signal.

These graphs give an indication of the relative differences that various geometries give to both gravity and gravitational gradient signals. Spherical or point sources fall off in signal strength with one on the distance squared while gradients fall off at one on the distance cubed. Large flat densities of mass however appear as static fields with lower gradient signals until the sensor is positioned further than one radius away from the mass disk. Figure 4.13 shows that the gravitational signal of real-world groundwater systems are likely to be flat when surveys are done close to the mass source relative to its size. Additionally, it clearly shows that while gradients are good at measuring small mass distributions that are more closely approximated as a sphere, the gradient signal quickly approaches zero for larger more homogeneous mass disks.

Actual systems are made up of a continuous variation of density with varying length scales and strengths. This results in signals like the more complicated gravitational fields simulated in the following section. However, it is still useful to keep these two simple cases in mind as they allow some assessment of the observed signals. For example, the dams show up strongly in the gradient signal as they are more point source like and generate a strong gravitational gradient compared to the streams that have a more constant and wide mass distribution similar to the disk source. By comparison the streams show a strong gravitational signal at all survey heights while the dams fall of in signal far more quickly.

4.2.2 Building the toy world

To simulate what a real-world groundwater gravitational signal would look like, a three-dimensional toy world of mass distributions was built based on the area around middle creek farm shown in Figure 4.14.

Figure 4.14: Aerial view of Middle Creek simulation area (5 km x 5km white outline). Borehole locations marked with green pins. Image taken from Google earth, 2021 Maxar Technologies, Latitude: -30.48856, Longitude: 150.122738, Date: 2021-07-02.

This area was chosen is it contains a number of borehole survey sights and has been well studied by UNSW hydrologists allowing us to make educated guesses on the likely distribution of groundwater in this system. The surface of the simulated world was informed using topographic data taken from the NSW Foundation Spatial Data Framework - Elevation and Depth - Digital Elevation Model (DEM). The voxel grid is made up of 1001(5000 m) by 1001(5000 m) by 75(74 m) voxels(distance) in the North, East and Vertical directions respectively. Figure 4.15 shows this world from the surface level followed by increasingly striped back voxel layers to the base groundwater level. The following subsections will detail the construction of this model.

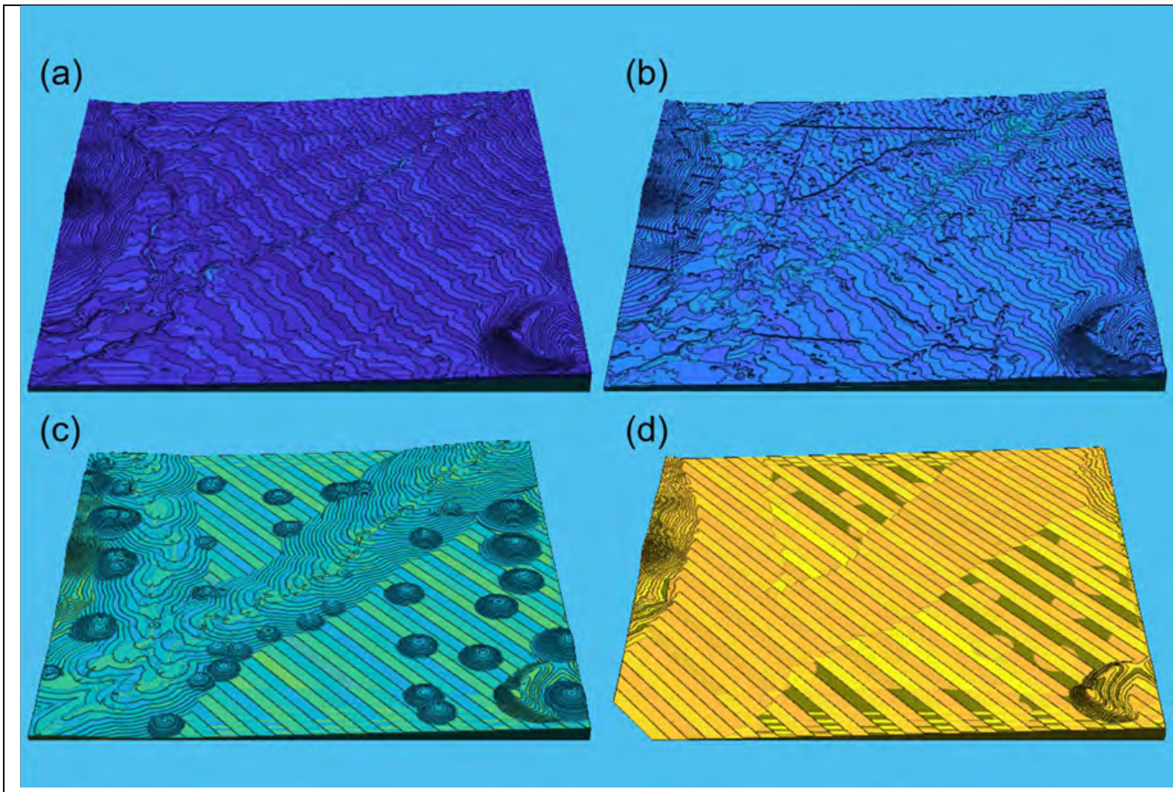


Figure 4.15: Model world based on Middle Creek area. The world is made up of 1001(5000 m) by 1001(5001 m) by 75(74 m) voxels(distance) in the North, East and Vertical directions respectively. The vertical axis is exaggerated by 10 in these figures for clarity, lines indicate 1m height contours. Section (a) shows ground level of the toy world where the topographic data was taken from the NSW Foundation Spatial Data Framework - Elevation and Depth - Digital Elevation Model. Section (b) shows the toy world with evapotranspiration areas made transparent these areas were generated based on the satellite imagery shown in Figure 4.14. Section (c) strips away the ground except for the streams and dam groundwater recharge areas and section (d) shows the baseline water level taken from the borehole groundwater level in drought years and limiting the depth below the surface to 20 m. This pulls up the effective base level water table in the higher areas.

4.2.2.1 Tree coverage and evapotranspiration

The effects of evapotranspiration are often brought up as one of the hardest to measure mechanisms for water loss in a system. To investigate the feasibility of using gravimetry to measure this loss of water a simple evapotranspiration mask was generated for the simulation from the aerial photography. This was achieved by colour filtering the aerial image to find the dark green pixels typical of treed areas and then spatially filtering the resultant image to smooth the results to a small radius around each treed area. This mask is shown overlayed on the aerial imagery in Figure 4.16.

Figure 4.16: Middle creek farm simulation showing approximate areas of dense tree coverage (white) and hence areas of greater evapotranspiration.

Within the simulation the mask was used to define areas of reduced mass down to 5 m below the surface. This was done to simulate the effects of trees and roots removing retained water from the soil. The result of this in the toy model can be seen in Figure 4.15(b).

4.2.2.2 Stream and dams

One of the identified contributors to groundwater recharge is from water streams. To add this effect to our model a simple algorithm to find surface waterflow was developed that is robust to the filtered NSW DEM data based on a D8 type method. Using this algorithm, the locations of Middle creek and Horesarm creek were identified in the simulation as shown in Figure 4.17. The dams in the simulation were identified by visual inspection of the aerial photography and their locations record and converted into the simulation coordinates.

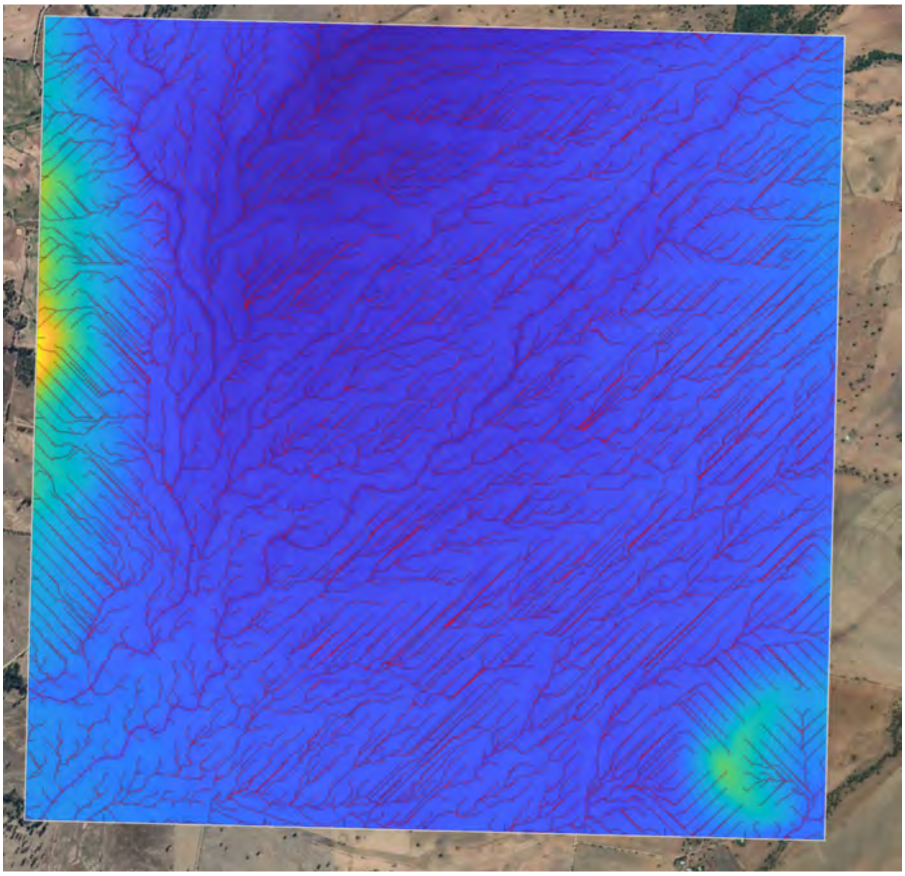


Figure 4.17: Water stream flows (red) from a grid start point over the Middle Creek farm simulation.

From each surface location of the streams and for each dam the toy world was altered to gently raise the water table up to this point. These new voxels were tagged in the simulation so they could be altered separately between dams and streams. The result of this can be seen in Figure 4.15(c) showing a raised water table along the streams and at each dam in the toy world.

4.2.2.3 Baseline Groundwater

The final layer of the model consists of the baseline groundwater level. This was generated by fitting a surface plane to the lowest borehole water table heights recorded at the peak of the previous drought. Due to a lack of boreholes across the full model area some approximations were made to keep the water table at a sensible depth. Additionally, a maximum water table depth from the ground surface was set at 20m resulting in a raised water table under some of the hills withing the model. This baseline level is shown in Figure 4.15(d).

4.2.2.4 Rainfall groundwater

An additional layer of simulation not shown in Figure 4.15 was to add an even distribution of mass to the top 5m below the surface to simulate an even dampening of the soil due to rainfall events. This layer of the model overlaps with both the evapotranspiration and the dam and streams groundwater voxels.

4.2.3 Gravitational simulation of toy world

Gravity surveys of the modelled toy world are preformed using parallel computing and approximating each voxel as a point source and summing across all sources. The gravity signal for a point $\vec{r}$ for a mass m at point $\overrightarrow{r'}$ is given by the following equation.

$$g(\vec{r}) = \frac{Gm\vec{r}}{\left|\vec{r} - \overrightarrow{r'}\right|^3}$$

Similarly for the gradient it is given by taking the derivative of the above equation for each cartesian direction.

The results of this simulation could be improved by altering the equation to be the true gravitational signal of a rectangular prism voxel however by using a sufficiently fine grid the signals for each voxel can be closely approximated to a point source and still return meaningful results. For now, the only result where this has significant impact is in the ground level gravitational gradient survey, where the simulated signals are higher than expected due to the proximity to the surface level voxels. Future work to the simulation is planned to implement the rectangular prism gravitational equations and correct this result. An additional limitation is that some of the simulated fields show anomalous effects at the edge of the simulation due to the finite simulated world that would not exist in a real survey.

One advantage of the way gravitational fields sum and scale is that for each gravitational survey, the results could be easily separated into the gravitational contribution of each modelled effect in the previous section. Additionally, the total amount of water for each of these effects could be checked to be reasonable and easily altered to show the signal some particular volume of water of that shape would have.

A number of surveys were simulated across the full toy world following the topography at ground level, 10m above ground level, 50m above ground level, 186m above ground level and 500m above sea level (average of 186 m above ground level without following topography). A program was written to view this data that could switch between all gravity fields and their gradients, add levels of simulated noise to each survey and calculate and adjust the total volume of water from each modelled effect.

4.2.3.1 Gravity survey results

As mentioned previously, the primary goal of this part of the projects was to assess the feasibility of using gravitational surveys to measure ground water. To do this the realistic toy model world was created with the intention of discovering the expected magnitude of gravitational signals and how good sensors will need to be to be able to detect these signals.

The results in Figure 4.18 show gravity in the g_z direction at various heights. At ground level, close to the source of the signal it is easy to differentiate between the various contributors, with the evapotranspiration areas showing a reduction in the gravity signal of $\sim 2 \times 10^{-7} m.s^{-2}$ and the dams and streams showing an increase in the gravitational signal of $\sim 12 \times 10^{-7} m.s^{-2}$. As the survey height is increased, simulating the use of a drone or similar, these signals begin to decrease gradually in strength and blur out in position. As expected, the smaller sources such as the isolated dam groundwater falls of in strength quickly as the survey altitude is raised while the larger stream recharge groundwater maintains its signals due to the large area it covers. This is consistent with the expected results shown in Figure 4.13. The result in Figure 4.18(d) shows that even with a greatly reduced water volume, a modest 1000 ML increase in groundwater recharge from the two streams would still be visible from a 50 m above ground level survey with a sensor that could measure down to $10^{-8} m.s^{-2}$.

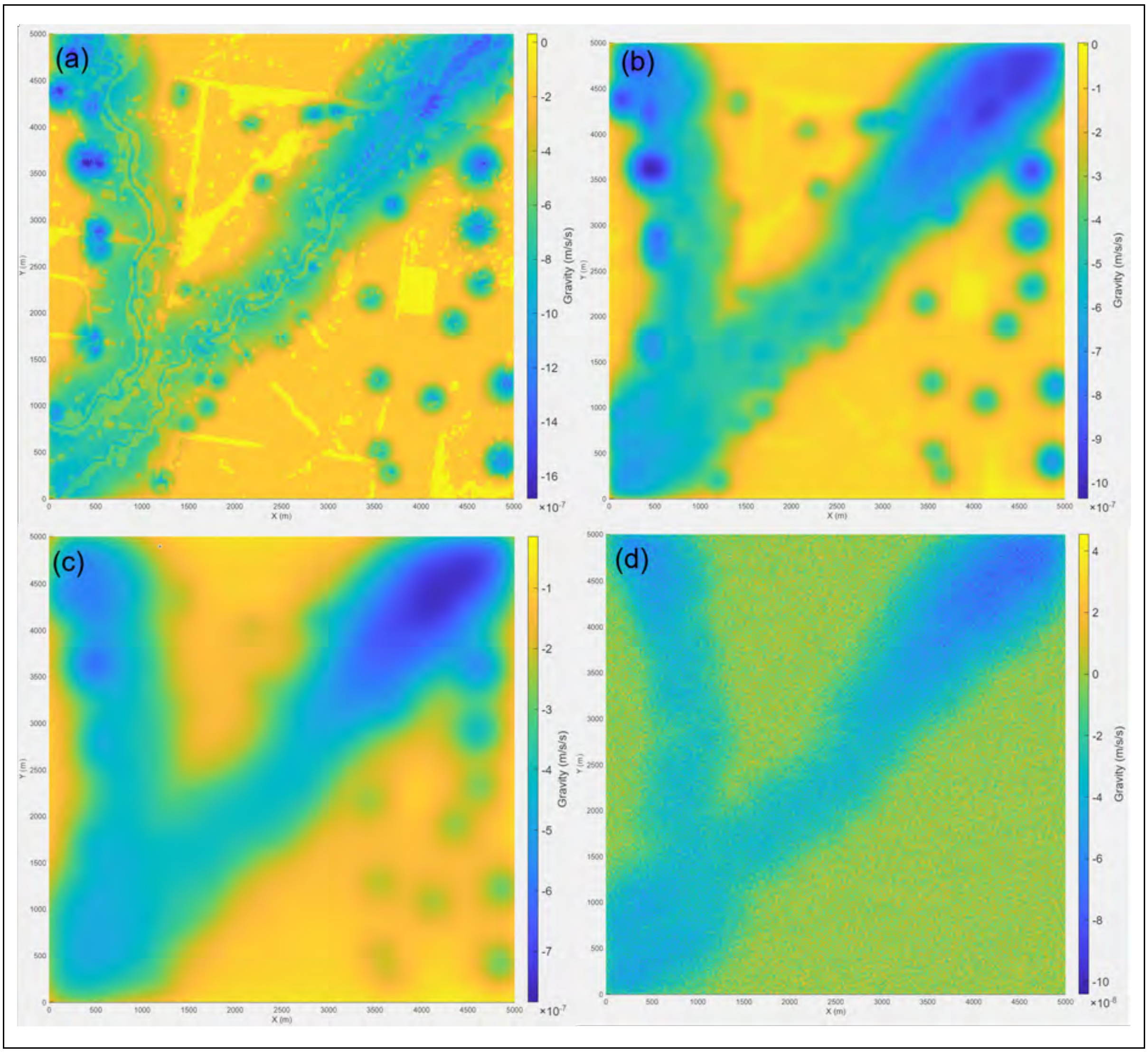


Figure 4.18: Gravity measured in the z (vertical) direction at a height above ground level of 0 m, 50 m and 186 m in (a), (b) and (c) respectively. In (a), (b) and (c) the system contains 21,481 ML of water 13,443 ML in the stream groundwater, 3,350 ML from dam groundwater, 6,263 ML from surface rainfall and -1,575 ML removed from evapotranspiration. (d) shows the simulation with only 1000 ML of water contained in the stream groundwater at a height of 50 m above ground level. All measurements include Gaussian noise with a standard deviation of 10^{-8} m/s/s.

This result is promising as portable sensors that can measure down to this level are already commercially available, meaning changes in gravity of this magnitude are feasibly measurable. However, currently all such available sensors are classical gravimeters that drift at rates of $\sim 2 \times 10^{-7} m.s^{-2}$ each day and as a result, the sensor may drift over the expected highest magnitude signals in as little as five days.

With the ongoing development of quantum gravimeters this drift parameter is expected to significantly improve as the sensors are locked to atomic references and will not require calibration.

4.2.3.2 Gravity gradient survey results

The vertical gravity gradient in the vertical direction G_{zz} is the next most utilised gravitational signal in existing surveys. Measuring gravitational gradients has benefits compared to gravity measurements as it removes any common vibrations or acceleration of the sensor. Because of this it is the typically preferred approach for most airborne surveys.

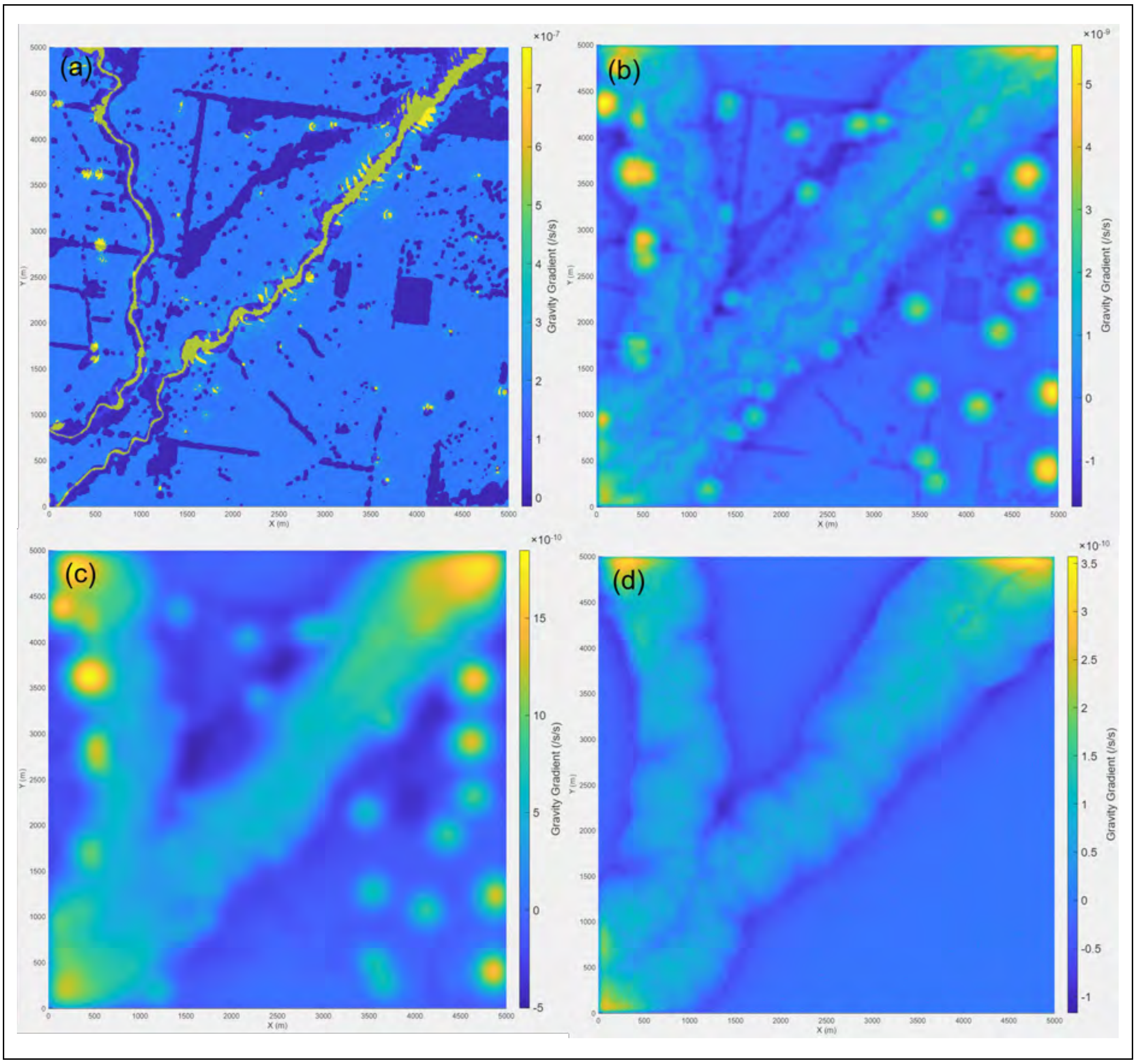


Figure 4.19: Gravity Gradient measured in the G_{zz} (vertical) direction at a height above ground level of 0 m, 50 m and 186 m in (a), (b) and (c) respectively. In (a), (b) and (c) the system contains 21,481 ML of water 13,443 ML in the stream groundwater, 3,350 ML from dam groundwater, 6,263 ML from surface rainfall and -1,575 ML removed from evapotranspiration. (d) shows the simulation with only 1000 ML of water contained in the streams at a height of 50 m above ground level. All measurements include Gaussian noise with a standard deviation of 10^{-11} /s/s.

The results shown in Figure 4.19 match the survey scenarios in the previous gravity measurement but show some expended different behaviour. Firstly, Figure 4.19(a) has the most unreliable data from the simulations and shows signals of a far greater magnitude that would be expected from the discussion and results in Figure 4.13 due to the point source approximation used. The results from the higher altitude surveys are sufficiently far from the computational grid and should be more reliable.

One of the primary points of difference of the gradient signal is that it is better at resolving smaller sources at higher altitudes with more sharply defined edges on dams and streams. It is also more sensitive to point sources and suffers from a lower signal over large flat surfaces like the streams.

4.2.3.3 Optimal quantum sensor and survey design

The simulation results from this study give minimum sensitivities required through the expected signal strengths, $1 \times 10^{-7}\,\mathrm{m}.s^{-2}$ for gravity and $1 \times 10^{-9}\mathrm{s}^{-2}$ for gradients. However further work is required on the inversion front to investigate the optimal measurement configuration and grid. Qualitatively the simulation shows that a large amount of data can be generated from a simple vertical gravimeter (the most likely first iteration of these devices) and as such is a good candidate for field trials.

Additional thought will be required for handling position and altitude corrections prior to field surveys. The gradient of earths field results in changes in gravity of $3 \times 10^{-6} m.\, s^{-2}$ over 1 meter of altitude difference. This gives a minimum required altitude precision of ~0.1 m for $3 \times 10^{-7} m.\, s^{-2}$ stability. Systems such as differential GPS may be required for reliable position and particularly altitude data, other alternatives include lidar positioning or local positioning systems.

4.2.3.4 Future works

A number of improvements to the simulation code need to be made, especially to clarify the results of gravity gradient signals.

Field trials and validation as discussed above.

4.3 Satellite gravity measurements of Australian water

Paul Tregoning
ANU, Canberra

4.3.1 Introduction

Space gravity observations from NASA's Gravity Recovery and Climate Experiment (GRACE) and the GRACE Follow-On (GRACE-FO) missions provide unique estimates of changes in total water storage on spatial scales from global/continental to basin/sub-catchment [14].

The distance between two satellites orbiting Earth is sensitive to changes in the strength of the gravity field, changes that are caused by adding/removing mass from locations on Earth, including water, ice, earthquake deformation and mantle convection. In Australia, changes in mass are dominated by hydrological processes. Thus, the changes in water resources in Australia can be estimated from the inter-satellite measurements of these space gravity missions. Although providing unique information, the estimates of changes in total water storage are of a spatial scale of hundreds of kilometres, which limits the immediate utility of the estimates of changes in total water for many specific hydrology applications.

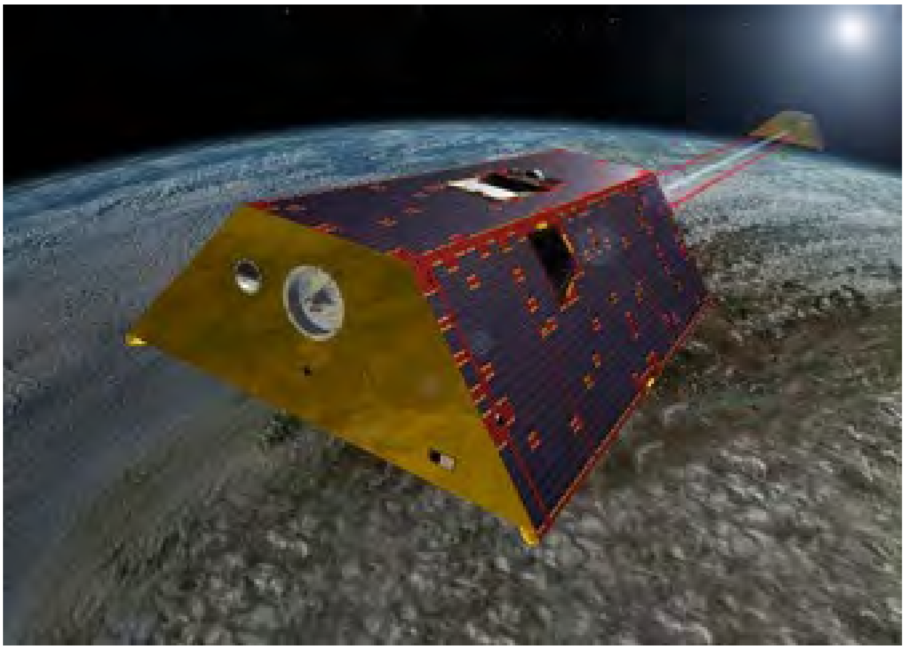


Figure 4.2320: Artist's impression of the GRACE Follow-On mission satellites (source: https://gracefo.jpl.nasa.gov/mission/overview/)

In this sub-project, an assessment was made of what space gravity observations can contribute to the understanding of managing water resources. The study included assessments of the ability to resolve a simulated change in water volume from space gravity observations, including the broad-scale mapping of how the significant rain events in 2019 and 2020 propagated across NSW. In addition, an assessment was made of whether it is possible to track floodwaters from southern Queensland as they flowed into NSW after the March 2020 flooding events.

4.3.2 Spatial Resolution

The inherent spatial resolution of estimates of changes in total water storage (TWS) from GRACE and GRACE-FO mission data is ~300-400 km [15] [16]. This is essentially governed by the fact that the satellites are typically orbiting Earth at an altitude (i.e. height above the surface of Earth) of 400-490 km. Reducing the altitude of the satellites in their orbit would increase the spatial resolution of the results, but also increases the atmospheric drag experienced by the satellites and, therefore, reduces the lifetime of the missions. Thus, users of GRACE/GRACE-FO data are likely limited to this level of achievable spatial resolution.

Nonetheless, there are many estimates of TWS that are provided at higher spatial resolution. Time varying gravity fields, in the form of spherical harmonics, are provided by the mission to degree/order 96, which equates to ~180 km, but this is not really the achieved spatial resolution. Mass concentration elements (mascons) are also used to represent the temporal gravity field by estimating a change in thickness of a plate of water across a tile of known area on the surface of the Earth. Mascon solutions [17] have been produced at spatial resolutions of 3° (~330 km, [18] 1° (~110 km, [19]; [20]) and 200 km [21]. The latter three studies state explicitly that this is not the actual spatial resolution of the estimates.

When the spatial resolution of the parameters used to estimate the changes in mass is smaller than the inherent accuracy of the GRACE/GRACE-FO data, high correlations occur between parameters. This causes "leakage" of signal, or a smearing of the mass changes between neighbouring parameters, which reduces the inherent accuracy of the solutions. For example, a significant mass change in one location may be under-estimated, with the remainder of the mass change signal being incorrectly assigned to neighbouring mascons.

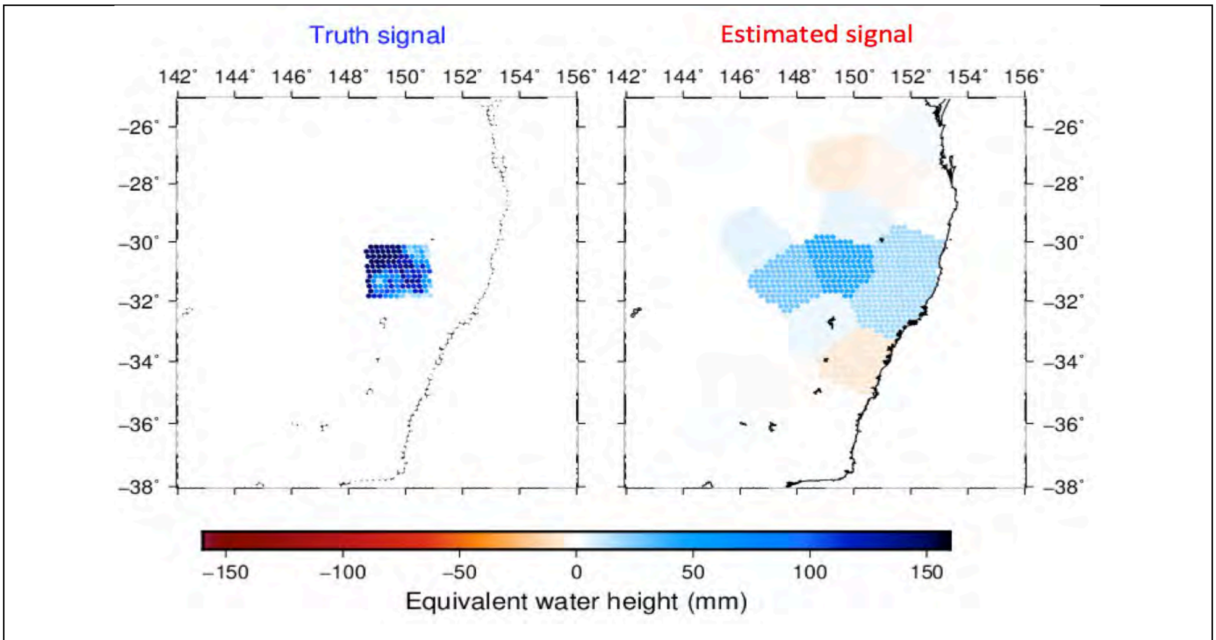


Figure 4.2421: Simulated signal in the Maules Creek region (left) and values estimated (right). The error in the estimated signals is caused by a) the mismatch in the location of the signals and the spatial pattern of the mascons used to estimate the signal, b) the correlations between the mascon parameters, and c) the fact that the signal magnitude changes on spatial scales smaller than the mascon parameters.

In an ideal case, the estimated signal in this simulation would match perfectly the actual signal that was used to create the truth orbits. This does not occur, for several reasons. First, the mascon estimates are made at a spatial scale of 200 km, which is below the inherent accuracy level of the GRACE/GRACE-FO data, meaning that there are high correlations between neighbouring mascon parameters. Second, the magnitude of the simulated signal varies at spatial scales of ~18 km, whereas the mascon estimates are really an average of such signals, made at the ~200 km scale. Third, the spatial pattern of the mascons used to estimate the signal doesn't match perfectly the spatial pattern of the simulated signal to be recovered.

The first issue can only be resolved by using larger mascons. However, doing so actually increases the importance of the second issue, since there will be greater variability of signal magnitude within larger mascons. The third issue cannot be resolved unless the spatial pattern of the mass changes is known in advance, which is unlikely to be the case.

4.3.2.1 Presence of Signal in the simulated observations

A recent study by Han et al. [22] identified an abrupt surge of water storage of around 60-70 trillion km3 during the March 2021 significant rain events. They assessed the inter-satellite measurements along the ground track of the flight of the satellites and used these measurements to quantify the changes in water volumes. This approach has the potential to provide estimates with a lower temporal latency and, possibly, higher spatial resolution.

To assess the viability of this new approach, the inter-satellite observations from the Maules Creek simulation were assessed along the ground tracks. The observations used to estimate the temporal gravity field shown in Figure 4.24 are the inter-satellite range acceleration (the double time derivative of the change in distance between the satellites – see Allgeyer et al. [21] Even though the water change signal on the ground is relatively small – in this case a maximum of 150 mm – the effects of it are present in the range acceleration observations. The magnitudes of the range acceleration observations are plotted on the ground tracks (the flight paths of the satellites over NSW) in Figure 4.25.

The spatial pattern of the simulated signal is not clearly evident. While large negative signals are present in the range acceleration observations when the satellites pass directly over the water volume changes (lower panel of Figure 4.25), significant positive signals are also present in advance of and after the satellites fly over the mass change locations. In addition, range acceleration observations on ground tracks that do not pass directly over the region of simulated mass change also contain non-zero signals. These relate to the volume changes that are physically located to the side of the ground tracks, not to changes directly below the satellites.

Therefore, the range acceleration observations along any ground track are the sum of water volume changes directly beneath the satellites plus changes within ~1000 km of the ground track of the satellite pass. The strong negative signal as the satellites pass over a location with a positive water volume change does dominate the observations, which may indicate some utility of this method for tracking water flows through the Australian landscape.

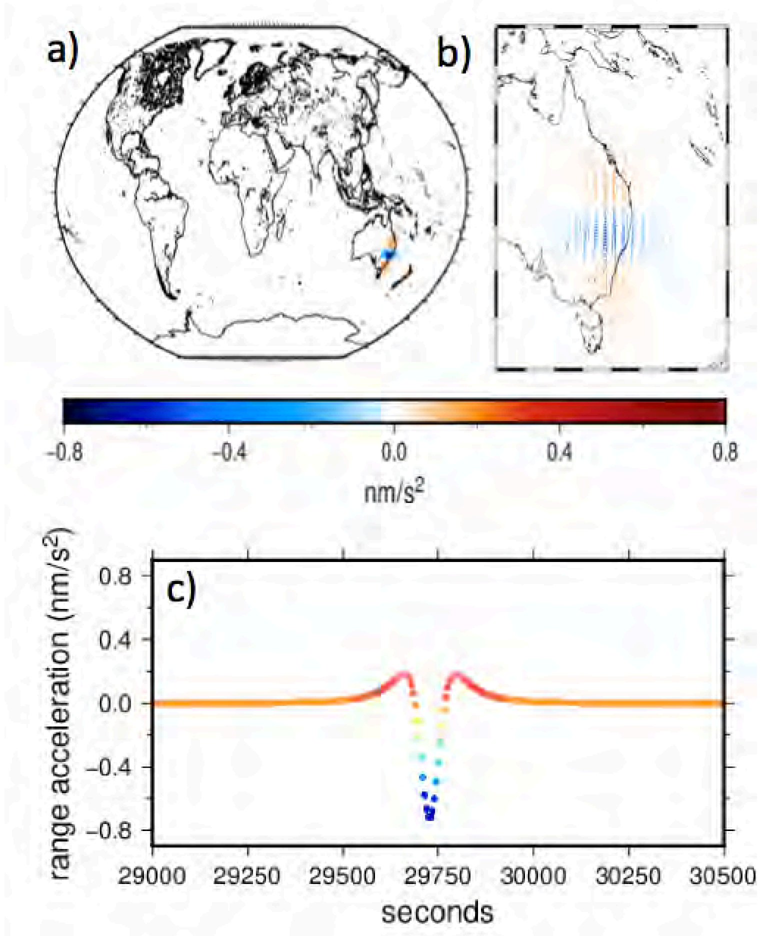


Figure 4.2522: Range acceleration residuals along the ground track of the satellites for the simulation of mass change in the Maules Creek region shown a) globally and b) over eastern Australia. Orbits were calculated using satellite locations in July 2016. c) time series of range acceleration residuals as the satellites passed over the simulated signals in the Maules Creek region. While there is a negative peak at the spatial location of the simulated mass change, there are also lobes of positive signal north and south of the mass change.

4.3.3 Analysis of GRACE/GRACE-FO data

To demonstrate the capability of space gravity data to estimate water signals in Australia, the GRACE-FO data were analysed for 2019 and 2020. The analysis was done using the ANU GRACE software, as described in Allgeyer et al. [21]. Mass concentration (mascon) tiles of ~40,000 km^2 (roughly 200 km x 200 km) were used, consistent with the geometry used in the Maules Creek simulations described above.

4.3.3.1 Spatial analysis

Using January 2019 as a reference month, changes in water (expressed as a thickness of equivalent water height) have been calculated across eastern Australia. January 2019 was a very dry period, being at the end of the ~3-year drought across eastern Australia; therefore, the changes in water seen across the landscape represent the replenishment of water resources.

The intense rainfall event that occurred around Townsville, 28 January to 8 February, is clearly evident and the progression of water southward into NSW is well captured. Note that the increases in NSW shown in Figure 4.26 will be the sum of both water flowing south from Queensland and rain that fell in NSW.

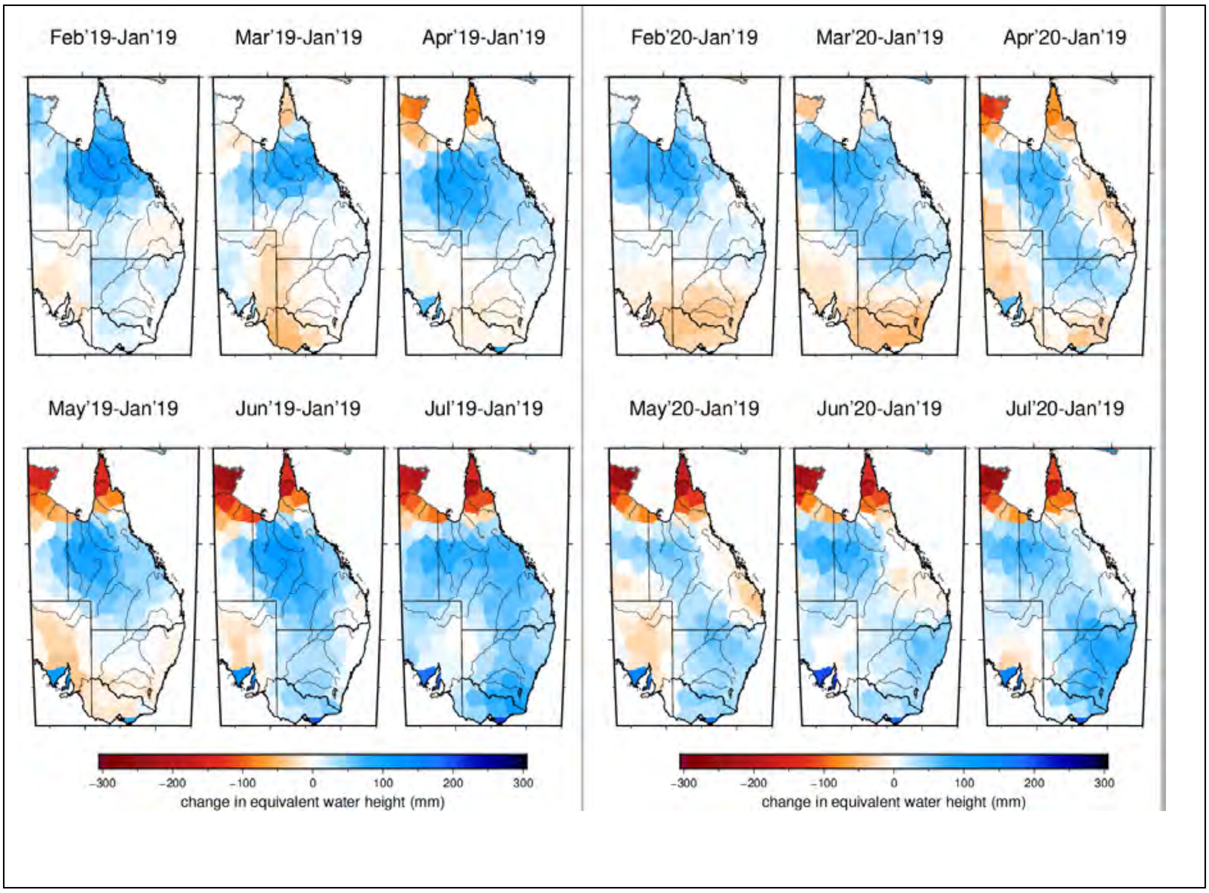


Figure 4.236: Left: Total water storage across eastern Australia February to July 2019. Right: February to July 2020. In both cases January 2019 is used as the baseline. Hydraulic activity such as the onset of the dry season in the Top End (darker red indicating a reduction in water volume), or the flow of water from monsoonal storms in Queensland, down the Channel Country to NSW, can be observed.

The la Niña season in 2020 also contributed significant rainfall over eastern Australia. Again plotted using January 2019 as a reference, the strong influx of water in southern Queensland, propagating southward into NSW, can be seen in

4.3.3.2 Along-track range acceleration analysis

Large rain events provide natural experiments where space gravity data might be able to track the movement of very large flows of water down river systems. The rain event in March 2020 was used to assess the information contained in the range acceleration residuals of the GRACE-FO data, using the method described in the simulations above but applied to real data. The information is shown as a series of 5-day ground tracks in Figure 4.23.

The pattern is quite complex to interpret. Because real observations have been used here, the values plotted along the ground tracks include not only the water signals of interest but also a variety of errors/noise from a number of different sources (e.g. mismodelling of the satellite orbits, actual instrument/observation noise, errors in background models used in the analysis etc). While noisy, these are the observations that were used to create the monthly map for March 2020 shown in Figure 4.29; therefore, the required signals of the water changes are present in these observations along the ground tracks. It is beyond the scope of this current project to determine with what spatial and temporal resolution such information can be extracted, but there is definitely potential here to be able to quantify water mass changes located within river systems.

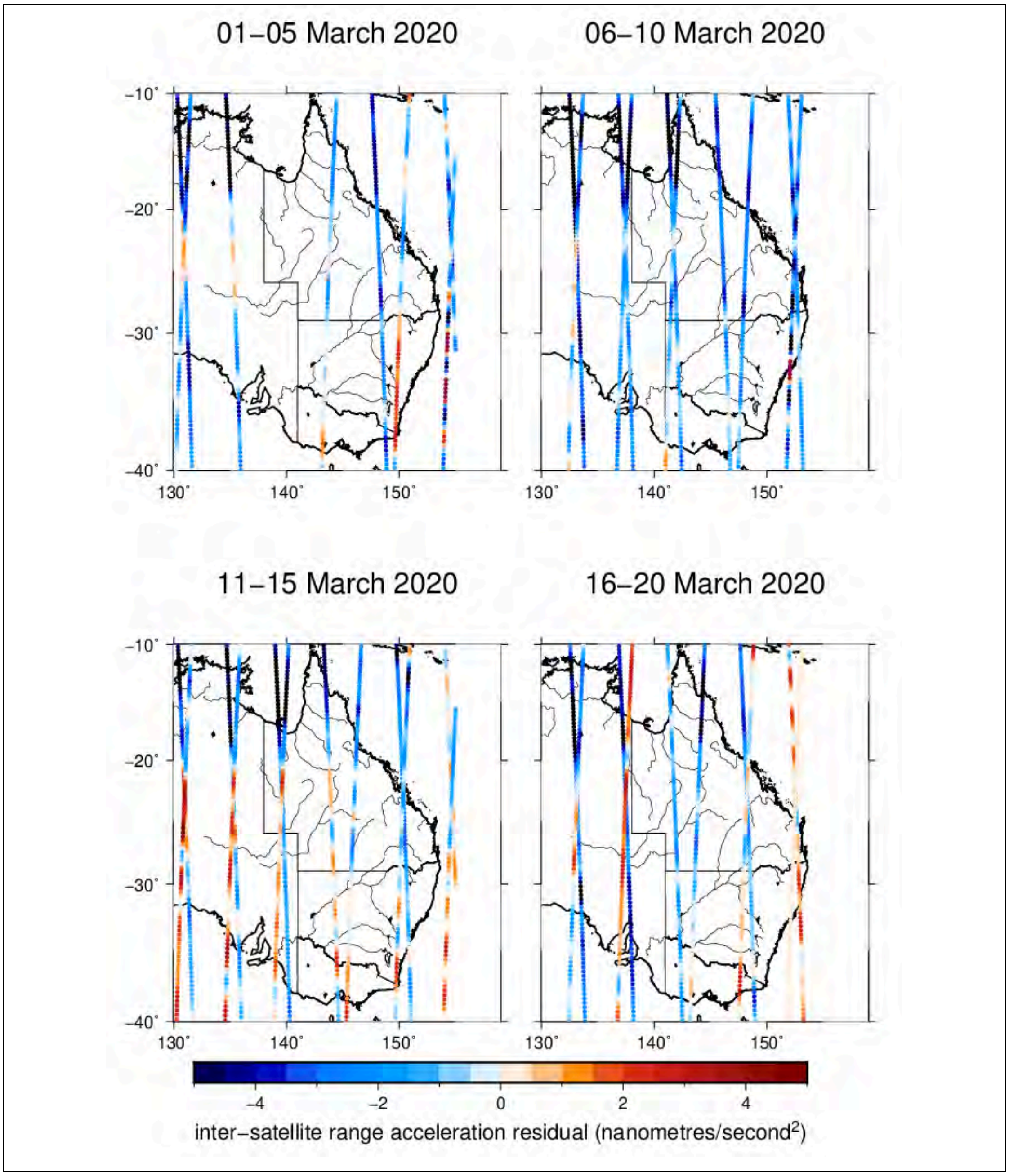


Figure 4.249: Inter-satellite range acceleration residuals in 5-day maps, plotted beneath the locations of the satellites. A negative residual indicates the presence of more water than average. Ground track coverage can only be densified by stacking additional days together, but this does not preserve the time-varying nature of the information that is required to track the flow of water from QLD to NSW during this time period.

4.3.4 Future Work

Terrestrial-based gravity measurements using cold atom gravimetry provide a means of sensing very localised changes in water stores, but this observing technique doesn't upscale easily to basin-scale or state-level studies. In situ groundwater bore measurements provide hydraulic head information but extrapolating point-wise measurements over large spatial areas can lead to errors. Knowledge of the porosity of the subsurface structures is also required to convert borehole measurements of groundwater level into groundwater volume.

A combination of space gravity data, satellite-based measurements of shallow soil moisture and in situ terrestrial measurements can provide both broad-scale and fine-scale quantification of changes in water. Integration/assimilation of space gravity data with in situ measurements has the potential to enhance the utility of all these measurement types. The combination could lead to detailed, small-scale knowledge from point-wise measurements in specific places (e.g. along river courses) infilled with broad-scale information from satellite measurements.

Use case: loss of environmental flow waters

One of the use cases considered in this project related to whether it was possible to determine what happens to the "missing water" that is released in environmental flows but doesn't arrive at the intended location. This is a difficult problem to resolve for a number of reasons. First, the flows are small in magnitude and spatial extent, making them difficult to detect with satellite observations. Second, there is often a relatively long path downstream along which water can be lost from the river. Third, it is not evident how to determine by observation whether water is lost through evaporation, theft or through recharging of groundwater.

Satellite measurements offer one of only a few possibilities for providing insights into this issue. Evaporation of water would cause a reduction in mass which, in principle, is detectable through measurement of gravity change. In contrast, recharge of groundwater aquifers is a change in vertical location of mass and would not cause a discernible change in gravity. Water theft would cause a horizontal translation of water from river to on-farm storage which, in principle, is detectable through gravity measurements and also by detection of changes in water levels in dams/reservoirs etc.

A careful study is needed to determine whether it is possible to use the space gravity measurements along the ground track of the satellites to detect changes in river flows. Australia has experienced a number of large flood events in the past few years, causing significant flows down many river systems. These events could be used as case studies, including comparisons with in situ river height measurements, to ascertain to what extent a combination of space gravity and satellite altimetry (used to measure water heights from satellites) can address the issue of the loss of environmental water flow. Existing satellite data could be used in a desktop study, since both space gravity and satellite altimetry data are available since 2002.

4.4 Where is All the Water?: Probabilistic Modelling Framework

R. Willem Vervoort, Joshua Simmons, Gilad Francis, Richard Scalzo
DARE ARC Training Centre, The University of Sydney
January 2022

4.4.1 Executive Summary

Previous NSW government reports have identified large-scale "unaccounted differences" in the water balance across numerous NSW catchments. These volumes are typically around 20% of inflows [23] ,but may occasionally reach 50% of inflows in some catchments [24]. These numbers tend to be comparable to overall licensed extractions, as well as to surface water recovery volumes under the Murray-Darling Basin Plan, making uncertainty in the water balance a major source of risk for decision-making in water management. The DARE component of the Where Is All The Water? (WIATW) work program develops a probabilistic modelling framework to explain and quantify unaccounted differences in the NSW General Purpose Water Accounting Reports (GPWARs) for major rivers up to catchment scales. The proposed probabilistic modelling framework is based around the application of well established Bayesian inference techniques. This report builds on the conceptual Bayesian inference approach to the probabilistic modelling framework outlined in the previous Milestone 1 report. In this pilot project, the uncertainty quantification of three components of the system is demonstrated focusing on the years 2019 - 2020.

- The uncertainty in estimating the evaporation and transpiration from river reaches and riparian zones.
- The uncertainty in estimating the surface water groundwater connection and groundwater flow given river reach and groundwater observations.
- The uncertainty in the estimated rating curves at gauges in the river. Furthermore the intended deliverables from future stages of this work focus around:
- A priority ranking for the various sources of uncertainty to guide investment in data collection aimed at reducing the overall uncertainty in the water accounting along this reach of river.
- A framework for determining the optimal configuration when designing new data collection programs. For example, to assist in deciding on the spatial placement of new sensors or the temporal frequency Lidar surveys to minimise uncertainty.
- A methodology that can be integrated into the NSW GPWARs to provide a better quantification of the uncertainty in the reported water balance terms.
- A methodology that can assist with the inversion of quantum sensor data to estimate the groundwater recharge and storage.

This report describes the overall water balance problem with a focus on several of the key physical parameters, giving them a mathematical description with associated uncertainty. Simplified proof of concept calculations are undertaken to quantify the uncertainty, which can be incorporated in future work into the probabilistic modelling framework for the overall water balance.

4.4.2 Introduction

The *Where Is All the Water?* (WIATW) work program aims to explain unaccounted differences in the General Purpose NSW water accounting for major rivers[8] (GPWARs) [25] [24] [23], and to determine the role of data from new sensor deployments (low-cost environmental sensors, space-based large-scale gravity and ground-based cold-atom gravity) in helping to close the water balance. The unaccounted differences in the river system accounting tend to be comparable to overall licensed extractions, as well as to surface water recovery volumes under the Murray-Darling Basin Plan, making uncertainty in the water balance a major source of risk for decision-making in water management.

Bayesian approaches in hydrology in general have concentrated on parameter inference [26]. In contrast, there have only been few examples of full Bayesian modelling of the water balance [27] [28] [29], although ground- water modelling [30] and spatial integration of runoff [31] has been done. The reason for this is most likely the difficulty in deriving a closed form solution of the water balance and associated differentiation to be able to account for uncertainties in both time and space. Bulygina and Gupta [27] concentrated on a data assimilation approach assuming an unknown model structure to make forward predictions of the streamflow. More recently, Ossandon [29] developed a hierarchical model of a river network for daily ensemble streamflow forecasting. In contrast, Smith and Gronewald [26] developed a model to describe the water balance of the Great Lakes in the USA and to estimate the uncertain inflows, which is likely the closest comparison to the work here.

In an earlier report [4] for this project (hereafter "Milestone 1 report"), highlighted that while uncertainty in complex systems is unavoidable, it can be quantified using different methods. A Bayesian probabilistic modelling framework to systematically quantify the uncertainties across the river system was proposed, and the overall time investment to develop for the Namoi catchment was presented. The probabilistic modelling framework is useful up to catchment scales, for the purpose of making water management decisions at the state level in NSW under uncertain conditions. Models made under this framework may be used to cost-optimally acquire new data sets, in order to reduce the prediction variance for future management decisions.

Development of the overall framework is well beyond the scope of the current project. However, as part of the *Where is All the Water?* project three potential "use cases" were developed for which different applications could be tested. These use cases are:

- Low flow (Dry) use case, focussing on a short river section between Lake Keepit and Gunnedah in the Namoi river for 2019;
- High flow (Wet or flooding) use case, concentrating on the Namoi catchment downstream of Narrabri and mostly for 2020 and 2021; and
- Large scale use case (flood), concentrating on the Darling from the Queensland border to Menindee Lakes, most likely also using 2021 as an example year.

The aim of this report from the University of Sydney's DARE ARC Training Centre is to produce a pilot study for the probabilistic modeling framework for the dry use case.

[8] https://www.industry.nsw.gov.au/water/allocations-availability/water-accounting/gpwar

4.4.3 The Model

The overall conceptual framework developed earlier in the Milestone 1 report conceptualises the river system as a space time model that includes all the different fluxes identified in Figure 4.30. In the sections below, the hydrological model is explained in more detail followed by a section outlining the general statistical modelling approach using Gaussian processes.

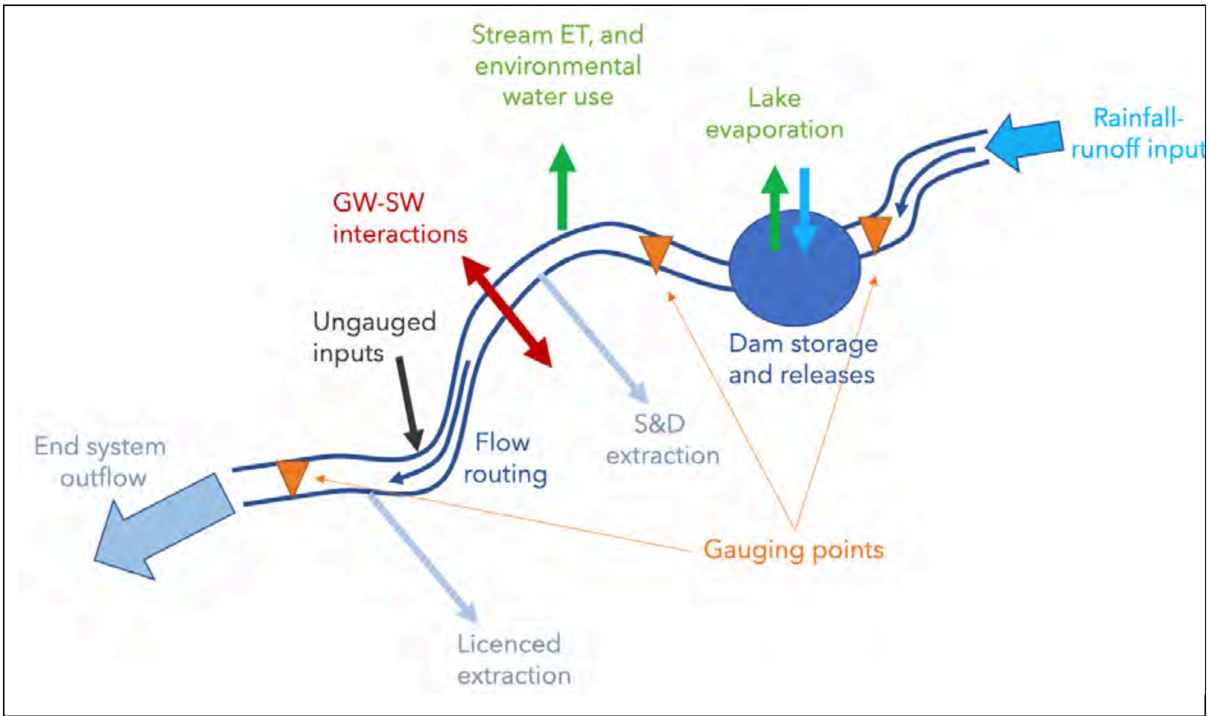


Figure 4.3025: Conceptual river system diagram for the framework.

4.4.3.1 Hydrology

4.4.3.1.1 Rainfall Runoff

Starting from the top of the catchment, the first input into a river is generated by a rainfall-runoff model. In essence this is:

$$Q(t) = P(t) - ET(t) - \frac{dS}{dt} + \epsilon = f(P, ET, S, t, \theta) + \epsilon \tag{11}$$

The general structure of hydrological models consists of the inputs and outputs Q(t) (the streamflow flux in mm/day), P (t) (the rainfall/precipitation flux in mm/day) and ET (t) (the combined evapotranspiration flux in mm/day); and parameters θ. In practice hydrological rainfall-runoff models can have any number of parameters to describe different delays within the system (linked to storage). Note that in this case P (t) S(t) and ET (t) are presented as "catchment average" values, representing the average of a distribution in space. As such, a description of the variability and uncertainty in space is required to characterise these distributions.

In hydrological modelling variables are generally reported as a "length" (mostly mm), representing area weighted averages. Within the ϵ term, any leakage out of the actual modelled catchment is also accounted for.

In essence this means that the overall "error" ϵ can be defined as

$$\epsilon = \sum(\epsilon_P(x,y), \epsilon_{ET}(x,y), \epsilon_S(x,y), \epsilon_\theta(x,y), \epsilon_{other}) \tag{22}$$

In other words, the overall "uncertainty" of any hydrological model is not well specified and includes the cumulative uncertainty of the different input variables and the miss-specification of the model [32].

From the top of the catchment, the water passes the first gauging station, which can be used to constrain θ in the rainfall-runoff model, providing the closure on the water balance. Moving down the catchment, sections of the river (i.e., river reaches) can be modelled individually knowing the input from upstream [29].

4.4.3.1.2 River Reach

From the first gauging station the water travels through the river reach, which can be considered a closed sub-system between two gauging stations. In other words, the basic water balance applies:

$$Q_{out(t)} = Q_{in(t)} + gains - losses - \frac{dS}{dt} + \epsilon \tag{33}$$

Here $Q_{out}(t)$ is the streamflow at the downstream station(s), while $Q_{in}(t)$ is the streamflow at the upstream station(s). Defining the gains and losses is more complicated as some of these are "known unknowns", but there may also be "unknown unknowns". In the latter case, the water balance cannot be closed, leading to an increase in the ϵ term visible in the uncertainty quantification. In Figure 4.30 the following processes are indicated:

- Groundwater and surface (GW-SW) interactions
- Evapotranspiration and environmental water use
- Ungauged inputs (a gain)
- Licensed extraction
- Stock and domestic extraction

Further processes that could also be considered are:

- In channel rainfall (a gain)
- Irrigation return flow
- Sewage Treatment Plant discharges
- Overbank flows and floodplain losses

All of these processes would require some sort of definition of the process or rate that determines the flux. Plenty of complex models exist in the hydrological literature, but as a first approach a simpler description might suffice. This can then subsequently be extended to more complex models. All these fluxes in turn have associated parameters and uncertainties which are summarised for the variables examined in this report in Table 4.6 and for other components of the overall probabilistic framework in the Milestone 1 report.

Table 4.6: Summary of processes which have been examined in this report (see Section 4), corresponding conceptual/mathematical models, data available to constrain them, and associated uncertainties.

Process	Conceptual model	Data	Uncertainties
Stream flow (length/time)	Rating curve fit relating river height to river flow	Continuous river levels; discrete flow measurements at gauge locations for rating curve generation	Uncertainty in the discrete measurements and the integration of the discrete measurements with the continuous data; uncertainty in the conceptualisation of the rating curve model and uncertainty in the rating curve model calibration (20 – 40%, see Tomkins, 2014); uncertainty in the model conceptualisation of the process
Reach length groundwater losses and gains (length/time)	Simple resistance (Darcy) or more complex gradient based loss equation using conductivity estimates	Groundwater level changes from "nearby" groundwater bores (length); Australian soil information	Uncertainty in groundwater level data; uncertainty in spatial integration of the Australian soil information; uncertainty in the connectivity between river and groundwater; uncertainty in the spatio-temporal model
Land-based evapotranspiration (ET) (length/time)	Penman-Monteith or derivatives	Measured continuous climate data: radiation, temperature, windspeed, humidity	Uncertainty in actual climate data; uncertainty in the Penman-Monteith model; uncertainty in the spatio-temporal model due to sparse data and spatial model fitting (including kernel/variogram estimation); uncertainty in the riparian fringe around the river; uncertainty in vegetation map or fractional cover derived from satellite data
Stream and lake evaporation (length/time)	Penman-Monteith or derivatives	Measured continuous climate data: radiation, temperature, windspeed, humidity	Uncertainty in the spatio-temporal model (see land-based ET); uncertainty in surface water area through time to understand total volumetric losses given only sparse water level measurements

4.4.3.2 Statistical modelling with Gaussian process regression

Many of the sources of model uncertainty can be described in terms of unknown or uncertain spatiotemporal functions. Although these can be described using familiar parametric forms such as polynomials or splines, with uncertainties on the coefficients, we will find it convenient below to look at non-parametric ways of specifying random functions. Gaussian process regression (aka kriging) provides a convenient way to do this, where the posterior predictive distribution (our preferred way to quantify uncertainty in random functions) can be calculated in closed form using linear algebra operations.

The Gaussian process formalism is used in multiple sections below so we briefly review them and develop some notation here. Gaussian processes can be thought of as distributions over functions to be fit to data. The properties of the functions drawn from the distribution are usually specified in terms of a symmetric, positive definite kernel function $K(x, x'; \Lambda)$, describing the covariance between two evaluations with different feature vectors x and x', possibly as a function of hyperparameters Λ. The feature vectors can be locations or time points in the case of spatiotemporal models, or can be other ordinates such as river heights or flows if the function being fit is, for example, a ratings curve. The hyperparameters Λ include important scales in the problem such as length scales, time scales, noise variances, and so forth. For example, a commonly used kernel is the square exponential kernel

$$K(x, x';\ S_0, \ell) = S_0^2\ exp\left(\frac{-\|x - x'\|^2}{\ell^2}\right) \tag{44}$$

where x and x' might represent spatial locations, S_0 is an amplitude and ℓ a characteristic length scale over which the function is smooth.

Given data y taken at arbitrary covariates $X = \{x_i\}$, assumed to have iid Gaussian noise with variance σ^2, and a kernel function $K(x, x')$, the prediction for the gridded measurements y' at locations X' is multivariate Gaussian with mean and covariance

$$\mu_{y'} = K_{x'x}(K_{xx} + \sigma^2\ I)^{-1}y \tag{55}$$

$$\Sigma_{y'} = K_{x'x'}\ -\ K_{x'x}(K_{xx} + \sigma^2\ I)^{-1}K_{xx'} \tag{66}$$

where $K_{XX'}$ is the matrix with elements $K(x, x')$ for each pair of station and grid locations. Any linear operation on gridded data described in terms of a matrix $\alpha = Ay'$, including area-weighted averages or numerical integration, is Gaussian-distributed with mean $A\mu_{y'}$ and variance $A\Sigma_{y'}A^T$.

4.4.4 Case Study

As a case study, a focus has been placed on the period of the 2019/2020 GPWAR (see 2 for water level and rainfall data over this period) which covers the "Dry" use case as mentioned in the Introduction (1). The specific river reach used, which is the Namoi River from Keepit dam to Gunnedah (see 3 for area map) including the Peel confluence, covers the following gauging stations:

Inflow: Gauge 419006, Peel River at Carroll Gap

Inflow: Gauge 419007, Namoi River at Downstream Keepit dam

Inflow: Gauge 419084, Mooki River at Ruvigne

Outflow: Gauge 419001, Namoi River at Gunnedah

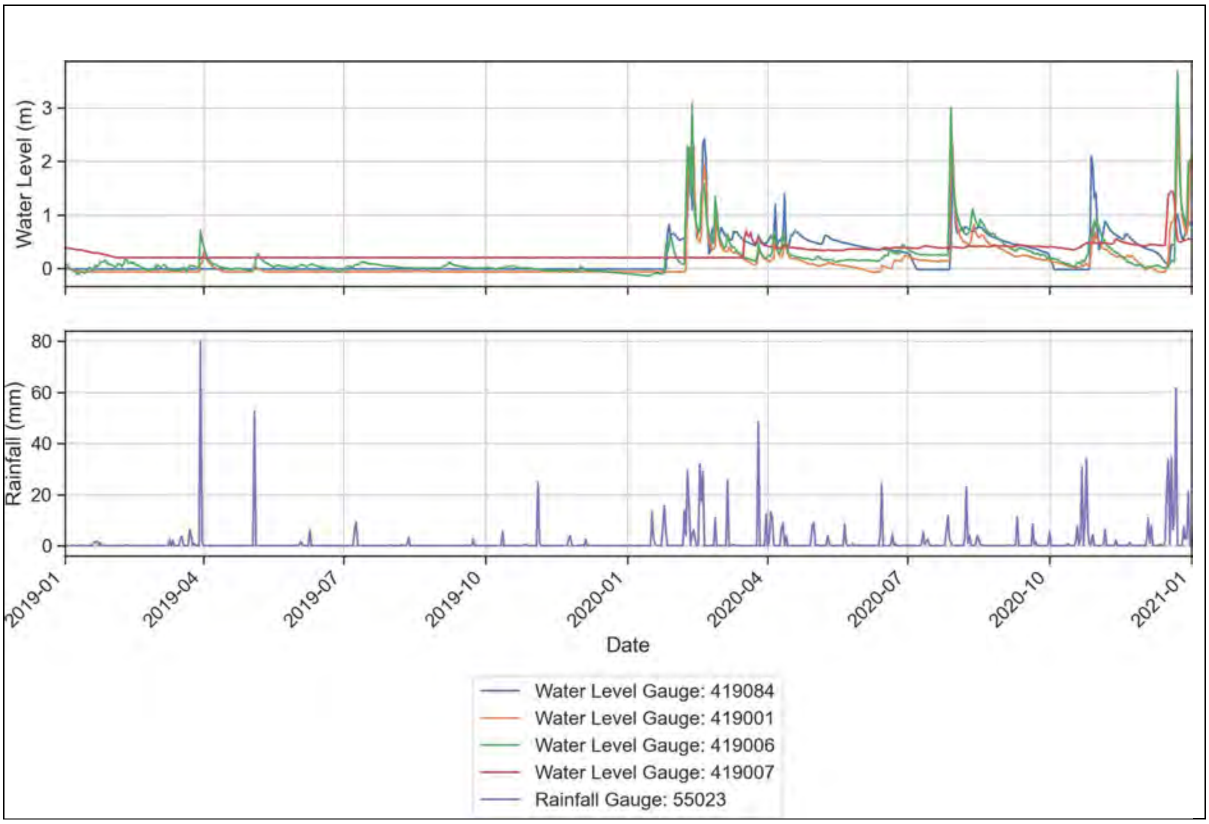


Figure 4.3126: Daily mean water level (m above gauge zero) at the inflow and outflow gauges and daily sum rainfall (mm) at the Gunnedah Pool station (55023) over the study period from 2019 to 2020.

Using the "Dry" use case and focussing on a section with no major urban development means that ungauged inputs, licensed extraction from urban utilities, and overbank flow and floodplain losses can be ignored. Furthermore, in-channel rainfall is assumed to be negligible.

In the following subsections the proposed model structure will be further described including possible priors for each of the remaining components for a low flow period between two gauging stations. Results of sub components that have been developed will be highlighted in each section. With these assumptions the overall balance equation becomes:

$$Q_{out(t)} = Q_{in(t)} - \left(E_{water(x,y)} + ET_{env(x,y)}\right) - Pump_{license} - Pump_{sd} - GW_{SW(x,y)} + \epsilon \quad (77)$$

In this equation ET_{env} is the evapotranspiration of the environment and E is the evaporation from open water in the river channel. These terms are discussed further in Section 3.1.

Licensed extractions along the river reach were obtained from WaterNSW/DPIE to develop a time series of the licensed surface water extractions (Pumplicence). However, including these will be part of future work and licensed extractions have been ignored in the present study. In addition, there is no detailed data for extractions for stock and domestic extractions. For this specific report, we are ignoring all extractions and are assuming that all stock and domestic water needs are addressed using groundwater extractions from deeper groundwater sources. In future work, this can be added to the model uncertainties. Stock and Domestic extractions could be assessed using a simple function that relates daily temperature to stock and domestic use ($Pump_{sd}$), i.e.:

$$Pump_{sd} = \beta_o + \beta_1 \times \text{Temperature}_{max} + \in \quad (88)$$

The means that the main processes to be estimated in the model in this report are the uncertainty in the gauging, evapotranspiration and environmental water use, and the groundwater - surface water interaction ($GW_{SW\,(x,\,y)}$). Note that the groundwater - surface water interaction can be both positive and negative, representing either a loss or a gain of water.

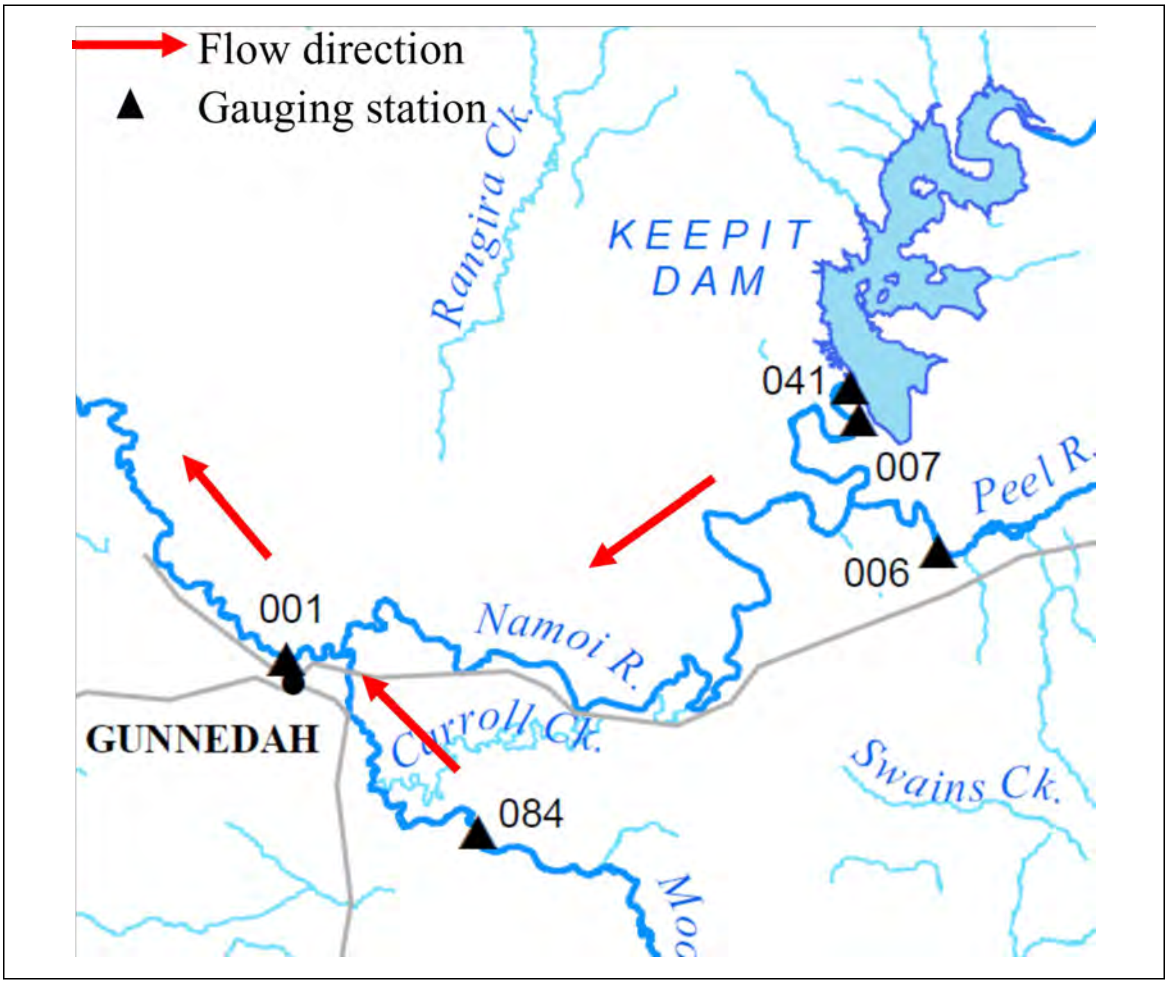


Figure 4.3227: Map of the case study area (Source: NSW Government). Note: gauging station 419041 has not been used as it is the storage gauge for Keepit Dam.

4.4.4.1 Evapotranspiration and environmental water use

This flux consists of two parts: 1) the open water evaporation from the river, and 2) the transpiration and evaporation from the ecosystems in the riparian zones and parts of the river bed with no standing water.

4.4.4.1.1 Open water evaporation

A first approximation of the open water evaporation, which only occurs if surface water is present in the river, is shown below::

$$E_{water}(\mathrm{x,y}) = Area_{river}(x,y) \times E_o(\mathrm{x,y}) + \epsilon \tag{99}$$

The point potential evaporation based on the location in the river $E_0(x, y)$ is calculated using:

$$E_o(\mathrm{x,y}) = (\mathrm{ETrx,y}) \times \mathrm{Kc} \tag{1010}$$

with the crop coefficient (Kc) taken as 1.05 for open water and the reference evapotranspiration (ETr) calculated using the FAO Penman-Monteith equation (Allen et al., 1994):

$$\mathrm{ETr} = \frac{0.408\delta(R_n - G) + \gamma \frac{C_n}{T + 273} u_2(e_s - e_a)}{\delta + \gamma(1 + C_d u_2)} \tag{1111}$$

where R_n is the net radiation, G is the soil heat flux, δ the saturation vapour pressure-temperature slope, u_2 the mean wind speed measured at 2 m, T the mean daily temperature, γ corresponding to the psychrometric constant, e_s to the saturation vapour pressure, e_a the actual pressure vapour, while C_d and C_n are constants for standard short crops, equivalent to 0.34 and 900, respectively. Future work could further analyse this the uncertainty stemming from this equation itself, defining each of the inputs R_n, u_2, T and e_a with their individual uncertainties.

For this phase of the study, a simplified approach has been adopted for obtaining the wetted area of the river for each point estimate $Area_{river}(x, y)$ using Digital Elevation Model (DEM) data (obtained from the Commonwealth of Australia (Geoscience Australia) ELVIS portal) and water level data from the available gauges.

$$Area_{river} = Length_{river} \times Width_{river} \tag{1212}$$

The river length ($Length_{river}$) has been determined for each of the three reaches in the study area (Namoi River, Peel River and Mooki River), using watershed analysis on the 5 m DEM data of the region (collected in 2011). The time-varying river width ($Width_{river}$) was determined at the available water level gauges mapping water level to width using cross-sections extracted from the available 1 m (collected in 2014 and 2019) and 5 m DEMs in the area (see Figure 4.33 and Figure 4.34 for examples). At this stage it is assumed that the river width is constant along each of the river sections corresponding to the available gauges, with planned additional work noted in Section 3.1. The presence of "open water" in the reach is based on the water level measured at the gauging stations, masked using a threshold level. Below this level zero open water evaporation is assumed.

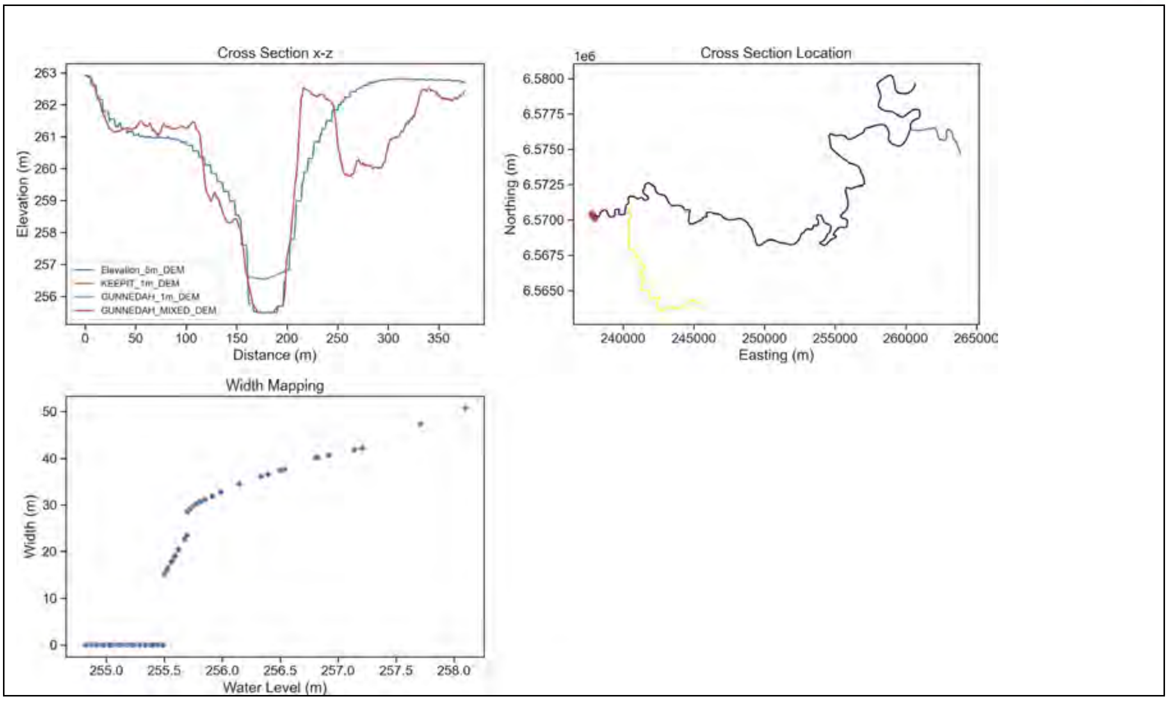


Figure 4.3328: An example of the water level to river width mapping at the downstream water level gauge 419001 (shown by the red marker in the top right plot). Multiple available DEMs show the variability stemming from measurement, processed resolution, and topography changes over time.

For this study, the point reference evapotranspiration values (ETr(x, y)) have been interpolated (with as- sociated uncertainty) using a GP regression fit to calculated values at available weather stations in the SILO database [5]. Spatial variability of ET over the study period of interest as output from the GP is shown below in Figure 4.. The spacing of the output points at which both $E_{water}(x, y)$ and $ET_{env}(x, y)$ were evaluated were determined using the length scales of the fitted GP kernel (approximately 0.13 degrees in both latitude and longitude).

Future work could significantly improve the estimation of wetted area using a combination of DEM and satellite data and allow for the quantification of uncertainty in this term. An example of using water detection algorithms is outlined in [3], which can be translated to Sentinel data from the original MODIS approach. This approach can be applied for all cloud free images, which in the low flow use case considered here will probably result in the majority of images. Other useful datasets that could inform the wetted area term include the "Water Observations from Space" dataset [12] and associated "Hydromorphological attributes for all Australian river reaches" dataset [8] if the resolution of these data are adequate.

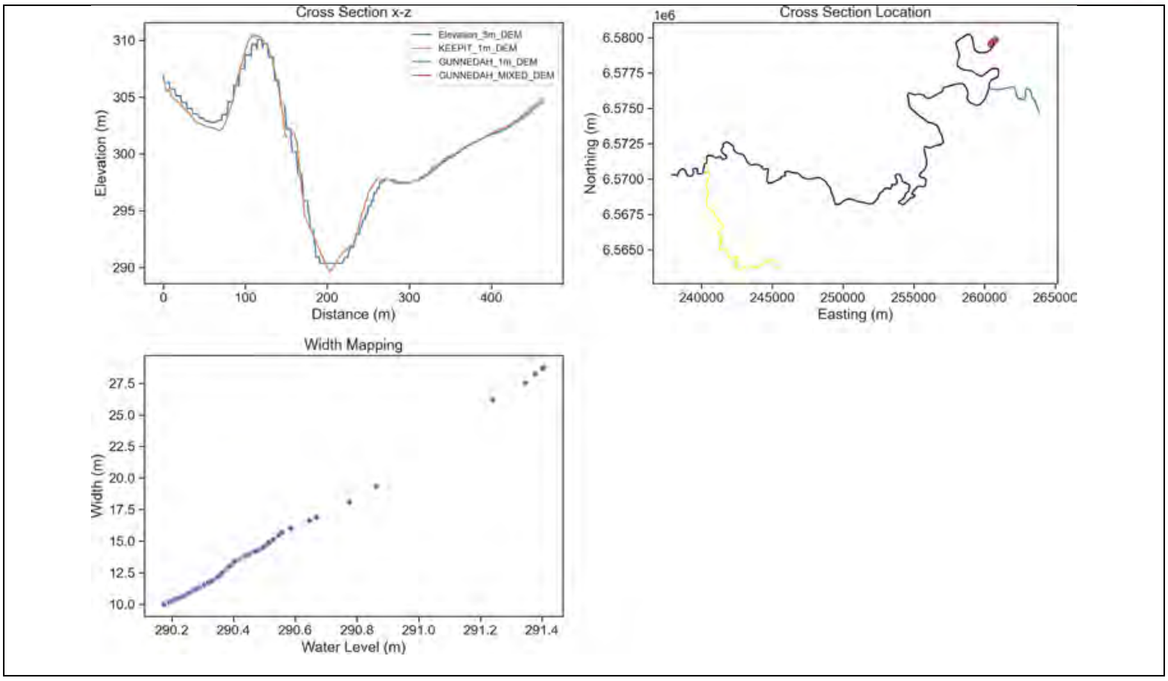


Figure 4.3429: An example of the water level to river width mapping at an available upstream water level gauge.

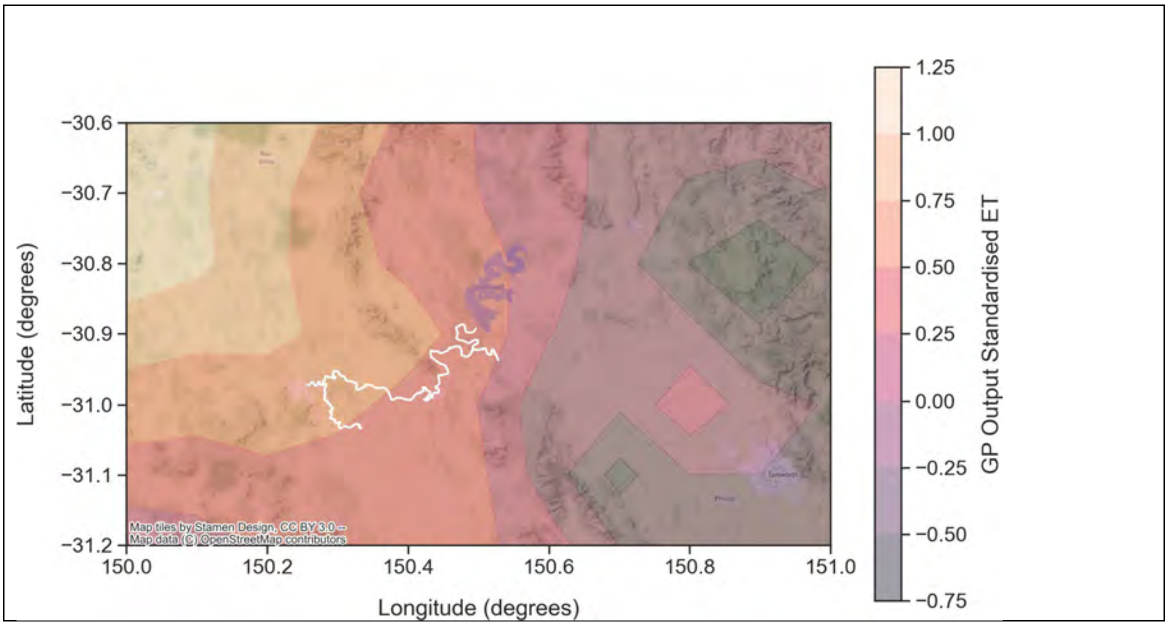


Figure 4.3530:Standardised output of the GP output for ET r(x, y) showing the spatial variability across the study area.

4.4.4.1.2 Environmental Water Use

Environmental water use is defined as the amount of water used by the riparian vegetation that surrounds the river. Using satellite imagery we can identify areas of riparian vegetation, but we an also simply take a number of pixels away from the river as the "riparian fringe". For example a buffer of 100 - 200 m around the river can be interpreted as a "riparian fringe". For this simplified study, a fixed width of 100 m was chosen, excluding the open water river area as calculated in Section 3.1.1.

Initial results for both the evaporation and evapotranspiration components, incorporating the uncertainty in the interpolation of ETr(x, y), are shown below in Figure 4.36. The uncertainty in this case represents only the spatial variability in ET, rather than that of the total ET flux. Future work to quantify the uncertainty in the river width and length terms (as discussed in Section 3.1.1) is required to further develop this component and this will likely have a significant impact on the magnitude of the total uncertainty for the ET flux.

To better estimate the evapotranspiration component and associated uncertainty, future work could incorporate the estimates of actual evapotranspiration at vegetated pixels in a buffer around the stream channel.

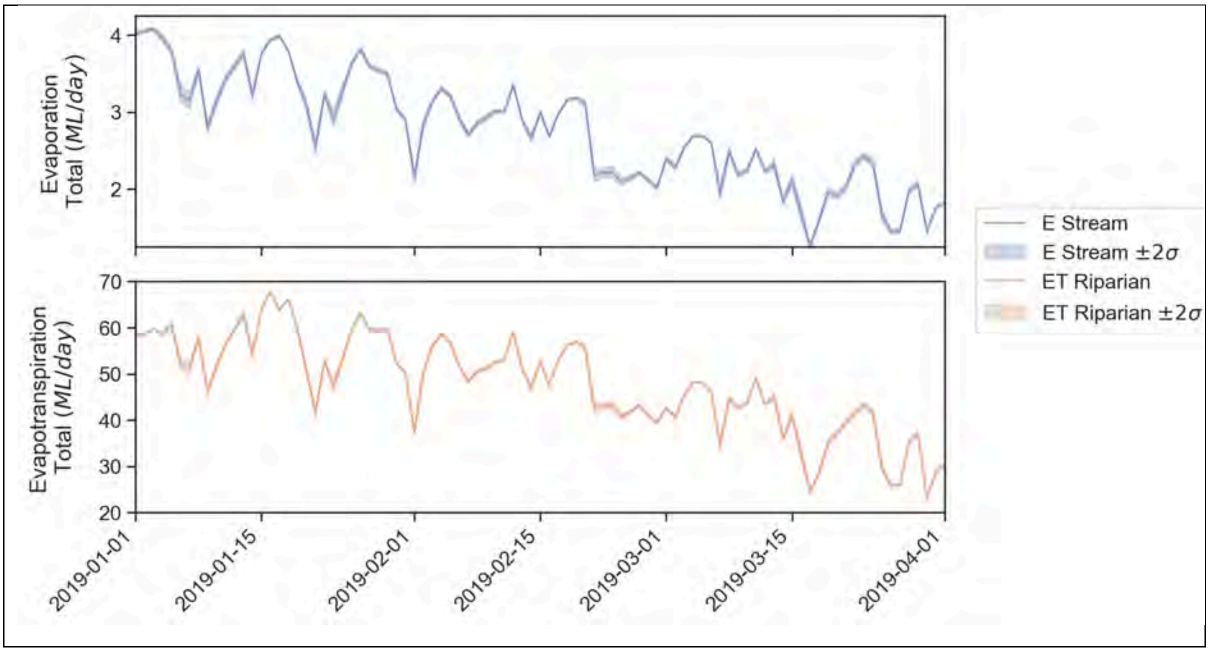


Figure 4.3631: Initial results for evaporation and evapotranspiration along the length of river reach in the study area for January to April 2019. As noted in the accompanying text, this first-pass study only incorporates a small component (the spatial variability in ETr(x, y)) of the overall ET uncertainty and as such the uncertainty bands are narrow.

This can be obtained, for example, using the CMRSET satellite evapotranspiration methodology [6] applied to Sentinel 2 and Landsat datasets (see [11]). This dataset would provide monthly AET data at 25 m resolution which could be used to generate time varying crop factors to be used with SILO daily ET data. Subsequently multiplying the derived AET by the area of the buffer will provide an estimate of the environmental water use. Lumping to larger areas will provide a variance of the estimate which can be also included in the uncertainty calculations.

The original CMRSET paper by [6] provides RMSE values for the regression fit of the actual ET prediction model from the satellite reflectance, but does not provide direct error bounds for the parameters. Similarly, the new data set [11] also does not specify any uncertainty, which formally would need to be estimated. Satellite products generally define the accuracy of the observations and the instruments which can be incorporated in the calculations of the reflectance equations.

4.4.4.2 Groundwater – Surface Water interaction

The simplest way to represent the groundwater/surface water interaction (F_{GWSW}) is using the effective waterconductance (K_{eff}) and the difference in potential (h) between the groundwater and the water level in the channel. Both of these water levels need to be scaled to the Australian Height Datum (AHD) or some other common datum. The flux of water flowing from the river into the ground is then usually calculated as

$$F_{GWSW(x,y)} = K_{eff}(x,y) \times \frac{\delta h(x,y)}{\delta x} + \in \quad (1313)$$

Here FGWSW(x, y) is the groundwater/surface water interaction for the subsection under consideration, Keff (x, y) is the effective conductivity (rate of transmission of water per unit time) of the subsection, and h(x, y) is the water level and x is the distance between the observed water level in the stream and the observed water levelin the river channel. This version of the equation is a finite-difference approximation to Darcy's law

$$F_{GWSW} = K_{eff} \nabla h + \in \quad (1414)$$

where now the groundwater head h is a differentiable function. In general h is well-known along the river, but for groundwater sites h is available only at sparsely situated groundwater bores. Equation 14 also assumes that the river is connected to groundwater, which will depend upon the form of the spatial conductivity field Keff . There could also be several groundwater layers, so for this application, locations near the river are preferred in order to observe the shallowest available groundwater.

The groundwater data can be easily obtained from https://realtimedata.waternsw.com.au/. Relevant groundwater bores for this case study along the reach would be:

GW093001, GW093000, GW036271, GW036268, GW030307, GW036236, GW030305, GW030304, GW030303, GW030302, GW036237, GW036238, GW039338, GW036239, GW021087, GW021086, GW021085, GW030300, GW030299, GW030298, GW036272, GW036289, GW965580, GW965581

These bores will be affected by groundwater pumping, which will affect the water table and therefore the gradient and groundwater surface water interaction.

Note that Equation 14 does not describe a change in overall storage with time over some unit area, to put it on equal footing with the other processes. Some kind of enclosing contour or surface is needed with respect to which the change can be calculated. Additionally, it is not possible to infer Keff from the hydraulic head h based on Equation 14 in the given form, which doesn't close the loop to derive any kind of likelihood for Keff in terms of the data for the hydraulic head h.

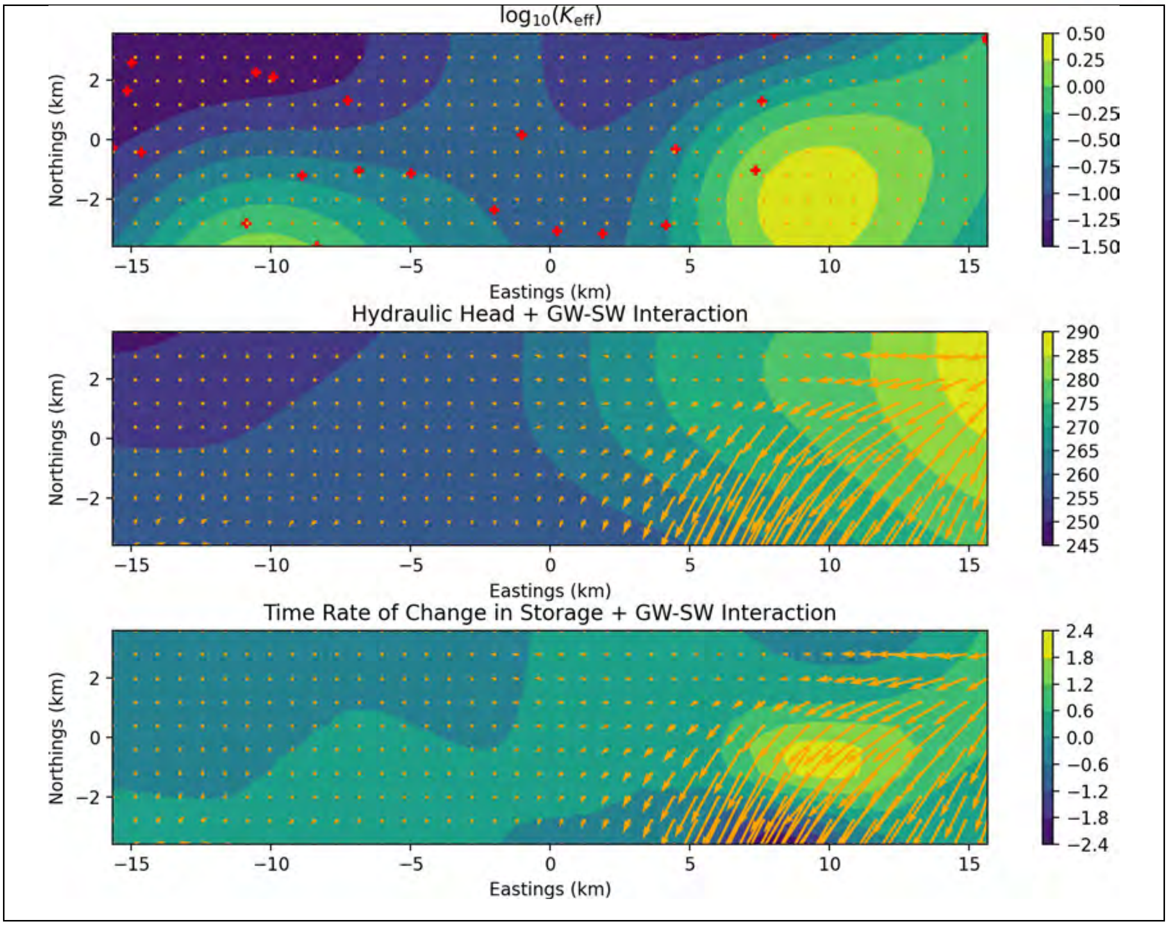


Figure 4.3732: Calculation of 2-D groundwater flux. Top: Hydraulic conductivity field Keff (assumed). Red crosses: training points (borehole positions). Orange dots: prediction points (for numerical integration). Middle: Gaus- sian process regression for hydraulic head (contours) with inferred groundwater flux (orange arrows). Bottom: Time derivative of hydraulic head based on Equation 15 (contours) with inferred groundwater flux (orange ar- rows).

To incorporate the time dimension explicitly, we can connect Equation 14 to the change in stored volume in a column of water underneath an infinitesimal area element dA on the ground as seen from overhead. The volume is Sh dA, where S is the specific storage (of order a few percent), and the rate of change of volume is S(dh/dt) dA. Using the divergence theorem then gives us

$$\frac{dh}{dt} + S^{-1}\nabla\cdot\left(K_{eff}\nabla \mathrm{h}\right) = \frac{dh}{dt} + S^{-1}\left(K_{eff}\nabla^2 \mathrm{h} + K_{eff}\cdot\nabla h\right) = 0 \qquad (1515)$$

which is a diffusion equation for h based on Darcy's law. All other constant factors with physical dimensions have been absorbed into the definition of Keff .

Spatial slices of h can be calculated from a GP regression on the data values taken at a given time t, and spatial derivatives can be obtained from the GP solution based on a cross-covariance kernel without having to re-train (see section 9.4 of Rasmussen & Williams 2006). This allows us to solve, analytically or using finite dif- ferences, for the spatial covariances of the derivatives and thus for the uncertainty in the flow of groundwater. Figure 4.37 shows the motion of groundwater under a synthetic assumed hydraulic conductivity field Keff , shown in the top panel. The interpolation is done from a time slice of real hydraulic head near the Keepit- Gunnedah reach of the Namoi River, with boreholes shown by the red crosses; the coordinates have been transformed to physical units of kilometers. The orange dots show the locations of a denser grid of prediction points to be used for numerical integration. In the middle panel, level set contours of the regression for the hydraulic head are shown, together with the inferred flux FGWSW. The vector field shows sensible behaviour, with high (low) flux in regions of high (low) Keff , and with the flux vectors oriented at right angles to level contours of h. The bottom panel shows contours for the time derivative term with FGWSW superimposed.

Since it is difficult to directly show uncertainties on the 2-D plots, Figure 4. shows a spatial slice through the regressions for groundwater head and its derivatives, including bands of uncertainties. In each case the bands of uncertainty are well-behaved. The solution is smooth on length scales of kilometers, reflecting the available groundwater bore data, and higher derivatives can be easily calculated with reasonable uncertainties attached.

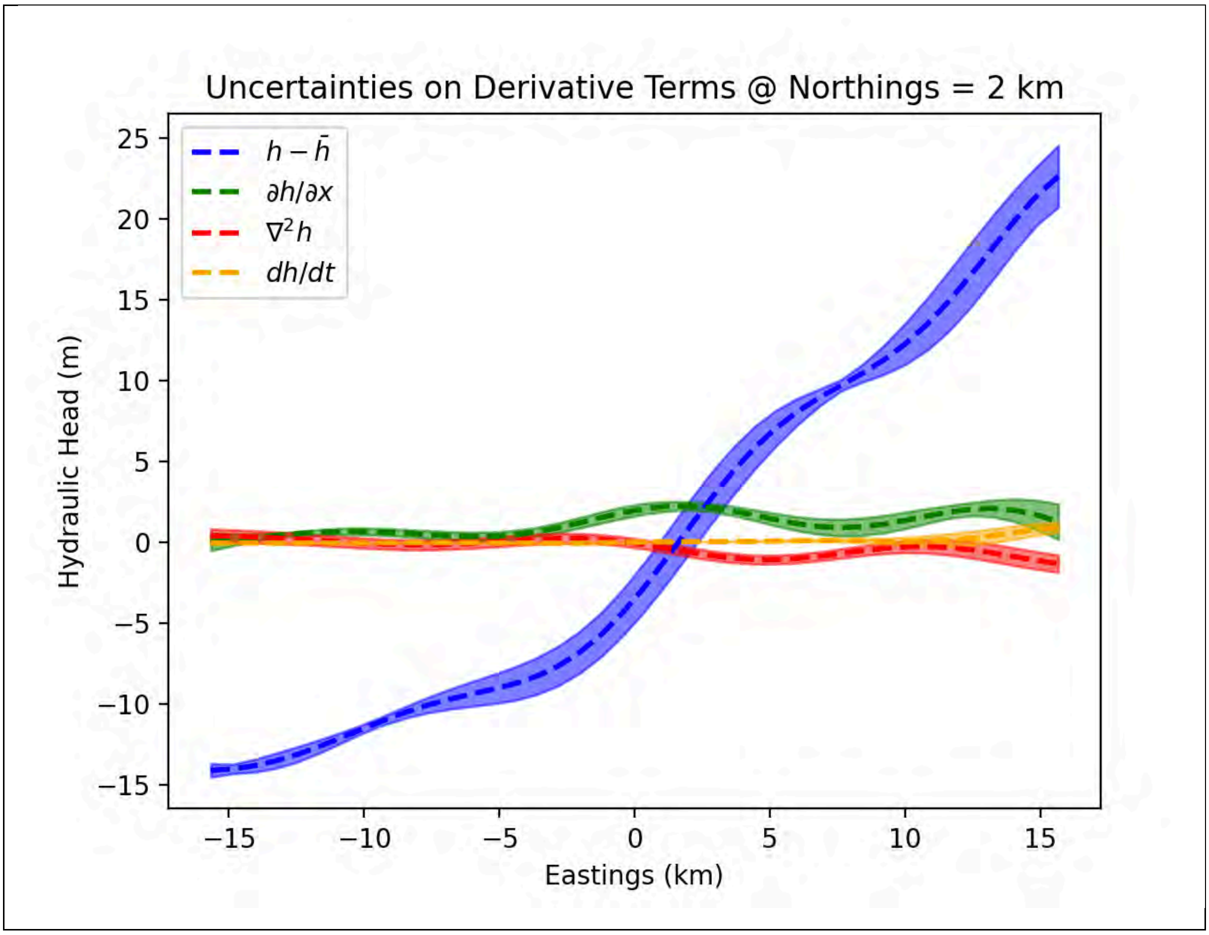


Figure 4.3833: A 1-D slice through the hydraulic head configuration from Figure 4.38. Blue: hydraulic head h relative to its mean value h¯. Green: latitude component of the hydraulic head gradient ∂h/∂x. Red: the Laplacian 2(h). Orange: the time derivative term dh/dt = S−1 (Keff h) calculated from Equation 15. The error snakes show 68% uncertainty contours.

4.4.5 Future Work and Estimated Effort

4.4.5.1.1 Future work: Inference over K_{eff}

Using Equation 15, we can predict the entire hydraulic head time series using GP regression, integrating each spatial slice forward in time to predict the following one. A forward Euler scheme is liable to destabilize quickly given the large time steps between seasonal measurements. A better choice is Heun's method, a second-order Runge-Kutta method, with step size Δt:

$$\tilde{h}_{i+1} = h_i + f(t_i, h_i)\Delta t \tag{1616}$$

$$h_{i+1} = h_i + \frac{\Delta t}{2}\left[f(t_i, h_i) + f(t_{i+1}, \tilde{h}_{i+1})\right] \tag{1717}$$

where f(ti, hi) = −S−1∇ (Keff∇hi) is formed from the spatial slice at time ti. Another possibility, if the condition number is small enough, would be to calculate A(δt) = exp(−δt S−1∇ (Keff∇)) and apply it directly to h. This can be done using the SVD decomposition to diagonalize a finite-difference matrix representation of the differential operator enclosed inside the exponential.

Although the closed-form derivation is cumbersome, the action of integrating Equation 15 for some time interval δt is a linear operator A = A(δt) on hi, and therefore we can be sure that the prediction will also be a Gaussian process (Ahi, AKh AT). In this way, a kernel for time correlations of a self-consistent ground- water diffusion process is induced, so that we don't need to specify it explicitly.

The time integration of Equation 15 allows us to back out Keff , using the likelihood

$$\log \mathcal{L} = -\frac{1}{2}\sum_{i=1}^{N_t-1} \Big\{(h_{i+1} - Ah_i)^T\left(AK_{h_i}A^T + \sigma_h^2 I\right)^{-1}(h_{i+1} - Ah_i) + N_r \log 2\pi + \log\det AK_{h_i}A^T\Big\} \tag{1818}$$

where Nr and Nt are the number of boreholes and the number of time slices respectively. We can read this in terms of conditional probability as log P (hi+1 hi, Keff) and use a GP prior on log Keff .

This whole calculation can then be placed in an optimization or sampling loop for the parameters of Keff (i.e. the latent conductivity values at the boreholes and any other inducing points of our choosing), as well as all the GP hyperparameters which enter implicitly into Equation 18. While challenging, this is the sim- plest method known to us right now to solve self-consistently for Keff . It is a mesh-free numerical solution, potentially cheaper than a detailed finite-volume model along the lines of MODFLOW, that also respects the uncertainties in the spatial fields.

4.4.5.1.2 Future work: Bayesian optimization of groundwater sensor locations

No mention has yet been made of the dependence of the groundwater solutions on particular sensor place- ments. Although this represents future work as well, a framework follows directly from our descriptions of uncertainty here in terms of Gaussian processes. One of the conveniences of Gaussian process regression is that the posterior predictive variance (or co-variance) depends only on the sensor locations and not directly on the measured values. As a result, any function of the posterior predictive variance of a Gaussian process — such as the uncertainty for a flux of groundwater in or out of a strategically important region — can be expressed as a function of additional future sensor locations. Sensors can then be appropriately placed to minimize the uncertainty. This is the most common overall strategy in the Bayesian optimization literature.

As written, the current groundwater flux model depends on direct hydraulic head measurements based on groundwater bores. Additional borehole instrumentation can be evaluated directly. To the extent that the conductivity Keff is correlated with the void fraction (or porosity) fV , inference over Keff may also help to constrain fV and provide a connection to gravity sensor measurements. Since the gravity sensors are also linear functions of the underlying groundwater density, the groundwater model here can directly predict the influence of saturated-zone groundwater on gravity signal — at least for scales larger than the interpolation scale. Including both Keff and fV in the model would then allow us to determine relative costs and benefits of ongoing gravity monitoring versus borehole instrumentation, and optimize locations for any needed sensors.

4.4.5.2 $Q_{out}(t)$ and $Q_{in}(t)$

Considering a river reach, the main direct measurements of flow would be based on river gauging data at the entry (Qin) and exit (Qout) points of the reach. River gauging data is based on a relationship developed between river height (which is fairly easy to measured) and river flow (which is harder to measure).

There are a series of "Australian standards" around gauging, for which AS3778.2.3 is the main one, and this is incorporated in the BOM standards 3. However, uncertainties in the rating curves are not commonly derived, even though an Australian standard exists (AS3778.2.4) that gives some guidance.

A rating curve depends on the the hydraulic properties of a stream channel. Quantifying these dependen- cies is hard in practice, hence rating curves are approximated with a fixed, usually small, set of calibration measurements (gauging). There are several challenges associated with deriving rating curves in this manner:

Different flow regimes are characterised by different rating curves. With a limited number of observa- tions, it is difficult to identify the transition point between flow regimes.

Extremes flow events, both high and low, are rare, leading to fewer calibration measurements and a higher uncertainty about the rating curve at these conditions.

The hydraulic conditions can change over time and after high flow events leading to non-stationarity of rating curve in time. This is currently mitigated by an ad-hoc schedule of re-calibration.

The gauging uncertainty is correlated to the the measured flow. A-priori, the relationship between them is unknown, making inference more difficult.

As with the other processes, we use GPs to represent the rating curves. GPs are flexible enough to capture changes of the rating curve at different flow regimes. However, to ensure a robust fit, we require the following three adaptions of the standard GP:

One. Power transformation. As stage-discharge data is often distributed over several orders of magnitude, it is difficult to identify a single meaningful lengthscale for the GP kernel. A simple approach to alleviate this issue is to perform a logarithmic transformation for both stage and discharge to re-scale data into a standard scale. before fitting the model. The use of non-linear transform results in a non-Gaussian likelihood, which might affect our GP regression. However, we assume that for the range of stage- discharge data, the likelihood of log-data is approximately Gaussian.

Two. Composite kernel function. To capture changes in the rating curve over time, we define a composite kernel

$$K_{RC}(h,t),h',t';\ S_{o,}(\ell_h,\ell_t)) \quad = S_o^2 \exp\left(K_h(h,h':\ell_h) \times K_t(t,t':\ell_t)\right) \qquad (1919)$$

that yields the covariance between the pairs of stage and time (h, t) and (h ′, t ′). Separate kernel functions Kh and Kt are used to represent potentially different dependencies in time and stage. Kh and Kt can take different functional forms. However, their hyperparameters (ℓh, ℓtS0) are learned jointly.

Three. Mean function. With limited number of river gaugings, especially of the extremes, a rating curve must provide meaningful extrapolation beyond the available data. In the absence of data, a GP reverts to its mean function m(x). In many cases, the mean is assumed constant. However, a single value cannot capture the rating curve behaviour at both low and high flow extremes. Instead, we use another GP mGP (h) as a mean function to capture a "typical" rating curve at that gauging location.

Formally, mGP (h) GP (p, Km) is a Gaussian process with a mean function p(h) and a covariance function Km(h, h ′), both functions of the stage h only. By fitting data from multiple years, ignoring time dependencies, mGP provides a "timeless" reference when there are big time gaps between river measure- ments. To ensure meaningful extrapolation at high flow events, p(h) is defined as a linear polynomial of the stage. p(h) also constrains mGP to a monotonic functional form.

Figure 4. plots the current (as of December 2021) GP rating curves of the selected stations given gauging data from 2010-2021. Clearly, the amount and distribution of available gauging data differs between stations. The distribution of data in stage and time affects the fidelity and performance of the GP model. For example, the large discharge variations of station 419006 result in a higher process uncertainty, indicated by a wider confidence interval. With all stations, the process uncertainty is mostly constant, though narrower closer to data points. Note that as data has been log-transformed prior to modelling, the true discharge variance (after exponentiation) will no longer be constant or symmetric around the mean. Therefore, the variance increases as the discharge increases.

To assess the effect of rating curve uncertainty on the water balance in our case study area, we examine data at the selected gauging stations collected during the 2019-2020 season. Stage data, as recorded by WaterNSW and shown in Figure 4.a, is converted into discharge using a prescribed rating curve. Figure 4.c depicts the hourly discharge estimates based on a GP rating curve conversion. WaterNSW's discharge estimates based on the official rating curves, presented in Figure 4.b as a comparison, mostly fall within the 95% confidence interval of the GP-based estimates (shaded regions). Differences at very low flows are due to reported negative water levels, which had to be replaced by a fixed value to ensure a positive input to the log-transform. This indicates that the GP model is consistent with WaterNSW's rating curves.

Another benefit of the GP representation is the explicit use of time as a parameter of the model. This means that every data point is converted with a slightly different rating curve depending on the date of measurement. This also affects the uncertainty estimates, such that uncertainty increases for predictions further in time from the actual gauging data . The current GP model assumes a smooth kernel over time, leading to gradual changes in rating curves with a rate of change that depends on the temporal lengthscale. While our framework can also accommodate non-smooth transition in rating curve via GP approximate inference, we leave this for future work.

The uncertainty about the rating curves propagate into the annual estimates of Qin and Qout. We estimate the probability distribution over p(Q) by sampling possible rating curves fRC from our GP model fRC GPRC. The entire level data series (as in 11a) is converted using fRC and summed over to produce samples of p(Q) fRC. Figure 4.a and Figure 4.b shows the distribution over the annual in and out flux during 2019 and 2020, respectively. The uncertain rating curve of 419006 affect the overall Qin uncertainty in both years. In contrast, the low uncertainty of 419001 and the regular flow at that station leads to a narrow distribution over Qout. While the distribution over Qin is quite larger, it is clear that 2019 was a losing year with additional losses of approximately 3000ML in that river reach. Rainfall during 2020 resulted in a gaining river with a surplus of roughly 40K ML.

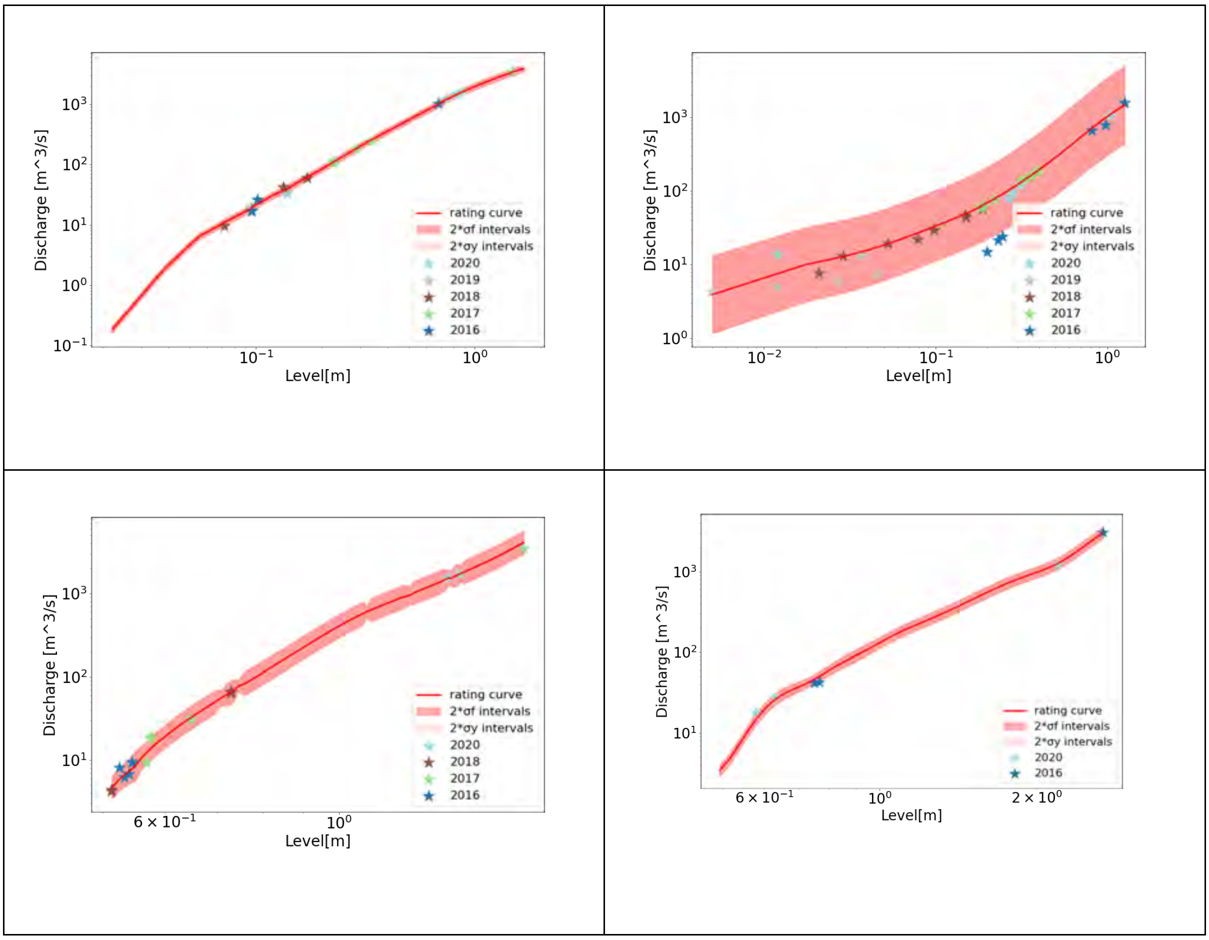


Figure 4.3934: GP Rating curves for gauging station along the Namoi river between Keepit dam to Gunnedah. All rating curves are based on a nested Gaussian process with a stage/time composite kernel and a linear mean function. All gauging data from 2010 has been used to fit the GP rating curve model. The solid line is the posterior mean of the GP, the dark shade is the process 2σ interval and the light shade the likelihood 2σ interval.

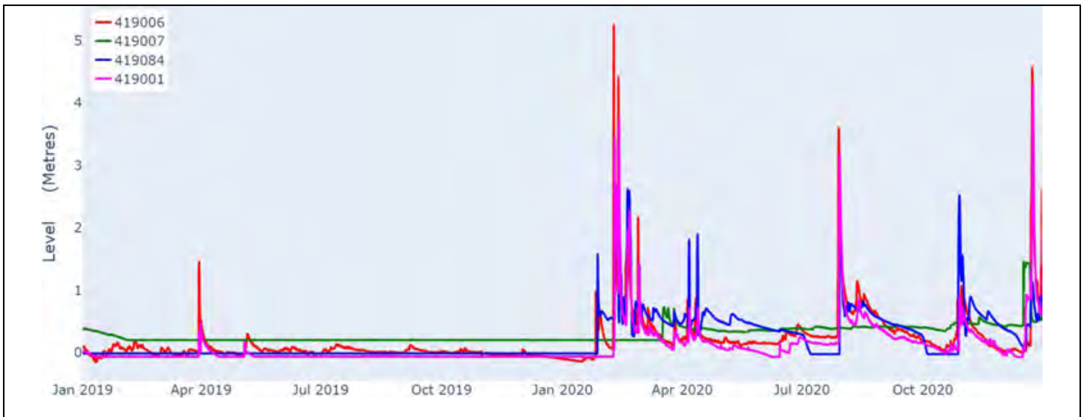
4.4.6 Discussion and Conclusion

In this report we develop a pilot for components of the probablistic modelling framework proposed in the Milestone 1 report. This demonstrates that a Gaussian Processes (GP) approach can be used to model several of the processes and quantify the uncertainty. This is because GPs can model the covariance in both space and time, and between space and time. Here we only demonstrate a pilot quantification of groundwater surface water interactions, evapotranspiration losses and rating curves. While a Bayesian approach to many of these processes is not new (e.g., [10, 18]), incorporating this as part of an overall probabilistic framework is a novel approach.

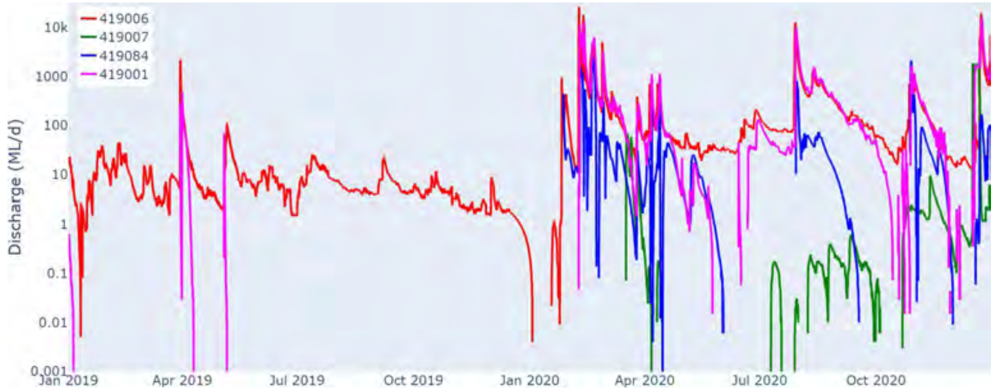
The current pilot only demonstrates three components of the overall system to highlight how the GPs can be used to deliver quantification of the uncertainty. For a full river reach system this needs to be integrated in a complete water balance, which can be constrained to sum to 0 using the likelihood (e.g., [16]).

In Figure 4.c, we show how the derived rating curves can be used to predict streamflow and uncertainties using the recorded waterlevels. These values then will feed into the evaporation modelling (3.1.2), in particular to define the variable river widths for each flow rate. It will also determine the river pressure heads (3.2) and subsequently the groundwater surface water quantification to develop the complete estimation of uncertainty.

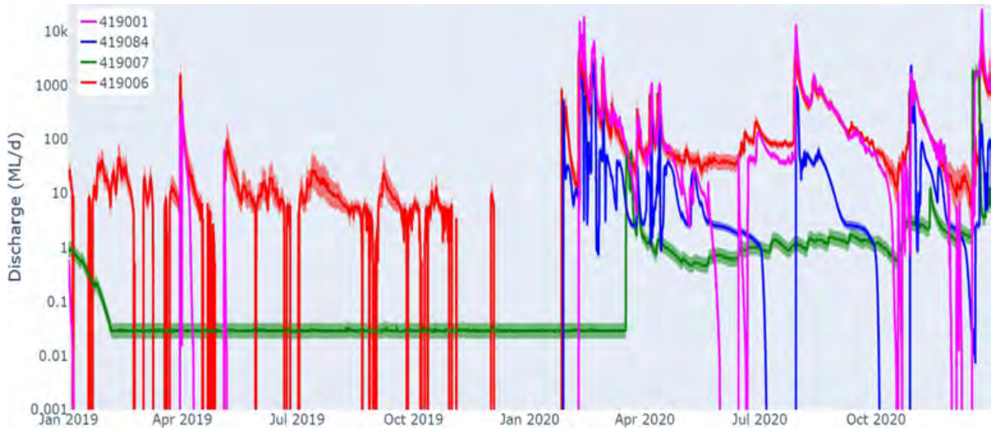
In conclusion, the results of the first pilot of the probabilistic modelling framework show a clear ability to quantify uncertainties based on the application of GPs to the spatial temporal problems. With well chosen priors and smoothing kernels these functions are flexible enough to capture the variation in the data, and fit both time and space covariances.



(a) Hourly level /stage measurements

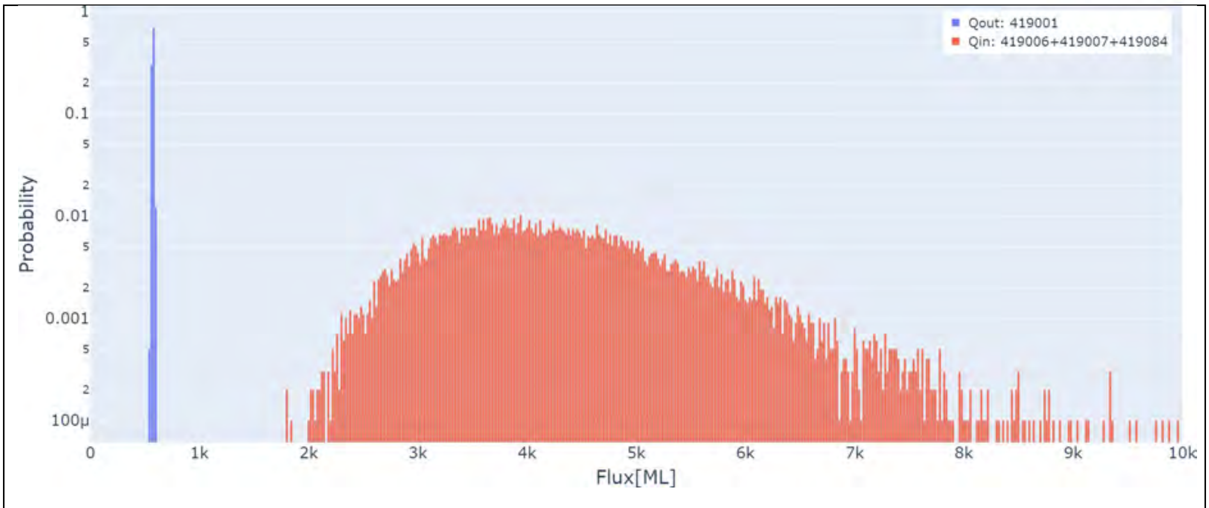


(b) WaterNSW discharge estimates

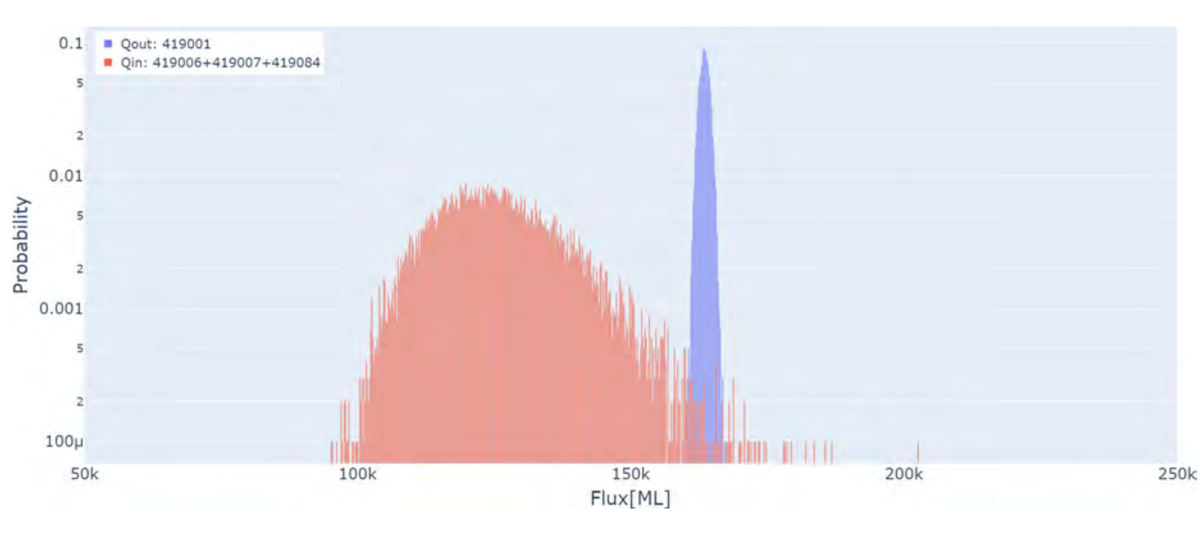


(c) Discharge estimate using GP rating cuves

Figure 4.C35: The Namoi river's flow during the 2019-2020 drought. (a) Hourly river level observations down- loaded via WaterNSW's webservices (https://realtimedata.waternsw.com.au/). (b) Hourly discharge esti- mates based on WaterNSW's rating curves (WaterNSW webservices). (c) Discharge estimates and uncertainty calculated using the derived GP rating curves. The mean of the GP (solid line) The Shaded region represents the 95% confidence intervals around the mean (solid line). The GP produces discharge estimates that are similar to the official discharge reported by WaterNSW's. Differences at very low flows are due to reported negative water levels (missing values in a due to logarithmic scale). These had to be replaced by a fixed value (1e − 9) to ensure a positive input to the log-transform.



(a) 2019



(b) 2020

Figure 4.a36: Namoi - Keepit to Gunnedah - Annual streamflow balance estimates during (a) 2019 and (b) 2020. Each estimate is calculated using a sampled rating curve drawn from the GP model. The distribution over Q_{out} is tight due to low uncertainty about the rating curve of 419001 and the low variability of the measured water level during 2019-2020. The variance of Q_{in} is much larger due to the uncertain rating curve of 419006.

Table 4.7: Table 1: Estimated effort on model components.

Process	**Weeks**	**Urgency**	**Comments**
Hierarchical conceptual model of catchment	6–9	A	Includes all river reaches and other hydrological sub-units as nodes, as well as link topology (which may include some un- certain edges); work with hydrologists

Rainfall input (gridded)	6–9	B	Mostly engineering; 1st pass model is GP but hold bet- ter models in reserve – suitable kernel may be in question; needed for the Macquarie University sensor contribution; nota dominant uncertainty on its own, but is a common factor for other uncertain flows
Land-based evapotranspi- ration (gridded)	6–9	B	Mostly engineering; use GP infrastructure but this time mean function and kernel may be non-trivial; needed for the Mac- quarie University sensor contribution; not a dominant uncer-tainty on its own, but is a common factor for other uncertain flows
Stream and lake evapora- tion (sensed area)	6–9	B	Mostly engineering; known remote sensing solution from [4] can most likely be used on rivers as well – extract rivercourses from BOM Geofabric, include reach data; significant single-factor contribution to uncertainty in some catchments
Ungauged inflow from rainfall-runoff model	12–18	B	Engineering and hydrology input; inference on model pa- rameters to include uncertainty; perhaps Bayesian model av- eraging approach if conceptual model uncertain; usually un- certainty is subdominant because it involves small volumes
Ratings curve for dam stor- age	6–9	A	Mostly engineering on existing ratings curve data; random function method such as GP or functional PCA is good ap- proach; this curve will introduce correlated uncertainty into all releases with dam at given water level, so assess further impact on estimates of downstream flows
Ratings curve for stream flow	6–9	A	Mostly engineering on existing ratings curve data; random function method such as GP or functional PCA is good ap- proach; this curve will introduce correlated uncertainty into all releases with dam at given water level, so assess further impact on estimates of downstream flows
Reach length groundwater losses	12–18	A	More challenging problem involving numerical groundwater models though can be lumped conceptual if not 3-D; concep-tualization may be uncertain so Bayesian model averaging may be necessary for this; data include groundwater bores, soil information; includes "river-aquifer" interactions flagged by UNSW, and ANU gravity could help constrain if hydroge-

			ology prior is secure
Reach length extraction	6–12	B	Regression problem; expert input from water management
Reach length environmental water use	12–18	B	Regression problem; data science + hydrology + remote sens- ing?
Groundwater use	6–9	C	Basic forward model problem; mostly engineering
Gravity inversion with strong geology prior	18–24	A	Engineering + data science + hydrology; effort to agree upon and parameterize suitable forward model
Gravity inversion with limited geology prior	24–36	B	Experimental; recommend water tableomography approach, test in well-instrumented area where poor coverage can be simulated by removing data

4.4.7 Further work and estimated effort

In the Milestone 1 report we outlined a draft program of work and time/cost investment to fully develop the Probabilistic modelling framework. Here we repeat this in Table 1 which represents the estimated amount of effort to deliver each of the sub-models in a larger probabilistic model of the water balance. We consider data science model-building effort, expert hydrology effort, and software engineering or infrastructure building. We also give each item a priority ranking of A (most urgent) to C (least urgent) if resources are scarce.

In this current report we have already delivered parts of Table 1, in particular:

- a pilot rating curve uncertainty quantification (4.1)
- a pilot uncertainty quantification of evapotranspiration from stream and riparian zone (3.1.2)
- a pilot uncertainty quantification of groundwater surface water interaction (3.2)

However, many of the elements in the table are clearly "future work" and therefore the estimates in the table can still guide further development.

Some items are relatively simple and can be delivered in under 2 months (2-3 sprints). These are mostly well-established data science methods (parametric regression models or Gaussian processes) being applied to quantify uncertainty in well-understood hydrological processes. The technical risk for these items is low, and the effort is mostly engineering time involved in collating data and validating the solution. These items also tend not to dominate the uncertainty in our initial investigations, although their uncertainty is still not at present rigorously quantified and thus the contribution of other uncertain items will become more strongly constrained if these are understood more precisely.

Of the quickly delivered items, ratings curve uncertainty is well-posed and relatively straightforward to implement, while also contributing substantially to the uncertainty in certain catchments with large volumes of water in dam releases or gauged flows. In this current report, we demonstrate a pilot of the ratings curve uncertainty (4.1). Uncertainties in the ratings curve will also introduce correlations between downstream flows that will inflate final uncertainties in the water balance, and will not average out as the number of observations increase. Ratings curve uncertainty therefore seems like a relatively low-risk, high-value investment.

Other items are necessarily more experimental and higher risk, fusing conceptual or numerical hydrological models with random functions and sampling the uncertainties via MCMC. These are usually addressing processes that are much less clearly understood and, unsurprisingly, are thus given longer time horizons. Groundwater models, including those for which cold atom gravity sensing will be relevant, fall into this class. While it is not difficult to generate gravity signals from a given model of rock properties, the amount of effort that goes into specifying priors on realistic rock property models and validating them even in a well- instrumented context is expected to be larger than for the other components.

Delivering a fully developed end-to-end working prototype, we recommend concentrating only on the priority rank A items. Each of these tasks can be pursued independently in any order, but are each scoped at around 2 FTE — one data scientist, one half-time hydrologist, and one half-time research engineer. Thus, for example, we expect 2 FTE of effort can deliver either the high-priority water balance hydrological submodels, or preliminary gravity inversions for groundwater systems, on that timescale. With 4 FTE of effort (2 data scientists, 1 hydrologist, 1 engineer) both of those programs could be attempted.

4.5 Groundwater Hydrological Processes

4.5.1 Introduction

Understanding the groundwater and surface water interaction is important for sustainable management of water. Groundwater recharge is one of the components of the water balance that is notoriously hard to determine as it is difficult to measure directly (experimentally) because it is taking place in the subsurface. Its estimation is also associated with large spatio-temporal uncertainties due to variability in climate and heterogeneity of soils and hydrogeology.

There are three fundamental mechanisms of groundwater recharge, diffuse recharge where the groundwater table is directly recharge by rainfall through the general land-surface, focussed recharge where the groundwater table is recharged by flow in stream/river channels and flood plain recharge during large floods (overbank flows). In this work we will focus on the two first mechanisms of recharge.

In the context of the WIATW project, we are considering different hydrologic scenarios labelled 'Use Cases'. These use cases relate to both stakeholder water balance issues defined by WaterNSW, NRAR and NSW-DPIE and physical climatic and hydrologic conditions. While in reality surface water flow conditions and interactions with the subsurface represent a continuum of climatic conditions (different degrees of antecedent dryness) and sizes of flow events (i.e., controlled dam-release to large floods), we are here simplifying the complexity by defining a 'low flow use case' and 'high flow use case' according to section 3.3.1.3 of this report. In summary the low flow use case explores transmission losses during very dry periods and the high flow use case "covers a flooding event or a very wet period". In our report we will attempt to relate our conceptual understanding of surface water groundwater interactions and recharge processes to these use cases. For the low-flow case, transmission losses can be due to groundwater recharge, vadose zone storage, and/or evapotranspiration. Hence our work around groundwater recharge can inform both use cases.

4.5.2 Methods

4.5.2.1 Study Sites and installations

For this project data from two NSW NCRIS sites were used, Wellington Research Station and Maules Creek Catchment (see Figure 4.40). Each of these sites contain an extensive network of piezometers where the groundwater levels have been monitored for over 10 years. A combination of self-contained (Solinst) and vented (In-Situ) groundwater level loggers were deployed in the boreholes and set for logging at 30 min intervals. The water level data from the non-vented loggers (Solinst) were corrected for barometric pressure variations using atmospheric pressure data from a Solinst baro-logger installed in the airspace of a borehole.

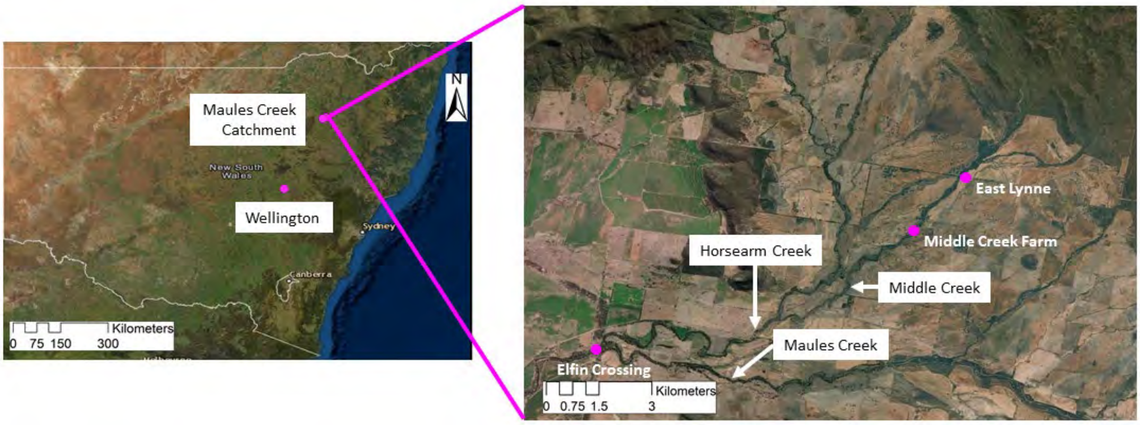


Figure 4.4037: Map of the NCRIS sites.

The Wellington NCRIS site is adjacent to the Macquarie River and consists of a range of piezometers on a profile that are located from the river flat up the hillside (see Figure 4.41). The geology consists of a consolidated bedrock of Silurian to Devonian meta-sediments and meta-basalts into which the Macquarie River has cut a valley. This Valley is partly infilled with up to 40 meters of unconsolidated alluvial sediments (gravels, sands, silts and clays). On the hillslopes there is a weathering profile of predominantly clays and colluvial deposits [33]. For this work the focus was on the hillslope bores as their only recharge mechanism is diffuse recharge, whereas the river flat bores have a mixture of diffuse and focused recharge (which is difficult to untangle). The construction details of relevant piezometers used in this work are given in the Data section (Table 8-1). Some of the hillslope piezometers are located in fractured rock and are uncased bores without screens and hence water enters these bores through the various fractures that are present in the basalt over various depths of the bores.

Figure 4.4138: Site map of Wellington NCRIS site. The Hillslope bores does not respond to changes in the river level (as opposed to the river flat bores) and hence can be used to analyse events of diffuse recharge.

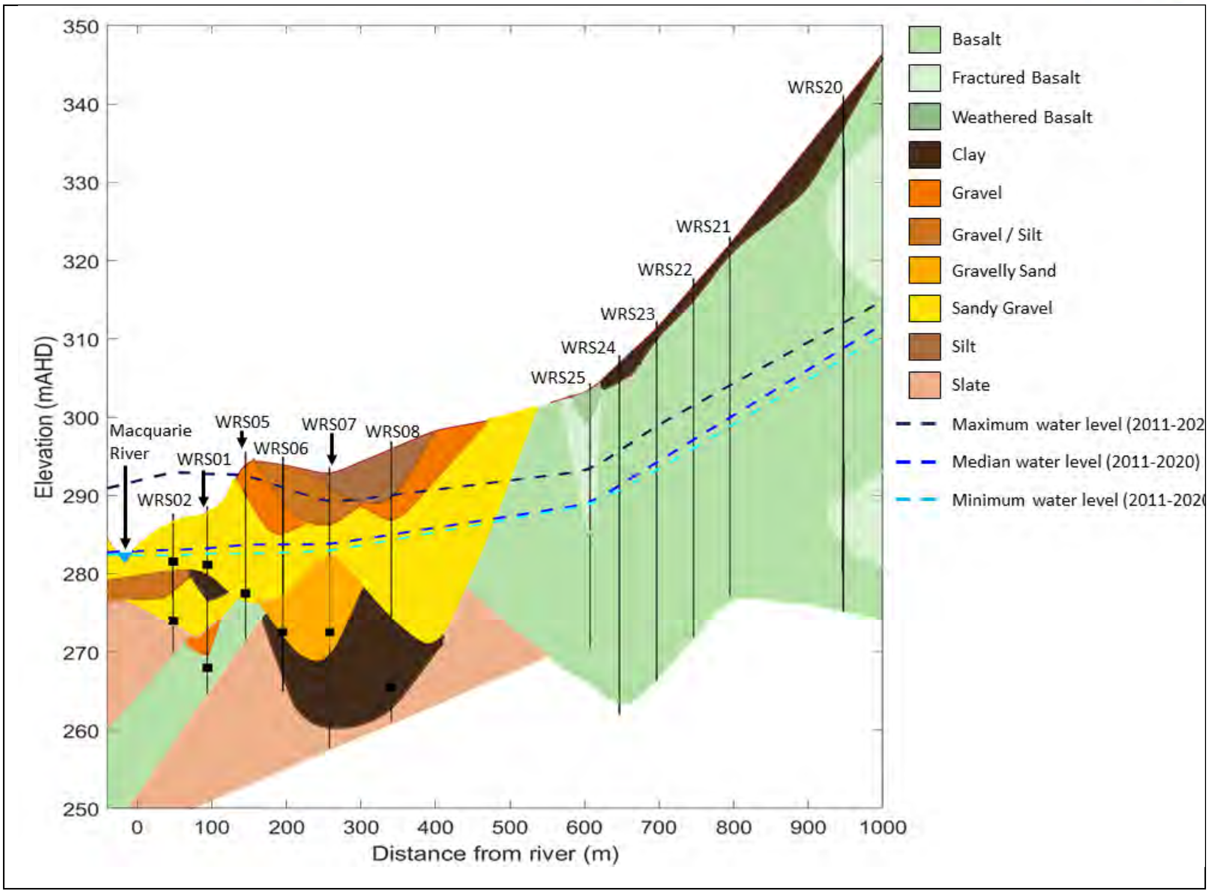


Figure 4.4239: Cross section of Wellington showing the piezometer locations and screens (for alluvium piezometers) and maximum, median and minimum groundwater levels recorded between 2011-2020.

The studied part of the Maules Creek Catchment contains three creeks, Middle Creek that joins Horsearm Creek which then joins Maules Creek that is a tributary to the Namoi River. There are three NCRIS sites, two which are adjacent to Middle Creek: East Lynne and Middle Creek Farm, and a third site, Elfin Crossing which is adjacent to Maules Creek as shown in Figure 4.40. Middle Creek is an intermittent creek with periods of flow lasting days to months after significant rainfall in the mountains upstream [34]. During periods of no surface flow the groundwater table recedes below the creek creating an unsaturated zone below the creek bed. Elfin Crossing is located at a perennial section of Maules Creek downstream of the confluence of Horsearm and Maules Creek. It has zones of persistent groundwater discharge further upstream maintaining flow all year round. Figure 4.43 shows the cross section of the sites and the locations of the piezometers utilised in this study, with their details given in the Data section (Table 8-2). The geology consists of up to ~30 m of unconsolidated alluvium (cobbles, gravels, sands, silts and clays) deposited on mainly Permian sandstones, claystones, conglomerates and coal-measures [35]

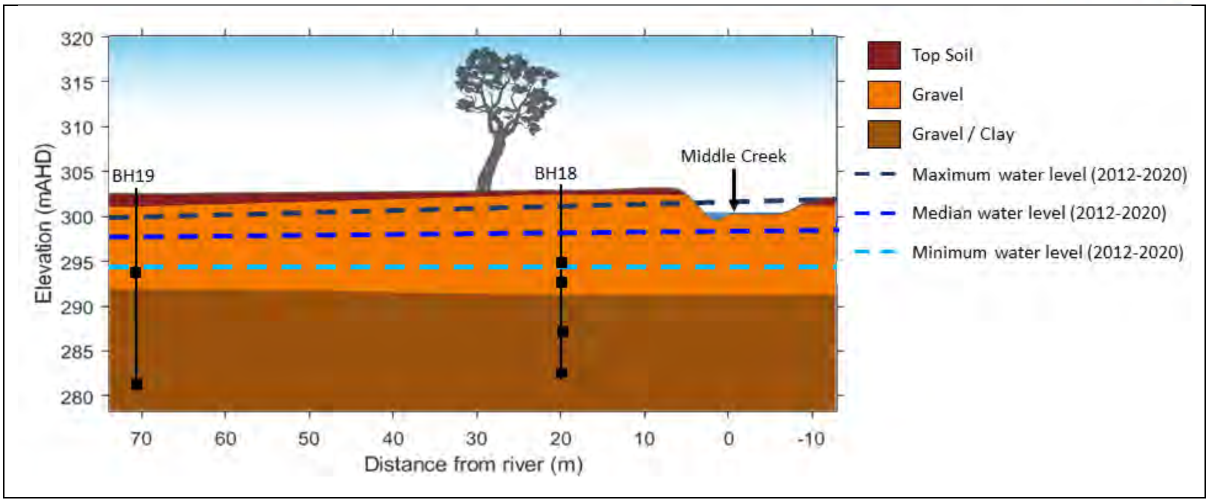


Figure 4.4340: Cross section of Middle Creek Farm showing the location of piezometer screens (black squares) and maximum, median and minimum groundwater levels recorded between 2012-2020.

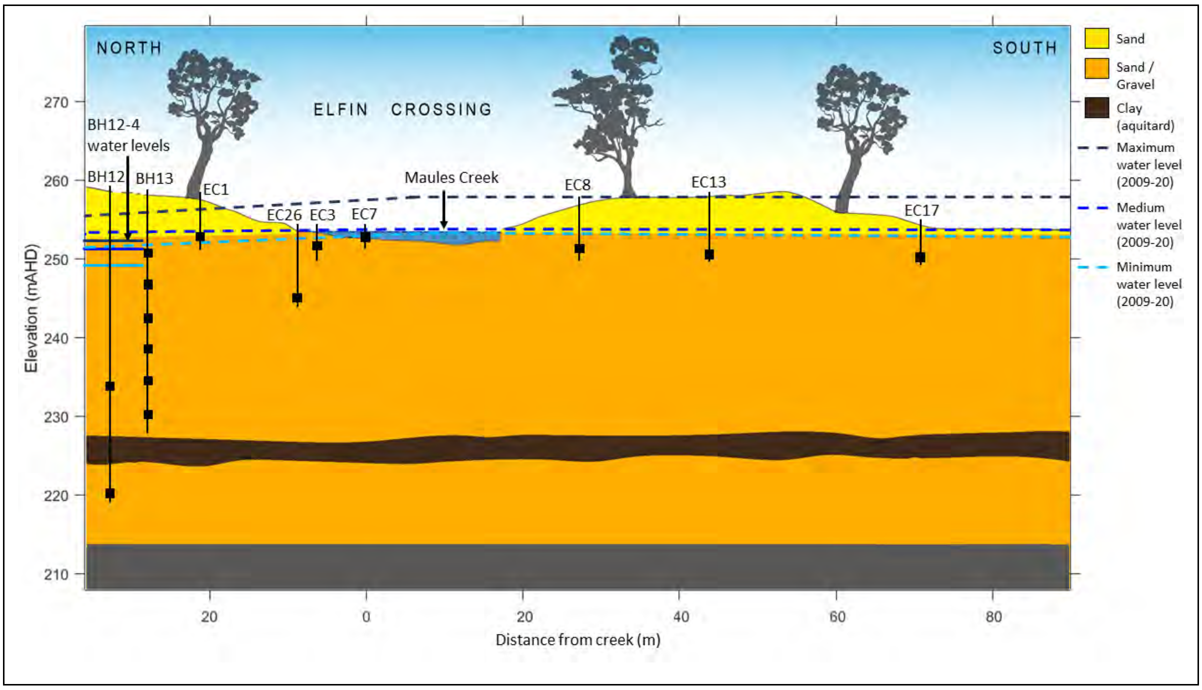


Figure 4.4441: Cross section of Elfin Crossing showing the location of piezometer screens (black squares) and maximum, median and minimum groundwater levels recorded between 2009-2020.

4.5.2.2 Soil Moisture Balance (SMB) Model

To explore a basic temporal water balance model and periods of potential groundwater recharge a soil moisture model was used. The model used was originally applied by Cuthbert, et al. 2014 [36]at Wellington Caves (about ~10 km from the Wellington Research Station) and is shown schematically in Figure 4.45, where for the purpose here Rch represents groundwater recharge and Dr drainage to the deeper system. For this project the shallow karst store (S2) was replaced with a shallow vadose zone. This model, also known as a bucket model, gives flexibility in how it is set up with the number of layers ('buckets') able to be chosen. The input data for this model are the timeseries of rainfall and calculated potential evapotranspiration. For the various soil related parameters in the model, the values that had been determined for the Wellington Caves site were used and are given in Table 3. These values were chosen as the Wellington Caves site has similar soil conditions to the Wellington Research Station and hence it was believed that these values would be also representative of the Research Station.

Table 4.8: The parameter values used for the Wellington SMB model. These values were taken from Cuthbert, et al. 2014.

Parameter	Description	Value
θFC	Field capacity of soil	25%
θWP	Wilting point of soil	5%
Ze	Soil depth subjective to evaporative drying	0.1m
Zr	Rooting depth	0.47m
Kc	Crop efficient	1
P	Readily Available Water to Total Available Water ratio	0.5
B	Proportional of bare soil	0.1

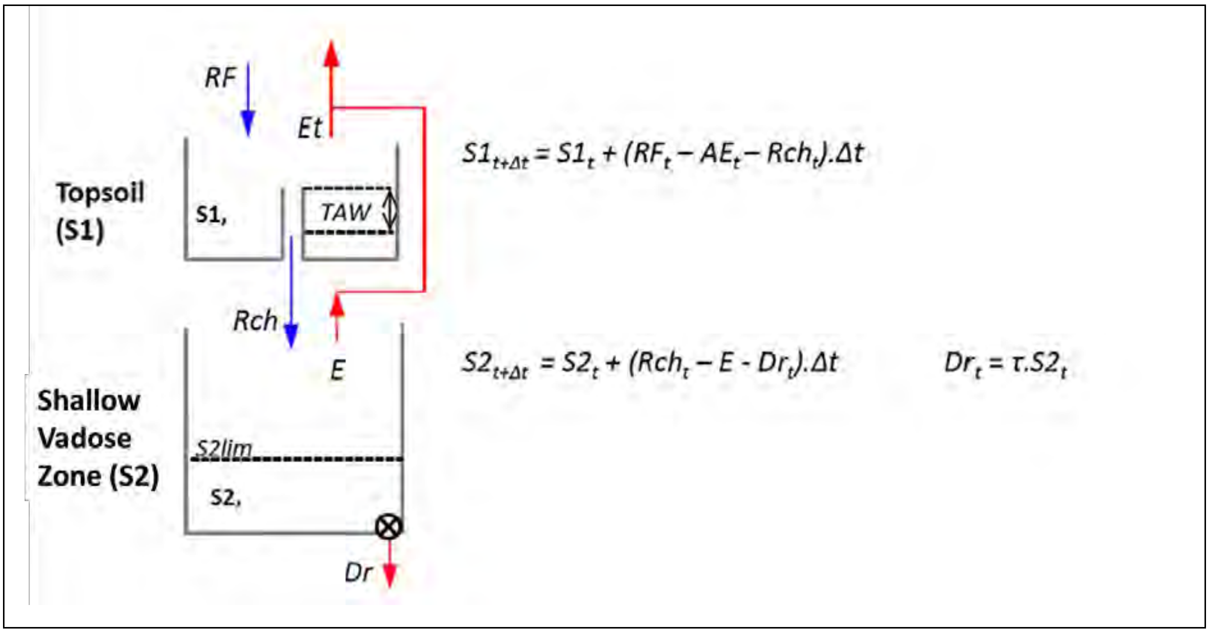


Figure 4.4542: Graphical representation of the Soil Moisture Balance model adapted from Cuthbert, et al. 2014.Where RF represents infiltration from rainfall, Rch groundwater recharge, Et evapotranspiration, E evaporation, Dr drainage to the deeper system, TAW total available water in soil, S1 soil store, S2 shallow vadose zone store.

4.5.3 Results

4.5.3.1 Wellington

Based on the timeseries of groundwater levels for the hillslope bores at Wellington Research Station, potential diffuse recharge events were identified as shown in Figure 4.46. The diffuse recharge events were identified based on an increase in the groundwater level in multiple bores and if they displayed a typical recharge profile (i.e., a sharp initial increase followed by a longer recession) and matched with observed rainfall events. There is one significant recharge event in 2016 across all bores, comprised of seven individual recharge events. There are four smaller events over the study period, though some of these are not observed widespread across the site, suggestive of spatial variability in diffuse recharge, where the recharge threshold is only exceeded at some sites.

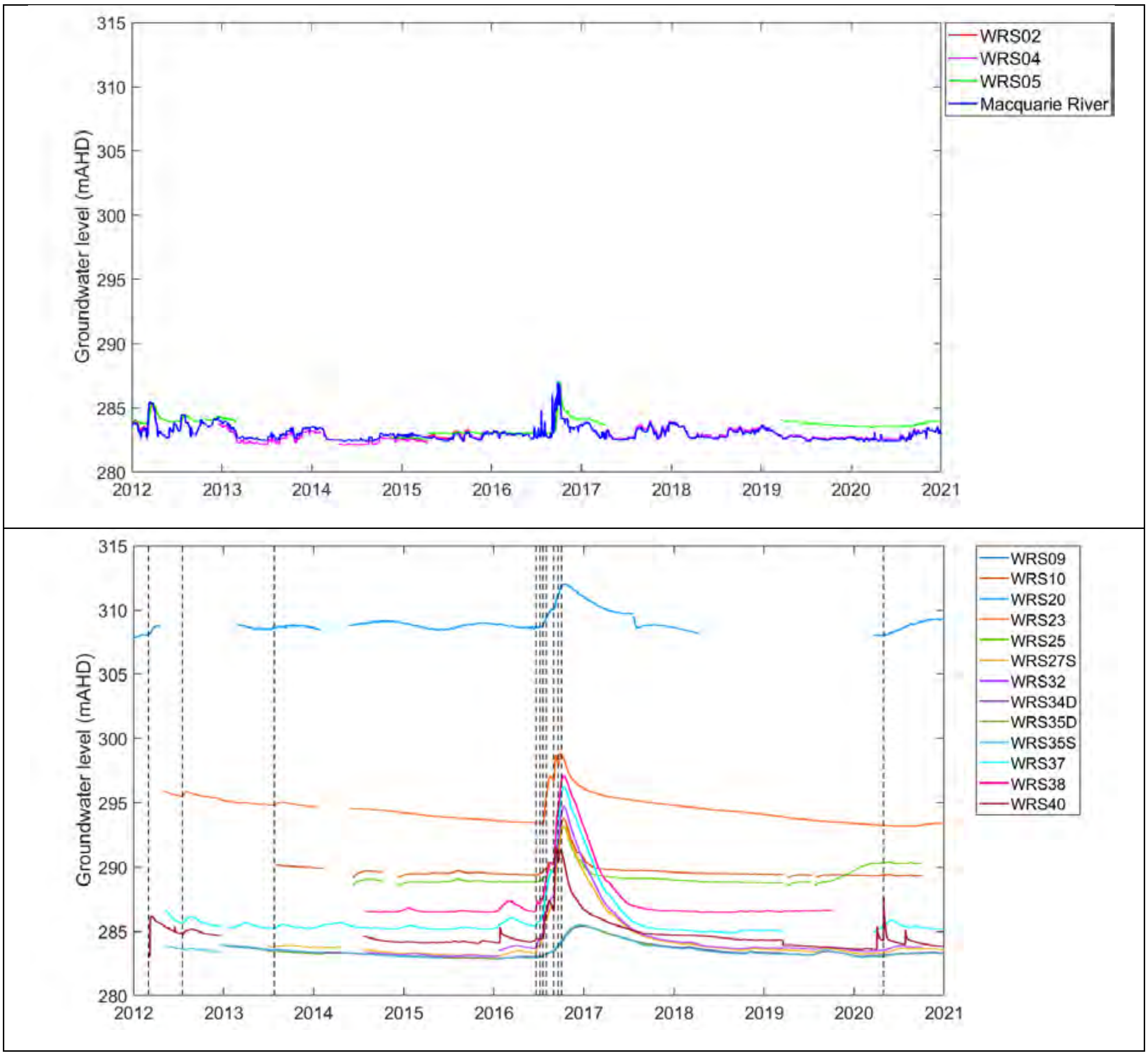


Figure 4.4643: Top: Surface water (data from WaterNSW) and groundwater levels for river flat monitoring bores at the Wellington Research Station. Bottom: Groundwater levels for hillslope monitoring bores at Wellington. Potential diffuse recharge events are indicated with a dashed black line.

4.5.3.2 Maules Creek Catchment

Middle Creek Farm Site

The measured timeseries of surface water and groundwater levels at Middle Creek Farm are shown in Figure 4.47. Two nested bores are located at this site BH18 which is about 20 m from the creek and BH19 which is about 70 m from the creek. Over this period from 2012 to 2020 there are ten occurrences of significant increase in the groundwater levels (ranging from 1 to 4 m), and all of these correspond with periods of sustained stream flow. BH19 shows a damped and lagged response to flow in the creek compared to BH18. During flow periods there are smaller recharge events where there is a high frequency (i.e., 'spikey') response in the groundwater to changes in surface water levels.

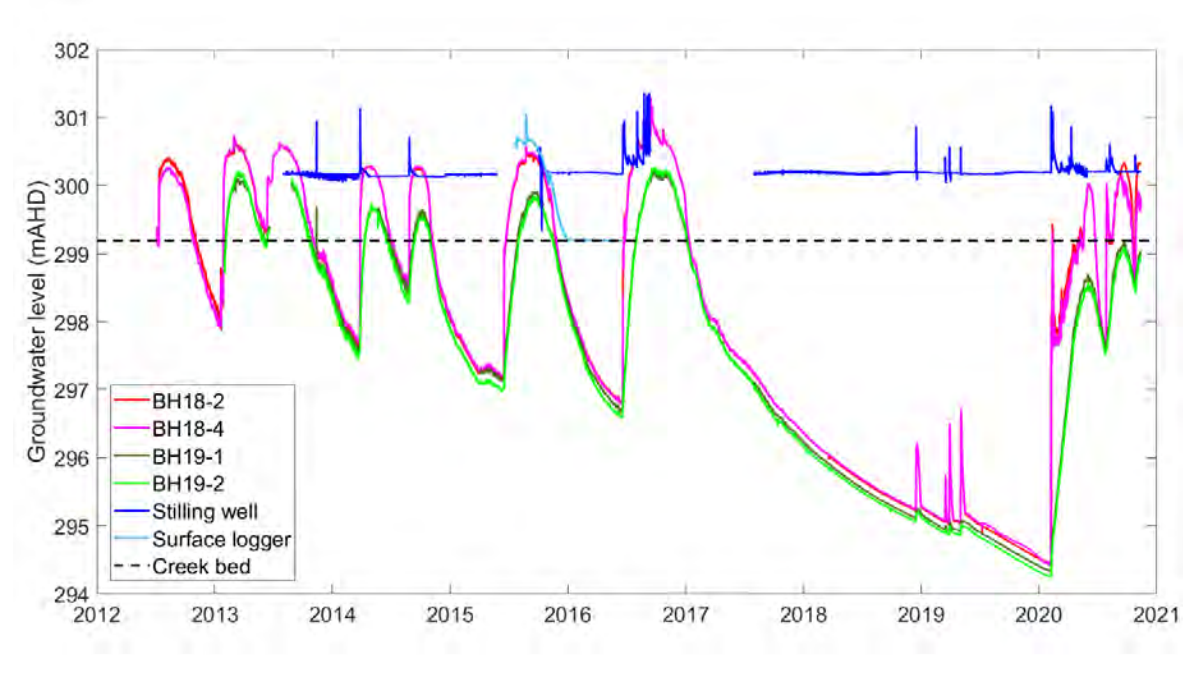


Figure 4.4744: Surface water and groundwater levels at the Middle Creek Farm site. Surface water levels are measured by a logger in the stilling well placed 1 m above the deepest part of the creek bed. Note that when the stilling well signature (blue line) is horizontal there is no flow in the creek.

Elfin Crossing

The measured groundwater levels at Elfin Crossing are shown in Figure 4.48. At this site, as reflected in the timeseries, the groundwater table of shallow bores remains above the creek bed throughout the year (except for a period between 2018 and 2020). Due to the general permanence of surface water, the stream at this section is classified as a perennial stream. The connection between the groundwater table and the creek water levels is therefore maintained over time resulting in more constant groundwater levels compared to Middle Creek Farm. These generally constant groundwater levels are interrupted by high frequency (i.e., spikey) groundwater level responses to flood events in the stream.

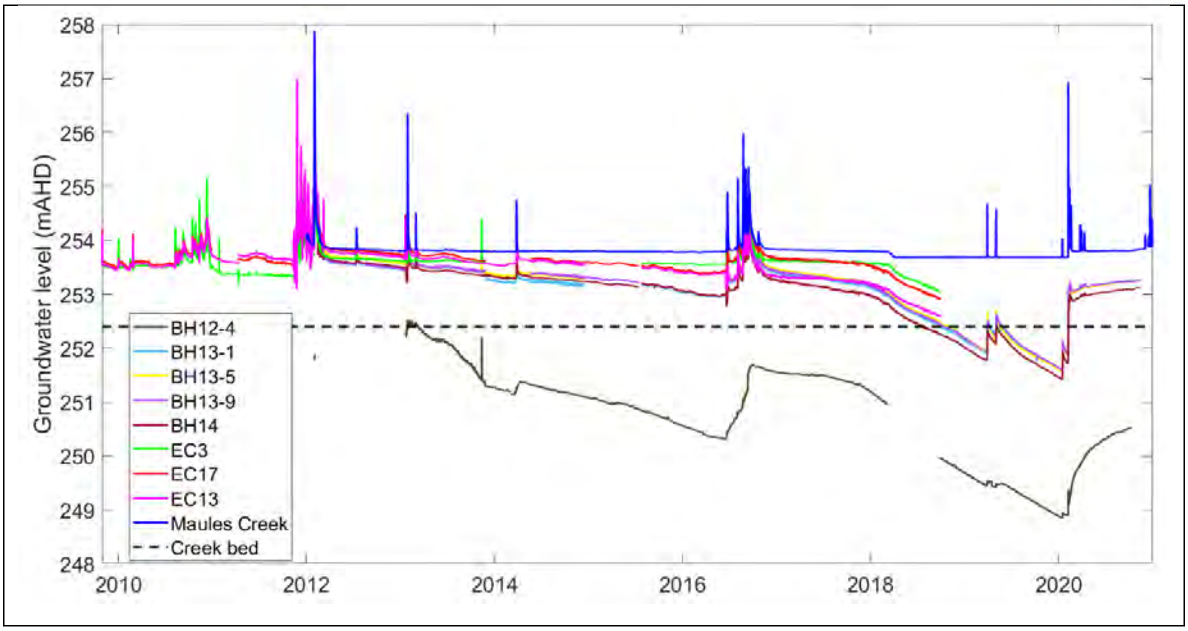


Figure 4.4845: Surface water levels (data from WaterNSW) and groundwater levels at Elfin Crossing on Maules Creek.

4.5.4 Discussion

4.5.4.1 Hydrological Conceptual Models

Based on the groundwater timeseries from the different sites, hydrological conceptual models were developed to represent different groundwater processes. These processes have been related to the different use cases explored in the WIATW project. In intermittent streams, such as Middle Creek, especially after a dry period without surface flow, the groundwater table is lower than the stream bed (i.e., disconnected), as shown in Figure 4.47. When stream flow is generated by rainfall upstream there is immediate response in both BH18 and BH19, however the peak in BH19 is damped and lagged. As BH19 is further away from the creek this is consistent with focussed recharge from the creek as the surface water infiltrates down to the water table and laterally away from the stream as represented in Figure 4.49a.

In perennial streams the surface water and groundwater levels are in a dynamic equilibrium, and different effects are observed to flow events. At Elfin Crossing there are high frequency groundwater levels responses to increases in the surface water level as the continuously saturated subsurface and the creek buffer each other, as a two-way connection exists. Therefore, for perennial streams a flow release may lead to bank storage effects with the stored water returning to the stream after the flow-release (unless large groundwater abstraction is happening at the same time), represented in Figure 4.49b and c. This same behaviour can be observed at Middle Creek in 2016 - during a period of sustained surface water flow - as the groundwater table during that period connected to the stream. So, while Middle Creek is classified as an intermittent stream, it can temporarily behave as a perennial stream connected to the groundwater table at times of substantially wetter conditions.

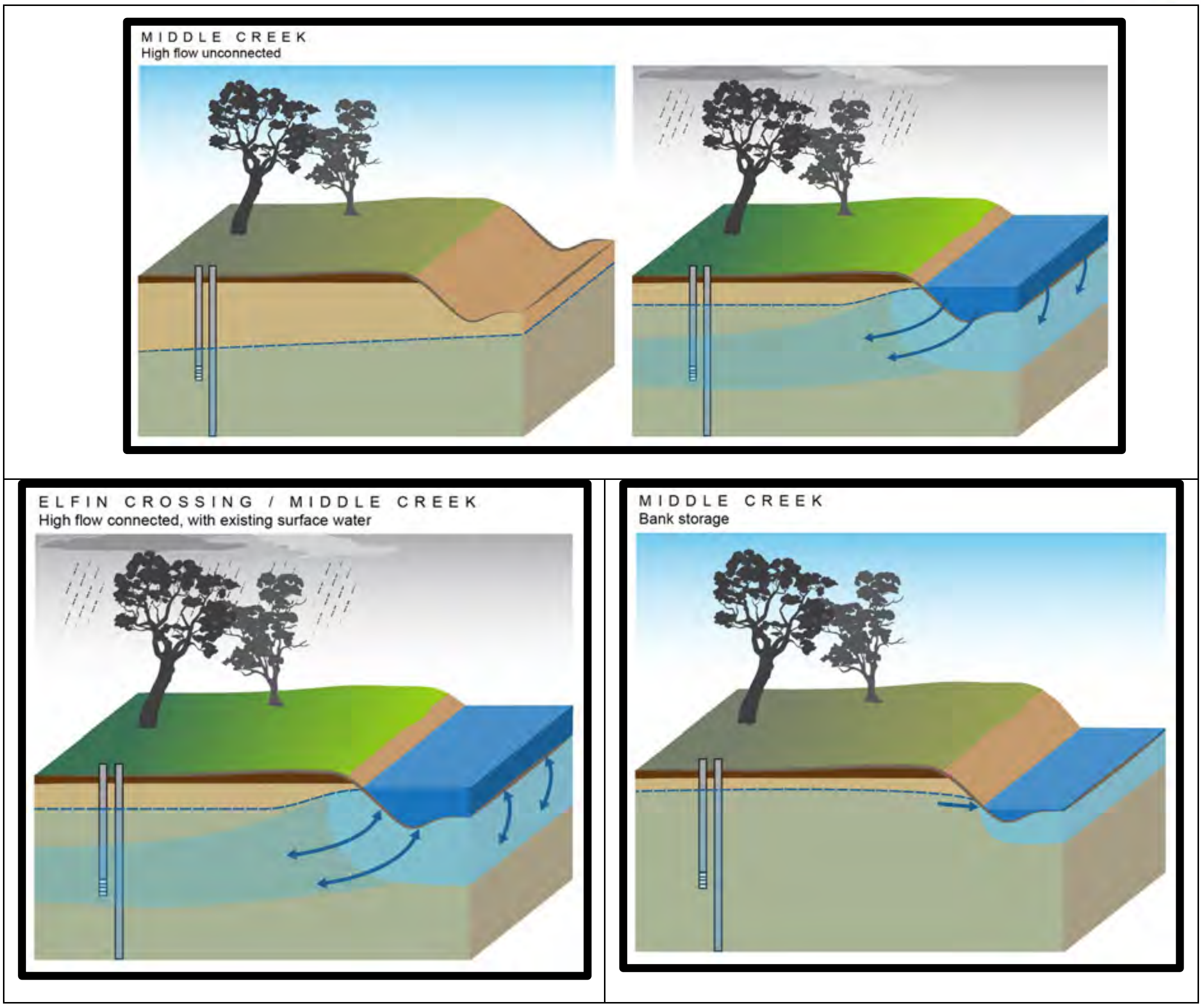


Figure 4.4946: Hydrological conceptual models for the high flow use case. A shows the groundwater response to a flow event with the groundwater table prior to the event some distance below the creek bed (i.e., as observed in intermittent creeks). B shows the groundwater response to a flow event where the stream level and groundwater levels were connected prior to the event. C shows the return of water stored in the creeks banks as the surface water flow event in B recedes. The behaviour seen in B and C are typical for perennial streams.

4.5.4.2 Soil Moisture Balance Model

Wellington

To determine times of diffuse recharge a soil moisture balance model was used. The results of the Wellington SMB model are shown in Figure 4.50. The calculated periods of diffuse recharge from the Wellington SMB model and the identified periods of recharge from the groundwater level timeseries are given in Table 4.9. Of the recharge events determined by the SMB model all, with the exception of the two in 2020, correspond to recharge events observed in the groundwater timeseries. There are three additional potential recharge events identified in the groundwater timeseries, though one of these is shortly (20 days) after the recharge events predicted by the SMB model in 2020.

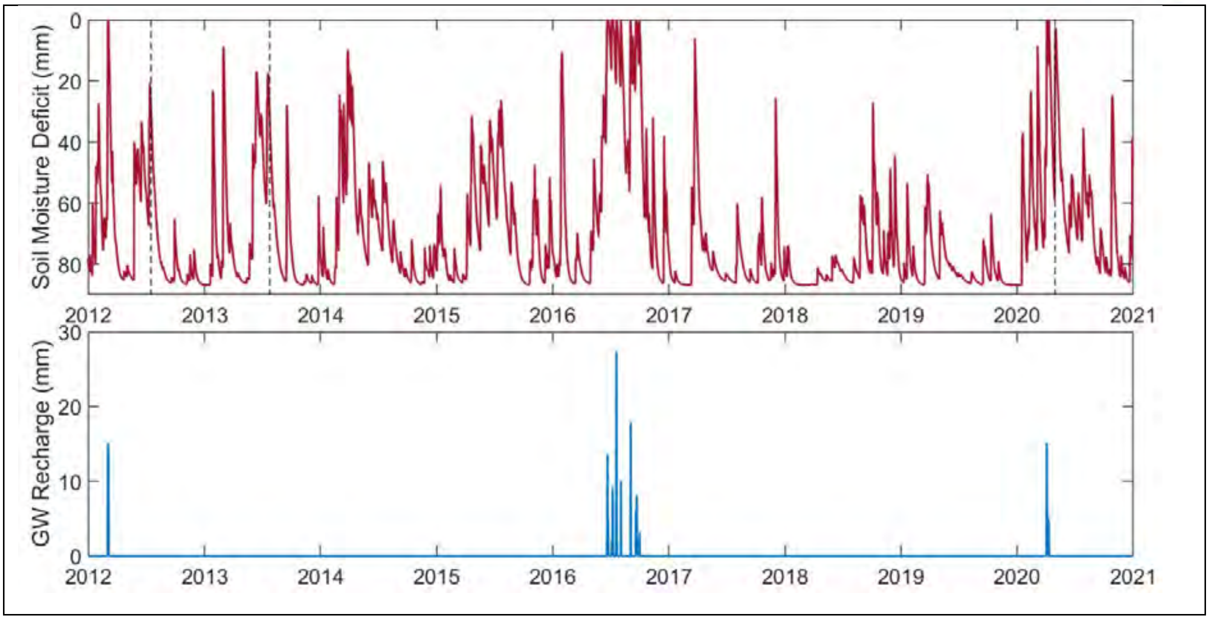


Figure 4.5047: The calculated soil moisture deficit and groundwater recharge from the Wellington Soil Moisture Balance model (SMB) model. According to the model, in principle, recharge can only happen when the calculated soil moisture deficit reaches zero and water is free to percolate to the water table. The potential diffuse recharge events that are in addition to those identified from this model are indicated with vertical dashed black lines.

Table 4.9: The identified recharge events for the Wellington Soil Moisture Balance model (SMB) and groundwater bores, with amount of rainfall for that day and 7, and 14 days prior.

Event number	Date of recharge (SMB)	Recharge (mm) (SMB)	Date of increase (GW)	14-day P (mm)	7-day P (mm)	Rainfall (mm)	Observed / bores with data
1	3/03/2012	29.4	6/03/2012	120.0	102.6	44.8	2/2
2			17/07/2012	53.4	53.2	0.0	4/4
3			25/07/2013	53.3	4.9	0.0	8/8
4	21/06/2016	23.1	21/06/2016	80.4	41.1	10.8	10/13
5	6/07/2016	16.3	7/07/2016	46.6	43.2	5.4	7/13
6	20/07/2016	30.1	19/07/2016	78.2	55.8	49.2	10/13
7	3/08/2016	10.0	2/08/2016	41.4	34.8	34.2	7/13
8	3/09/2016	17.9	31/08/2016	94.4	81.8	21.6	13/13
9	19/09/2016	14.4	20/09/2016	61.1	39.4	26.8	9/13
10	1/10/2016	3.2	1/10/2016	83.2	27.2	11.1	5/13
11	4/04/2020	15.2		113.8	102.8	56.0	
12	10/04/2020	7.0		124.4	88.6	21.4	

Middle Creek

The same SMB model was also applied for the Middle Creek Farm site, with the results shown in Figure 4.51. Five recharge events were identified over the period 2012-2020. These also corresponds with flow events in Middle Creek (see Table 4.10). In the groundwater level timeseries, there are seven additional potential recharge events which corresponds to times of lower rainfall totals than the diffuse recharge events and hence these recharge events are likely focussed recharge events. The monitoring bore locations at Middle Creek makes it harder to clearly differentiate between diffuse and focused recharge, than for the Wellington hillslope bores, as there are no bores that are not potentially influenced by the creek.

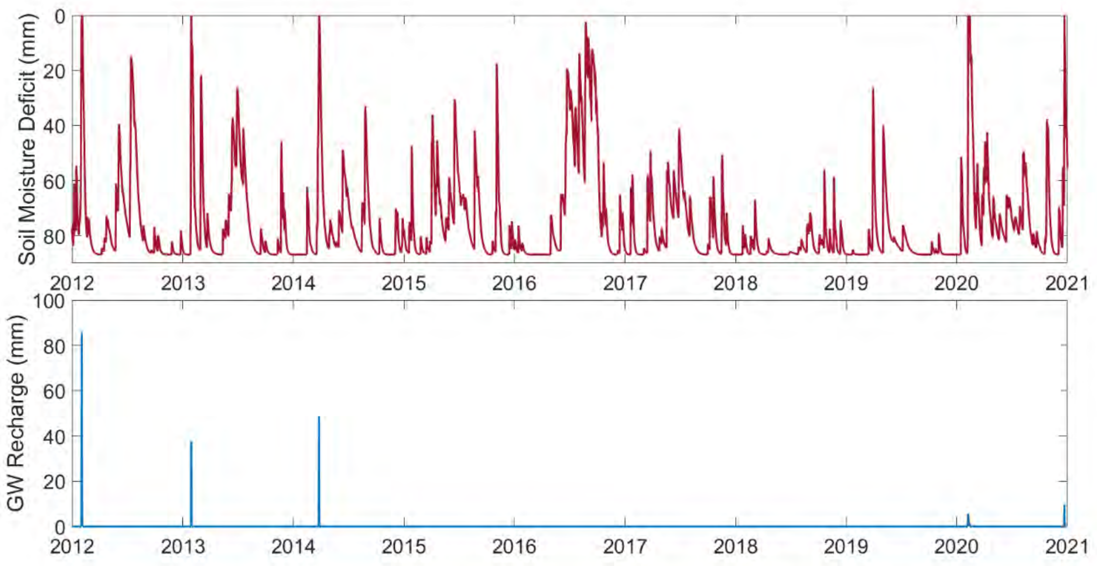


Figure 4.5248: The calculated soil moisture deficit and groundwater recharge from the Middle Creek SMB model.

Table 4.10: The identified recharge events for the Middle Creek Soil Moisture Balance model (SMB) and groundwater bores, based on BH18, with amount of rainfall for that day and 7, and 14 days prior. N/A indicates that groundwater data does not exist for that time period.

Event number	Date of recharge (SMB)	Recharge (mm) (SMB)	Date of increase (GW)	14-day P (mm)	7-day P (mm)	Rainfall (mm)
1	2/02/2012	97.3	N/A	200.1	190.6	92.4
2			11/07/2012	26.2	24.9	24.7
3	29/01/2013	37.7	28/01/2013	140.5	137.9	65.9
4			12/06/2013	35.5	15.5	1.2
5	27/03/2014	49.2	27/03/2014	171.8	167	75.8
6			27/08/2014	75.4	48.4	4.8
7			17/06/2015	33.9	33.8	33.6
8			19/06/2016	35.2	30.6	18.6
9	8/02/2020	13.5	8/02/2020	105.6	104.4	87.6
10			28/07/2020	16.7	15.7	8.8

11			25/10/2020	75.8	68.1	35.4
12	22/12/2020	9.6	N/A	128.2	128.1	82.4

To verify the SMB model, monitoring bore BH22 at East Lynne was used as it is approximately 1 km away from the stream and therefore there would be a lag for focused recharge events while there would be a more immediate response to diffuse recharge events. The timeseries for groundwater levels in BH22-3 is shown in Figure 4.53, along with groundwater levels in BH18-2, which is an analogy for stream flow. There is a delayed and subdued response in BH22-3 to responses in BH18 and hence flows in Middle Creek, with an average lag of 69 days. Unfortunately, BH22-3 only contains data that covers one of the potential diffuse recharge events. However, for this event there is slight increase in the groundwater levels which is consistent with this being at least partially a diffuse recharge event.

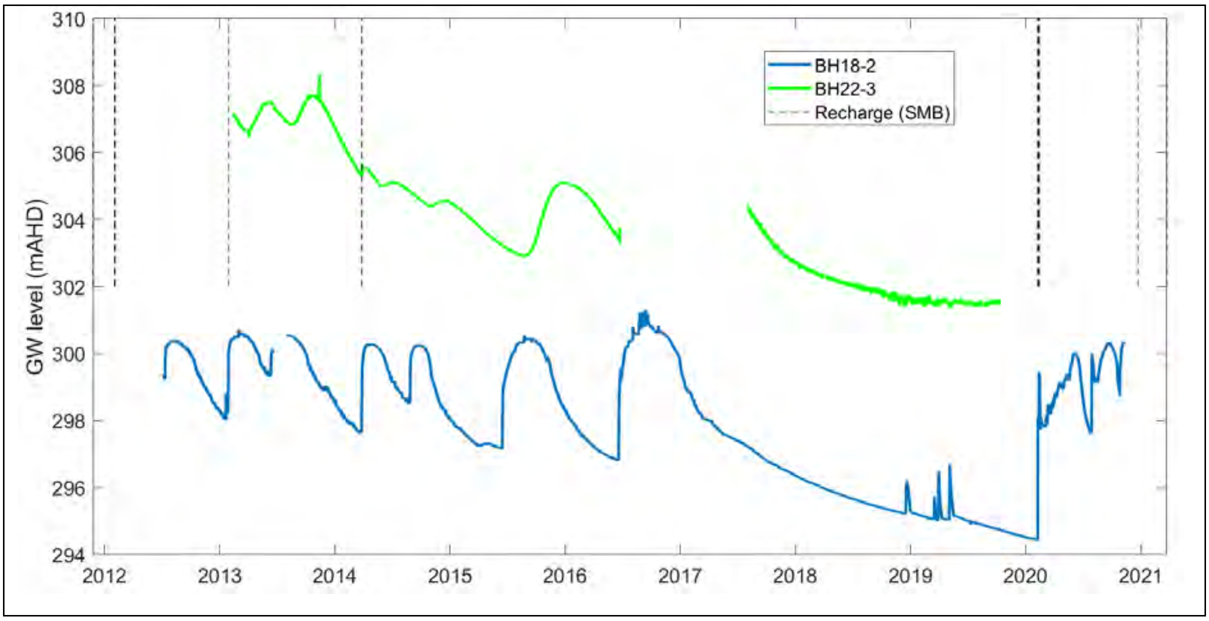


Figure 4.5349: Comparison between groundwater levels in BH18-2 (as an analogue for stream flow) and BH22-3 groundwater level timeseries. The diffuse recharge events identified by the soil moisture balance (SMB) model are indicated with dashed black lines.

To confirm the other identified diffuse recharge events from the SMB model, the response in BH19 to different recharge events was investigated. Of the five recharge events observed in BH19-1, two were potentially combined diffuse and focused recharge events (27/03/2014 and 8/02/2020), and the remaining three were focussed recharge only. When the groundwater response after each flow event is compared (see Figure 4.54), the recharge event on 27/03/2014 shows a steeper response which would fit with both diffuse and focussed recharge occurring. This further supports the diffuse recharge occurring at this time as predicated by the SMB model. The response to other predicted diffuse recharge event though appears very similar to the focussed recharge events. This could be due to the fact the predicted diffuse recharge is low compared to the focused recharge component (see Table 4.10) and/or the parameters of the model need fine-tuning.

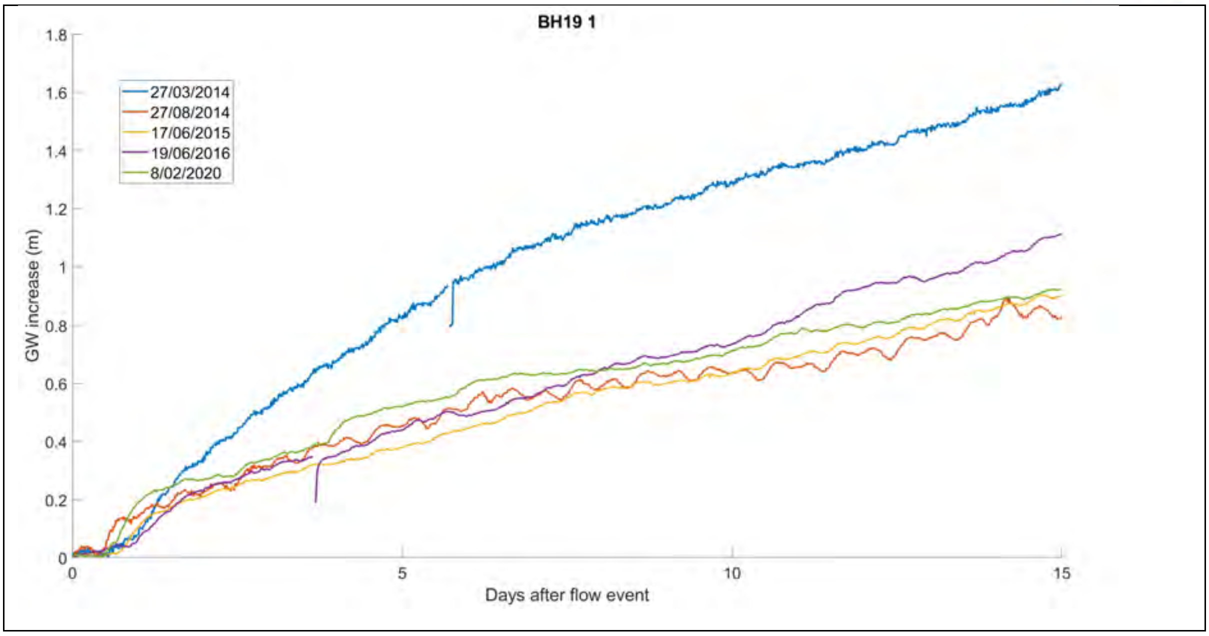


Figure 4.5450: Comparison of response in groundwater levels at BH19-1 after a flow event. The zero time on the x-axis is the time at which the event started in the BH18-2 (as an analogue for stream flow).

4.5.4.3 Low Flow Use Case

The SMB models show that there are times when rainfall infiltrates into the soil and vadose zone, with a decrease in the soil moisture deficit but this does not result in groundwater recharge. In dry periods when the unsaturated zone is larger, the rainfall that infiltrates does not reach the groundwater table but is removed by evapotranspiration, as represented in Figure 4.56. This means in dry and warm climates; the recharge is often much smaller than the infiltration. This demonstrates how vadose zone storage and evapotranspiration can contribute to transmission losses for flow releases in dry streams. Groundwater recharge in intermittent streams can also contribute to surface water transmission losses. The amount of loss will partly be controlled by the location of the groundwater table at the time of the release, as the lower the groundwater table (the greater the unsaturated zone), the greater the surface water losses.

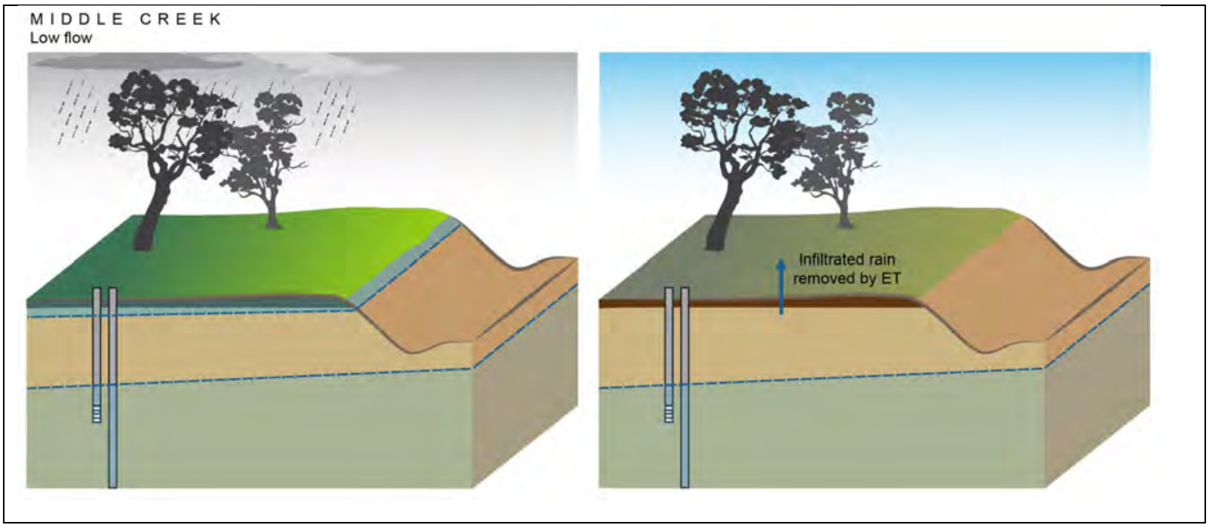


Figure 4.5651: Hydrological conceptual model for the low flow use case.

4.5.4.4 Groundwater Recharge

At Middle Creek over the 9 years of monitoring there was more focussed recharge events (10 events) than diffuse recharge (2-3 events at most) and the magnitude of the focussed recharge events was larger. Which shows that focussed recharge is the predominate recharge mechanism for intermittent creeks in semi-arid/arid regions. The occurrence of diffuse recharge events was also similar at Wellington, with only one significant event over the same time period. At Wellington the focused recharge is limited as this section of the Macquarie River is perennial and therefore, like for Elfin Crossing on Maules Creek, most streambed and bank infiltration during surface water flow events will return as bank storage to the creek as the surface water flow recedes.

At the Middle Creek Farm site, the focussed recharge events, in general, were associated with lower rainfall than the diffuse recharge event. At this site stream flow is generated by significant rainfall upstream in the mountains, which does not always correspond with significant rainfall locally at the Middle Creek Farm site. This potentially means that direct rainfall at such sites may not be a determining factor for groundwater recharge, especially if stream flow is generated by rainfall further upstream.

4.5.4.5 Limitations

Using piezometers/boreholes to gain understanding of groundwater processes is limited to locations where they exist. The cost of drilling does limit the installation of boreholes and hence groundwater level data is as a rule spatially sparse. Which means that determining the spatial variability of recharge remains a challenge. While boreholes can give insights into groundwater processes in the saturated zone, water movement through the unsaturated zone can only be inferred. Hence a knowledge gap exists around water movement in the unsaturated zone, which nascent terrestrial quantum gravity sensor technology may be able to provide insights into.

Some aspects of this work have been limited by lack of data due to logger failure (e.g., verification of the SMB model at Middle Creek). In addition to data gaps due to logger failure, loggers can also suffer from drift over time. Therefore, manual dip readings of the water levels need to be carried out on a regular basis to detect any drift.

4.5.5 Conclusions

This work has shown that focussed recharge is the predominate recharge mechanism in areas adjacent to intermittent streams, with on average one to two recharge events per year. This process [17]also potentially means that direct rainfall at such sites may not be a determining factor for groundwater recharge, especially if stream flow is generated by rainfall further upstream. In contrast, diffuse recharge is more episodic, with only one to two significant events, depending on location over a ten-year period.

5 Works Cited

[1] H. Durrant-Whyte, “Review of water-related data collections, data infrastructure and capabilities,” 2020. [Online]. Available: https://www.chiefscientist.nsw.gov.au/__data/assets/pdf_file/0003/337503/Water-Data-Review.pdf. [Accessed December 2020].

[2] D. McCallum et al, “Advanced Pipe Sensing to Reduce Leaks and Breaks,” A report prepared by the NSW Smart Sensing Network for a collaborative project led by Sydney Water, 2020.

[3] F. Chen and Y. Wang, “Advanced Pipe Sensing to Reduce Leaks and Breaks - Data Analytics: Multi-modal approaches data analytics to prioritise sensing and developing predictive analytic,” Data Scince Institute, UTS, 2020.

[4] R. Scalzo, “Where is All the Water - Milestone 1 report,” 2021.

[5] CSIRO, “Aquawatch Australia,” [Online]. Available: https://www.csiro.au/en/about/challenges-missions/aquawatch.

[6] Weathershack. [Online]. Available: https://www.weathershack.com/static/ed-tipping-bucket-rain-gauge.html.

[7] A. Yudhana, J. Rahmayanti, S. A. Akbar, S. Mukhopadhyay and I. R. Karas, “Modification of manual raindrops type observatory ombrometer with ultrasonic sensor HC-SR04,” *International Journal of Advanced Computer Science and Applications,* vol. 10, no. 12, 2019.

[8] Sunjray. [Online]. Available: http://sunjray.com/product/jdz05-1a-stainless-steel-tipping-bucket-rain-gauge/.

[9] FleetSpace. [Online]. Available: https://fleetspace.com/ .

[10] J. Pudashine, A. Guyot, A. Overeem, V. Pauwels, A. Seed and R. Uijlenhoet, “Rainfall retrieval using commercial microwave links: Effect of sampling strategy on retrieval accuracy.,” *Journal of Hydrology,* 2021.

[11] L. Halloran, G. Rau and M. Andersen, “Heat as a tracer to quantify processes and properties in the vadose zone: A review.,” *Earth-Science Reviews,* no. 159, p. 358–373, 2016.

[12] R. Huang, J.-x. Huang and C. Zhang, “Soil temperature estimation at different depths, using remotely-sensed data,” *ournal of Integrative Agriculture,* vol. 19, no. 1, 2020.

[13] t. T. Network. [Online]. Available: https://www.thethingsnetwork.org/docs/lorawan/spreading-factors/.

[14] B. Tapley and and twenty others, “Contributions of GRACE to understanding climate change,” *Nature Climate Change,* no. https://doi.org/10.1038/s41558-019-0456-2., 2019.

[15] B. D. Tapley , S. Bettadpur, M. Watkins and C. Reig, “The gravity recovery and climate experiment: Mission overview and early results,” *Geophysical Research Letters,* vol. 31, no. doi:10.1029/2004GL019920, 2004.

[16] H. Save, S. Bettadpur and B. D. Tapley, “Reducing errors in the GRACE gravity solutions using regularization,” *Journal of Geodesy,* vol. 86(9), no. doi:10.1007/s00190-012-0548-5., pp. 695-711.

[17] P. Tregoning , R. McGirr , J. Pfeffer, A. Purcell, H. McQueen, S. Allgeyer and S. C. McClusky, “ANU GRACE data analysis: characteristics and benefits of using irregularly shaped mascons,,” *Journal of Geophysical Research: Solid Earth,* vol. 127, no. 10.1029/2021JB022412, 2022.

[18] M. M. Watkins, D. N. Wiese, D. N. Yuan , C. Boening and F. W. Landerer, “ Improved methods for observing Earth’s time variable mass distribution with GRACE using spherical cap mascons,” *Journal of Geophysical Research,* vol. 120, pp. 2648-2671, 2015.

[19] H. S. Save, S. Bettadpur and B. D. Tapley, “ High-resolution CSR GRACE RL05 mascons,” *Journal of Geophysical Research,* vol. 121, pp. 7547-7569, 2016.

[20] B. D. Loomis , S. B. Luthcke and T. J. Sabaka, “ Regularization and error characterization of GRACE mascons,” *Journal of Geodesy,* vol. 93, pp. 1381-1398, 2019.

[21] S. Allgeyer, P. Tregoning, H. McQueen, S. C. McClusky, E. K. Potter, J. Pfeffer, R. McGirr, A. P. Purcell, T. A. Herring and J. P. Montillet, “ANU GRACE data analysis: orbit modelling, regularization and inter-satellite range acceleration observation,” *Journal of Geophysical Research: Solid Earth,* vol. 127, no. 10.1029/2021JB022489, 2022.

[22] S. C. Han , I. Y. Yeo, M. Kakhi , C. McCullough , E. Lee and J. Sauber, “Novel along-track processing of GRACE Follow-On ranging measurements found abrupt water storage increase and land subsidence during the 2021 March Australian flooding,” *Eart and Space Science,* no. , 2021.

[23] I. Reinfeld, “Where is the missing water in NSW: the known unknowns, the unknown unknowns and the role of NRAR,” Natural Resources Access Regulator, 2021.

[24] D. Berry, “Where is All the Water?,” WaterNSW, 2021.

[25] S. Hamilton, “Groundwater in NSW,” DPIE, 2021.

[26] T. Smith, A. Sharma, L. Marshall and Y. Tang, “Tools for investigating the prior distribution in bayesian hydrology,” *Journal of Hydrology,* vol. doi: https://doi.org/10.1016/j.jhydrol.2016.04.032. URL https://www.sciencedirect.com/science/article/pii/S0022169416302232, p. 538:551–562.

[27] N. Bulygina and H. Gupta, “Estimating the uncertain mathematical structure of a water balance model via bayesian data assimilation,” *Water Resources Research,,* vol. 45(12), no. doi: https://doi.org/ 10.1029/2007WR006749. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/ 2007WR006749., 2009.

[28] A. D. Smith and J. P. Gronewold, “Development and analysis of a bayesian water balance model for large lake systems,” 2018.

[29] B. Ossand, U. Rajagopalan, J. Lall, S. Nanditha and V. Mishra, “A bayesian hierarchical network model for daily streamflow ensemble forecasting,” *Water Resources Research,* vol. 57(9), no. doi: https://doi.org/10.1029/2021WR029920. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021WR029920., 2021.

[30] A. D. Woodbury and T. J. Ulrych, “A full-bayesian approach to the groundwater inverse problem for steady state flow,” *Water Resources Research,* vol. 36(8), no. doi: https://doi.org/10.1029/2000WR900086. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000WR900086., p. 2081–2093.

[31] T. Roksvog, I. Steinsland and K. ngeland, “A two-field geostatistical model combining point and areal observations— a case study of annual runoff predictions in the voss area,” *Journal of the Royal Statistical Society: Series C (Applied Statistics),* vol. 70(4), no. doi: https://doi.org/10.1111/rssc.12492. URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/rssc.12492., p. 934–960, 2021.

[32] E. Jeremiah, L. Marshall, S. A. Sisson and A. Sharma, “Specifying a hierarchical mixture of experts for hydrologic modeling: Gating function variable selection,” *Water Resources Research,* vol. 49(5), no. doi: https://doi.org/10.1002/wrcr.20150. URL https://agupubs.onlinelibrary.wiley. com/doi/abs/10.1002/wrcr.20150., p. 2926–2939, 2013.

[33] P. Graham, M. S. Andersen, M. McCabe, H. Ajami, A. Baker and R. I. Acworth, “To what extent do long duration hihg volume dam releases influence aquifer interations,” *Hydrogeology Journal,* vol. 23, pp. 319-334, 2015.

[34] G. C. Rau, L. J. Halloran, M. O. Cuthbert, M. S. Andersen, R. I. Acworth and J. Tellam, “Characterising the dynamics of surface water-groundwater interactions in intermittent and ephemeral streams using streambed thermal signatures,” *Advances in Water Resources,* vol. 107, pp. 354-369, 2017.

[35] M. S. Andersen and R. I. Ackworth, “Stream-aquifer interactions in the Maules Creek catchment, Namoi Valley, NSW, Australia,” *Hydrogeology Journal,* vol. 17, pp. 2005 - 2021, 2009.

[36] M. O. Cuthbert, A. Baker, C. N. Jex, P. Graham, M. S. Andersen, P. Treble and R. I. Acworth, “Drip water isotopes in semi-arid karst: implications for speleothem paleoclimatology,” *Earth and Planetary Science Letters,* vol. 395, p. 194–204, 2014.

[37] J. Kodikara, “Advanced Condition Assessment & Pipe Failure Prediction Project 2011 – 2017, Final Technical Report,” Monash University, Melbourne, 2017.

[38] S. Arcos, S. Read, K. Tyson and D. Dent, “LCS Array for Water Pipe Leal Detection Trial Report 61257007_087-1 Revision A,” Thales, 2019.

[39] ReportLinker, "Global Pipeline Monitoring Systems Industry," 31 August 2020. [Online]. Available: https://www.globenewswire.com/news-release/2020/08/31/2086065/0/en/Global-Pipeline-Monitoring-Systems-Industry.html.

[40] D. Nicholas and G. Moore, "Corrosion of ferrous watermains: past performance and future prediction–a review," *Corrosion and Materials,* p. 34(2): 33–40., 2009.

[41] R. E. Melchers, "Models for prediction of long-term corrosion of cast iron water mains," *Corrosion,* pp. 76(5) 441-450, 2020.

[42] D. McCallum, J. Close and S. Cripps, "Where is All the Water?," NSSN, 2020.

[43] N. McCarthy, "Some Cities Suffer Enormous Water Losses Every Year," 21 March 2016. [Online]. Available: https://www.statista.com/chart/4528/some-cities-suffer-enormous-water-losses-every-year/.

[44] J. Valls Miro, "Review of Acoustic Sensing for the Prediction of Leaks and Breaks in Water Mains," NSW Smart Sensing Network, 2019.

[45] K. Hardman and J. Close, *Presentation: Advanced Pipe Sensing to Reduce Leaks and Breaks,* 2019.

[46] K. Hardman, P. J. Everitt, G. D. McDonald, P. Manju, P. B. Wigley, M. A. Sooriyabandara, C. C. N. Kuhn, J. E. Debs, J. D. Close and N. P. Robins, "Simultaneous Precision Gravimetry and Magnetic Gradiometry with a Bose-Einstein Condensate: A High Precision, Quantum Sensor," *Phys. Rev. Lett.,* vol. 117, no. 13, p. 138501, 2016.

[47] W. Gao, R. Shamshiri and L. Yu, "Review of research progress on soil moisture sensor technology," *International Journal of Agricultural and Biological Engineering,* vol. 14, no. 10.2, pp. 32-42, 2021.

6 Glossary

DEM: Digital Elevation Model data.

GRACE: Gravity Recovery and Climate Experiment. An initial and follow on mission using satellites to quantify changes in mass on Earth by measuring Earth's gravity field. Operated by NASA and the German Aerospace Centre, commencing 2002.

GRAMM: Gravitational Modelling and Measurement. The program name for the team at ANU working on gravity sensors

GW: Groundwater

LoRaWAN: Communication protocol 'long-range wide area network'

Mascon: Mass Concentration elements

MATLAB:

MQ – Macquarie University.

SW: Surface Water

TRL: Technology Readiness Levels (TRLs) track progress of the development of an innovation on a scale of one to nine. Using TRL scales as a performance indicator recognises that projects often operate in pre commercial development phases, but still make significant progress towards a technology development and later commercial outcomes. (Australian CRC Definitions)

TRL 1	TRL 2	TRL 3	TRL 4	TRL 5
Basic principles observed	Technology concept formulated	Experimental proof of concept	Technology validation in lab	Technology valid in relevant environment
TRL 6	TRL 7	TRL8	TRL 9	
Demonstration in relevant environment	Demonstration in operational environment	System complete and qualified	Successful mission operations	

Voxel: a single sample, or data point, on a regularly spaced, three-dimensional grid

7 Data

Table 7.1: : Relevant bore details for Wellington

Bore	Latitude	Longitude	Screen/s	Screen interval (m)	Maximum depth (m)
WRS09	-32.5723493	148.9847819	WRS09	17-20	
WRS10	-32.5738335	148.9848462	WRS10		
WRS20	-32.5749458	148.9813574	N/A		65.62
WRS23	-32.5741507	148.9838527	N/A		45.47
WRS25	-32.5738639	148.9847431	N/A		32.77
WRS27	-32.5732363	148.9846283	WRS27S	20-23	
WRS32	-32.5729188	148.9839737			43.4
WRS34	-32.5724148	148.9843344	WRS34D	32-38	
WRS35	-32.5724880	148.9838227	WRS35S	22-25	
			WRS35D	34-37	
WRS37	-32.5726309	148.9828227	N/A		45
WRS38	-32.5726937	148.9823897	N/A		65.65
WRS40	-32.5757548	148.9852061	N/A		45

Table 7.2: : Relevant piezometer details for Maules Creek Catchment

Location	Bore	Latitude	Longitude	Distance to nearest water course (m)	Screen/s	Screen depth (m)
East Lynne	BH22	-30.4501398	150.1635433	1075	BH22-3	29.8
Middle Creek Farm	BH18	-30.4658620	150.1628821	20	BH18-1	9.36
					BH18-2	11.46
					BH18-3	17.28
					BH18-4	22.28
	BH19	-30.4656196	150.1624323	71	BH19-1	9.33
					BH19-2	22.87
Elfin Crossing	EC17	-30.4960467	150.0834755	70.8	EC17	5.79
	EC13	-30.4958587	150.0832788	44.0	EC13	7.505
	EC3	-30.4956370	150.0828367	6.3	EC3	3.105
	BH13	-30.4954425	150.0827965	28.0	BH13-1	8.57
					BH13-5	16.76
					BH13-9	24.75
					BH13-11	28.855
	BH12	-30.4954078	150.0827614	32.6	BH12-1	25.825
					BH12-4	40.7
	BH14	-30.4953685	150.0827221	38.5	BH14	26.0

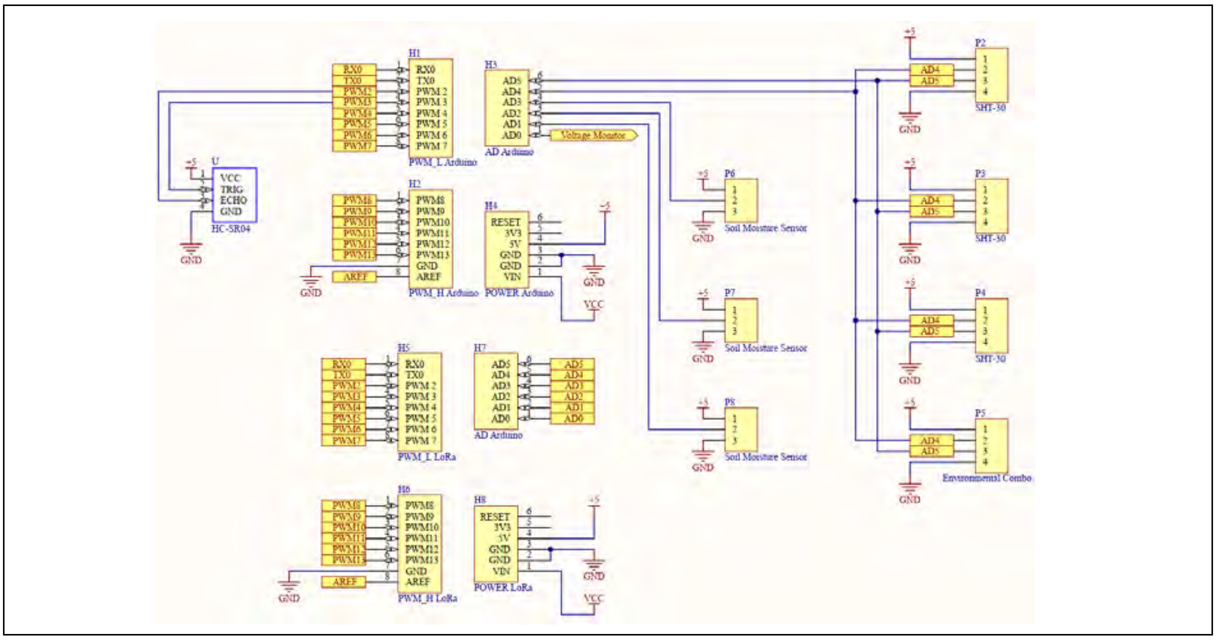


Figure 7.1: Designed system electronic connection diagram.

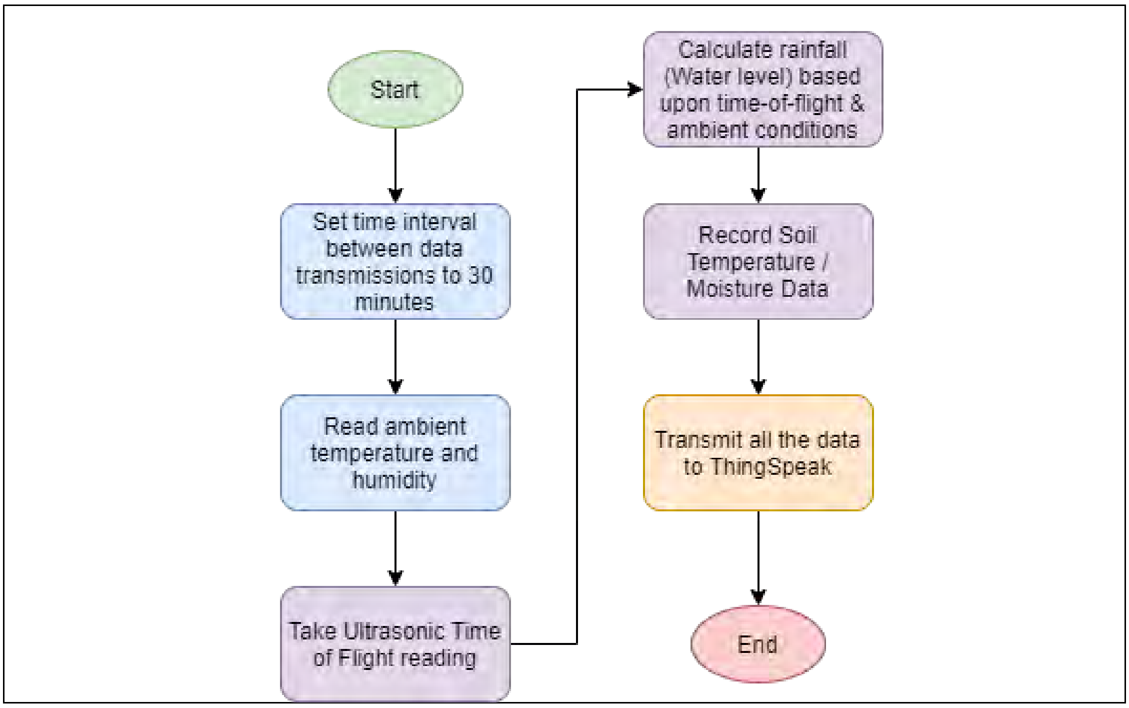


Figure 7.2: Flow Chart diagram for sensor system operation.