

DEVELOPMENT OF COMPUTER CODE TO OPTIMISE THE PRODUCTION OF HYDROGRAPHS IN SYDNEY, AUSTRALIA

Cassandra Murphy
Environmental Engineer,
 AECOM, PO Box Q410, QVB PO, Sydney, 1230, AUSTRALIA
 Cassandra.murphy@aecom.com

ABSTRACT

Long term groundwater level monitoring has been conducted in support of the environmental impact statement (EIS) prepared for the WestConnex M4-M5 Link project (RMS, 2017). The M4-M5 Link will connect two predominately underground twin drained motorway tunnels and multiple interchanges which will be excavated from the Hawkesbury Sandstone and Ashfield Shale. In order to assess the impacts of the tunnels and associated infrastructure on the local groundwater systems, a groundwater level baseline data set is being developed to characterise natural seasonal and tidal variations in the existing hydrogeological regimes.

A network of data loggers have been installed in groundwater wells screened in alluvium, Ashfield Shale and Hawkesbury Sandstone and are programmed to monitor groundwater levels through hydrostatic pressure fluctuations at an hourly interval. Barometric monitoring was also undertaken to correct for atmospheric pressure variations. Due to the large amount of data generated by the project and the shortcomings of Microsoft Excel as a streamlined graphing tool, innovation was required on the traditional method of using Excel for data analysis. As such, a custom R code was created to automate the process of hydrograph generation. The code imports folders of raw data logger files, barometrically corrects the data, identifies the groundwater well location of each raw data logger file, matches each file to the manually recorded standing water levels, and converts the groundwater level to m AHD from the raw pressure files. The code then produces hydrographs, plotting the reduced standing water levels against rainfall, exporting a fully formatted combined PDF of each hydrograph. This process significantly reduces the time spent on data manipulation and graphing and the potential for human error during the data manipulation phase.

This visualisation of the data has allowed groundwater trends within the measured time to be observed, including the tidal and rainfall dependant nature of recharge. In the Hawkesbury Sandstone, groundwater fluctuations shown in the hydrographs display a high correlation with rainfall patterns observed over the monitoring period. The alluvium is primarily controlled by local recharge and discharge conditions within the measured time. Several hydrographs also show tidal oscillations, suggesting hydraulic connections with nearby canals, creeks and Sydney Harbour. The creation of this data processing code has greatly reduced the time required for data manipulation in assessing groundwater patterns in the area of the M4-M5 Link WestConnex project. For other large groundwater monitoring projects, investing time early in the development of data processing codes could add similar benefits.

1 INTRODUCTION

Long term groundwater monitoring was conducted in support of the EIS prepared for the WestConnex M4-M5 Link project. The project location and attributes are shown in Figure 1 and comprise of:

- Twin mainline tunnels between the M4 East at Haberfield and the New M5 at St Peters;
- The Rozelle interchange;
- Connections and tunnels between the M4 East and New M5 and the proposed future Western Harbour Tunnel and Beaches Link;
- The Iron Cove Link between Victoria Road and Anzac Bridge; and
- Tunnel ventilation systems (RMS, 2017).

The project will be constructed in alluvium, Hawkesbury Sandstone and Ashfield Shale below the water table. In order to assess the impacts of the tunnels and associated infrastructure on local groundwater systems, a monthly groundwater monitoring program was conducted between June 2016 and June 2018. This program aimed to characterise the existing hydrogeological regimes in the three aquifers by collecting standing water levels and groundwater quality parameters.

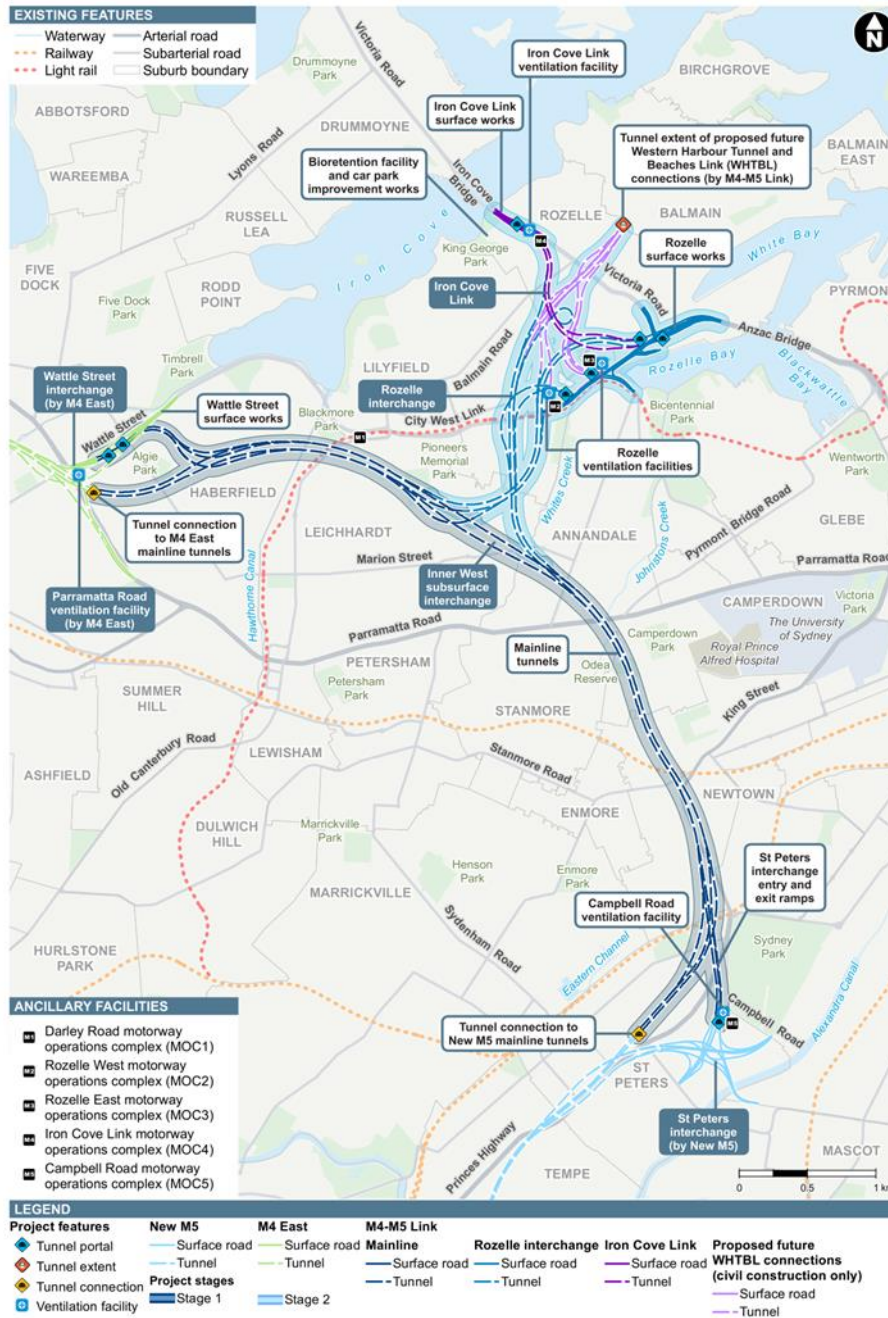


Figure 1: Overview of M4-M5 Link (RMS, 2017)

To assess changes in groundwater levels over time, hydrographs were created for each monitoring well. Hydrographs are time series graphs showing groundwater elevation as a line graph and rainfall as a bar graph. Hydrographs are most commonly used for visually determining hydrologic flow patterns in river and stream systems, but have also been used to present groundwater data and estimate groundwater recharge patterns globally (Moon et al. 2004; Cai and Offerding 2016). For this project, hydrographs were used to identify a correlation between rainfall and groundwater levels and to aid groundwater model development.

The collection of long term groundwater monitoring data is crucial as an input to groundwater model calibration. However, manipulating and visualising large datasets is often time consuming and can potentially be prone to human error. Due to the large amount of data generated by the groundwater monitoring program and the shortcomings of Microsoft Excel as a streamlined graphing tool, innovation was required to replace the traditional method of using Excel. As such, a custom R code was created to automate the process of hydrograph generation.

2 GROUNDWATER MONITORING NETWORK

The groundwater monitoring network comprised of 58 monitoring wells installed between May 2016 and May 2017 in four lithologies: alluvium (10), Mittagong Formation (1), Ashfield Shale (8) and Hawkesbury Sandstone (39). The monitoring wells were screened at secondary structural features in the expected tunnel zones and sealed with bentonite above and below the well screens. Once installed, these wells were developed to remove drilling mud and establish connection with the aquifer.

The locations of the monitoring wells are shown in Figure 2.

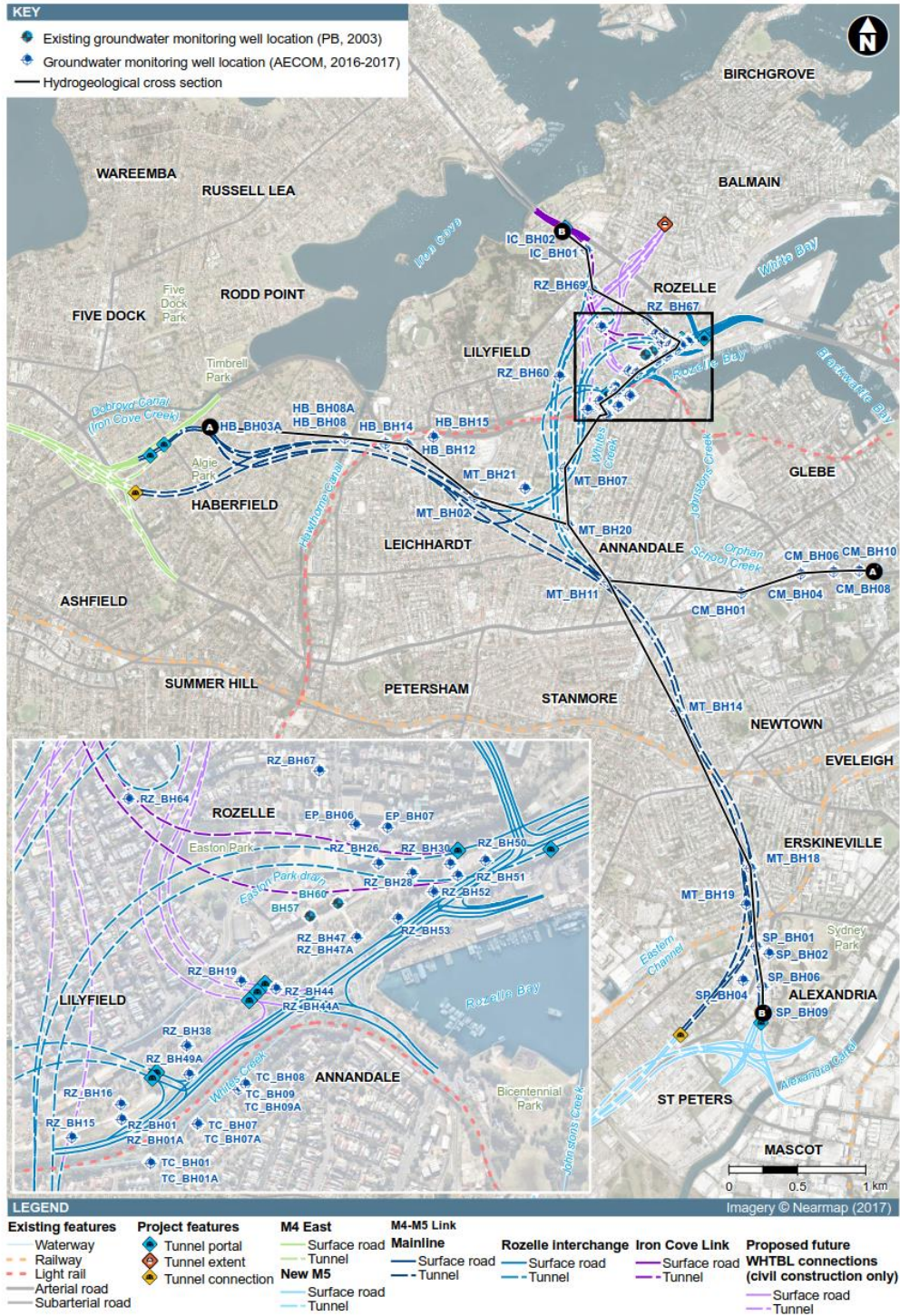


Figure 2: Groundwater Monitoring Well Locations (RMS, 2017)

Monthly groundwater gauging and sampling occurred over two to three days each month. Standing water levels were measured, along with the field parameters of electrical conductivity, temperature, dissolved oxygen, pH, and redox potential. Samples were routinely analysed for a suite of inorganic and organic analytes including heavy metals, nutrients, cations and anions.

A network of groundwater level data loggers were installed in the monitoring wells to automatically monitor groundwater level trends. Manual groundwater levels were monitored monthly with an electronic dip meter to calibrate the logger traces. The data loggers monitored groundwater levels through hydrostatic pressure fluctuations and were programmed to record at hourly intervals.

Barometric monitoring was also undertaken to correct for atmospheric pressure variations. The barometric pressure logger was placed within the extent of the network and recorded ambient air pressure throughout the program.

The climate within the project area is temperate. Rainfall is measured daily at Bureau of Meteorology Station 66062 located at the Sydney Observatory. More rainfall is recorded in the late summer and early autumn months, with March receiving the maximum monthly rainfall of 129.6 mm and August receiving the minimum monthly rainfall of 68.4 mm. Over the time period of this project, yearly rainfall levels were above average. The rainfall residual mass, or difference between expected annual rainfall and experienced rainfall, can be seen in Figure 3.

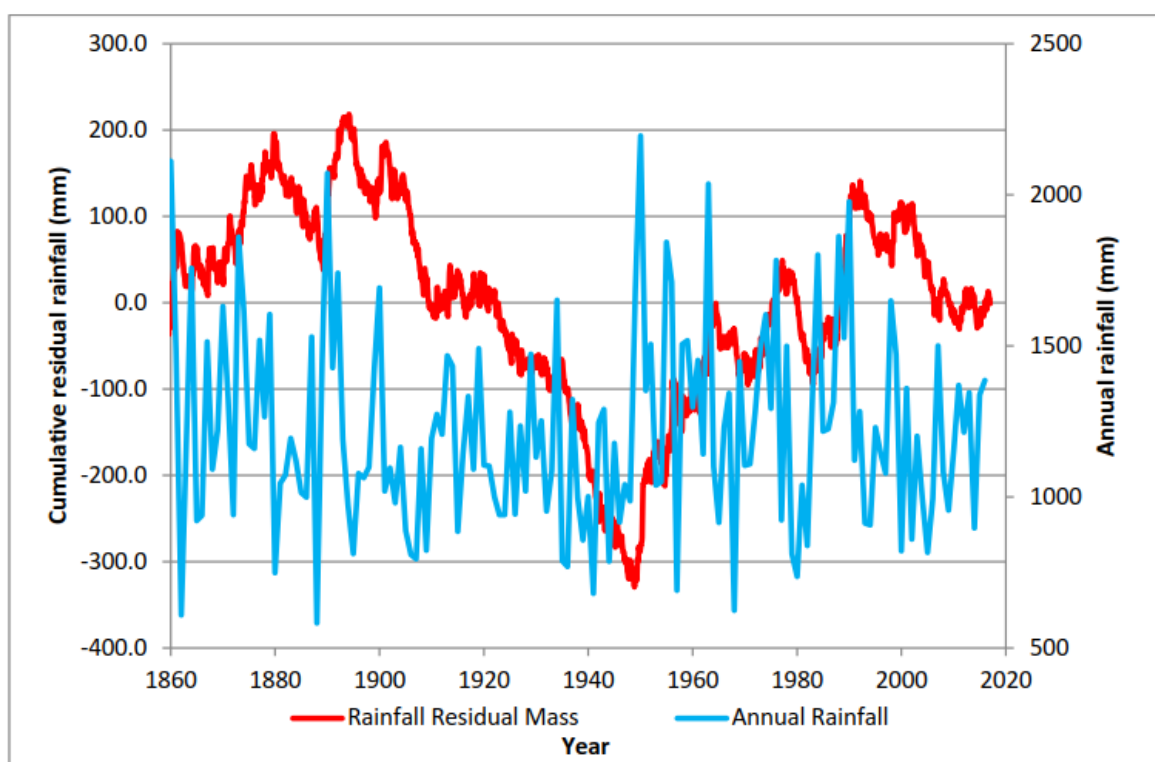


Figure 3: Sydney Residual Rainfall Mass (Station 66062) (RMS, 2017)

3 R SCRIPT

To streamline data manipulation and graphing, a custom R script was written to automate the process of hydrograph generation. R is a programming language that specialises in statistical computing and graphics and an R script is a file containing all the commands that would be typed into R as a combined document to allow for more complex data manipulation (R Core Team, 2017).

The script imports folders of raw data logger files, barometrically corrects the data, identifies the groundwater well location of each raw data logger file, matches each file to the manually recorded standing water levels, and converts the groundwater level to m AHD from the raw pressure files. The script then produces hydrographs, plotting the standing water levels against rainfall, exporting a fully formatted combined PDF of each hydrograph. This process is described in more detail in the following sections.

3.1 FILE IMPORT

The initial step is to import the relevant data. Data manipulation in Excel begins with formatting the data, writing formulas or the input of data into a template. In contrast, R is able to import and read files of data directly into the program as tables ready to be edited. R is also able to read the file names within a specified folder and import them when prompted.

Rainfall measurements were extracted from the publicly available climate records at Sydney Observatory. These are presented in CSV files with days, months and years separated into different columns. Once imported into R, these were able to be transformed into one column containing the date.

R is also able to read a variety of date formats during import. This removes the need to format time series into a common date format and allows data to be combined by time and date in future steps.

3.2 DATA MANIPULATION

Data manipulation can occur in R by applying operations to whole columns instead of copying formulas down to each row as in Excel, reducing mistranslations of data or typographical errors in formulas, like the lack of cell reference locks when they are copied.

Pressure readings, recorded as the length of the water column above the logger, were translated to standing water levels with barometric correction and the monthly groundwater gauging data that accompanied the download of the data loggers.

Barometric correction of the data loggers is required to gain accurate readings of groundwater height. As hydrostatic data loggers record the pressure above them and translate that into length of a water column, additional pressure experienced by the logger may be erroneously attributed to a groundwater flux. Barometric correction removes the pressure associated with atmospheric changes from the pressure recorded above the logger. Commercially available data loggers often have dedicated software that is able to barometrically correct the data logger files. This process is able to be done in R, as pressure can be described as a length measurement of a fluid as shown in the formula below:

$$WC = P_{\text{Data logger}} - P_{\text{Barometric logger}} \quad (1)$$

where WC = Water Column (cm H₂O);
P = pressure (cm H₂O)

Barometric correction can therefore be applied by subtracting the pressure of the atmosphere from the total pressure experienced by the logger. As R is able to merge two data sets based on a shared column of variables, in this case date and time, a combined data set with both relevant pressures matched for each time step is able to be quickly and accurately produced.

Once the pressure data has been corrected, it can then be translated into a standing water level elevation in m AHD. The groundwater elevation recorded just prior to the data logger removal and download will reflect the height of water measured above the logger. As the data logger is stationary, changes to the height of water detected can be applied to the measured groundwater elevation. This can be seen in the formula below:

$$SWL = \frac{(WC_i - WC_e)}{100} + GWL_m \quad (2)$$

where SWL = Standing Water Level (m AHD);
WC_i = Water Column of time step under inspection (cmH₂O)
WC_e = Water Column of the last time step (cmH₂O)
GWL_m = Manual Groundwater Level (m AHD)

For example, if the height of water above the data logger increases by 2 cm from recording just prior to the data logger removal, the groundwater elevation will also increase by 2 cm from the groundwater elevation measured just prior to data logger removal. Using this reference point, the whole data set is able to be transformed.

A loop, or programming tool that repeats a section of code, can be created in R to import and manipulate each file sequentially, without the need for each file to be handled independently. This ability is key to saving time during data manipulation.

If multiple data sets have been imported into R, matching the groundwater elevation to the data logger can either be done manually or with a search function. The raw CSV files exported by commercial data logger programs typically contain some form of data logger identification, like the monitoring well identification number. If these can be isolated for each data set they can be used to match to a corresponding manual groundwater level measurement in a combined reference table, automating the process and allowing the correction to occur within the loop.

3.3 GRAPH GENERATION

Within Excel, plotting bar and line graph types on one graph is possible. However, both data sets rarely align correctly without the creation of a secondary horizontal axis and the selection of only relevant rainfall data for the bar graph. As this usually has to be hidden for formatting, checking that the graphs are displaying the correct data is time consuming and not intuitive, as the axis has to be unhidden for viewing within each Excel file.

As previously mentioned, R can combine multiple data sets based on one column of data. If matched by the date and time column, data sets of transformed standing water levels can be matched to the corresponding rainfall recordings for the time periods that the data loggers were recording. The timescale of this merged data set will be set to when the data logger was operational for both variables.

Graphs can be overlapped, allowing both standing water levels and rainfall to be plotted. The creation of these figures can also occur within a loop, saving time and creating consistency. Features like titles can be numbered during the looping process, allowing them to be labelled with figure numbers in a sequence or the name of the monitoring well specific to data set being transformed.

3.4 BARRIERS TO USE

Although R allows the rapid manipulation of raw data files, it also has several limitations. As the steps of data manipulation are not as readily visible as in Excel, small errors may occur during the development of the code. To prevent this from occurring, data can be exported from each step to allow for checking. The format of R also may present a barrier to its use, as a knowledge of the R programming language is required. Detailed code comments or potentially making a user interface could aid the learning process for inexperienced users.

4 HYDROGRAPH INTERPRATATION

The production of the hydrographs has allowed groundwater trends within the measured time to be observed in the three aquifers. This can be seen in more detail in the M4-M5 Link EIS (RMS, 2017), with a small subset shown below.

4.1 HAWKESBURY SANDSTONE

In the Hawkesbury Sandstone, groundwater fluctuations shown in the hydrographs appear to respond to the rainfall patterns observed over the monitoring period. Groundwater elevations in EP_BH06, located in Rozelle, can be seen to increase after all large rainfall events, with a large increase in February 2017 to March 2017 following several rainfall events of over 20 mm/day (Figure 4). RZ_BH16, located at Rozelle also shows these general trends (Figure 4).

While EP_BH06 appears to experience small daily fluctuations, RZ_BH16 experiences more pronounced daily fluctuations in groundwater elevation of approximately 150 mm. These are attributed to tidal influences from Sydney Harbour.

Over the time period of the project, the groundwater elevations vary less than 1 m, indicating that the rainfall infiltrating and recharging the aquifer is approximately equal to the discharge to adjacent water bodies, including Sydney Harbour, and evapotranspiration.

4.2 ALLUVIUM

The alluvium is primarily controlled by local recharge and discharge conditions. In both TC_BH01S and TC_BH08, groundwater elevations increase after rainfall and decrease over the drier period between September 2016 to February 2017 (Figure 4). The pronounced tidal oscillations of TC_BH08 suggests hydraulic connections with the nearby Sydney Harbour. The relationship between the groundwater elevations of both TC_BH01S and TC_BH08 with rainfall events indicates a strong connection between infiltration and recharge.

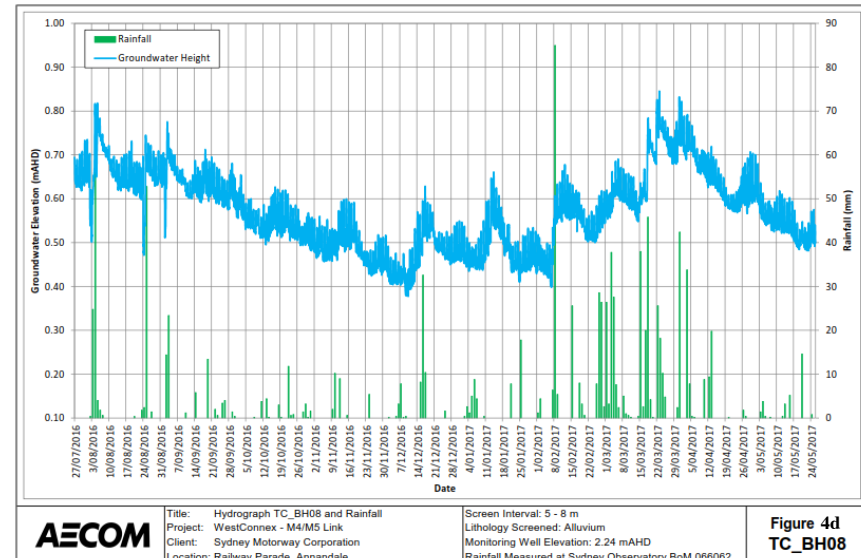
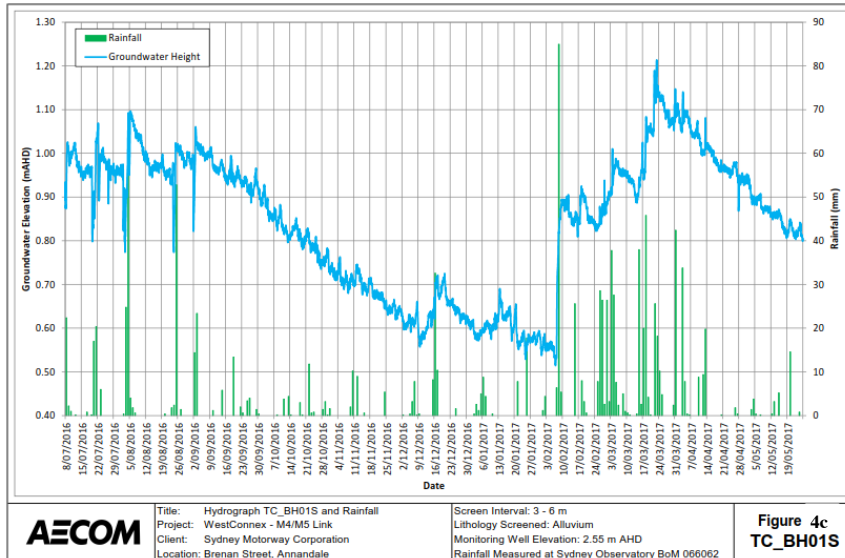
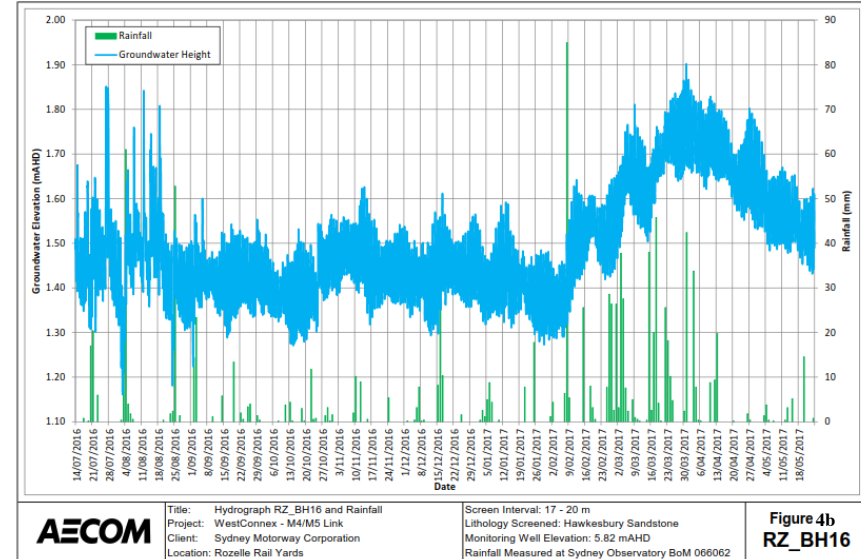
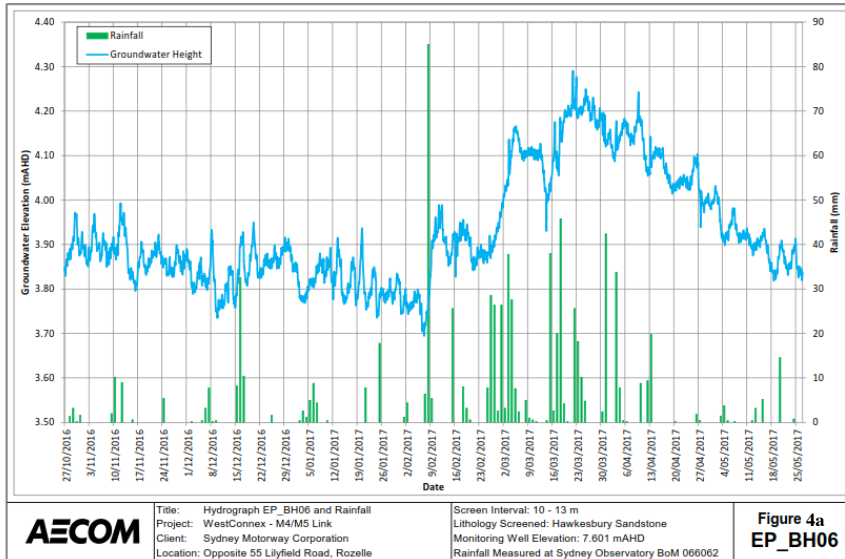


Figure 4: Hydrographs of EP_BH06 (a), RZ_BH16 (b), TC_BH01s (c) and TC_BH08 (d) (RMS, 2017)

5 CONCLUSIONS

As part of the routine groundwater monitoring that was undertaken to support the EIS for the M4-M5 Link motorway project, a network of hydrostatic pressure data loggers were installed in groundwater wells screened in alluvium, Ashfield Shale and Hawkesbury Sandstone. The ongoing monitoring associated with this project has resulted in a large amount of data that is cumbersome to efficiently manipulate in Excel. As such, an R script was developed for the purpose of data manipulation and visualisation.

The script imports, barometrically corrects, and transforms the raw data logger data into standing water levels in m AHD. It then plots the hydrographs of groundwater elevations against rainfall and exports a pdf of the figures. These figures have aided in the interpretation of groundwater level trends in the three aquifers, revealing that both rainfall recharge and tides influence groundwater level fluctuations.

The creation of this data processing code has greatly reduced the time dedicated to data manipulation as well as aided in assessing groundwater patterns within the measured time in the area of the M4-M5 Link. For other large groundwater monitoring projects, investing time early in the development of data processing codes could add similar benefits.

6 ACKNOWLEDGEMENTS

The Author wishes to thank the NSW Roads and Maritime Service for permission to publish some of the results of this investigation. The experience and professionalism of our subcontractors also contributed to the success of this investigation, namely TerraTest, Hagstrom Drilling and Numac Drilling (drilling contractors). The Author also gratefully acknowledges the efforts of RMS project managers Angelique Easton and Verity Humble-Crofts, AECOM Technical Director Andrew Cook and AECOM hydrogeologists Brad Eismen and Graham Hawkes.

7 REFERENCES

- Cai, Z., & Ofterdinger, U. (2016) *Analysis of groundwater-level response to rainfall and estimation of annual recharge in fractured hard rock aquifers, NW Ireland*. Journal of Hydrolog, **535**: pp. 71-84.
- Moon, S. K., Woo, N. C., & Lee, K. S. (2004). *Statistical analysis of hydrographs and water-table fluctuation to estimate groundwater recharge*. Journal of Hydrology, **292(1-4)**: pp. 198-209.
- R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL: <https://www.R-project.org/>.
- Roads and Maritime Services (2017). *WestConnex – M4-M5 Link Environmental Impact Statement Appendix T - Technical Working Paper: Groundwater*, August 2017.