23 June 2017

WRL Ref: WRL2017018DJA L20170623

Mr Clay Preshaw
Executive Director, Resource Assessments
The NSW Department of Planning and Environment

Submitted Online via the Major Projects Portal

UNSW SYDNEY

Water Research Laboratory

Dear Mr Preshaw,

Hume Coal Project SSD 15_7172: Peer Review of Conceptual and Numerical Modelling that Predicted Likely Groundwater Impacts

The Water Research Laboratory (WRL) of the School of Civil and Environmental Engineering at UNSW Sydney (the University of New South Wales) has peer reviewed the available groundwater models and related reporting for the Hume Coal Prospect. The focus of the review was a critical review of the site hydrogeological conceptualisation and its numerical representation. The review objective was to facilitate the NSW Government's interpretation of the EIS model prediction. WRL's review was undertaken and documented by WRL Principal Groundwater Engineer, Mr Doug Anderson. Mr Anderson declares no conflict of interest in providing this advice.

The following documents were specifically included in this review:

1. Coffey (2016a), "Groundwater Assessment, Volume 1: Data Analysis", *Hume Coal Project Environmental Impact Statement*, Appendix H, Volume 4B, 17 November 2016.
2. Coffey (2016b), "Groundwater Assessment, Volume 2: Numerical Modelling and Impact Assessment", *Hume Coal Project Environmental Impact Statement*, Appendix I, Volume 4B, 17 November 2016.
3. Pells and Pan (2017), "Groundwater Modeling of the Hume Coal Project", *Pells Consulting Technical Report #S025.R1*, 17 May 2017.
4. Hume Coal (2016) "Bore Logs", *Hume Coal Project Environmental Impact Statement*, Appendix J, Volume 4B, Prepared by Parsons Brinkerhoff.
5. Kalf (2016), "KA Peer Review of Coffey Groundwater Modelling Assessment of the Hume Coal Project", *Hume Coal Project Environmental Impact Statement*, Appendix J, Volume 4B, 12 September 2016.
6. Merrick (2016), "Hume Coal Project – Groundwater Impact Assessment Peer Review", *Hume Coal Project Environmental Impact Statement*, Appendix J, Volume 4B, 14 September 2016, Hydro Algorithmics Report HA2016/8
7. Pells and Pells (2013), "Three Dimensional Groundwater Model of Hume Coal Prospect, Southern Highlands NSW", *Pells Consulting Technical Report #P029.R1*, 2013.
8. IESC (2017), "*Advice to decision maker on coal mining project, IESC 2017-083: Hume Coal Project (EPBC 2015/7526) – New Development*", 8 May 2017.

This letter summarises the review findings.

Water Research Laboratory
School of Civil and Environmental Engineering | UNSW SYDNEY
110 KING ST, MANLY VALE, NSW, 2093, AUSTRALIA
T +61 (2) 8071 9800 | F +61 (2) 9949 4188 | ABN 57 195 873 179 | www.wrl.unsw.edu.au
Quality System Certified to AS/NZS ISO 9001
Innovative answers for tomorrow's water engineering questions, today | Since 1959

GREEN GLOBE Winner 2015

This review document is structured as follows:

1. The Peer Reviewer
2. Overview of Previous Studies
3. Peer Review Summary
    a. Model fitness for purpose
    b. Predicted impacts
    c. Economic implications
    d. Recommendations
4. Summary of key issues with the EIS Numerical Model and reporting
5. Discussion
    a. Unbiased estimation
    b. Calibration to Berrima Mine inflows
    c. Recharge
    d. Stratigraphic layer properties and hydrogeological properties
    e. Comparison of numerical models
    f. Which model is right?
    g. Compliance with SEARs

A tabular summary of this review letter is provided as an Attachment in the checklist format requested by the *Australian Groundwater Modelling Guidelines* (2012). For convenience the answers to the corresponding review questions by Kalf (2016) and Merrick (2016) are reproduced in the Attachment A so the reader can compare the opinions of the different reviewers on the models. Please note that not all items raised in this letter are summarised in the review checklist table and it is recommended that this letter be read prior to the checklist.

1. The Peer Reviewer

Mr Anderson is a member of the International Association of Hydrogeologists and the Institute of Engineers, Australia. He has previously assisted the Department of Planning (herein the Department) in the capacity of independent groundwater reviewer on other state significant road and tunnel development projects. He has considerable experience and expertise in groundwater modelling and hydrogeological site characterisation. His CV is attached to this review.

Mr Anderson adopts a position that is neither for nor against resource development and supports engineering and hydrogeological practice that avoids future environmental, social and economic externalities. On past projects, Mr Anderson has supported the development of resource models for industry including the extraction of water for tar-sands in Alberta and obtaining Lithium brines in Argentina. He has also previously been engaged by other community groups to provide critical review comment of site conceptualisation and modelling work for Australian coal mine projects.

2. Overview of Previous Studies

In summary, there is one groundwater impact assessment (the EIS) supported by a conceptual and numerical model prepared by Coffey (2016a/b) and at least five reviews of that assessment, one of which is supported by a conceptual and numerical model (Pells and Pan, 2017). The work and predictions by these experts highlight considerable variability and uncertainty regarding subsurface conditions and some differing opinions of how simply this uncertainty should be represented in a model. Subsurface conditions significantly influence the impacts of underground mining operations. Modelling assumptions of subsurface conditions significantly influence modelling predictions.

3. Peer Review Summary

While some very good characterisation and conceptualisation work has been undertaken for this project by Coffey (2016a), this review finds that there are some potentially significant issues with the site numerical model and concludes that the predicted impacts of the proposed development may be significantly more unreliable, uncertain and potentially much larger than predicted by the proponent.

This opinion was developed by examining the reporting of the numerical modelling work presented in Coffey (2016b) and carefully contrasting it with the reviewer's modelling experience and the other review documents listed above. Contrasts were drawn between the numerical model and the site characterisation and conceptualisation work reported by Coffey (2016a), Hume Coal (2016) and Pells and Pan (2017). The critical review was also supported by a number of independent data analyses. The reviewer is prepared to revise his opinion if more information is made available for review, such as the MODFLOW-SURFACT numerical model input files, digital copies of all site geological logs, the electronic databases of site property measurements and the associated technical reporting.

In summary, the reviewer finds that the groundwater impact assessment is incomplete because:

1. Key technical details and sufficient justifications for the numerical modelling approach and numerical model parameterisation are missing from the EIS. This includes, but is not limited to presentations of model geometry (elevations), properties and boundary conditions in plan and cross-section view to allow exact reproduction and proper examination of the model;
2. Some of the proponent's modelling assumptions, simplifications and calibration workflows appear inconsistent with best practice and too simplistic and given the nature of the geology at the site, the magnitude of the development and the potential impacts to groundwater;
3. Modelled site properties such as permeability and porosity are decreased significantly with depth but not varied spatially and this appears substantially inconsistent with the characterisation work completed at the site. It is also inconsistent with the workflow needed to reliably predict the spatial distribution and uncertainty of mining impacts on groundwater and the likelihood and risks of large water inflows into specific mining-panels;
4. The layer-averaged model properties established through trial-and-error adjustments are insufficiently justified and appear significantly inconsistent with the reviewer's assessment of the median values of the actual properties measured at the site by the proponent (the reviewer examined the comprehensive data-sets collated by the proponent including the geological logs and the hydraulic test data such as pumping and packer tests);
5. The numerical model has one or more numerical problems that may impact the prediction;
6. The model calibration may under-predict mine inflows and impacts because of 2-5 above;
7. An appropriate uncertainty analysis has not been presented;
8. Sensitivity analysis by third party experts predict the potential for much larger impacts; and
9. Full consideration of Secretary's Environmental Assessment Requirements (SEARs) are, in the opinion of this reviewer, given the significance of the project risks, not demonstrated.

On this basis, in the opinion of the reviewer, the proponent's predictions of likely impact are not scientifically defended. This also appears to be the view of the advice and questioning provided by the Independent Expert Scientific Committee (IESC, 2017). When asked if the groundwater model provided reasonable estimates of the likely impacts to water-related resources, the IESC stated, "*there are gaps in the documentation that hinder independent verification of potential impacts*".

The reviewer acknowledges that this opinion is also shared by the Pells and Pan (2017) review but not by the Kalf (2016) and Merrick (2016) peer reviews. The potential reasons for this disagreement are examined in more detail below.

3.1 Fitness for Purpose

The Kalf (2016) and Merrick (2016) peer review checklist tables (see Attachment A) conclude that the EIS model is "fit-for-purpose" for examining the likely impacts of the project. The basis for that conclusion is, however, not clear to the present reviewer since:

- Complete reasoning is not provided in the checklist table and the reader is referred to justifications in Section 4.3.5 of Appendix I, Vol 4B of the EIS (Coffey, 2016b); and
- The reviews appear to have been completed some months prior to finalisation of the EIS modelling work and make a number of recommendations for changes that do not appear to have been addressed in the final EIS.

The present reviewer does not accept the reasoning for model fitness reported in the EIS because:

1. Coffey (2016b) do not substantiate their self-assigned Class 3 model confidence level as suggested by Merrick (2016) in the documents sighted for this review; and
2. The argument is selective as demonstrated by Section 4 of this review, the IESC (2017) advice on the project and the 70 page report by Pells and Pan (2017).

Following detailed examination of the EIS documents and the confidence level classification table presented in the Australian Groundwater Modelling Guidelines (Barnett et al., 2012), the present reviewer established that the EIS model was actually of Class 1 confidence level having about 30% of Class 1 elements, 60% of Class 2 elements and 40% of Class 3 elements. The guidelines state the overall confidence level classification is determined by the lowest class. The Pells and Pan (2017) model which was developed by the community as a check of the EIS model was also assessed to be a Class 1 confidence level model (having about 50% of Class 1 elements, 30% of Class 2 elements and 30% of Class 3 elements).

The elements of modelling that lower the EIS model to a Class 1 classification include:

- "*Cumulative mass-balance closure error exceeds 1%*";
- "*Model parameters outside the range expected by the conceptualisation with no further justification*" (with respect to median values indicated by the field data);
- *"Calibration is based on an inadequate distribution of data*" (e.g. at depth);
- *"Model… key calibration statistics do not meet agreed targets*" (no targets set);
- Limited model validation

According to the *Australian Groundwater Modelling Guidelines* (Barnett et al, 2012):

- A Class 1 confidence model is suitable for:
  - *"First pass estimates of extraction volumes and rates required for mine dewatering";*
  - *"Developing coarse relationships between groundwater extraction… and associated impacts";*
  - "*Understanding groundwater flow processes under various hypothetical conditions*".
- A Class 2 confidence model is suitable for:
  - *"Prediction of impacts of proposed developments in medium value aquifers";*
  - *"Estimates of dewatering impacts for mines and excavations and the associated impacts".*
- A Class 3 confidence model is suitable for:
  - "*Evaluation and management of potentially high-risk impacts*"; and
  - Design of "*complex mine-dewatering schemes*".

NSW Government (Pritchard et al., 2004) has assessed that groundwater resources in the Southern Highlands are "highly valued". While an exact dollar value for this classification is not known to this reviewer, the above referenced guidelines appear to recommend a Class 2 model for estimating mine impacts in medium value aquifers and a Class 3 model for evaluating high-risk impacts.

3.2 Predicted Impacts

Both the Coffey (2016b) and Pells and Pan (2017) models predict some significant impacts to groundwater users in excess of the minimum harm criteria of the *NSW Aquifer Interference Policy* (2012). We refer you to those documents to gain an appreciation for the range of impacts that might be associated with the project.

It is our view that none of the modelling work completed to date has generated a single prediction which can be considered an unbiased estimate of the aquifer properties or the likely groundwater impacts of undertaking the project. This may become clearer with improved reporting but it is likely that additional modelling work is required to improve the confidence in the predictions.

3.3 Economic Implications

The reviewer has not undertaken or reviewed an economic assessment to calculate the value of GDP of groundwater to primary production within the potential zone of influence of this project, nor the cost of mitigation measures (possibly in perpetuity) or lost production. If not already completed, it is recommended that this work be undertaken. The assessment should be probabilistic assigning likelihoods to the impacts predicted by the modellers. This will likely require more modelling work to ensure the model base case (null scenario) is the best possible, unbiased representation of the available field data. It would also require a formal calibration constrained uncertainty analysis. The findings of this exercise should then be compared to the direct economic benefits of the mine and the need to maintain electricity generation capacity in NSW.

3.4 Conclusion

In the opinion of the reviewer, the purpose of a groundwater model for environmental impact assessment is to predict the likely impacts of the project for the benefit of end users who make decisions about the project. Decision makers cannot make reliable decisions when the accuracy, uncertainty and limitations of the groundwater model are not quantified and fully reported in the context of project impacts and outcomes.

Groundwater models are also very useful tools for proponents for forecasting project costs associated with mine water management and make good measures for a particular mine location and geometry. If model parameter values are not appropriately conservative, spatial variability in key parameters is not simulated, the model has difficulties converging / conserving mass or uncertainty is not fully quantified in terms of mine inflows and groundwater drawdown, then the costs of mine water management may be under-estimated and this could jeopardise the financial viability of the project, environmental management and community outcomes.

3.5 Recommendations

Significant additional budget should be allocated to modelling and economic assessment. The missing technical details, uncertainty analysis and justifications for the current modelling should be presented for review and acceptance by experts. Until this occurs the modelling information presented in the Pells and Pan (2017) submission should be considered the best available sensitivity and uncertainty analysis.

Consistent with the precautionary principle as defined in the NSW *Environmental Planning and Assessment Act 1979*, the reviewer recommends against approving and conditioning the development proposal on the basis of the information provided to date. The reviewer is prepared to revise this recommendation upon consideration of the proponent's responses to submissions.

4. Summary of Key Issues with the EIS Numerical Model and Reporting

There are a number of specific technical aspects of the EIS for which there is a lack of justification and for which there is no corresponding or adequate analysis to explore uncertainty. This includes but is not limited to:

1. A mass balance error of negative 4% (calibration) to 7% (scenarios) suggesting the numerical model was having problems converging to a stable solution or that some components of the water balance have not been reported. More water appears to exit the model domain than enter (see Table 1). The numerical stability and mass balance of the model should be improved to meet the Class 1, Class 2 and Class 3 model confidence level listed in the Australian Groundwater Modelling Guidelines 2012. This is <5% for Class 1, <1% for Class 2 and <0.5% for Class 3. Grid convergence and consistent calibration and simulation time-steps needs to be confirmed;
2. Normalised / Scaled root mean square errors (N/SRMS) for hydraulic head predictions and observations of about 12% (See Figure 1). This is larger than the 5% to 10% typically suggested in various guidelines. The EIS does not document any discussions with DPI Water or the Department of Planning regarding the acceptable calibration metrics for the model. DPI Water should be consulted and meeting minutes should be published. Similar plots for baseflows, hydraulic conductivity, specific yield and storage are not provided;

Table 1: Predictive Model Mass Balance – Table 10 of Appendix I, Volume 4B (Coffey, 2016b)

Table 10. Modelled average flow budget over the period of mine inflow (mining and simulation years 1 to 22 inclusive) for the case of active Hume mining.

IN (ML/day)		OUT (ML/day)	
Rainfall Recharge	34.9	Baseflow to Wingecarribee River	10.2
Release from Media Storage	7.4	Baseflow to other Rivers and Creeks	10.9
Leakage from Reservoirs	2.4	Berrima Mine Inflow (to river)	2.5
Leakage from Medway Dam	0.6	Loch Catherine Mine Inflow (to river)	0.4
		Inflow to Other Mines (to rivers)	0.6
		Private Pumping	11.0
		Hume Mine Inflow	2.6
		Evapotranspiration	10.4
TOTAL	45.4	TOTAL	48.6
Discrepancy: -3.2 ML/day (-6.8 %)			

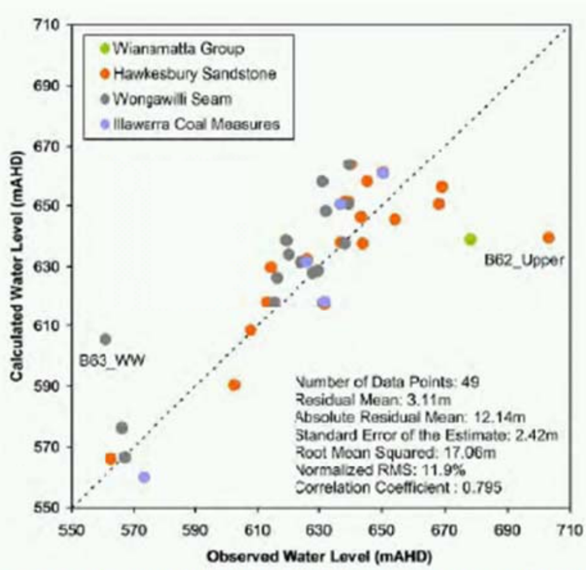

Figure 4.2. Comparison of last available observed water levels (March to July 2015) and model-calculated water levels for 27 August 2015.

Figure 1: Model Calibration to Hydraulic Head Targets - Figure 4.2 of Appendix H, Volume 4B of the EIS (Coffey, 2016a)

3. A trial and error manual calibration to some of the observation data rather than best practice calibration to all of the observation data with a software tool that provides an unbiased estimate of the aquifer properties with minimum error variance. In the reviewer's opinion, the model should be calibrated with a software tool that provides this functionality (e.g. PEST) and an uncertainty analysis should be undertaken as substantiated by many of the following points.
4. Most groundwater level hydrographs appear quiescent (Merrick, 2016), i.e. dormant, in response to stresses and all but one transient observation location (HU_38) has large calibration errors (See Figures B3 in Appendix I, Volume 4B of the EIS). Note that HU_38 is matched very precisely which is surprising for a manual calibration and may suggest that a manual calibration focussed on this well only or some unreported model optimisation was attempted for this one location. The reviewer also noted that HU_38 appears to be located over the Hume Coal lease and not Berrima Mine area which is reported to be the focus of the calibration efforts to Berrima Mine inflows. Given that model properties were reported as constant across the model domain layers (Merrick, 2016), evidence needs to be provided to demonstrate that HU_38 is a representative location to calibrate model parameters across the entire model domain. Based on figures in the EIS such as Figure C4 of Appendix H in Volume 4B of the EIS (Coffey, 2016b) it appears that HU_38 is located in a sandstone outcrop area with below average hydraulic conductivity. Based on many of the transient calibration hydrographs, groundwater levels were predicted to decline by around 5m during calibration. This may point to the model initial condition being insufficiently developed in several parts of the model domain;

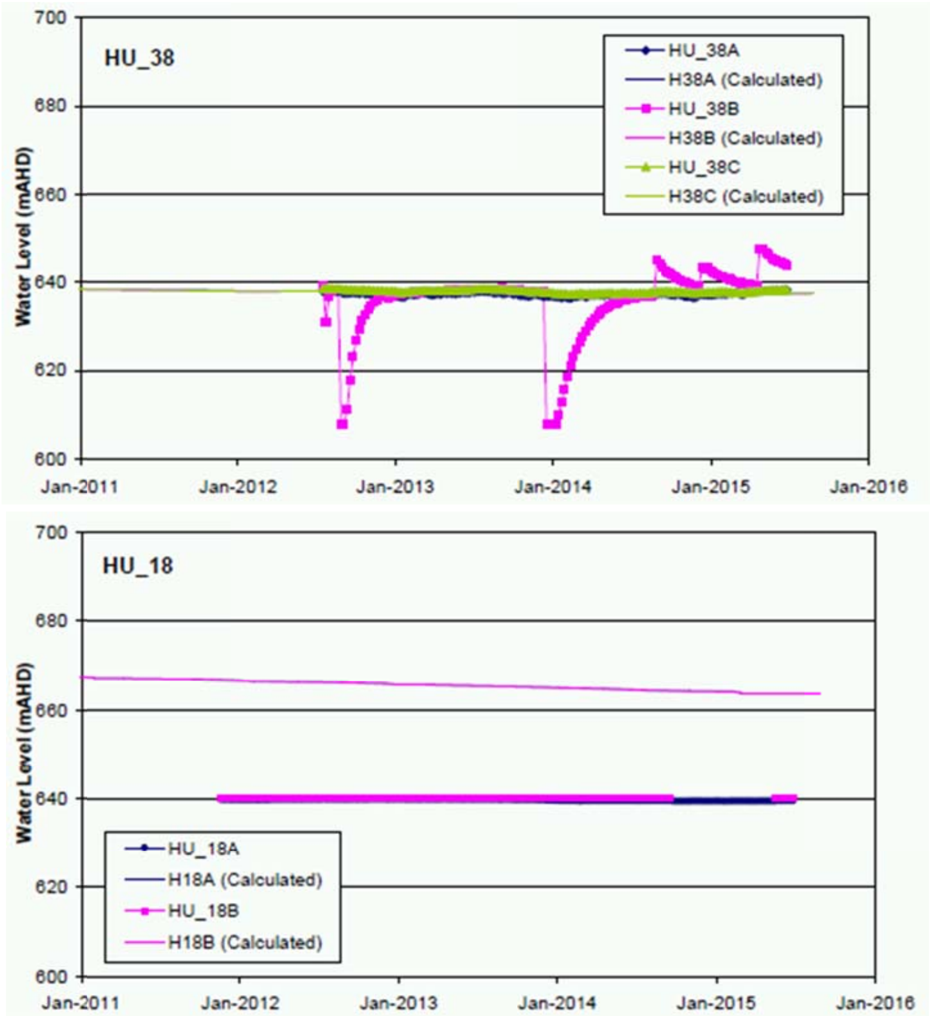

Figure 2: Two Examples of Transient Calibration Performance - Figure B3 of Appendix H, Volume 4B of the EIS (Coffey, 2016a)

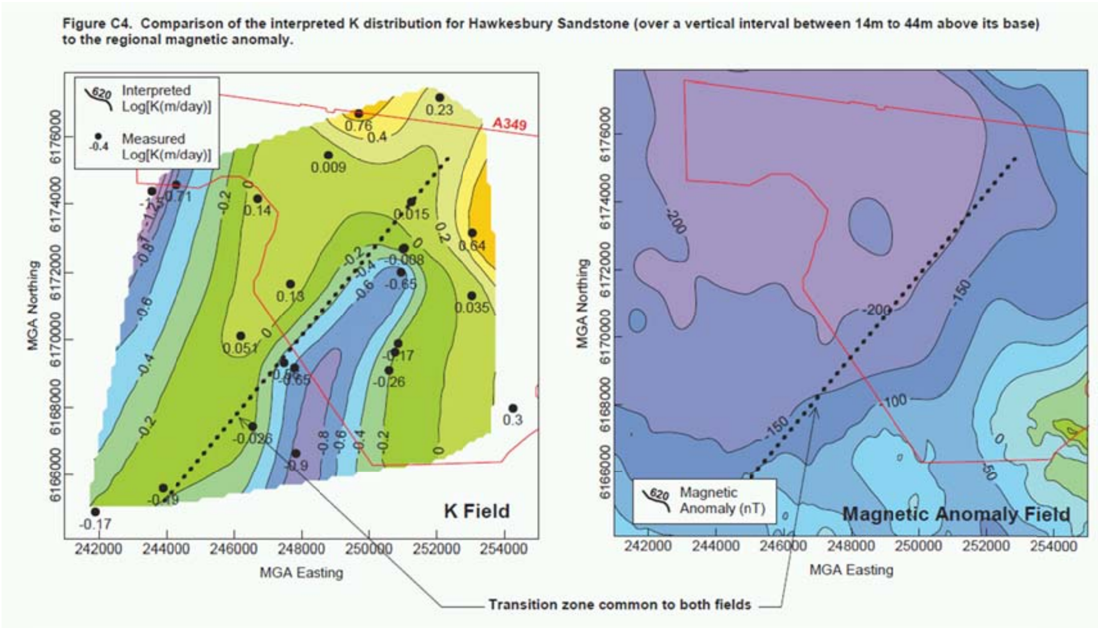

Figure 3: Figure C4 of Appendix H, Volume 4B of the EIS (Coffey, 2016a)

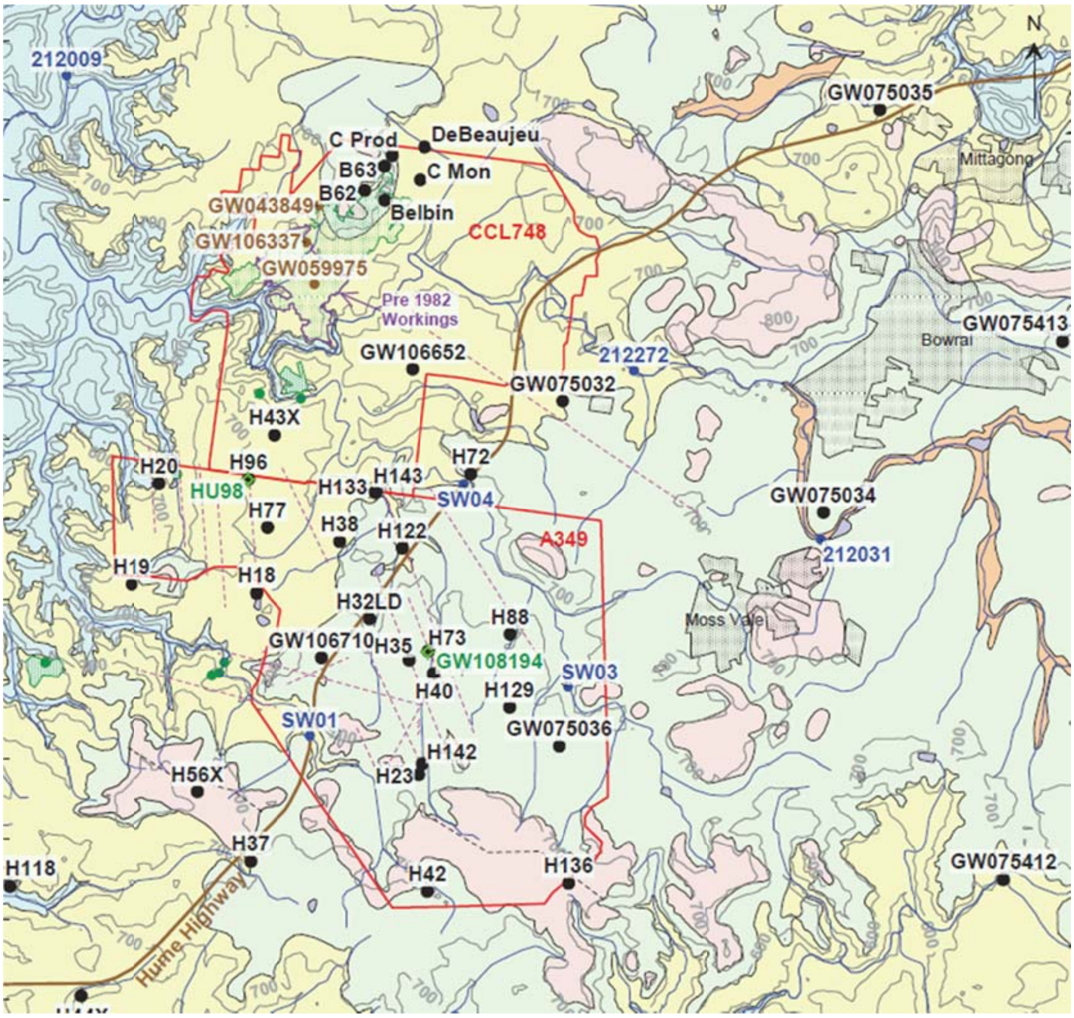

Figure 4: Locations of Project Bores from Appendix H, Volume 4B of the EIS

5. Model predictions being insensitive (Merrick, 2016) or relatively insensitive (Coffey, 2016b) to the drain conductance parameter that controls inflows of groundwater into the mine workings but 28% sensitive to rock properties much further away (up the stratigraphic sequence). The Pells and Pan (2017) model exhibits a significant sensitivity to the drain conductance parameter. In the reviewers opinion, for a calibrated model, the drain conductance parameter should be amongst the most sensitive of all model parameters. The lack of sensitivity could be caused by numerical instability, a poor choice of drain conductance parameter or an unrepresentative initial condition / choice of aquifer parameters. This needs to be examined, reported and resolved.
6. No clear description and justification of the physical basis for the chosen mine drain conductance parameter and the limitations of the chosen parameter on model prediction and calibration. For example, the mine drain conductance parameter might be intended to represent the reduction in hydraulic conductivity caused by partial desaturation in the zone of relaxation, a low permeability (hydraulic conductivity) layer within the first model layer above and below the mine workings or some combination of both. If the former, the model would under-predict mine inflows and drawdown at early time. It is also unclear how the drain conductance parameter is used to represent min-inflow from above and below given that rock properties above and below the mine workings may be different. All of these aspects of modelling should be clarified and justified before the modelling can be accepted.
7. A lack of detail on how inflows through the roof of the Berrima Mine workings were represented in the numerical model with drain cells given the anecdotal observations reported in Coffey (2016a/b) that "*void inflow rate appeared to be approximately proportional to the area of seam roof exposed, with no obvious lateral inflow from the Wongawilli seam?*". The proponent should present a water balance documenting the proportion of inflows into Berrima Mine works from (a) above the coal seam, (b) below the coal seam, and (c) from storage within the coal seam to demonstrate that the model is well calibrated.
8. The drain conductance parameter of 0.05 m^2/day. This is nominally equivalent to a hydraulic conductivity of 4 x 10^{-10} m/s in the first two metre average thickness model layer in the zone of relaxation immediately surrounding the mine voids. As stated in (6) above the basis for this value is not described. The Pells and Pan (2017) report provides an analogy that this is equivalent to tanking the entire mine with a thick layer of clay. It is also equivalent to assuming that the entire mine is surrounded by a thin layer of competent (defect free) and continuous siltstone or claystone.

   More specifically, the value appears:

   a. 30 times smaller than the modelled saturated hydraulic conductivity values applied to coal seam and adjacent model layers representing interburden (siltstone and claystone).
   b. Some 300 times smaller than the true saturated hydraulic conductivity values for those locations where the geological logs (Appendix L, Vol 4B of the EIS) and conceptualisation (Figure 4.3 of Volume 4B, Appendix H of the EIS) indicate that interburden siltstone and claystone are absent.

   Such a low mean value for the drain conductance parameter needs to be substantiated by clear conceptualisation and hydrogeological analysis supported by site data including geological logs, field and lab hydraulic conductivity data and measurement of water retention curves. Water retention curves in rock can be measured by centrifuge as demonstrated by McCartney (2007).

   In the absence of detailing reporting of the thickness and properties of interburden between the coal seam and the productive Hakwesbury Sandstone aquifer for each of the proposed mine panels, in the reviewer's opinion an average drain conductance parameter of about 1.0 m^2/day would be more consistent with the data presented to date with values varying spatially between 0.05 m^2/day and 20 m^2/day depending on the local geological logging, mapping and facies transmissivity data. In the reviewer's opinion the model should be recalibrated with a spatially variable parameter set. A more suitable sensitivity / uncertainty analysis for this parameter should be provided.

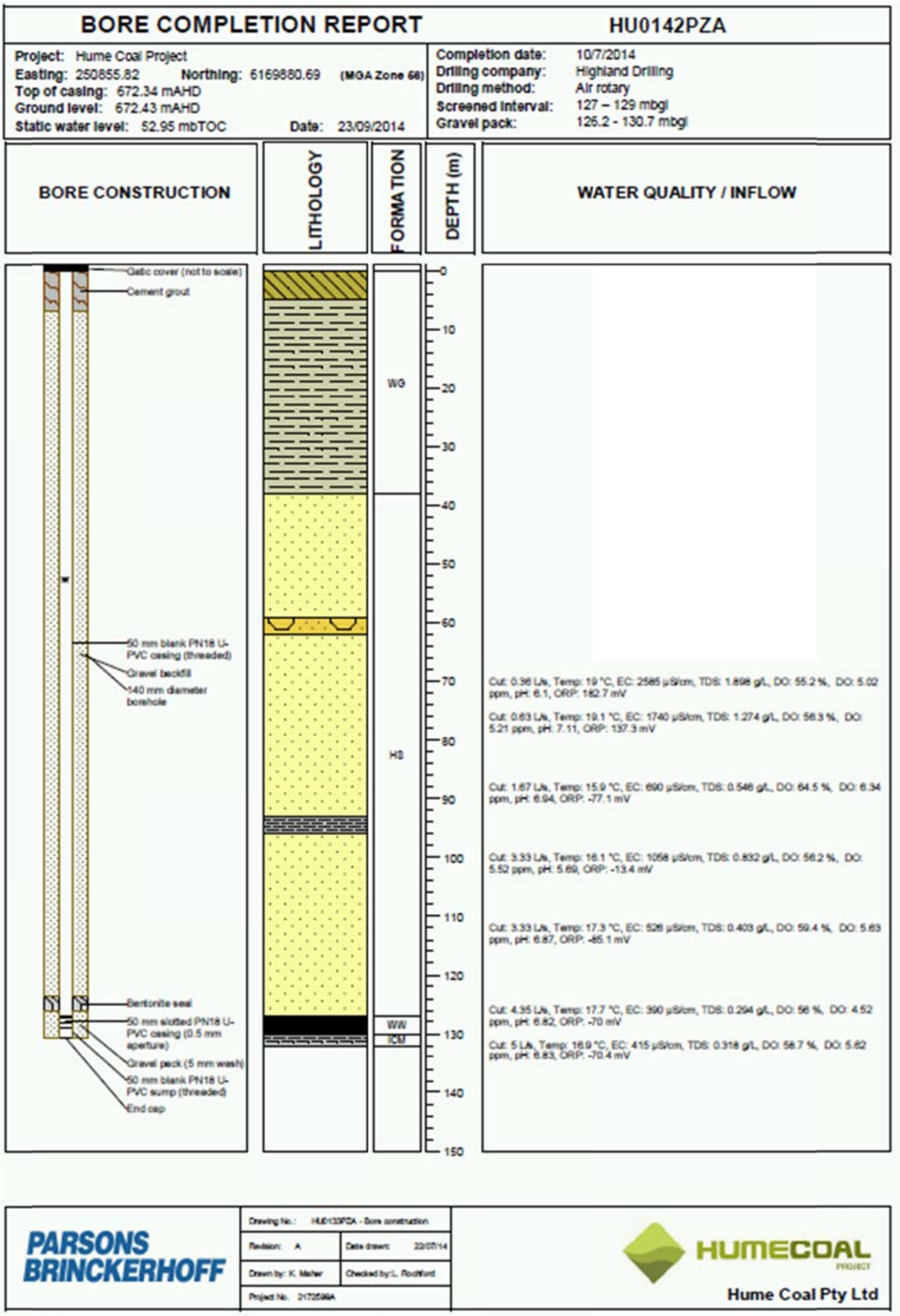

Figure 5: Example Bore Log from Appendix L, Volume 4B of the EIS (Coffey, 2016a) showing Hawkesbury Sandstone (HS) in direct contact with the coal measures (WW)

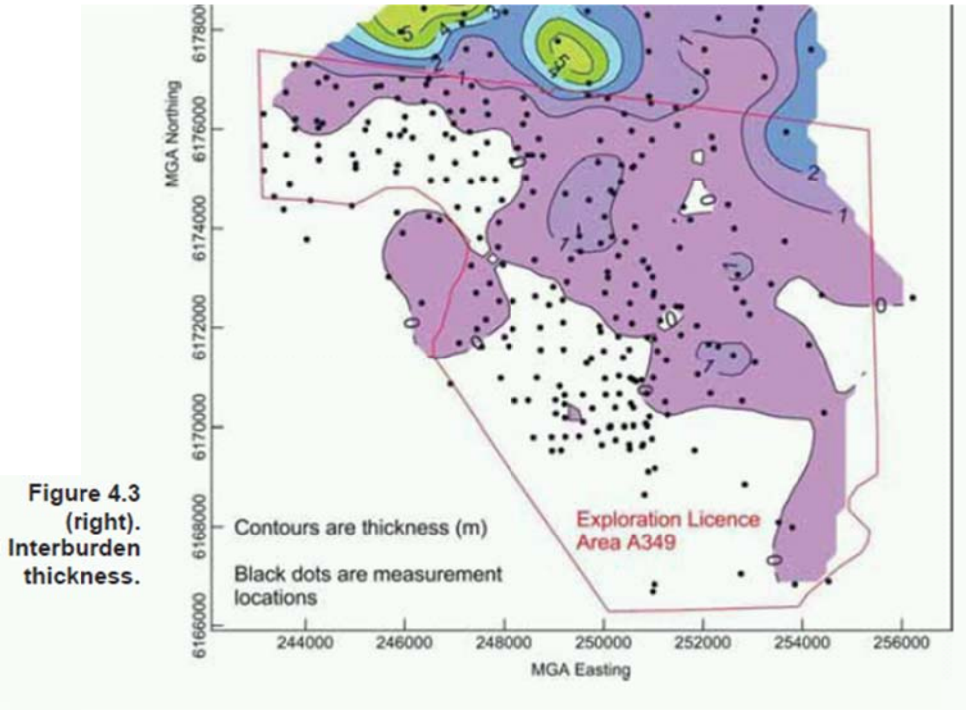

Figure 6: Proponents mapping of low permeability interburden between Hawkesbury Sandstone and Coal Measures – Figure 4.3 of Appendix H, Volume 4B of the EIS (Coffey, 2016a)

9. No conceptualisation of the unsaturated hydraulic conductivity and storage parameters of the site or how these were represented in the model (Pells and Pan, 2017). Further information needs to be provided to substantiate the model predictions and describe the numerical scheme and its input parameters throughout the model domain. For example, it needs to be specified whether groundwater drainage was simulated in MODFLOW-SURFACT with pseudo-soil functions or some unsaturated model?[1]. If the later, the proponent needs to provide evidence that the simulated reductions in saturated hydraulic conductivity due to partial desaturation around the mine-voids were not double counted in the mine drain conductance parameter. This reporting should include details of the unsaturated model fitting parameters and examples of how well these accord with the site data and measurements. The Pells and Pan (2017) model provides one example of how unsaturated model parameters can be reported;
10. The model is reported to be calibrated to match simulated inflows to the Berrima mine voids to the observed outflows from the Berrima mine voids (Coffey, 2016b) but the model is not reported to be calibrated to the time-dependency between rainfall and mine-inflow. Coffey (2016a) report that the highest correlation between rainfall and mine inflows is a lag time of about 12-months. The proponent should undertake particle tracking and provide arrival time statistics to demonstrate that the model can replicate this lag-time. The effective porosity used in the particle tracking calculations should be reported and must be consistent with the modelled aquifer storage values.
11. Limited validation is presented to support the statement that the model is calibrated to Berrima Mine inflows. The ability of the model to predict recent and historical mine outflows should be demonstrated with smaller stress periods consistent with the stress periods simulated for forward prediction.
12. The model initial condition for calibration at Berrima is generated by the model, however, no calibration statistics of hydraulic heads for the initial condition are presented for review, there

[1] This has been a point of concern to the Department's independent groundwater reviewer on previous state significant development projects for large coal mines where some significant additional modelling was requested of the proponent after EIS exhibition when public comment was closed (e.g. Shenhua Watermark).

are no reports of the initial condition being regenerated after model calibration and no reporting of the sensitivity of the calibration or the prediction to the initial condition. This information needs to be presented. Cross-sections presented in the EIS (e.g. Figure 4.3 of Coffey 2016b) suggest that the model has difficulty predicting the groundwater levels above the Berrima workings. The reviewer acknowledges the difficulty in matching this data given the very limited data available at Berrima mine and recommends a sensitivity and uncertainty analysis;

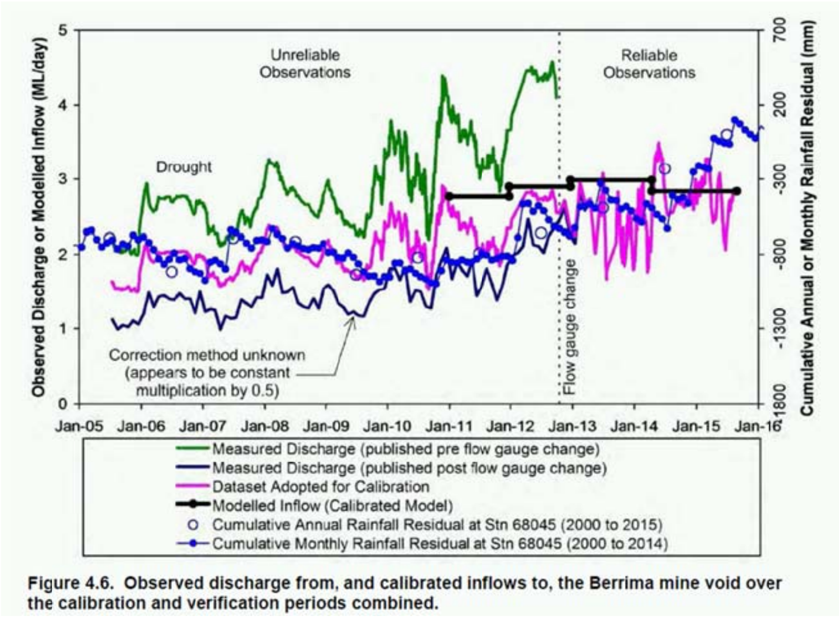

Figure 7: Proponent's calibration to Berrima Mine Outflows for one-two year stress periods – Figure 6.1 of Appendix I, Volume 4B of the EIS (Coffey, 2016a)

13. Application of a linear drift correction of minus 30% to 41% to recorded outflows from Berrima Mine between mid-2009 and October 2012 to achieve calibration to predicted Berrima Mine inflows. Noting that the rating curves for any weir might only be accurate to ±5% to 20%, the proponent should present an analysis of error for the flow gauging weirs used at Berrima Mine, justify the physical basis for the time dependent drift in the rating curve and report the sensitivity of the model calibration and prediction to uncertainty regarding mine inflows and outflows, including volumes of water that may leak into the workings, only to flow back into the formation;
14. The green line showing measured discharge from Berrima mine in Figure 4.6 of Coffey (2016b) does not match the raw data submitted in EPL returns for Berrima mine. The proponent should report the duration and details of the moving average filter that has been applied to the green line shown in Figure 7 above.

3/04/2012	4.22
4/04/2012	4.22
11/04/2012	5.66
17/04/2012	5.66
19/04/2012	7.37
23/04/2012	4.22
24/04/2012	4.22

Figure 8: Extract from EPL reporting of Berrima Mine Outflows

15. Hydraulic conductivity values in Hawkesbury Sandstone that are modelled by Coffey (2016b) as spatially invariable parameters (Merrick, 2016) despite clear and comprehensive mapping evidence by the proponent to the contrary as shown in Figure C4 of Coffey (2016a) – reproduced as Figure 3 above. In the reviewers opinion, a spatially variable distribution of

hydraulic conductivity must be modelled to achieve calibration to inflows to both Berrima Mine and the Hume Coal Prospect. One of the modellers of the Boral model for Berrima mine, Katarina David has also stated (pers comms, 21 June 2017): "*I believe the two sites are different as recharge and catchment areas differ significantly, so the inflows will be different even if all else was the same... The only very accurate information at Berrima are the inflows and wq at the discharge point, all else is very limited, including geology, bores and water levels. The strata properties are all calibrated in the model to inflows, no field data exists*";

16. Groundwater recharge is represented as a spatially invariable percentage of rainfall, set in the reviewers opinion at the lower bounds of what could be considered reasonable on the basis of the available data reported in Pritchard et al. (2004). The values are also lower than some of the site specific conceptualisation presented in Appendix H, Volume 4B of the EIS (Coffey, 2016a), e.g. piezometer H44XB. The chosen values also appear inconsistent with and lower than the recharge rates used by Boral to calibrate to observed outflows at Berrima Mine (PSM, 2016; David, 2016). Given that Boral's model is also stated to be calibrated to Berrima mine inflows and given the value of the resource and the predicted impacts of the project, it is recommended that the Hume Coal model employ a spatially variable distribution of recharge consistent with the available data, land cover and land use. This will require re-calibration of the hydraulic conductivity and storage values used within the model, possibly to higher values. Given the complexity of this task, it is recommended that this work be undertaken in PEST using pilot points and regularisation. Pilot points and regularisation allow the modeller to specify preferred values of hydraulic conductivity and recharge at locations constrained by data. The calibration software then automatically interpolates and adjusts the values across the model domain to match all the available observations and data while minimising the error. Undertaking this work also enables reliable sensitivity analysis and calibration constrained uncertainty analysis to be undertaken.

17. Page 54 for of Vol 4B, App I of the EIS and elsewhere argues that the model is well calibrated because it matches inflows to Berrima Mine inflows and is "*Coupled with reliable apriori estimates of rainfall recharge…"* to determine an average, spatially invariable value of groundwater recharge of 1.8%. Reliable field estimates of rainfall recharge collected in the model domain should be documented. Available tracer tests or mass balance analyses for data in the model domain to support the 1.8% value should be documented. Further water table fluctuation analyses other than the one reliable observation point (H44XB) reported by Coffey (2016a) for which a groundwater recharge rate of approximately 3.5% of rainfall was calculated should be provided.

    Confined aquifer water storage values (specific storage) in Hawkesbury Sandstone that are one order of magnitude lower than the best field measurements presented by the proponent and inconsistent with mathematical equations applied to typical rock porosity and compressibility data for the Sydney Basin (Pells and Pan, 2017).

18. The aquifer pump testing analysis reported in the EIS states that data interpretation was based on unconfined aquifer conditions and that optimised specific yield values were 1.5% and specific storage was 3 x 10^{-6} m/s. Only one value was reported despite a number of tests being undertaken. In their numerical model, Coffey (2016b) adopt specific yield values for Hawkesbury Sandstone ranging from 1% at shallow depths to 0.3% at about 120m depth and specific storage values decreasing from 1×10^{-6} m^{-1} to 5×10^{-7} m^{-1} at around 120m depth. This appears inconsistent with the pumping test data. The aquifer tests are reported to have been conducted at a range of depths across this depth interval. Model calibration should be re-attempted with values consistent with data and mathematical theory. Uncertainty analysis should be undertaken;

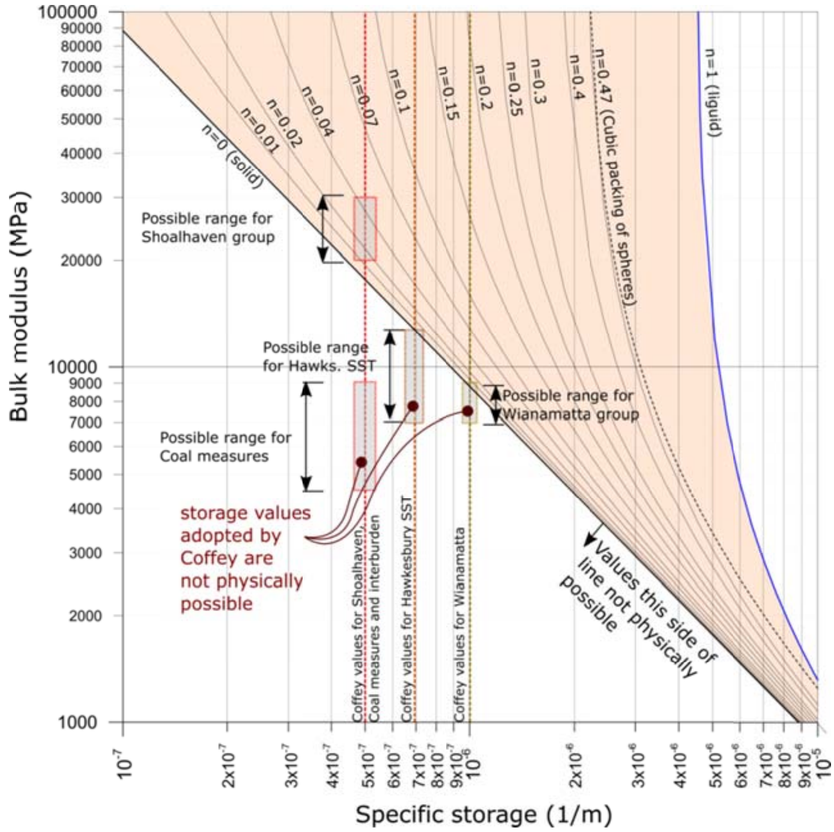

Figure 9: Figure 2.8 of Pells and Pan (2017) – Comparison of EIS modelled specific storage values reported in Coffey (2016b) to typical values of rock bulk modulus and porosity

19. Unconfined aquifer water storage values (specific yield) in Hawkesbury Sandstone that appear to assume that the pores of the rock mass never drain, only the defects (e.g. joints / fractures). This may have some basis given work at other sites overseas as noted by the modellers but applying this without site data could bias model predictions. Clear field and laboratory evidence that substantiates the validity of this modelling assumption is required given that these values appear about twenty (20) times lower than those modelled by Boral at the Berrima Mine as reported in David (2016) and PSM (2016). In centrifuge tests of drainage from rock core McCartney (2007) emphasises that it is very important to use site-specific, experimentally derived water retention curves instead of schematic curves representative of groups of soil or rock types as no single numerical solution can accurately define the behaviour of unsaturated materials;
20. Hydraulic conductivity values in Hawkesbury Sandstone that appear more consistent with matrix (pore space) hydraulic conductivity than bulk / defect hydraulic conductivity despite modelling of drainage being based mainly on defect porosity (item 19 above). This appears logically inconsistent and the extent of propagation of drawdown away from the mine may be under predicted. Clearer evidence is required to substantiate the validity of this modelling approach;
21. Figure 4.3 of Volume 4B, Appendix H of the EIS which suggests that Hawkesbury Sandstone overlies sandstone across only 50% of the lease area. The Pells and Pells (2013) assessment using a different data set found that Hawkesbury Sandstone directly overlaid coal measures in 90% of locations. The reviewer's assessment finds that 80% of the subset of available geological logs reproduced in Appendix L of Volume 4B of the EIS showed no significant low-permeability interburden unit between coal measures and the overlying Hawkesbury Sandstone Formation. The basis and working for the proponent's mapping needs to be demonstrated. A spatially variable hydraulic conductivity distribution needs to be simulated for calibration and prediction. A model sensitivity and uncertainty analysis to account for subjective mapping decisions is required.

22. The properties assigned to the five (5) numerical model layers between the mined Wongawilli coal seam and the Hawkesbury Sandstone (Layers 6-10) of unclearly conceptualised thickness (either 12 m or 10 m on average according to Table 2 or 3 of Volume 4B, Appendix I of the EIS). Geological logs in the EIS (Volume 4B, Appendix L) show varying thicknesses of coal, siltstone, conglomerate and sandstone facies in these intervals. Sometimes these facies are interbedded and sometimes some facies are absent. A number of issues have been identified with the current conceptualisation of layers and model properties:
    a. The proponent calls one of these layers (Layer 8) interburden and states that it is meant to represent a thin layer of low permeability siltstone or claystone. As stated above, mapping data provided by the proponent (Figure 4.3 of Volume 4B, Appendix H of the EIS) suggests this layer is absent from 50% of the lease area and is less than 1m thick on average, yet the modellers report to represent this unit as a continuous layer of 2m average thickness across the entire mine lease. This does not appear appropriate and may result in under-prediction of drawdown impacts and mine inflows.
    b. Review of the bore logs in Appendix L, Volume 4B of the EIS finds almost no claystone. Siltstone and claystone appeared to be absent from Model Layer 8 in about 80% of the logged holes and the average thickness of this layer was calculated to be 0.4m, not 2m.
    c. Regardless of the actual facies present in the geological logs, the proponent modelled all five (5) of these layers with exactly the same properties and exactly the same properties as the underlying coal seam. In the reviewer's opinion this is not an appropriate, physically based model conceptualisation.
    d. In addition, the coal seam layers were assigned hydraulic conductivity values of 6×10^{-8} (horizontal) and 1×10^{-8} m/s (vertical) and this was flagged by the IESC as possibly being too low (i.e. at the lower bound of uncertainty).
    e. The reviewer conducted his own upscaling exercises on the geological log data (without access to the proponent's laboratory and field testing data-sets or all their geological logs) and estimated that the average vertical and horizontal hydraulic conductivity values in these layers could be anywhere up to 200 times what was actually modelled.
    f. The EIS notes that when absent, geological units were represented with layers of 0.1m minimum thickness. This is standard practice, however, there was no reporting to confirm that the hydraulic properties of the overlying layers were copied into these model layers to ensure correct representation and modelling of the system. If these updates were not undertaken, the modelling work should be revisited.

    On this basis, the following information should be presented by the proponent to demonstrate that they developed an unbiased estimate of the properties of each numerical model layer:

    i. Elevations of the numerical model layers;
    ii. Cross sections of model layers showing geological logs and the locations and values of field and laboratory measurements of hydraulic conductivity, porosity, specific yield and specific storage;
    iii. A model upscaling analysis to demonstrate that unbiased estimates of average hydraulic conductivity and storage have been assigned in each model layer;
    iv. A sensitivity / uncertainty analysis of different interpolation / scaling algorithms;
    v. Details of how minimum thickness (0.1m) model layer properties were represented;
    vi. Measurements of coal seam permeability to demonstrate that the modelled coal seam properties are a good representation of the site conditions.
23. Hydraulic conductivity values in the overlying Hawkesbury Sandstone at two (2) to two hundred (200) times lower than best available field measurements in the public domain and about ten (10) times lower than what was modelled by Boral at the nearby Berrima Mine and quite different to those values modelled by Pells and Pan (2017):

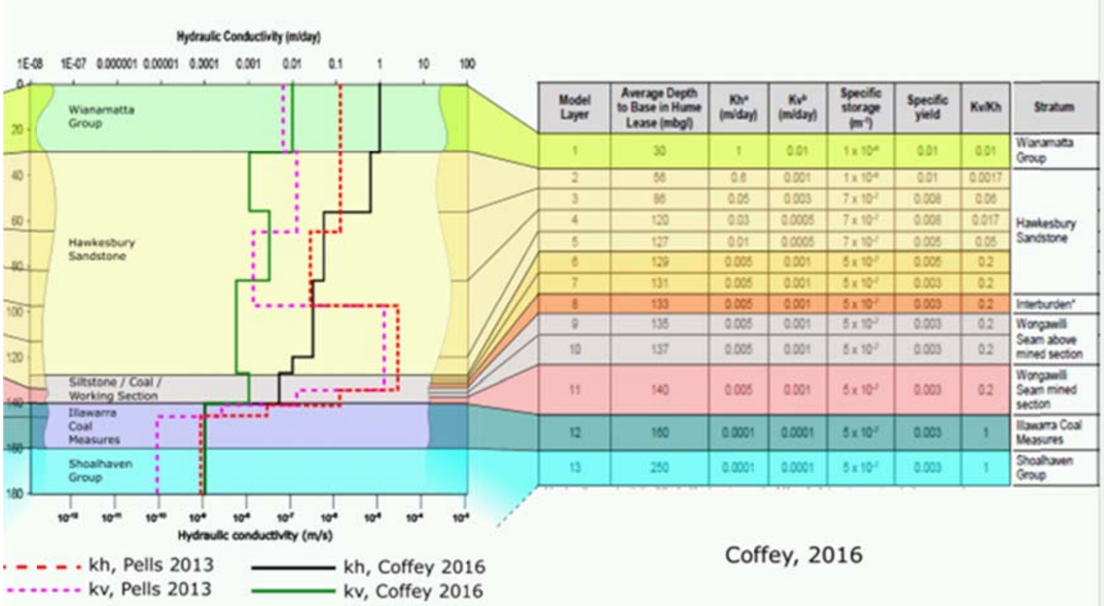

Figure 10: Coffey (2016b) Model Parameters as Summarised by Pells and Pan (2017)

More specifically:

i. Horizontal hydraulic conductivity (Kh) values consistent with "long duration pumping tests" reported for distant sites in Tammetta (2012) but apparently inconsistent with analysis for local pumping wells in the model domain including H98, GW108194, Belbin, Culpepper M, Summer Dell, Ravenswood and Wongonbra 1 and 2 which yielded much higher K values (see figure below);

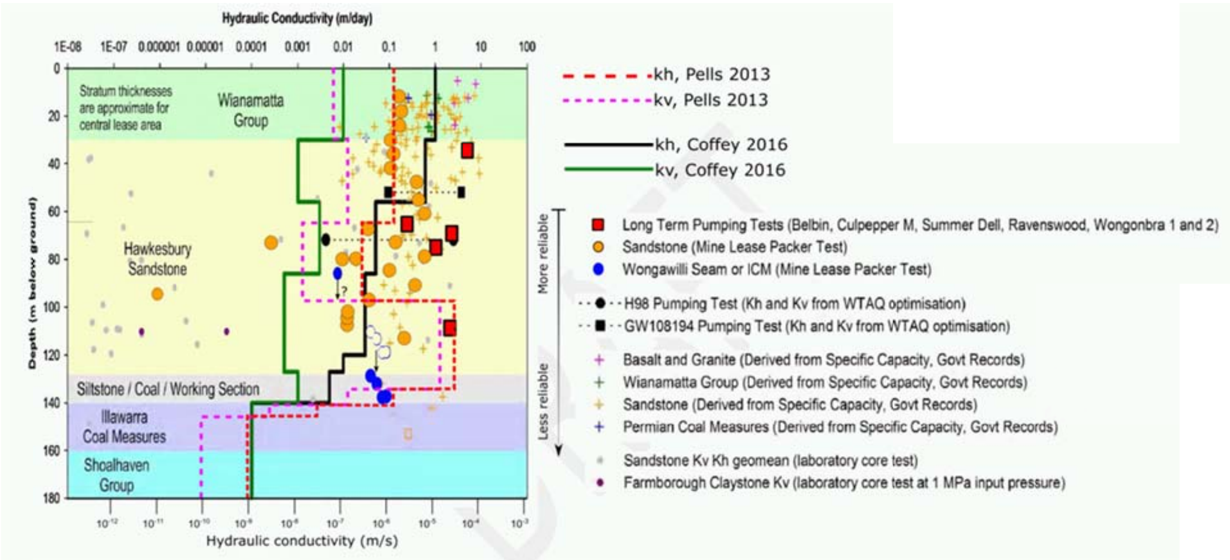

Figure 11: Coffey (2016b) and Pells and Pan (2017) Model Parameters vs Site Data (reproduced from Pells and Pan, 2017)

ii. Vertical hydraulic conductivity values that are consistent with laboratory gas-permeameter measurements of rock cores for distant sites reported in Tammetta (2012) (compare Figure 10 and Figure 12) but inconsistent with generally accepted up-scaling principles (Person et al., 1996) and suggestions in Tammetta (2012) that above 200m depth the bulk hydraulic conductivity is dominated by defects and fracture flow and not the hydraulic conductivity of the rock matrix;

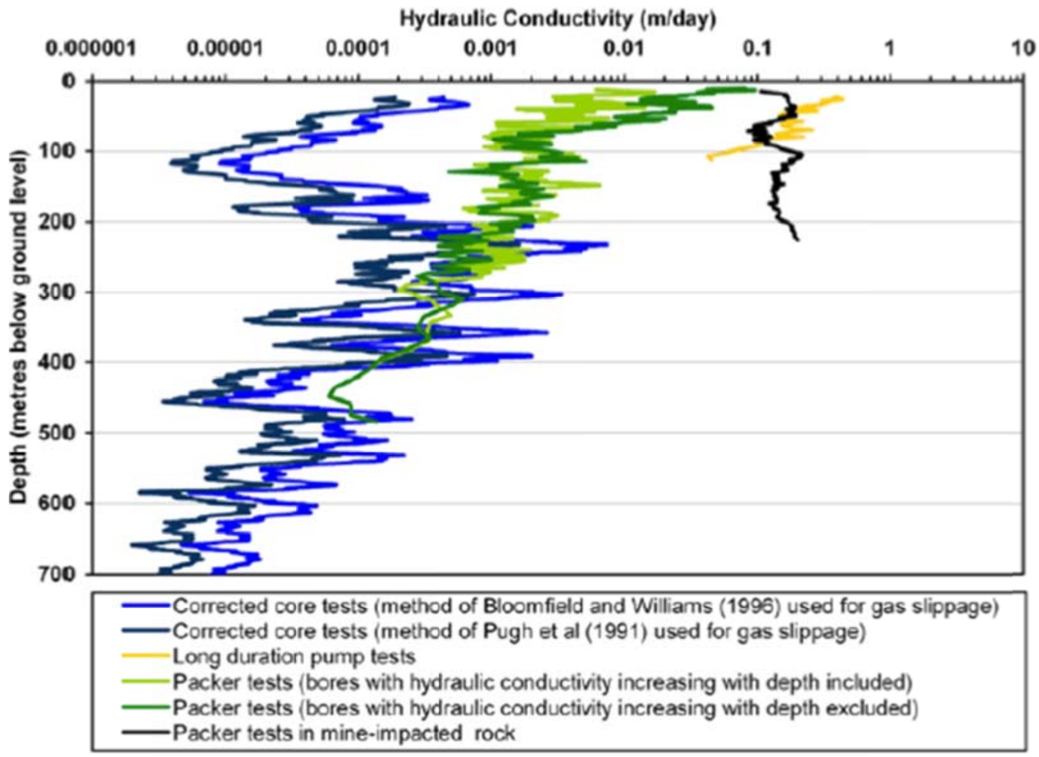

Figure 12: Hydraulic Conductivity Data for a Different Site (Tammetta, 2012)

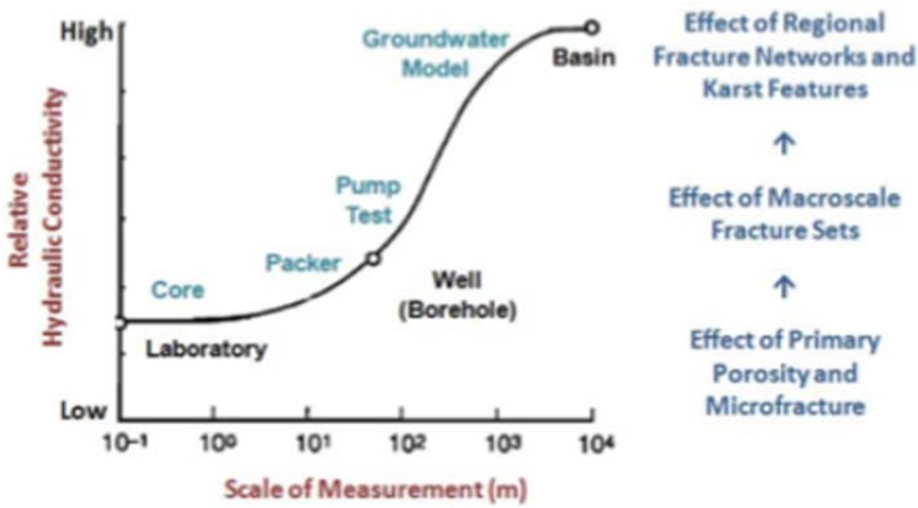

Figure 13: Scale Dependence of Hydraulic Conductivity (Person et al., 1996)

iii. Hydraulic conductivity is modelled to decrease significantly with depth which is inconsistent with the site data as presented in Figure C3 of Appendix H, Volume 4B of the EIS (compare Figure 10 and 14). It is acknowledged that some decreasing trend is likely and could be inferred in the upper panel of Figure 14 for outcrop sandstone, however, it must also be acknowledged that there are insufficient data below 40m depth to infer any statistically significant conclusions about the nature of the trend.

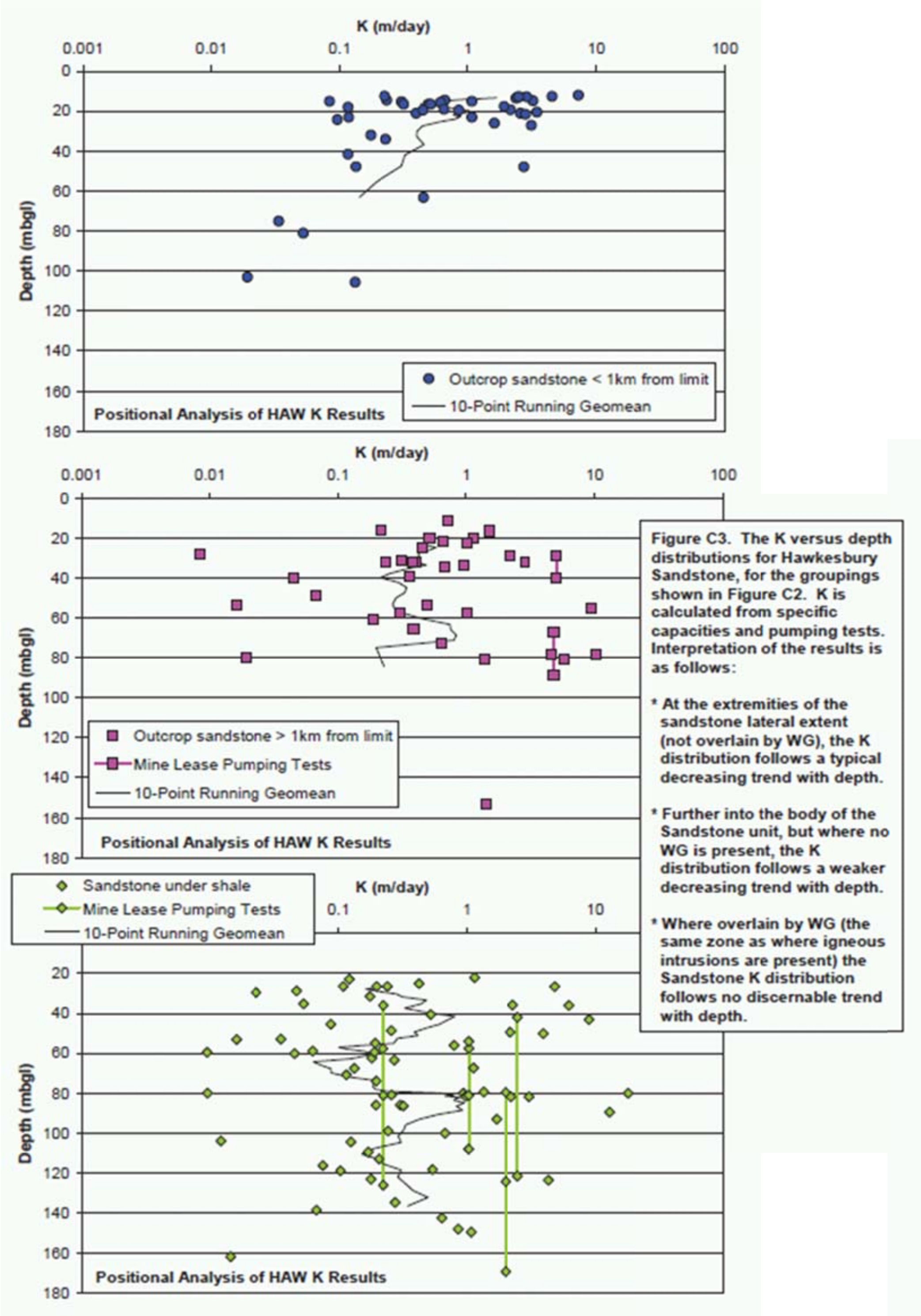

Figure 14: Figure C3 from Vol 4B, App H, App C of the EIS

Calibration and uncertainty analysis should be reattempted with spatially variable values of hydraulic conductivity and larger values of storage and recharge as the data dictates. The model should be recalibrated with calibration optimisation software that minimises an objective function of residuals calculated as the difference between observed and modelled aquifer hydraulic properties, groundwater levels and flows. Each of these residual groups should be weighted based on worth and each data point weighted based on data-point reliability. Identifiability and sensitivity of all parameters should be reported;

24. The vertical to horizontal hydraulic conductivity ratio of the geology between the Hawkesbury Sandstone has been determined from draft pump test analysis to be 0.0017 from the differences in pump test water level responses at H96C in sandstone around 70-90m depth and H96B in Wongawilli Seam at 90-100m depth. Justification of the following issues is required:
    a. Exactly the same ratio applied to layer 2 of the numerical model for Hawkesbury Sandstone between about 30 and 56m depth? Please present the geological evidence that substantiated this assumption.
    b. The ratios in all the numerical model layers below this 30-56m depth been set to values 2-100 times larger than the pump test analysis values
    c. The Kv to Kh ratio in the numerical model layers between Hawkesbury Sandstone and the Wongawilli Seam appears to be set at 0.2.
25. It appears that an unconfined aquifer model was utilised to interpret the aquifer test data, yet there appears to be some suggestion of very low permeability layers in the Hawkesbury sandstone with inferred Kv values that are two to four orders of magnitude lower than the inferred Kh values. The proponent should discuss whether parts of the aquifer system are confined or unconfined. If parts of the aquifer are actually confined, then the proponent should describe how an unconfined aquifer analysis might impact the aquifer test interpretations of aquifer properties. If this effect is uncertain, the aquifer test data should be reinterpreted with a geological model and set of assumptions that match the available field data. Given that the numerical model does not appear to rely on the aquifer test interpretations at the site, it is also recommended that the proponent demonstrates how well their numerical model parameters reproduce the available pumping test data. This work may require some refinement to the model (or development of a sub-model) as most of the pumping and observation wells appear to be located in the same numerical model grid cell.

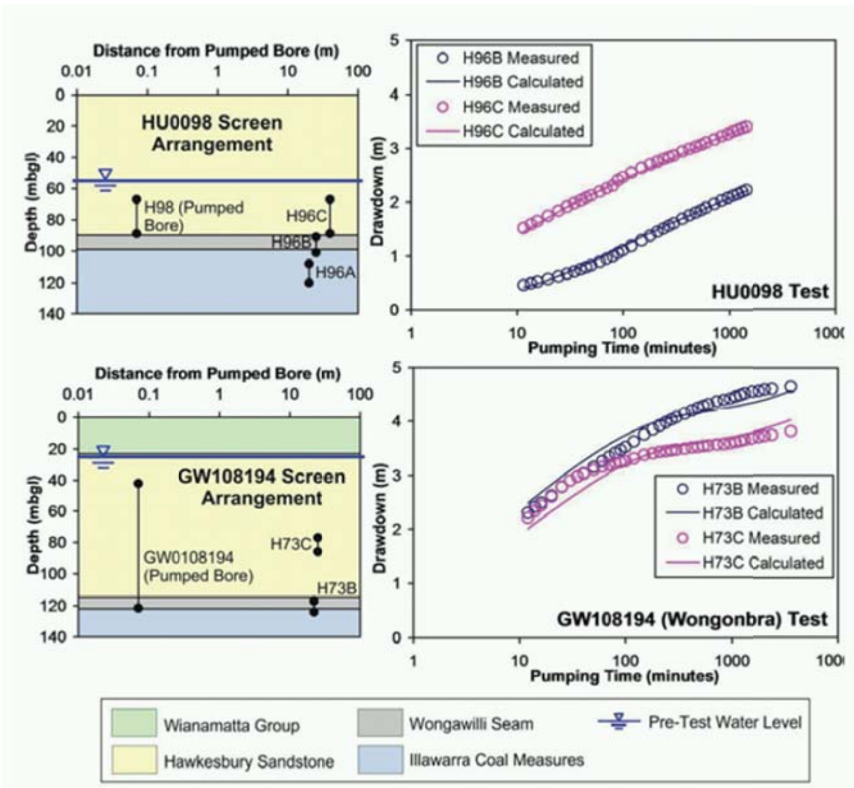

Figure 5.7. Calculated and observed drawdowns using the WTAQ algorithm in an optimisation capacity, for pumping tests at HU0098 and GW108194 on the Hume lease.

Figure 15: Figure 5.7 from Vol 4B, App H of the EIS

26. Suggestions in the EIS that one pump test (GW108194) was run for seven (7) days (10,000 minutes) but presenting and analysing data for only about 3-4 days (about 5,000 minutes) just as boundary conditions in the aquifer appeared to start influencing the test results[2]. All of the data from the test needs to be reported;
27. Calling the seven (7) day pumping test (reduced to 3-4 day for unknown reasons) a "*long duration pumping test*". Given the presence of boundary conditions, the pump test should have been run for longer to characterise these boundaries;
28. Failure to present the draft pumping test interpretation report and other laboratory reports for rock properties for independent peer review. Laboratory and field data sets, interpretive reports and sources of data that informed the site models should be made available;
29. Failure to present the pump test response and fits of the pump test interpretation model for monitoring well H96A for the pumping test at HU0098. The proponent should present:
    a. plan view maps of all monitoring wells monitored during pump testing;
    b. data-records for all monitoring piezometers prior to, during and after testing;
    c. predictions of the aquifer test model for all of the locations near the test well;
    d. justifications for the duration of the pump testing.
30. More drawdown in the coal seam than the Hawkesbury Sandstone for pumping test GW108194 and the opposite for pumping test HU0098. Also pump test interpretations at GW108194 that over-predict drawdown in the Wongawilli Coal Seam at early time and under-predict drawdown at late-time. The proponent should discuss which pumping tests provide the most accurate assessment of the horizontal, vertical hydraulic conductivity, specific storage and specific yield of the Hawkesbury Sandstone and the Wongawilli Seam. The proponent should justify why all of these values were not included in the numerical model. The proponent should report the ability of the numerical model to reproduce all the available pumping test data with appropriate refinement and assumptions regarding well losses.
31. No reports of upscaling the geological and geophysical logs with facies hydraulic conductivity and storage data to support the modelled values of hydraulic conductivity, aquifer confined storage (specific storage) and unconfined storage (specific yield). This was best practice in overseas resource assessments of lithium brine reserves during 2010 where the value of the liquid resource and risk to fresh groundwater resources were both high. It may now be standard practice. It has also been undertaken for some past CSG projects in Queensland. In the reviewers opinion, if an economic analysis of project risks and benefits supports this, or is not undertaken, it is recommended that an upscaling analysis be undertaken to support the argument that the modelled values represent an unbiased estimator of the likely impacts and for determination of suitable input values for an uncertainty analysis;
32. The statement in Section 3.2.1 of Vol 4B App I that: "*Elevations for the inverts of these and other channels over the model domain are based on digital elevation information available from the Australian Government, checked against LiDAR topographic survey data for the Hume Lease*". The EIS does not clearly define what checks were performed and, in the event of any inconsistencies, the corrections that were applied. This should be provided as the EIS model utilises topography data from NASA's Shuttle Radar Topography Mission (STRM) flown in 2000. It is well known that the average elevation errors in this data set are +-10m with significantly greater errors in close proximity to channels and incised valleys. If LIDAR data was not used for all river boundary conditions, the modelling of surface-water / groundwater interactions and calibrated model properties may be incorrect. The proponent should clarify in detail with attributed and labelled plan view maps, the elevations assigned to the model river and drain cells.

---

[2] The US Ohio EPA (2006) has the following to say about duration of pumping: "*Economic factors and time constraints also may be influential; however, economizing the period of pumping is not recommended… Though not absolutely necessary, it is recommended that tests be continued until the cone of depression has stabilized and does not expand as pumping continues. Such a steady state or equilibrium can occur within a few hours to weeks or never… Plotting drawdown data during tests often reveals anomalies and the presence of suspected or unknown boundaries, and assists in determining test duration.*"

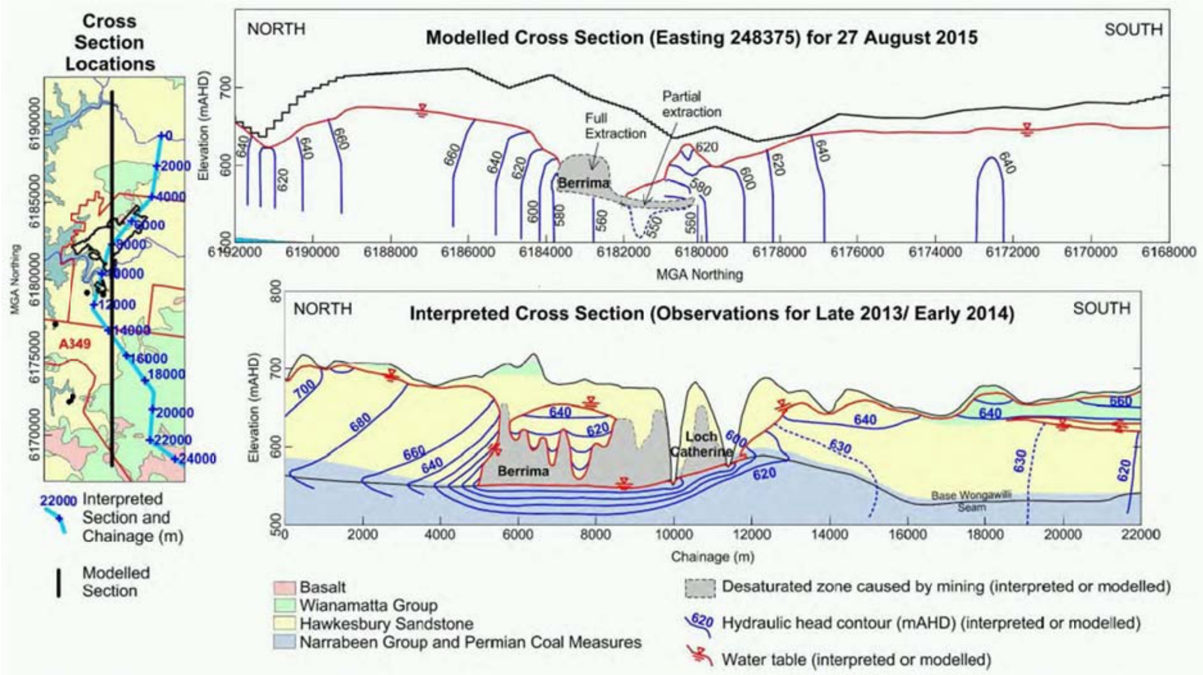

Figure 4.3. Modelled hydraulic heads along a north-south cross section through Berrima mine for 27 August 2015, compared with interpreted hydraulic heads for late 2013 / early 2014 along a nearby cross section.

Figure 16: Comparison of STRM and LIDAR cross sections – Figure 4.3 Appendix I, Volume 4B

33. The EIS does not report or show how the geology is defined around the river channels and the Merrick (2016) review suggests some river drain conductance values seemed to be set too high. Merrick (2016) and Coffey (2016b) suggests this would be conservative as stream base flow impacts would be over-predicted. The present reviewer cannot agree with this conclusion as insufficient information on model construction has been reported and several lines of evidence in the model reporting point to potentially different conclusions. To demonstrate that the river drain conductance parameter is appropriate and conservative, the modellers should detail how the surfaces in the model were created, present these surfaces in plan-view and in cross-section view through several of the creeks and streams, particularly the most incised ones, showing model layers and hydraulic conductivity.

    In support of this recommendation it is noted that a common problem in many simple numerical models is the creation of model layers by subtracting surfaces of sparse data from the Digital Elevation Model (DEM). If this DEM has large errors (like STRM data) and/or the geological picks and properties describing these computedgeological surfaces are not modified then low permeability layers of variable thickness may be “wrapped around” the streams resulting in incorrect predictions of groundwater and surface water exchange. Figure 17 (next page) shows what this might look like for the proponent’s groundwater model based upon the average model layer elevations reported by the proponent.

    Figure 17 highlights that if this is not corrected:

    a. predictions of drainage to escarpments can be artificially limited;
    b. predictions of surface water – groundwater interactions can be limited;
    c. model calculations may become unstable near river boundary conditions due to alternating high and low values of properties in adjacent model cells;
    d. modellers may struggle to achieve model calibration with symptoms including:
        a. problems with convergence and mass balance errors;
        b. incorrect reductions to groundwater recharge to achieve calibration to heads;
        c. incorrect alterations to the properties of deeper aquifer layers because groundwater flow in those model layers becomes incorrectly sensitive to the properties in the overlying shallow aquifers and aquitards;

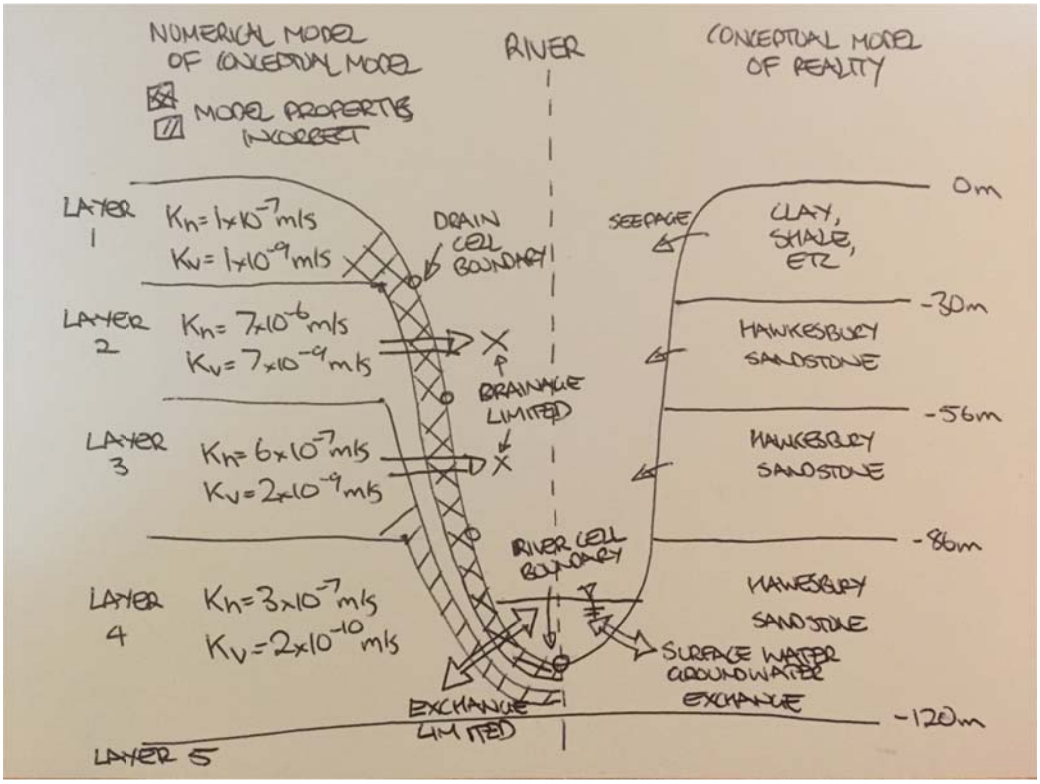

Figure 17: A Common Problem with Numerical Models of Surface-Water Groundwater Interaction

This common modelling limitation can be readily corrected by copying the correct hydraulic properties to the model cells in the vicinity of the streams, however, this does require a model with spatially variable hydraulic conductivity and storage values assigned in every layer of the model around all surface water depressions in the landscape.

Since Merrick (2016) reports that the model employs constant hydraulic properties in all the model layers and the drain conductance parameter appears to be set too high, the model has convergence problems, and no peer reviewers report inspecting the MODFLOW-SURFACT model files, the present reviewer believes that this essential work may not have been undertaken by the proponent. If that was the case the model calibration to river heads and base flows and indeed all groundwater predictions from the model would be incorrect.

34. The assumption of no-flow boundary conditions at assumed surface-water no-flow divides. Groundwater does not necessarily honour topographic divides in the land surface in the same way as groundwater. It flows from high pressure (e.g. high elevation) to low pressure (e.g. low elevation) as constrained by the geological formations and recharge. The proponent reports no sensitivity testing of these assumptions. Work by Pells and Pells (2013) recommended by the present reviewer demonstrated that models of the area are quite sensitive to the no-flow assumption. Pells and Pells (2013) reports being able to increase recharge by a factor of five (5) to accommodate uncertainty in the boundary condition and properties of the Shoalhaven Group. A sensitivity and/or calibration constrained uncertainty analysis based on a general head boundary condition in the Shoalhaven Group informed by elevations of more distant surface water bodies and bulk hydraulic conductivity of the Shoalhaven Group formation is recommended
35. Inconsistencies in the hydrogeological flow system conceptualisation between the groundwater flow model and geochemical impact assessments in relation to coal seam permeability (see the review by Jewell, 2017 for additional details).

5. Discussion

In summary there appear to be many potential issues with the modelling including:

1. Disconnects between the simulation of matrix and defect flow;
2. Simulating what should be a spatial variable distribution of hydraulic conductivity and aquifer storage with constant hydraulic conductivity and storage in individual model layers;
3. Model layers and properties that may not properly honour the geology or hydraulic data;
4. Potentially unrepresentative boundary conditions including model topography, rivers, recharge and lateral boundary flows that may under-represent groundwater inflow into the mine workings; and
5. Problems with model convergence.

The proponent appears to justify their approach and many of the discrepancies described above on the basis that their model matches and is calibrated to observations of inflows at Berrima mine. This argument is questioned in the key issues raised above. Section 6 (page 73) of Pells and Pan (2017) presents a one page summary rejecting this justification. This justification is supported by 70 pages of conceptual and numerical modelling evidence using a range of numerical models with similar, but slightly lower, confidence level classification to the EIS model. In addition to these points, the following commentary is provided.

5.1 Unbiased Estimation

As a general scientific principle, the application of expert judgement to subjectively neglect one set of measurements in favour of another, such as inflows to another mine at a different location over hydraulic conductivity within the mine lease (which appears to be the case for the present EIS) should be avoided. Best practice requires an unbiased least squares error estimator. Such algorithms are provided in many readily available model calibration optimisers and uncertainty analysis software such as *PEST*. However, there is no evidence of such algorithms being used to develop the predictions presented in this EIS. Given the predicted impacts and issues raised above it is recommended that PEST or another similar tool be used to refined the current model calibration and present an uncertainty analysis.

5.2 Calibration to Berrima Mine Inflows

In the reviewer's opinion, it is not appropriate to develop a model of the Hume Coal Prospect and report that it is calibrated on the basis that modelled inflows to a more distant mine (Berrima) match observed outflows from that mine when:

1. The model does not adequately reproduce the reliable pumping, specific capacity and laboratory test data that have been obtained to inform hydraulic conductivity values within the proponent's mine lease;

2. The EIS model adopts constant hydrogeological properties and boundary conditions in each of its computational layers. The proponent should demonstrate how the model can be reliably calibrated to predict both water levels at HU_38, inflows into the Hume Coal mine voids and inflows into the Berrima mine voids if the same model properties are used everywhere in the model and those properties do not match the calibrated properties presented by consultants of Boral in their Berrima Mine models.

3. There is an absence of inflow monitoring data to the Berrima mine voids, only outflows since 2005 and some question as to the representativeness of the available data.

4. The initial conditions, inputs and assumptions into the Hume Coal model calibration to Berrima mine inflows are not clearly documented in the EIS. The proponent should demonstrate how the model was calibrated, what the calibration and verification statistics of

the initial groundwater level condition for the transient model calibration were and how the initial condition after transient model calibration was regenerated. The sensitivity of the calibration to the initial condition should be assessed.

The basis for accepting the proponent's decision to implicitly reject the available hydraulic conductivity measurement data or the Boral model properties and boundaries in favour of a new set of spatially invariant model properties without sufficient sensitivity testing or any uncertainty analysis therefore remains unclear.

5.3 Recharge

In Vol 4B App I of the EIS (Coffey, 2016b) it is stated that: "*calibrated rainfall recharge rate is 1.8% of incident rainfall*". In Vol 4B App H of the EIS (Coffey, 2016a) it is stated : "*Recharge to the groundwater system occurs mainly by rainfall infiltration. Recharge may also occur from drainage channels wherever the stream stage is higher than the water table. Annual recharge to the water table is estimated to be about 2% of annual rainfall for the Hume area. Annual baseflow to drainage channels is estimated to be about 1.5% of rainfall from baseflow analysis*."

In a peer review of the Berrima groundwater model developed by David (2015a), the peer reviewer from PSM (2016) noted that:

- "*The local recharge catchment area overlying the Berrima Colliery is about 60 km^2…recharge rates recharge to the local catchment based on 2 per cent of annual average rainfall broadly reflect long-term inflow to the colliery of about 3,000 kL/day*"
- "*…recharge estimates may vary significantly dependent on catchment area characteristics. The application of 2 to 4 per cent of annual average rainfall as recharge to a catchment area of 60 km2 would generate a watershed of 0.94 to 1.87 GL/annum (2,560 to 5,130 kL/day). The former value compares closely to the long-term inflow to the colliery (CDM Smith, 2014) of 3,000 kL/day based on simulated recharge rates that typically ranged from 1 to 4 per cent and 8 per cent in areas overlying the colliery*"

On this basis, how can the Hume coal model be calibrated to inflows to Berrima mine if:

- the recharge required to match inflows in Boral's model at Berrima mine is 1% to 4% of rainfall and 8% over the colliery;
- the recharge in the Hume Coal Model is a constant 1.8% of rainfall everywhere;
- 1.5% of rainfall is base flow to streams; and
- The one reliable observation point examined in Hawkesbury Sandstone(H44XB) was assessed to have a groundwater recharge rate of approximately 3.5% of rainfall (Coffey, 2016a)?

The proponent should justify the modelled water balance and calibration in more detail.

5.4 Stratigraphic layer model and hydrogeological properties

The EIS contains about 60 pages of geological and geophysical logs and makes reference to the model predictions being supported by numerous unpublished field and core scale hydraulic tests. No supporting technical report is attached to describe and demonstrate how this data supports the numerical model, e.g. the process that was followed to map these logs into the chosen model layers elevations or the hydraulic tests to model layer properties. No details of the model elevations were provided. This analysis should be presented for independent verification.

Having reviewed the geological logs and undertaken some preliminary hydraulic conductivity upscaling exercises, the reviewer believes that the geological layers and properties in the EIS numerical model may have been developed somewhat subjectively without full reliance upon all the collected site data. While this may not be the case, confidence in the assessment would be greatly improved by the publication of the technical reporting of the upscaling exercises that were undertaken on the detailed geological logging and core / field scale hydraulic conductivity / storage tests. Upscaling exercises of this nature are required to ensure that appropriate spatially variable or un-biased average parameter values are assigned to each layer of the numerical model and the uncertainty in parameters is fully and properly assessed.

To support this review argument we refer to Table 2 below. This presents the reviewer's analysis of the geological logs reported in Appendix L, Volume 4B of the EIS. The statistical summary demonstrates there is little basis for modelling layers 6-10 with exactly the same hydraulic properties and an average layer thickness of 2m (refer to Figure 6 for model properties).

Table 2: Reviewer's Summary of Geology shown in Appendix L, Volume 4B of the EIS based on visual picks from the printed logs that attempted to honour the average model layer thicknesses reported by the proponent in Table 3, Appendix I, Volume 4B of the EIS

EIS Model Layer	Model Formation	Thickness Stats				Percentages					
		Avg	Stddev	Min	Max	Coal	Siltstone	Claystone	Conglomerate	Sandstone	Interbe[d]
5	Hawkesbury Sandstone	5.3	2.3	1	7	0.0%	1.9%	1.9%	0.0%	96.1%	0.0%
6	Hawkesbury Sandstone	2.1	0.5	1	4	0.0%	5.0%	0.0%	6.7%	85.0%	3.3%
7	Hawkesbury Sandstone	2.1	1.2	0.5	8	0.0%	0.0%	0.0%	15.8%	80.8%	3.3%
8	Narabeen Group	1.0	1.5	0	5	7.1%	65.1%	0.8%	0.0%	1.6%	22.2%
8	WWR Ply	0.1	0.4	0	2						
8	Farmborough Claystone	0.0	0.0	0	0.25						
9	Wongawilli above mined	1.5	1.3	0	5.25	79.3%	11.8%	0.0%	0.0%	0.0%	8.9%
10	Wongawilli above mined	1.5	1.2	0	5.25	84.4%	6.9%	0.0%	0.0%	0.0%	8.7%
11	Wongawilli mined	3.2	0.8	0.5	3.5	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Note that the statistics of Narrabeen Group (including Siltstone, WWR Ply and Farmborough Claystone layers) are only for those wells where a low permeability unit was picked in the geological logs between sandstone and coal measures. The reviewer's interpretation of the logs was that a lower permeability unit within Narrabeen Group rocks – Model Layer 8 - was absent 80% of the time and this reduces the average thickness of the interburden units in Layer 8 of the model to 0.4m.

The comments by the reviewer that the model may under-predict mine inflows and mine impacts are further supported by an upscaling exercise using typical facies values of horizontal and vertical hydraulic conductivity applied to the underlying Table 2 dataset using geometric and harmonic averaging. These data, calculations and results are detailed and have not presented in this review document.

5.5 Comparison of numerical models

It is relatively easy to create a model that generates a single prediction of groundwater impact, however, a single model and single prediction rarely provides an informative assessment. Due to the inherent uncertainty in our understanding of hydrogeological systems, multiple models and predictions are typically required.

The Pells and Pan (2017) model is predicated on accurate LIDAR topography, a model layer structure that matches key major formations based on publically available information prior to EIS publication with some minor subsequent updates, hydraulic conductivity distributions with depth that match publically available data, recharge values that match published literature values, realistic boundary conditions based on assumptions and a basic calibration to publically available hydraulic head measurements. Some of their publically available information is not as current as that relied upon in Coffey (2016a/b) as it is difficult and very timing consuming to extract this information from the EIS documents rather than electronic data files.

The Coffey (2016a/b) models are predicated on some unknown combination of inaccurate STRM and accurate LIDAR topography. They employ a model layer structure with more computational layers than Pells and Pan (2017) including thin stratigraphic layers close to the coal seam. The delineation of numerical model layers in the Hawkesbury Sandstone (HS) Formation appears more arbitrary than Pells and Pan (2017). There is no suggestion that Coffey (2016a/b) considered the differences in the geological facies and downhole geophysical logs when delineating their model layers or model layer properties, which does not appear to be appropriate.

The calibrated[3] hydraulic conductivity (K) and aquifer storage (S) values in the Coffey (2016a/b) model are much lower than those in the Pells and Pan (2017) model and a significant portion of the published measurement data. There are a number of reasons for these differences:

1. To match the available observation time-series the hydraulic diffusivity of the system needs to be properly represented. Thus if Pells and Pan (2017) say K is large, S must also be large. In contrast if Coffey (2016a/b) say K is small, S must be proportionally smaller.
2. Coffey (2016a/b) calibrated their model K and S parameters to the reported inflows to the Berrima Mine Void and assumed constant K and S values across their model domain. In doing so they appeared to disregard many field measurements of K which showed clear spatial trends in that data that were hypothesised to relate to geological structures.
3. Coffey (2016a/b) based their predictions on an unclear model of ultra-low permeability in the zone of desaturation above the mine-voids.
4. Coffey (2016a/b) based their predictions on a model in which water (or most water?) only drains from the rock defects (e.g. fractures) and not the rock matrix (i.e. pores) but in which groundwater flow appears to be dominated by hydraulic conductivity values more akin to matrix rather than defect hydraulic conductivity. This model was not clearly defined or substantiated with relevant data and references.
5. The groundwater recharge (R) values in Coffey (2016a/b) are notably lower than in Pells and Pan (2017) and David (2016) possibly because there are very limited measurements of groundwater recharge at the site and both modellers make assumptions based on previous work at other distant sites and different initial conditions. In addition Coffey (2016b) does not model a spatially variable recharge distribution whereas Pells and Pan (2017) do. Differences may also arise because Coffey (2016a/b) appear to assume no flow boundary conditions at some edges of their model domain, particularly in the Shoalhaven Group geology whereas Pells and Pan (2017) have chosen to undertake sensitivity testing with a general head boundary condition.

There are other differences between the two models in terms of calibration to baseflow data but the present reviewer is not prepared to provide comments on these just yet due to ambiguities in the sources of model topography data and numerical model layer definitions about key surface water drainage features that interact with groundwater.

[3] Calibration is the act of modify the input parameters of a model to match outputs of the model to historical observations such as groundwater levels and stream flows. The process of calibration is sometimes referred to as history matching. The verb "calibrated" indicates that some attempt was made to make the model better represent historical observations by changing the values of unknown model inputs. These unknown model inputs are most commonly hydraulic conductivity, aquifer storage and boundary condition values such as recharge and hydraulic heads. When calibrating models, some modellers elect to disregard some historical observations in favour of others. Observations that are often disregarded included hydraulic conductivity and aquifer storage. This can sometimes be justified on the basis of measurement error and/or geological variability.

5.6 Which model is right?

Disagreements between experts about hydrogeological site conceptualisation and modelling are relatively common. Disagreements typically arise when there are gaps in reporting, inadequate field data and/or insufficient data analysis. These gaps are then addressed with different, sometimes equally valid assumptions about the groundwater system, to arrive at vastly different predictions. We believe this impact assessment is no exception.

The present reviewer finds that the base case input parameters and assumptions adopted by both modellers (Coffey and Pells Consulting) are not unreasonable for the geological environments described, but only if they are clearly justified by explanation and supporting data and observations. For example, Coffey (2016a/b) may have very good reasons for adopting their very low drain conductance values above the workings, their low hydraulic conductivity values and their very low drainable porosity and specific storage values in the Hawkesbury Sandstone, but the absence of justification, modelling details and apparently conflicting site data does not support their case.

Based on the information presented by the proponent and all the reviewers to date, the present reviewer believes it is precautionary to assume that the Coffey (2016b) prediction represents a possible lower bound prediction of project impact and the Pells and Pan (2017) prediction represents a reasonable measure of uncertainty in the project impact predictions. If these differences are environmentally and economically significant, the key to resolving these differences will be clearer reporting, further modelling analyses and possible additional site investigations to constrain uncertainty.

5.7 Compliance with SEARs

The reviewer has not performed a comprehensive review of the SEARs as this is ordinarily performed by the Departments independent groundwater reviewer. However, he notes that the SEARs suggest a number of environmental planning instruments, guidelines, policies and plans may be relevant to the environmental assessment, in particular the Australian Groundwater Modelling Guidelines (2012) and the NSW Aquifer Interference Policy (2012).

The SEARs also required "*an assessment of the likely impacts of the development on the quantity and quality of the region's surface and groundwater resources, having regard to the EPA's, DPI's and Water NSW's requirements and recommendations*" as per SEARs Attachment 2. In SEARs Attachment 2, DPI Water recommended to the Department of Planning that the EIS be required to include a number of items including: "*A detailed assessment against the NSW Aquifer Interference Policy (2012) using DPI Water's assessment framework*" and "*Full technical details and data of all surface and groundwater modelling, and an independent peer review of the groundwater model*".

Assuming these guidelines and policies are required to be followed by the proponent and the recommendations provided by DPI Water are accepted by Department of Planning, it would appear that the SEARs have not been fully met by the proponent because:

1. The NSW Aquifer Interference Policy (2012) requires an independent peer review of the modelling work in accordance with the Australian Groundwater Modelling Guidelines (2012)
2. The proponent has only conducted peer reviews of the groundwater modelling on the basis of the much older Murray Darling Basin Groundwater Modelling Guidelines (2001) and a number of review questions from the Australian Groundwater Modelling Guidelines (2012) appear not to have been answered (see Attachment B).
3. Full technical details and data of all groundwater modelling have not been provided as summarised in this letter and as detailed in Attachment A.

4. The Australian Groundwater Modelling Guidelines recommends uncertainty analysis and no uncertainty analysis has been performed.

The importance of uncertainty analysis and reporting is highlighted in the following documents:

1. Section 1.5.5 of the Australian Groundwater Modelling Guidelines (Barnett et al., 2012);
2. Currell et al (2017) in the May 2017 edition of the Journal of Hydrology which discusses data gaps, model conceptualisation issues and court challenges for the Carmichael Coal Mine;
3. Recent comments by the QLD Land Court who rejected the Acland mine expansion proposal.

Barnett et al. (2012) states:

> *The level of effort applied to uncertainty analysis is a decision that is a function of the risk being managed. A limited analysis, such as an heuristic assessment with relative rankings of prediction uncertainty, or through use of the confidence-level classification, as described in section 2.5, may be sufficient where consequences are judged to be lower.* ***More detailed and robust analysis (e.g. those based on statistical theory) is advisable where consequences of decisions informed by model predictions are greater.*** *Because uncertainty is an integral part of any model, it is recommended to consider early in the modelling project the level of effort required for uncertainty analysis, the presentation of results and the resources required.*

Currell et al. (2017) reports:

> *Despite the large scale of the project, it appears that critical scientific data required to resolve uncertainties and construct robust models of the springs' relationship to the groundwater system were lacking at the time of approval, contributing to uncertainty and conflict. For this reason, we recommend changes to the approval process that would require a higher standard of scientific information to be collected and reviewed, particularly in relation to key environmental assets during the environmental impact assessment process in future projects"*

In 2004 the NSW Department of Infrastructure, Planning and Natural Resources (Pritchard et al., 2004) issued a report entitled *A review of the status of the groundwater resources in the Southern Highlands* in which the groundwater resources were stated to be "highly valued". Given that the EIS model (Coffey, 2016b) predicts very large impacts to a number of wells over a small area and the Pells and Pan (2017) model presents sensitivity analysis and predictions that are concerning, we recommend that the advice of the Australian Groundwater Modelling Guidelines highlighted above in bold is followed and the overall costs and benefits are assessed with triple bottom line accounting.

Acknowledgement

The Australian Groundwater Modelling Guidelines recommend that all modelling work be peer reviewed by independent third parties. In providing the commentary attached to this letter, WRL acknowledges that this review work was funded by the Coal Free Southern Highlands Inc. The modelling work has now been reviewed by a range of experts funded by both the proponent and project objectors. The reviewer does not object to the project per se, but would if the proponent cannot subsequently demonstrate that the model predictions are suitable and appropriate. The reviewer acknowledges that this is a critical review undertaken to identify potential gaps and issues with the proponent's model.

The reviewer also wishes to acknowledge the significant efforts of Coffey Geosciences who have worked on the site characterisation and conceptualisation for the Hume Coal Prospect for about the last 18-months. This is quite a short period of time in which to conceptualise site processes and properties and to develop and model groundwater inflows and impacts for such a complicated site and mine-plan given the available data, reporting limitations and value of the groundwater resource. The reviewer congratulates Coffey Geosciences for their efforts to date and recommends that they be funded to progress the modelling work and reporting to the next level. This work should address uncertainty and 'least-squares' updates to calibration to include spatially variable hydraulic conductivity and recharge parameter fields.

Thank you for considering Mr Anderson's comments on this project. If the Department's reviewers or the modellers at Coffey Geosciences have any questions regarding this review please contact Mr Anderson in the first instance (email: d.anderson@wrl.unsw.edu.au).

Yours sincerely,

G P Smith
Manager

Attachments:
References
Attachment A – Peer Review Checklists

References

Barnett B., T.L., Post V., Evans R.E., Hunt R.J., Peeters L., Richardson S., Werner A.D., Knapton A. and Boronkay A. (2012). Australian groundwater modelling guidelines, *Waterlines Report No. 82 - June 2012*. National Water Commission, Canberra

David, K. (2016) ADDENDUM TO HYDROGEOLOGICAL ASSESSMENT FOR THE CLOSURE OF BERRIMA COLLIERY FOR Boral Cement Limited, July 2016.

Ohio EPA (2006). Chapter 4 Pumping and Slug Tests. *Technical Guidance Manual for Ground Water Investigations*. State of Ohio Environmental Protection Agency, Division of Drinking and Ground Waters, December 2006.

Pritchard, S., Russell, G. and Hehir, W. (2004). A review of the status of the groundwater resources in the Southern Highlands, NSW : ensuring the sustainability of the water source New South Wales. Dept. of Infrastructure, Planning and Natural Resources

PSM (2016). RE: BERRIMA COLLIERY CLOSURE – PEER REVIEW OF GROUNDWATER STUDY, Letter to Boral by David Piccolo with reference : PSM2691-009R Rev0, Pells Sullivan Meynink, 1 April 2016.

McCartney, J. S. 2007. Determination of the hydraulic characteristics of unsaturated soils using a centrifuge permeameter. 3277571 Ph.D., The University of Texas at Austin.

Neuzil, C.E. (1994) “How Permeable are Clays and Shales?” *Water Resources Research*. 30(2):145-150

Person, M., Raffensperger, J.P., Ge. S. and Garven, G. (1996). “Basin-Scale Hydrologic Modelling”. *Reviews of Geophysics*, 34, 61-87.

Pritchard, S., Russell, G. and Hehir, W. (2004). A review of the status of the groundwater resources in the Southern Highlands, NSW : ensuring the sustainability of the water source New South Wales. Dept. of Infrastructure, Planning and Natural Resources

Tammetta, P. (2012). Groundwater Study, Area 3B Dendrobium Coal Mine, Numerical Modelling. Prepared by Coffey Geotechnics for BHP Billiton Illawarra Coal.

Tammetta, P. and Hewitt, P. (2004). Hydrogeological properties of Hawkesbury Sandstone in the Sydney region. Australian Geomechanics Journal, 39, 93-108.

Hume Coal Project SSD 15_7172

UNSW-WRL Peer Review of Conceptual and Numerical Modelling that Predicted Likely Groundwater Impacts

ATTACHMENT A – PEER REVIEW CHECKLISTS

Review Questions AGMG (2012) *MDBC (2001)**	**EIS Model (Coffey 2016a/b)**						**Pells and Pan (2013-2017) Model**	
	Kalf (2016)		***Merrick (2016)***		***Anderson (2017) – This Review***		***Anderson (2013,2017)***	
	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
1. Planning								
1.1 Are the project objectives stated? *(1.1) Is there a clear statement of project objectives in the modelling report?*	Very Good		Very Good	For both documents.	Yes	"*assess impacts on the groundwater system and groundwater users due to the proposed mining*"	Yes	To highlight gaps in the EIS assessment. To provide an appropriate sensitivity analysis of the likely impacts of the project proposal to address gaps in the EIS. To provide Department of Planning and DPI Water with sufficient information to assess the uncertainty associated with the impacts of the project proposal.
1.2 Are the model objectives stated?					No	Model objectives are not clearly stated in a single section of the document.	Yes	The purpose of the model is to: - test the modelling assumptions, inputs, outputs and predictions provided by the proponent of the EIS - provide sufficient sensitivity analysis to allow an uncertainty analysis to be performed.
1.3 Is it clear how the model will contribute to meeting the project objectives? *(1.4) Has the modelling study satisfied project objectives?*	Very Good		Very Good	AIP and WSP requirements are met. Drawdown mitigation options are explored.	Yes	The model provides one possible prediction of project impacts. Other predictions are possible and this is not reported in sufficient in detail.	Yes	The Pells and Pells (2013) and Pells and Pan (2017) reports provide a very good assessment of the uncertainty and identifies several possible issues with the EIS model.
1.4 Is a groundwater model the best option to address the project and model objectives?					Yes		Yes	
1.5 Is the target model confidence-level classification stated and justified? *(1.2) Is the level of model complexity clear or acknowledged?*	Yes		Yes	70% Class 3 & 30% Class 2 confidence classification. What is achieved could be substantiated by ticking checklist attributes for all classes, as models will bridge several classes.	Yes	Groundwater flow in Hawkesbury Sandstone is complex due to dual porosity (matrix and defect flow) and is difficult to model. For the Westconnex New M5 project which tunnelled in the same geological basin, the modellers elected not to develop a transient model because of these issues and because of limited ground water level and inflow data for other tunnels in the basin. The subjective checklist provided in the guidelines is not presented by Coffey (2016b) to demonstrate how these classifications (previous column) were arrived at. The rating seems a little high given our assessment of the data, model setup and calibration. We assess the model satisfies about 40% of Class 3 criteria, 60% of Class 2 and 30% of the Class 1 criteria. This makes it a Class 1 model with some elements of Class 3 but mainly Class 2. Some of the most important aspects such as model convergence fall within Class 1.	No	A model confidence-level classification of Class 1 was assigned by this peer reviewer in 2013. The confidence level has been improved since then and the model is appropriate for the intended model purpose. Less effort was expended on the discretisation and calibration than in the EIS model but reporting also demonstrates that considerably more effort was expended than in the EIS to ensure: • the topography was valid; • the aquifer storage values were valid; • the parameters matched the available field data, upscaled lab / geology data and theory; • numerical model solutions and simplifications were suitable, converging and stable; • sensitivity analysis was thorough. If the EIS model was to demonstrate these aspects and concern remained about the range of predictions presented then there would be reason to update this model to a higher confidence level classification.
1.6 Are the planned limitations and exclusions of the model stated?					Yes	It is acknowledged that the assessment has been prepared generally in accordance with the Australia Groundwater Modelling Guidelines and the NSW Aquifer Interference Policy. Coffey (2016b) state: "*In fractured media with large mining stresses, the modelling results will not exactly represent conditions on a local scale but are more representative on a medium to regional scale. Actual observations made in the future, during Hume mine operation, may differ from predictions made herein*"	Yes	The model cannot be developed to the same extent as the EIS model because many key data-sets collected by the proponent have not been published in the public domain or published in a readily accessible electronic format.
1.7 (1.3) Is the level of model complexity clear or acknowledged?	Very Good		Very Good	Doc B: Tables 4, 10, 11.				
1.8 (1.5) Are the model results of any practical use?	Yes		Yes	Adequately calibrated to multiple observation datasets. Reliability is thereby enhanced.	Yes	On the basis of the data and analysis presented and the information and reporting gaps, the model results appear to provide a lower bound estimate of the potential impacts of the project.	Yes	The results are suitable for exploring the sensitivity of the model prediction to uncertainty in the model inputs.
1.9 (1.6 – Merrick, 9.2 – Kalf) Is the model "fit-for-purpose"?	Yes		Yes	Purpose is assessment of potential environmental impacts due to mining method and mine plan. Fitness is defended in Section 4.3.5.	No	There appears to be biases in the input parameters utilised for calibration. Sensitivity analysis is limited and appears problematic and no uncertainty analysis is presented.	Yes	The model explores the potential range of impacts based on sensitivity testing of numerous model inputs and numerical modelling assumptions.

Review Questions AGMG (2012) *MDBC (2001)**	EIS Model (Coffey 2016a/b)						Pells and Pan (2013-2017) Model	
	Kalf (2016)		***Merrick (2016)***		***Anderson (2017) – This Review***		***Anderson (2013,2017)***	
	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
2. Conceptualisation								
2.1 Has a literature review been completed, including examination of prior investigations?					Yes	A number of statements and interpretations do not make reference to the original sources of the data.	Yes	The range of historical data sources identified in the preparation of the model indicate that a suitable literature review has been undertaken.
2.2 Is the aquifer system adequately described?								
2.2.1 hydrostratigraphy including aquifer type (porous, fractured rock ...)					No	Detailed hydrostratigraphy layer elevation information is not presented. It is customary to present maps of key model layer elevations and properties. No reporting of upscaling exercises on the available geological and geophysical logs with facies hydraulic conductivity and storage data to support the modelled values of hydraulic conductivity and aquifer confined storage (specific storage) and unconfined storage (specific yield); Specific storage values for modelled formations appear inconsistent with mathematical equations applied to typical rock porosity and compressibility data for the Sydney Basin;	Yes	The figures and text provide a good indication of the hydrostratigraphy
2.2.2 lateral extent, boundaries and significant internal features such as faults and regional folds					Yes, No, No	No flow boundaries along inferred structures are specified without presenting or clearly discussing the supporting hydraulic evidence for the compartmentalisation;	Yes	Locations of known faults are shown in cross section view (but not in plan). No base-flow information identified for rivers and streams.
2.2.3 aquifer geometry including layer elevations and thicknesses					No	This should be provided. The limited information providing is conflicting. Refer to Section 7 of the peer review letter.	Yes	Elevations shown in report figures and described in text. Thickness must be extracted from the model files.
2.2.4 confined or unconfined flow and the variation of these conditions in space and time?					No	Given the layering of the formations at this site and the time-series data that is available, this aspect of the conceptualisation could be significantly improved.	No	From the data present in the report it is apparent that the Hawkesbury Sandstone aquifer varies between confined and unconfined flow regimes.
2.3 Have data on groundwater stresses been collected and analysed?								
2.3.1 recharge from rainfall, irrigation, floods, lakes *(2.3) Have all potential recharge data been collected and analysed? (rainfall, streamflow, irrigation, floods, etc.)*	Very Good		Very Good	Rainfall recharge is well constrained by baseflow analysis.	Yes, No, ?, ?	• Rainfall: Groundwater recharge appears to be set at the lower bounds of what could be considered reasonable on the basis of the available data. There appear to be some issues with interpretation of rainfall data in Figure 3.2, e.g. 2.2m of annual rainfall in NSW?. Refer to Section 7.4.3 of this peer review for more details. • Irrigation recharge: A small bias is introduced because irrigation is excluded from the conceptual and numerical model but irrigation pumping is included.	Yes, No, ?, ?	Recharge from rainfall based on literature data and is in the range of 5 – 50 mm/a for various scenarios. No discussion of irrigation.
2.3.2 river or lake stage heights	Very Good		Very Good	Groundwater and creek water levels are examined to infer gaining/losing status.	?	The model is based on less accurate STRM rather than more accurate LIDAR topography. This could result in a number of issues. For further details see the review letter.	Yes	The model is based on LIDAR topography.
2.3.3 groundwater usage (pumping, returns etc) *(2.4) Have all potential discharge data been collected and analysed? (abstraction, evapotranspiration, drainage, springflow, etc.)*	Very Good		Very Good	Substantial private abstraction – difficult to estimate.	?	It is suggested in the EIS that one pump test was run for seven (7) days (about 10,000 minutes) but presenting and analysing data for only about 3-4 days (about 5,000 minutes), just as boundary conditions in the aquifer appeared to start influencing the test results. The draft pumping test interpretation report was not provided for independent peer review;	No	The model simulates changes from existing conditions.
2.3.4 evapotranspiration							No	The model simulates changes from existing conditions.
2.3.5 other? *(2.5) Have the recharge and discharge datasets been analysed for their groundwater response?*	Very Good		Adequate	Limited illustration of cause-and-effect analysis (e.g. Figure 3.3 and commentary in Section 6.1). Hydrographs in Appendix D are not compared with rainfall residual				

Review Questions AGMG (2012) *MDBC (2001)**	**EIS Model (Coffey 2016a/b)**						**Pells and Pan (2013-2017) Model**	
	Kalf (2016)		***Merrick (2016)***		***Anderson (2017) – This Review***		***Anderson (2013,2017)***	
	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
				mass. Vertical head gradients are considered (e.g. minor in Figure 6.6).				
2.4 Have groundwater level observations been collected and analysed? *(2.1) Has hydrogeology data been collected and analysed?*	Very Good		Very Good	Large groundwater database: 59 sampling points at 24 sites. Extensive aquifer/aquitard characterisation by packer tests (28), core tests (59) and pumping tests (2). Good coverage of hydrostratigraphy and water quality.	Yes	Much better than many past assessments but there appear to be some issues and several issues with the translation of this data and analysis into the numerical model. For details refer to the peer review letter. Much of the raw data and supporting work has not been tabulated or exhibited to allow third parties to replicate the modelling.	Yes/No	
2.4.1 selection of representative bore hydrographs					?	Unclear, could be justified with reasons.	No	Transient water level records are not identified.
2.4.2 comparison of hydrographs					Yes		No	
2.4.3 effect of stresses on hydrographs					Yes / Maybe Not	The aquifer pumping test reports are in draft status and not exhibited. There appears to be some undiscused issues with the tests and the data / interpretations. Some aquifer stress tests are interpreted with confined assumptions. Others are interpreted as unconfined. For further details see the review letter	Yes	Results of four single well tests with pumping stresses resulting in drawdown of 15 – 50 m.
2.4.4 watertable maps/piezometric surfaces? *(2.2) Are groundwater contours or flow directions presented?*	Very Good		Adequate	Flow directions should be added to Figure 6.3: Wianamatta Group and Upper Hawkesbury Sandstone head surfaces. Similar maps in Appendix E for Lower Hawkesbury Sandstone and Wongawilli Seam.	Yes		Yes	Collation of static water level data with accuracy of approximately ± 20 m. A selection of bore standing water levels (SWL) is contoured in a water-table / piezometric map in Figure 19. Based on Figure 48, the map in Figure 19 is a composite of SWL data from multiple geological formations.
2.4.5 If relevant, are density and barometric effects taken into account in the interpretation of groundwater head and flow data?					?	Unclear, this should be clarified	No	Not required for this model
2.5 Have flow observations been collected and analysed?					Yes	Utilised and analysed outflow data from Berrima Mine. Some aspects should be clarified. See the review letter.	No	
2.5.1 baseflow in rivers					Yes		No	
2.5.2 discharge in springs					No		No	
2.5.3 location of diffuse discharge areas?						Topography was based upon low resolution, low accuracy Shuttle Radar Topography Mission (SRTM) data rather than the higher resolution, more accurate, LIDAR data that may result in erroneous predictions of groundwater-surface water interactions and GDE impacts	No	
2.6 Is the measurement error or data uncertainty reported?					No			
2.6.1 measurement error for directly measured quantities (e.g. piezometric level, concentration, flows)					No		Yes	The accuracy of the groundwater level data compilation is identified as ± 20 m. Accuracy for pump test measures are not identified.
2.6.2 spatial variability/heterogeneity of parameters					No	A very good map of horizontal hydraulic conductivity variability in Hawkesbury Sandstone is presented for a very small section of the model domain. Additional maps and variability for other parameters and geological layers could be described on the basis of the geological logs and parameter upscaling exercises. Hydraulic conductivity values in Hawkesbury Sandstone that are spatially invariable despite clear and significant mapping	Yes/ No	Uncertainty in the hydraulic conductivity parameter is estimated for each model layer. Anisotropy, porosity, specific yield and specific storage are estimated at reasonable values. Heterogeneity of parameters within a model layer are not identified.

Review Questions AGMG (2012) *MDBC (2001)**	EIS Model (Coffey 2016a/b)						Pells and Pan (2013-2017) Model	
	Kalf (2016)		***Merrick (2016)***		***Anderson (2017) – This Review***		***Anderson (2013,2017)***	
	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
						evidence by the proponent's experts to the contrary; Hydraulic conductivity values in Hawkesbury Sandstone at two (2) to two hundred (200) times lower than best available field measurements in the public domain and about ten (10) times lower than what was modelled by Boral at the nearby Berrima Mine.		
2.6.3 interpolation algorithm(s) and uncertainty of gridded data?					No		No	Combination of hand contouring and digital interpolation techniques for the base of Wianamatta Shales and the top of the Illawarra Coal Measures. Other surfaces are derived conceptually using engineering judgement.
2.7 Have consistent data units and geometric datum been used?	Yes		Yes		Yes		Yes	
2.8 Is there a clear description of the conceptual model? *(3.2) Is there a clear description of the conceptual model?*	Yes		Yes	Thorough development of conceptualisation.	Yes	Some aspects could be clearer. Refer to the peer review letter.	Yes	
2.8.1 Is there a graphical representation of the conceptual model? *(3.3) Is there a graphical representation of the modeller's conceptualisation?*					Yes	Appendix H is a conceptual model and it contains several different figures describing various aspects of the conceptual model. There is a revised conceptual model in Appendix I in relation to numerical model parameters.	Yes	Figures of data that comprised the input into the numerical model and one cut-away figure of the 3D model stratigraphy. Report would benefit from a reproduction of the modelled stratigraphy at the cross-sections identified in Figure 16.
2.8.2 Is the conceptual model based on all available, relevant data?					No	The geological logs and different scales of field and laboratory data have not been combined and upscaled to support suitable parameter distributions.	Yes	
2.9 Is the conceptual model consistent with the model objectives and target model confidence level classification? *(3.1) Is the conceptual model consistent with project objectives and the required model complexity?*	Yes		Yes	Thorough development of conceptualisation.	No	The conceptual / numerical model parameters for the base case scenario appear to gravitate towards lower bound values for what might be plausible. Some calibration to larger recharge, hydraulic conductivity and storage parameters with updated boundary conditions should be attempted in a calibration optimiser. Irrigation returns excluded.	Yes*	*See comments below.
2.9.1 Are the relevant processes identified?					Yes		Yes	
2.9.2 Is justification provided for omission or simplification of processes?					No	Examples include the representation of hydraulic conductivity and aquifer storage processes in relation to the dual porosity of the Hawkesbury Sandstones, no-flow boundary conditions etc. For full details see the review letter.	Yes	Simplifications to the model geometry and boundary conditions are described and justified in the various reports and examined through sensitivity testing.
2.10 Have alternative conceptual models been investigated?					No	No the established parameter model appears to be at the lower bounds of reasonable hydraulic conductivity, storage and recharge values and calibration to larger values of these parameters is not disproven. There are inconsistencies in the hydrogeological flow system conceptualisation between the groundwater flow model and geochemical impact assessments	Yes	- Three alternative hydraulic conductivity-recharge cases (low, medium and high) - Two sets of alternate boundary condition concepts (fixed head, general head / seepage). - With and without subsidence induced fracturing.
2.11 (3.4) Is the conceptual model unnecessarily simple or unnecessarily complex?	No		No	Some unnecessary features are justifiably excised	Yes	To be commensurate with the data collected, the value of the groundwater resource and the magnitude of impacts predicted, the model should have: • a spatially variable percentage of recharge assigned based on land use / cover • a spatially variable hydraulic conductivity represented in Hawkesbury Sandstone and in the model layers above mine workings.	No	The model simplicity / complexity is suitable for the stated modelling purpose.

Review Questions AGMG (2012) *MDBC (2001)**	EIS Model (Coffey 2016a/b)						Pells and Pan (2013-2017) Model	
	Kalf (2016)		*Merrick (2016)*		*Anderson (2017) – This Review*		*Anderson (2013,2017)*	
	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
3. Design and construction								
3.1 Is the design consistent with the conceptual model?					No	See the review letter for further details.	Yes	A detailed check of the elevations of the numerical model layers against field data was not undertaken
3.2 Is the choice of numerical method and software appropriate (Table 4-2)? *(4.3) Is the software appropriate for the objectives of the study?*	Yes		Yes	MODFLOW SURFACT v3. There is a later v4 that allows time-varying properties; however, not necessary for this mine model.	Maybe Not	The proponent has not reported on the exact numerical methods used within MODFLOW SURFACT v3 or the inputs to these numerical methods.	Yes	
3.2.1 Are the numerical and discretisation methods appropriate?					Maybe Not	See above	Yes	Numerical methods whilst not discussed in the report text are identified in the model files. Discretisation is appropriate.
3.2.2 Is the software reputable?					Yes		Yes	
3.2.3 Is the software included in the archive or are references to the software provided					No	No electronic archive of model files or software was distributed for public review.	No*	The 2013 model files were made available for review. The 2017 model files have not been inspected.
3.3 Are the spatial domain and discretisation appropriate? *(4.1) Is the spatial extent of the model appropriate?*	Yes		Yes	Dimensions 38 km x 32 km. Area 752 km2. 379 columns x 425 rows x 15 layers. Cell sizes 50m to 200m. 6 layers in Hawkesbury Sandstone. Separate layers for caving. Local model: 6 km x 4 km. Area 15 km2. 100 columns x 40 rows x 2 layers. Cell size 100m.	Maybe Not	No sensitivity tests reported to demonstrate grid convergence. No reporting of how model was calibrated to pump test data given that all pumping wells and observation points would fall within one model grid cell.	Yes	
3.3.1 1D/2D/3D					3D		Yes	3D
3.3.2 lateral extent					Yes	If boundary conditions are appropriate	Yes	For the chosen model extents general head boundary conditions are most appropriate.
3.3.3 layer geometry					Maybe Not	See the review letter for full details. In summary: • The layer geometry is summarised in a table but maps are not provided. There is no way to check the validity of the geometry and there is conflicting information presented in Table 2 and 3 in Appendix I of EIS Vol4B. • The model is based on STRM topography rather than LIDAR data. This might distort layer geometry and modelled properties near key river / drain boundary conditions.	Yes	
3.3.4 Is the horizontal discretisation appropriate for the objectives, problem setting, conceptual model and target confidence level classification?					Yes		Yes	
3.3.5 Is the vertical discretisation appropriate? Are aquitards divided in multiple layers to model time lags of propagation of responses in the vertical direction?					Maybe Not	A number of model layers are represented immediately above the mine workings and this is good. However, no maps of layer thickness or elevation are presented and our analysis of Table 2 and 3 in Appendix I, Appendix L and Figure C4 in Appendix H of Volume 4B of the EIS raises a number of issues regarding the nature of the model layers that were actually modelled. For further details see the review letter.	OK	Additional model layers were added following early peer review in 2013. There are not as many layers as the current EIS model. The discretisation could be improved if the proponent released all their borelogs in electronic format into the public domain and they publish elevation maps and electronic data-sets defining their model layer elevations and geological layer picks.
3.4 Are the temporal domain and discretisation appropriate?					Maybe	The proponent states that there is a correlation lag between rainfall and inflows at Berrima Mine. The calibration of the model to this time lag is not reported in the EIS.	Yes	
3.4.1 steady state or transient					Yes	Transient	Yes	Transient
3.4.2 stress periods							Yes	
3.4.3 time steps?							Yes	

Review Questions AGMG (2012) *MDBC (2001)**	EIS Model (Coffey 2016a/b)						Pells and Pan (2013-2017) Model	
	Kalf (2016)		***Merrick (2016)***		***Anderson (2017) – This Review***		***Anderson (2013,2017)***	
	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
3.5 Are the boundary conditions plausible and sufficiently unrestrictive? *(4.2) Are the applied boundary conditions plausible and unrestrictive?*	Very Good		Adequate	Generally no-flow at distant borders, fixed head at Wingecarribee Reservoir, reference heads at streams (RIV and DRN), escarpment (DRN) and mines (DRN). All appropriate. RIV leakance of 0.01 /day is high, hence likely to overestimate mining induced losses from streams (conservative assumption). Not included: Irrigation recharge; Cement Works Fault; Robertson Basalt; Mt Gingenbullen intrusion.	Maybe Not	See comments below	Yes	
3.5.1 Is the implementation of boundary conditions consistent with the conceptual model?					No	The recharge appears to be at the lower bounds of the reported water table fluctuation analysis data and inconsistent with the recharge values reported in the region by other authors. For further details see the review letter.	Yes	A range of recharge values were tested to loosely devise several hydraulic conductivity (K) calibrations to the available data.
3.5.2 Are the boundary conditions chosen to have a minimal impact on key model outcomes? How is this ascertained?					No	The modellers assume boundary conditions based on assumed no-flow divides. There is no sensitivity testing with general head boundary conditions. No reporting of sensitivity on recharge is reported. For further details see the review letter.	Yes	A range of groundwater recharge rates was considered. Sensitivity testing was undertaken to demonstrate that the model was sensitivity to the assumption of no-flow boundary conditions in the Shoalhaven Group and modelling proceeded with general head rather than no-flow boundary conditions in these layers.
3.5.3 Is the calculation of diffuse recharge consistent with model objectives and confidence level?					No		Yes	
3.5.4 Are lateral boundaries time-invariant?					No	Not stated	Yes	
3.6 Are the initial conditions appropriate?					?	Not clearly stated. Model was started in 1979 in the middle of mining at Berrima and run (warmed up) for 32 years to 2011 to create an initial condition after which time a period a transient calibration was attempted. It is unclear if the model warm up was repeated after the model was calibrated to verify the initial condition and calibration against the available field observations of heads and mine inflows.	Yes	For the model objectives
3.6.1 Are the initial heads based on interpolation or on groundwater modelling?					Model		Yes	Groundwater modelling
3.6.2 Is the effect of initial conditions on key model outcomes assessed?					No	This could be a significant gap in the assessment.	No	Not stated
3.6.3 How is the initial concentration of solutes obtained (when relevant)?					No	Not modelled	No	Not modelled
3.7 Is the numerical solution of the model adequate?					No	The model has large mass balance errors. This suggests non-convergence. The modellers report conducting sensitivity tests of the mine drain conductance parameter which should be the most sensitivite parameter in the model but report it is 0% sensitive to perturbation which is difficult to believe. For further details see the review letter.	Yes	The mass balance error is less than 0.1%
3.7.1 Solution method/solver					No	Not stated	No	Not stated
3.7.2 Convergence criteria					No	Not stated	No	
3.7.3 Numerical precision					No	Not stated	No	
4. Calibration and sensitivity								
4.1 Are all available types of observations used for calibration? *(2.6) Are groundwater hydrographs used for*	Yes		Yes	Extensive monitoring network.	No	Pumping test hydraulic conductivity data appears to be ignored. Calibration is not demonstrated for all data types. For further details see the review letter.	No	Much of the data collected by the proponent is not published in a readily accessible electronic format. The model is calibrated to pumping test hydraulic conductivity data.

Review Questions AGMG (2012) *MDBC (2001)**	**EIS Model (Coffey 2016a/b)**						**Pells and Pan (2013-2017) Model**	
	Kalf (2016)		***Merrick (2016)***		***Anderson (2017) – This Review***		***Anderson (2013,2017)***	
	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
calibration?								
4.1.1 Groundwater head data					Yes		Yes	DEM, SWL and aquifer test data.
4.1.2 Flux observations					Yes	Baseflow fluxes are reportedly used, however, calibration is not demonstrated.	No	
4.1.3 Other: environmental tracers, gradients, age, temperature, concentrations etc.					No	No calibration to horizontal or vertical gradients for natural or pump testing gradients incorporating the geometry utilised for the 3D numerical model. No calibration to chemistry or tracer data to constrain the recharge parameter.	No	
4.2 Does the calibration methodology conform to best practice?					No	The proponent appears to have attempted observation history matching (model calibration) with a trial and error approach. A software tool that facilitates unbiased and semi-automated calibration to the data was not used. Such software examines residuals and ensures that the sum squared of residuals is minimised with minimum bias. The model calibration (Figure 4.2 in Appendix I, Vol 4B) from trial and error shows some bias.	No	
4.2.1 Parameterisation					No	There appears to be several issues with parameterisation. For details refer to "summary of specific issues" in cover letter and Section 7 of the detailed peer review comments.	No	
4.2.2 Objective function					No	The objective function used to define and determine calibration is not reported. For example, what are the weights and residuals for heads, baseflow fluxes, hydraulic conductivity data, specific storage, specific yield and mine inflows for the model?	No	
4.2.3 Identifiability of parameters					No	Identifiability of parameters, which highlights the most important parameters in the model, is not reported.	No	
4.2.4 Which methodology is used for model calibration?					Yes	Manual trial and error based on the experience of the numerical modeller at Coffey.	Yes	Manual trial and error based on the experience of Mr Steven Pells.
4.3 Is a sensitivity of key model outcomes assessed against? *(8.1) Is the sensitivity analysis sufficiently intensive for key parameters?*	Adequate		Adequa te	A: caving height; B: vertical hydraulic conductivity; and C: mine drain conductance.	No	For further details see the review letter.	Yes	A fairly comprehensive assessment of parameter sensitivity.
4.3.1 parameters					No		No	The report summarises the most sensitive variables (hydraulic conductivity and perimeter boundary conditions), however, the results of incremental changes to parameter, boundary and initial conditions are not presented in the report. This is acceptable for the stated model purpose and confidence level classification.
4.3.2 boundary conditions					No		No	
4.3.3 initial conditions					No		No	
4.3.4 stresses					No	The range of values tested for the drain conductance parameter is too small.	No	
4.4 Have the calibration results been adequately reported? *(5.1) Is there sufficient evidence provided for model calibration?*	Very Good		Very Good	Several lines of evidence: scattergram; performance statistics for verification; hydrographs plots; spatial pattern; section head pattern; mine inflow; baseflows; K values. Did not use PEST. No indication of spatial distribution of residuals. No scattergram for full calibration period.	No	Calibration to baseflow gauging data is not reported.		
4.4.1 Are there graphs showing modelled and observed hydrographs at an appropriate scale?							Yes	
4.4.2 Is it clear whether observed or assumed vertical head gradients have been replicated by the model?					No		Yes	Vertical head gradients between the Hawkesbury Sandstone and the Shoalhaven Group and from parts of the Wiannamatta Shales to the Hawkesbury Sandstones are under-estimated.
4.4.3 Are calibration statistics reported and					Yes	For h	Yes	RMS error values are not calculated.

Review Questions AGMG (2012) *MDBC (2001)**	**EIS Model (Coffey 2016a/b)**						**Pells and Pan (2013-2017) Model**	
	Kalf (2016)		***Merrick (2016)***		***Anderson (2017) – This Review***		***Anderson (2013,2017)***	
	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
illustrated in a reasonable manner?								
4.4.3 (8.2) Are sensitivity results used to qualify the reliability of model calibration?	Yes		Missing	Not reported.	No			
4.5 Are multiple methods of plotting calibration results used to highlight goodness of fit robustly? Is the model sufficiently calibrated?					Yes, Maybe not	• There is a 1:1 calibration plot for hydraulic heads which shows a mean bias towards underprediction of heads. • There is one cross section of observed (inferred) and modelled heads which shows some significant differences, including the inability of the model to predict a water table above the Berrima mine workings. • There is a comparison of "calibrated" and observed hydraulic conductivity which does not match the proponents pumping test data and ignores the publically available pumping test results reported by Pells and Pells (2013) and reproduced by Pells and Pan (2017) • The calibration mass balance error is 4% • A better calibration might be achieved with spatially variable parameter fields derived from Coffey (2016a) and analysis and upscaling of the geological logs using collected field and laboratory data. For further details see the review letter.	Yes, Maybe not	• 1:1 plots in Pells and Pells (2013) have large residuals than the Coffey (2016b) model, however, • The comparison of "calibrated" and observed hydraulic conductivity is a much better than the Coffey (2016b) model, and • The model mass balance error is less than 0.1% which is much better than 4 in the Coffey (2016b) model
4.5.1 spatially *(5.2) Is the model sufficiently calibrated against spatial observations?*	Very Good		Adequate	Spatial head pattern in Figure 4.4 is as expected. Section head pattern in Figure 4.3 is reasonable. Vertical head gradients on hydrographs – some good, some poor.	No	No plan view spatial distribution of residuals is shown. A significant amount of effort is required to interpret the model calibration. Many of the time-series calibration plots show predictive errors as large as 10m to 50m and quiescent (dormant) water level responses suggesting significant predictive error across much of the model domain.	No	No plan view spatial distribution of residuals is shown. A significant amount of effort is required to interpret the model calibration.
4.5.2 temporally *(5.3) Is the model sufficiently calibrated against temporal observations?*	Very Good		Adequate/Very Good	Mine inflow very good. Hydrographs for all bores are presented for comparison. Reasonable groundwater level matches and trends – simulated responses are quiescent. Some vertical gradients are reproduced.	No	Mine outflows at Berrima Mine are shown to be used for calibration of mine inflows at Berrima Mine: • Mine outflows are not mine inflows. • The time-step of the output appears to be greater than 1 year when the model stress periods are monthly. What are the residuals for the monthly predictions and is the lag time of about one year from significant rainfall to significant inflow reported by Coffey (2016a) matched by the model? Many of the time-series calibration plots show predictive errors as large as 10m to 50m and quiescent (dormant) water level responses. The only observation point that appears to be fitted very well or exactly is monitoring well HU_38. It is unclear how such a perfect fit was achieved at this well by trial and error, what modelled pumping stresses this observation well is responding to and whether this is a representative well to calibrate to over all others given its location (below average horizontal hydraulic conductivity and what appears to be zero interburden thickness). This aspect of calibration needs to be discussed and justified in detail.	No	See 4.8 below.
4.6 Are the calibrated parameters plausible? *(5.4) Are calibrated parameter distributions and ranges plausible?*	Yes		Yes	Figure 4.5	No	There are no distributions of parameters modelled within model layers and the values do not appear to match the available site data or other models in the area. For further details see the review letter.	Yes	Consistent with literature and field test data
4.7 Are the water volumes and fluxes in the water balance realistic?							Yes	A predicted mine inflow of 60 ML/d over a 45 km² area is equivalent to a seepage rate of 1.3 L/d per m² or a seepage velocity of 1.3 mm/d.
4.8 has the model been verified?					Some	There appears to be no verification to the Berrima Mine inflow data, no reproduction of pumping test drawdown or K values within a model with the same statigraphy and properties.	Yes / No	Results of four single well aquifer tests (see detailed comments in the body of this letter).
4.8.1 (6.1) Is there sufficient evidence	Only Just Adequate	Preliminary but much longer	Adequate	Several lines of evidence: scattergram; performance	No	Some of the comments provided by Merrick (2016) in previous column pertain to model calibration, not model verification.	No	

Review Questions AGMG (2012) *MDBC (2001)**	**EIS Model (Coffey 2016a/b)**						**Pells and Pan (2013-2017) Model**	
	Kalf (2016)		***Merrick (2016)***		***Anderson (2017) – This Review***		***Anderson (2013,2017)***	
	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
provided for model verification?		period would be desirable with ongoing monitoring.		statistics; hydrographs plots; spatial pattern; section head pattern; mine inflow; baseflows. No comparison offered with calibration performance. Short period of time (6-8 months).				
4.8.2 (6.2) Does the reserved dataset include stresses consistent with the prediction scenarios?	NA / Unknown		Yes	Short period of time (6-8 months).	No	Validation / verification period was from January 2015 to August 2015 which is many decades after mining commenced at Berrima Mine so model is calibrate to late time inflows when there is a significantly reduced hydraulic head above the workings compared to the commencement of mine.		
4.8.3 (6.3) Are there good reasons for an unsatisfactory verification?	NA / Unknown		Adequate	Reasons given in Section 4.3.1 for VWP groundwater levels.				
4.9 (5.5) Does the calibration statistic satisfy agreed performance criteria?	Adequate		Adequate	3.1m residual mean; 17mRMS; 12 %RMS.	No	The Australian Groundwater Modelling Guidelines (2012) guiding principle 5.4 states that "*performance mesuares should be agreed prior to calibration, and should include a combination of quantitative and non-quantitative measures*". No calibration performance targets appear to have been specified by Department of Planning or DPI Water prior to modelling. The scaled root mean squared error (SRMS) of 5% or 10%, not 12% is usually considered acceptable		
4.10 (5.6) Are there good reasons for not meeting agreed performance criteria?	NA / Unknown	Performance criteria have been met	Adequate	Reasons given in Section 4.3.1 for VWP groundwater levels.	No	No evidence of performance criteria being agreed with regulators based upon reviewed documentation.		
5. Prediction								
5.1 Are the model predictions designed in a manner that meets the model objectives?							Yes	
5.2 Is predictive uncertainty acknowledged and addressed?							Yes	Acknowledged and addressed through consideration of upper and lower bound variations to the hydraulic conductivity parameters and through variations to the recharge values and model boundary conditions.
5.3 Are the assumed climatic stresses appropriate?							Yes	
5.3.1 (7.1) Have multiple scenarios been run for climate variability?	No		Adequate	Average climate only.	No			
5.4 Is a null scenario defined?					No		No	Model predicts changes from mining operations relative to three similar but different initial conditions (the null scenarios). These scenarios are based on a combination of static groundwater level data from a number of dates and a range of realistic recharge values.
5.5 Are the scenarios defined in accordance with the model objectives and confidence level classification?							Yes	
5.5.1 Are the pumping stresses similar in magnitude to those of the calibrated model? If not, is there reference to the associated reduction in model confidence?					No	Not stated	Yes	The ultimate stresses induced by the simulation of mine void are approximately 2-6 times the reference stresses considered during model validation
5.5.2 Are well losses accounted for when estimating maximum pumping rates per well?					No	Not stated	N/A	The model does not simulate pumping wells, except for validation. Well losses are considered in the validation exercise.

Review Questions	**EIS Model (Coffey 2016a/b)**						**Pells and Pan (2013-2017) Model**	
AGMG (2012)	***Kalf (2016)***		***Merrick (2016)***		***Anderson (2017) – This Review***		***Anderson (2013,2017)***	
*MDBC (2001)**	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
5.5.3 Is the temporal scale of the predictions commensurate with the calibrated model? If not, is there reference to the associated reduction in model confidence? *(7.3-Merrick) Is the time horizon for prediction comparable with the length of the calibration / verification period?* (7.3 - Kalf) Is the time period for prediction comparable with the duration of the calibration period?	Greater than		No	19 years' prediction. 5 years combined calibration and verification.			No Yes	The calibration and validation is based on static water levels and four 2-day single well tests and the model makes predictions 40 years into the future. This is appropriate for the state model purpose and confidence level.
5.5.4 Are the assumed stresses and timescale appropriate for the stated objectives?							Yes	
5.5.5 (7.2) Have multiple scenarios been run for operational /management alternatives?	No		Very Good	Mitigation options. But only the base case is reported.				
5.6 Do the prediction results meet the stated objectives?							Yes	
5.7 Are the components of the predicted mass balance realistic?								
5.7.1 Are the pumping rates assigned in the input files equal to the modelled pumping rates?					?	Not stated		Pumping wells not simulated, except for validation. Not reviewed.
5.7.2 Does predicted seepage to or from a river exceed measured or expected river flow?					?	Not clearly stated for all surface water features		Modelling report does not identify seepage rates to or from the rivers or estimate river flows. Not reviewed.
5.7.3 Are there any anomalous boundary fluxes due to superposition of head dependent sinks (e.g. evapotranspiration) on head-dependent boundary cells (Type 1 or 3 boundary conditions)?					?	Not stated	No	
5.7.4 Is diffuse recharge from rainfall smaller than rainfall?					Yes		Yes	
5.7.5 Are model storage changes dominated by anomalous head increases in isolated cells that receive recharge?					?	Not stated	No	
5.8 Has particle tracking been considered as an alternative to solute transport modelling?					No	Particle tracking should be undertaken to provide a better understanding of the model behaviours and impacts.	N/A	
5.9 (7.4) Are the model predictions plausible?	Yes		Yes	Thorough investigation.	Maybe	Based on the available reporting this reviewer believes the predictions may gravitate towards under-predicting mine inflows and impacts.	Yes	The model highlights the uncertainty and sensitivity of the model prediction and will be an invaluable aid in assessing the impacts of the project.
5.10 (8.3) Are sensitivity results used to qualify the accuracy of model	Adequate		Adequate	Only for mine inflow: 0-28% increase. Not reported for changes in drawdown and	No	For further details see "summary of specific issues" in cover letter and Section 5 of the detailed peer review comments.		

Review Questions AGMG (2012) *MDBC (2001)**	**EIS Model (Coffey 2016a/b)**						**Pells and Pan (2013-2017) Model**	
	Kalf (2016)		***Merrick (2016)***		***Anderson (2017) – This Review***		***Anderson (2013,2017)***	
	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
prediction?				baseflow impacts.				
6. Uncertainty								
6.1 Is some qualitative or quantitative measure of uncertainty associated with the prediction reported together with the prediction? *(9.1) If required by the project brief, is uncertainty quantified in any way?*	Adequate		Maybe	Based on sensitivity analysis for three model parameter perturbations. Not a formal uncertainty analysis. Only reported for mine inflow: 0-28% increase. Not reported for changes in drawdown and baseflow impacts.	No	No uncertainty analysis is presented, only a base case scenario with a set of parameters that may be biased towards under-predicting impacts. The few sensitivity tests for just a couple of model parameters employed variations that do not appear to reflect the uncertainty in those parameters (i.e. the sensitivity tests were strictly sensitivity tests and not calibration constrained uncertainty analysis).	Yes	A qualitative / quantitative envelope of predictions is reported based on sensitivity tests with realistic changes to account for parameter uncertainty. This is not as good as a formal, calibration constrained uncertainty analysis.
6.2 Is the model with minimum prediction-error variance chosen for each prediction?					No		N/A	Not applicable for this model.
6.3 Are the sources of uncertainty discussed?					No		Yes	
6.3.1 measurement of uncertainty of observations and parameters					No		Some	Uncertainty in the anisotropy, porosity, specific yield and specific storage values are not discussed.
6.3.2 structural or model uncertainty					No		No	Structural uncertainty is not assessed. Boundary condition uncertainty is assessed.
6.4 Is the approach to estimation of uncertainty described and appropriate?					No		Yes	
6.5 Are there useful depictions of uncertainty?					No		Yes	
6.6 (9.2) Is the model 'fit-for-purpose'?	Yes				No	It is evident that a significant amount of time and budget has been expended on characterisation and modelling work, however, the amount of effort on numerical modelling appears (on face value) to be less than that expended by Shenhua on their Watermark Coal Mine which would have less impact on groundwater users. In addition, the impacts that are predicted appear quite large in some places and difficult to accept in others. There are also many questions about the suitability of various aspects and parameters of the model. Further work is required to justify model assumptions and develop improved model inputs that better represent the available data. More detailed sensitivity testing and uncertainty analysis is required to demonstrate the potential range of impacts from undertaking the project.	Yes	The modelling provides a good understanding of how the system responds to parameters, boundary conditions and numerical model settings which is suitable for understanding the potential range of impacts from the development. The model could be improved if: • additional data sets were made available by the proponent in electronic form • justifications for various model parameters selected by the proponent were provided and substantiated
7. Solute transport								
7.1 Has all available data on the solute distributions, sources and transport processes been collected and analysed?							N/A	See review by Jewell (2017)
7.2 Has the appropriate extent of the model domain been delineated and are the adopted solute concentration boundaries defensible?							N/A	See review by Jewell (2017)
7.3 Is the choice of numerical method and software appropriate?							N/A	See review by Jewell (2017)
7.4 Is the grid design and resolution adequate, and has the effect of the discretisation on the model outcomes been systematically evaluated?							N/A	See review by Jewell (2017)
7.5 Is there sufficient basis for the description and							N/A	See review by Jewell (2017)

Review Questions AGMG (2012) *MDBC (2001)**	**EIS Model (Coffey 2016a/b)**						**Pells and Pan (2013-2017) Model**	
	Kalf (2016)		***Merrick (2016)***		***Anderson (2017) – This Review***		***Anderson (2013,2017)***	
	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
parameterisation of the solute transport processes?								
7.6 Are the solver and its parameters appropriate for the problem under consideration?							N/A	See review by Jewell (2017)
7.7 Has the relative importance of advection, dispersion and diffusion been assessed?							N/A	See review by Jewell (2017)
7.8 Has an assessment been made of the need to consider variable density conditions?							N/A	See review by Jewell (2017)
7.9 Is the initial solute concentration distribution sufficiently well-known for transient problems and consistent with the initial conditions for head/pressure?							N/A	See review by Jewell (2017)
7.10 Is the initial solute concentration distribution stable and in equilibrium with the solute boundary conditions and stresses?							N/A	See review by Jewell (2017)
7.11 Is the calibration based on meaningful metrics?							N/A	See review by Jewell (2017)
7.12 Has the effect of spatial and temporal discretisation and solution method taken into account in the sensitivity analysis?							N/A	See review by Jewell (2017)
7.13 Has the effect of flow parameters on solute concentration predictions been evaluated, or have solute concentrations been used to constrain flow parameters?							N/A	See review by Jewell (2017)
7.14 Does the uncertainty analysis consider the effect of solute transport parameter uncertainty, grid design and solver selection/settings?							N/A	See review by Jewell (2017)
7.15 Does the report address the role of geologic heterogeneity on solute concentration distributions?							N/A	See review by Jewell (2017)
8. Surface water–groundwater interaction								
8.1 Is the conceptualisation of surface water–groundwater interaction in accordance with the model objectives?					?	Insufficient documentation. See review letter.	N/A	The conceptualisation of surface water–groundwater interaction is in accordance with and appropriate for the stated model purpose. The groundwater model is not coupled with a surface water model and is represented with constant head river model nodes along the river / creek boundaries.
8.2 Is the implementation of surface water–groundwater interaction appropriate?					?	Insufficient documentation. See review letter.	N/A	
8.3 Is the groundwater model coupled with a surface water model?					No		N/A	

Review Questions AGMG (2012) *MDBC (2001)**	**EIS Model (Coffey 2016a/b)**						**Pells and Pan (2013-2017) Model**	
	Kalf (2016)		***Merrick (2016)***		***Anderson (2017) – This Review***		***Anderson (2013,2017)***	
	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment	Yes/ No	Comment
8.3.1 Is the adopted approach appropriate?					?	Insufficient documentation. See review letter.	N/A	
8.3.2 Have appropriate time steps and stress periods been adopted?					?	Insufficient documentation. See review letter.	N/A	
8.3.3 Are the interface fluxes consistent between the groundwater and surface water models?					N/A		N/A	

*Where reviews have been completed in accordance with the MDCB (2001), answers have been transferred to the nearest equivalent AGMG (2012) review question. Where review questions did not match an additional item has been added in italics.

AGMG (2012): Merz, S. K. (2012). Australian groundwater modelling guidelines. *Waterlines Report Series*, (82).

MDCB (2001): Murray–Darling Basin Commission (MDBC) 2001, *Groundwater flow modelling guideline,* report prepared by Aquaterra, January 2001.

Hume Coal Project SSD 15_7172

UNSW-WRL Peer Review of Conceptual and Numerical Modelling that Predicted Likely Groundwater Impacts

ATTACHMENT B – PEER REVIEWER CV

Water Research Laboratory

Doug Anderson BE (Hons 1), MEngSc, MIEAust, MIAH
Principal Engineer

Never Stand Still | Faculty of Engineering | School of Civil and Environmental Engineering

Doug has 16 years of technical experience in groundwater - surface water resource and impact assessment. He designs and manages field investigations and groundwater monitoring programs in addition to undertaking environmental process and modelling studies. Doug delivers environmental assessment results and strategic environmental management advice. He helps his clients tackle challenging water issues to achieve their environmental and engineering objectives.

Doug maintains a strong background in hydrogeological site characterization, data management automation, numerical modelling, programming and geo-spatial data analysis. Doug is an expert groundwater modeller with several years of FEFLOW modelling experience. His expertise is complemented by background skills in civil engineering hydraulics and physical modelling. Doug employs a considered and practical approach to projects, working in a team environment to deliver quality project outcomes. His eye for detail in flow system conceptualisation provides decision makers with appropriate assessments of project risk and uncertainty.

Qualifications

BE Hons 1 (Environmental Engineering), UNSW, 2000
MEngSc (Groundwater Studies), UNSW, 2001

Professional History

2001-2009 : Project Engineer – Water Research Laboratory, UNSW
2010-2013 : Groundwater Modelling Specialist - AquaResource / Matrix Solutions Inc. (Canada)
2013- : Principal Engineer – Water Research Laboratory, UNSW

Specialist Fields of Expertise

- Hydrogeological site characterisation
- Groundwater flow and transport modelling
- Water resources management and protection
- Geo-spatial data analysis
- Information management and computer programming
- Coastal Imaging (Machine Vision)
- Civil engineering hydraulics

Summary of Relevant Experience

Doug has worked as a consultant for a range of industry and government clients to support environmental impact assessment, mineral resource development and site closure planning. He has accumulated groundwater resources expertise in fractured rock, coastal sand aquifers, moraines and salars in Australia, Argentina, Canada and the United States. Doug's experience includes all aspects of hydrogeological site characterization, e.g. conceptual model development, monitoring program design, drilling supervision, field data collection, data analysis, groundwater flow and transport modelling, environmental impact assessment and peer review.

Doug has worked at a number of waste disposal sites where contamination risks to groundwater and surface water must be investigated, modelled and managed with great care. This includes the radioactive waste disposal facilities at Ranger Mine in the Northern Territory and Little Forest Legacy Site at Lucas Heights. Doug's project experience also includes: the design and commissioning of effluent reuse monitoring programs; the assessment of groundwater contamination from urban and industrial landfilling; groundwater modelling for water protection studies; peer review of groundwater models; resource and reserve assessment for mineral brine projects; and the feasibility assessment of municipal extraction projects, wastewater disposal and effluent reuse.

Groundwater Resources Management and Protection

- Site investigations, data analysis, groundwater modelling and closure planning for a low-level radioactive waste facility at Little Forest Legacy Site, Lucas Heights (2016-)
- Groundwater monitoring at John Fisher Park legacy landfill site for Northern Beaches Council (2016)
- Aquifer test analysis and peer review of contamination monitoring for Burra Rd, Gundagai Landfill (2016)
- Measurement of landfill clay cap permeability by gas permeameter and UNSW geotechnical centrifuge (2016)
- Measurement of drill core permeability by geotechnical centrifuge for WA Department of Water Perth Confined Aquifer Capacity Study (2015)
- Confidential groundwater desktop and numerical model study for Newcastle City Council (2015-2016)
- Groundwater Impact Assessment adequacy review for Lynwood Quarry (2015)
- Groundwater and surface water flow modelling to support Ranger Mine Pit #1 and Pit #3 closure plans (2009, 2014)
- Groundwater modelling for the Municipality of Waterloo in Ontario to establish well head protection areas to help plan secure drinking water supplies for 500,000 residents (2010)
- Updated Lake Conjola Groundwater Monitoring Program and Response Plan for the NSW Public Works (2008)
- Groundwater investigations and modelling to assess the feasibility of a horizontal collector well system and

Water Research Laboratory
School of Civil and Environmental Engineering
UNSW AUSTRALIA | 110 KING ST MANLY VALE NSW 2093 AUSTRALIA | d.anderson@wrl.unsw.edu.au
T +61 (2) 8071 9800 | F +61 (2) 9949 4188 | ABN 57 195 873 179 | www.wrl.unsw.edu.au | Quality System Certified to AS/NZS ISO 9001

desalinization plant proposed as an emergency solution to drought-proof the for Wyong Shire Council supply (2004-05)

Managed Aquifer Recharge (MAR)
- Entry Level Assessment of MAR for Sydney Water for the townships of Galston and Glenorie (2013)
- Site selection, borehole drilling and groundwater monitoring program to assess the operational performance of the Moree Plains Shire Council effluent reuse scheme (2006)
- Groundwater investigations, numerical modelling and concept designs for EGIS Consulting and Department of Commerce to support the feasibility assessments for the effluent reuse scheme at Iluka (2001-2005)
- Groundwater modelling of virus transport in coastal sand aquifers for DLWC for the proposed effluent reuse scheme at Hat Head (2002)

Mining and Coal Seam Gas – Studies
- Measurement of drill core permeability by centrifuge for OGIA's Walloon Interconnectivity Research Project (2014)
- Background paper on groundwater resources and CSG for the NSW Office of the Chief Scientist and Engineer (2013)
- Measurement of drill core permeability by centrifuge for a coal mining / CSG client (2013)
- Technical advice and support to the United States Forestry Service, ERO and mining company groundwater groundwater modellers to developed an EIA model of the Rock Creek underground Copper Mine Prospect (2012)
- Hydrogeological characterization and groundwater modelling of brine deposits to support a NI-43101 reserve estimation for TSX / Lithium Americas Corp (2010-2012)
- Groundwater modelling for Tier II source water applications for proposed SAGD oil shale projects in Alberta (2012)

Mining and Coal Seam Gas – Peer Review
- OWS science and literature review to identify coal seam gas and coal mining knowledge gaps (2014)
- Peer review of Shenhua's groundwater model for the proposed Watermark Coal Prospect (2014)
- Peer review of the SHCAG's groundwater model of the Hume Coal Prospect (2013)

Mining and Coal Seam Gas – Peer Review
- Technical review of OWS's critical science review on coal seam gas and aquifer connectivity (2013)
- Technical review of OWS's critical science review on coal seam gas and groundwater modelling (2013)

Linear Infrastructure
- NSW Department of Planning Groundwater Peer Reviewer for Westconnex and Northern Beaches Hospital road upgrade (2015-)

Civil Engineering Hydraulics – Wastewater
- Monitoring of Warriewood STP Secondary Clarifiers (2005-06)
- CFD modelling of FL2000 wastewater separator (2002)
- WA Setting Tank Desktop Assessments (2003)
- Modelling for Christchurch Ocean Outfall (2003)

Education and Training
- Federal Office of Water Science Surface Water Training Course on Large Coal Mines and Coal Seam Gas (2015)
- FEFLOW Training Course (LAC, 2012; UNSW, 2002)
- Sydney Coastal Councils' Groundwater Workshops (2007)
- Australian Cotton CRC's Groundwater Workshops (2007)

Civil Engineering Hydraulics – Flooding and Coastal
- Eidsvold Weir Physical Model, QLD
- East Arm Port Ship Interaction Physical Model, NT (2002-03)
- Penrith Lakes desktop, numerical and physical modelling (Penrith Lakes Development Corporation, 2003-2009)

Information Management and Coastal Imaging
- Real-time web based coastal monitoring (Gold Coast City Council, Tweed River Entrance Sand Bypassing Project, Warringah Shire Council, 2001-2008,2013-)
- Hong Kong North Western Waters Database, Warringah Council Water Quality Database, Australian Councils' St Sweeping Database, ACO Polycrete Scheduling Software (2001 - 2003)

Publications

Timms, W. A., Crane, R., **Anderson, D. J.**, Bouzalakos, S., Whelan, M., McGeeney, D., Acworth, R. I. (2016). Accelerated gravity testing of aquitard core permeability and implications at formation and regional scale. *Hydrology and Earth System Sciences*, *20*(1), 39-54. doi:10.5194/hess-20-39-2016

Timms, W. A., Crane, R., **Anderson, D. J.**, Bouzalakos, S., Whelan, M., McGeeney, D., Acworth, R. I. (2014). Vertical hydraulic conductivity of a clayey-silt aquitard: accelerated fluid flow in a centrifuge permeameter compared with in situ conditions. *Hydrology and Earth System Sciences Discussion*, *11*(3), 3155-3212. doi:10.5194/hessd-11-3155-2014

Anderson, D.J., Timms, W.A. and Glamore W.C. (2009) "Optimising Subsurface Well Design for Coastal Desalination Water Harvesting", *Australian Journal of Earth Sciences*, 56, 53-60

Turner, I.L. and **Anderson, D.J.** (2007) "Web-Based and 'Real-Time' Beach Management System", *Coastal Engineering*, 54, 555-565

Glamore, W.C., Timms, W.A. and **Anderson, D.J.** (2007) "Injection or Release: Innovative Technologies for Disposing Recycled Water in Coastal Environments", *Proc. 16th NSW Coastal Conference*, Yamba, NSW, 7-9 November

Turner, I.L. and **Anderson, D.J.** (2006) "CZM Applications of Argus Coastal Imaging in Eastern Australia", *Proc. 15th NSW Coastal Conference*, Coffs Harbour, NSW, 7-9 November

Glamore, W.C., **Anderson, D.J.** and Timms, W.A. (2006) "Coastal Groundwater Intakes: Numerical Modelling of Coastal Wells for Desalination Source Water", *Proc. 30th International Conference on Coastal Engineering*, San Diego, USA, 3-8 September

Anderson, D.J., Timms, W.A. and Glamore W.C. (2005) "Optimising Subsurface Well Design for Coastal Desalination Water Harvesting", *Proc. NZHS-IAH-NZSSS Conference*, 28 November-3 December, Auckland (CD rom)

Anderson, D.J., Frazer, A., Jancar, T. and Miller, B.M. (2004) "The Implementation of PIV-PTV Techniques for Measurement of Velocities in Large Scale Physical Models", *Proc. 8th National Conference on Hydraulics in Water Engineering*, Gold Coast, 13-16 July

Anderson, D.J., Turner, I.L., Dyson, A., Lawson, S. and Victory S (2003) "Tweed River Entrance Sand Bypassing Project: 'Real-Time' Beach Monitoring and Analysis System via the World-Wide-Web", *16th Australasian Coasts and Ports Conference*, Auckland, 9-12 September

Water Research Laboratory
School of Civil and Environmental Engineering
UNSW AUSTRALIA | 110 KING ST MANLY VALE NSW 2093 AUSTRALIA | d.anderson@wrl.unsw.edu.au
T +61 (2) 8071 9800 | F +61 (2) 9949 4188 | ABN 57 195 873 179 | www.wrl.unsw.edu.au | Quality System Certified to AS/NZS ISO 9001