N

Groundwater Peer Review

BYLONG COAL PROJECT
Environmental Impact Statement

Hansen Bailey
ENVIRONMENTAL CONSULTANTS

Heritage Computing Pty Ltd ● ABN 75 392 967 126 ● T/A **HydroSimulations**
PO Box 241, Gerringong NSW 2534. Phone: (+61 2) 4234 3802

noel.merrick@heritagecomputing.com

DATE: 17 July 2015

TO: Nathan Cooper
Senior Environmental Scientist
Hansen Bailey
John Street
Singleton NSW 2330
Tel: (02) 6575 2007

FROM: Dr Noel Merrick

RE: **Peer Review – Bylong Coal Project Groundwater Impact Assessment**

OUR REF: HS2015/20

---

1. Introduction

This report provides a peer review of the groundwater assessment undertaken for open cut and underground mining for the Bylong Coal Project (the Project). The assessment has been done by Australasian Groundwater and Environmental (AGE) Consultants Pty Ltd for KEPCO Bylong Australia Pty Ltd (KEPCO). The Project is located about 55 km north-east of Mudgee western New South Wales (NSW).

The groundwater assessment is based on field investigations (by Douglas Partners Pty Ltd) and a regional numerical groundwater model. The groundwater modelling forms an important component of the environmental assessment for the project. The main purpose of the modelling is to assess potential impacts on groundwater levels on the Project Site and in the surrounding area, and also to quantify the incidental capture of streamflow and alluvial groundwater associated with the Bylong River and Goulburn River as required by the Aquifer Interference Policy (AIP). The model also provides an assessment of likely groundwater inflow to the open cut pits and to the underground voids as mining progresses in time.

The scope of work was limited to a peer review of AGE's groundwater report and completed model. The author previously reviewed the groundwater assessment submitted to the Gateway Panel. Since 2012, the reviewer has participated in three meetings and several teleconferences with the AGE modelling team on this Project. Electronic model files have not been examined by this reviewer.

The reviewer also conducted a high-level review of work in progress in January 2015, with particular attention paid to:

- the methodology being adopted for the modelling;
- unsaturated zone assumptions; and
- sensitivity analysis scenarios.

2. Documentation

The following report comprises the current documentation for the groundwater assessment:

1. AGE, 2015, Bylong Coal Project Groundwater Impact Assessment. Report prepared for Hansen Bailey Pty Ltd, 10 June 2015. 162p + 6 Appendices.

As there was a stated need in the reviewer's scope of work to consider opinions expressed by government peer reviewers on the modelling conducted for the Watermark Coal Project (Watermark), the following documents also were examined:

2. Mackie Environmental Research Pty Ltd, 2014, Independent Expert Advice to the (Planning Assessment) Commission by Dr Colin Mackie. Letter report to PAC, 14 August 2014, 13p.

3. AGE, 2014, Re: Proposal for Watermark Project - Support to Planning Assessment Commission Process. Letter report to Shenhua Australia, 22 July 2014, 15p.

4. Kalf and Associates Pty Ltd, 2014, Watermark Coal Project: KA Comments on the MER Model Audit and AGE Revised Modelling. Letter report to NSW Department of Planning and Environment, 23 October 2014, 5p.

Document #1 has 15 sections:

1. Introduction
2. Project description
3. Objectives and scope of work
4. Legislation, policy and guidelines
5. Regional setting
6. Field investigation program
7. Hydrogeological regime
8. Numerical model design
9. Model calibration and verification
10. Model predictions and impact assessment
11. Model uncertainty
12. Compliance with Government Policy
13. Management / mitigation measures
14. References
15. Glossary.

The Appendices are:

A. Study Requirements
B. Field Investigation Reports
C. Bore Surveying Results
D. Calibration and Hydrographs
E. Van Genuchten Parameters Literature Review
F. Predictive Sensitivity and Uncertainty Analysis

3. Review Methodology

There are two accepted guides to the review of groundwater models: (A) the Murray-Darling Basin Commission (MDBC) Groundwater Flow Modelling Guideline[1], issued in 2001,and (B) newer guidelines issued by the National Water Commission in June 2012 (Barnett *et al*., 2012[2]). Both guides also offer techniques for reviewing the non-modelling components of a groundwater impact assessment. The 2012 national guidelines build on the 2001 MDBC guide, with substantial consistency in the model conceptualisation, design, construction and calibration principles, and the performance and review criteria, although there are differences in details. The new guide is almost silent on coal mine modelling and offers no direction on best practice methodology for

1 MDBC (2001). Groundwater flow modelling guideline. Murray-Darling Basin Commission. URL: www.mdbc.gov.au/nrm/water_management/groundwater/groundwater_guides

2 Barnett, B, Townley, L.R., 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 82, National Water Commission, Canberra.

such applications. There is, however, an expectation of more effort in uncertainty analysis, although the guide is not prescriptive as to which methodology should be adopted.

The Bylong model type is Moderate Complexity (under the MDBC guidelines). The NWC 2012 guide has replaced the model complexity classification by a "model confidence level". The AGE report gives a thorough defence of the model's Class 2 classification (the middle category) in terms of data, calibration, prediction and key indicator checkpoints. A Class 2 model would be suitable for "prediction of impacts of proposed developments in medium value aquifers" and for "providing estimates of dewatering requirements for mines and excavations and the associated impacts". This is the appropriate level for a groundwater impact assessment for a mining development.

The groundwater guides include useful checklists for peer review. For this review, the 2-page Model Appraisal checklist[3] in MDBC (2001) has been used for groundwater model review. This checklist has questions on (1) The Report; (2) Data Analysis; (3) Conceptualisation; (4) Model Design; (5) Calibration; (6) Verification; (7) Prediction; (8) Sensitivity Analysis; and (9) Uncertainty Analysis. Non-modelling components of the impact assessments are addressed by the first three sections of the checklist.

The detailed assessment of the groundwater modelling is recorded in the peer review checklist in **Table 1**. Supplementary comment is offered in the following sections of this review.

4. Commentary

4.1 Report Matters

The groundwater assessment report is particularly lucid, being very readable with no verbose text. The illustrations are of high quality.

Previous review comments of a technical nature have been addressed satisfactorily, except for a few instances that are explored further in following sections. Earlier recommendations for editorial corrections have been addressed in full.

One earlier recommendation was for inclusion of a residual mass curve (showing rainfall trend) on several figures displaying groundwater hydrographs, stream flow and/or rainfall. Apparently this could not be done, as the graphs were produced with Microsoft Excel which allows only a limited number of plot axes and data types.

Overall, there are no matters of concern in the report as to structure or depth of coverage, other than a fairly brief treatment of data management in Section 13.8.

4.2 Data Matters

An extensive monitoring network has been installed, with 62 monitoring bores and 13 VWP sites. The baseline record is generally 2-3 years in duration, while there is a large number of government monitoring bores dating back 10-15 years.

While the authors of Document #1 are conscious of providing a cause-and-effect analysis, the description is hampered by the resolution scale adopted for hydrograph plots. The scale disguises correlations of head change with rainfall trend (e.g. Figure 7.7). Vertical hydraulic gradients have been explored in separate formations (Figures 7-12 and 7-13).

Groundwater quality is examined through full ionic analyses at 24 sites, and comparison with surface water analyses at nine sites. Relationships are explored through EC histograms and Piper plots.

To inform river/aquifer interactions, seepage meters were installed to indicate gaining/losing conditions. This is rarely done.

[3] The new guidelines include a more detailed checklist but they do not offer the graded assessments of the 2001 checklist, which this reviewer regards as more informative for readers.

Overall, data acquisition and data analysis have been thorough and they provide a firm foundation for the conceptual model depicted in Figure 7.22. The reviewer endorses this conceptualisation.

4.3 Model Matters

Modelling objectives are stated clearly (Sections 8.1 and 10.2.1) and model design has been tailored to meet these objectives. A Class 2 model is appropriate, as defended in Section 9.4.

Model extent is sufficiently broad to counteract possible boundary effects. Model layering gives an adequate representation of stratigraphy, with separate inclusion of the Ulan and Coggan coal seams.

The model holds 1.16 million cells, at the limit of what is tractable in practice, with a minimum cell resolution of 50 m.

4.3.1 *Modelling Methodology*

MODFLOW-SURFACT software is used, as is usual, and each model run comprises a number of time-slices. Most other modellers who use MODFLOW-SURFACT now use TMP (time-varying materials package) for progressive updating of physical properties in open cut infill and underground fractured zones. Although the time-slice approach can generate more spikiness in calculated fluxes than TMP, it is a legitimate approach. The use of custom software to batch the sub-models offers a versatility that is missing from the more standard approach.

There are no stand-out issues with the modelling methodology. There are a few features where numerical stability dictated adoption of some model parameters which might appear unrealistic - namely, low evapotranspiration rate, and unsaturated zone parameterisation outside physical range expectations. The model would not converge readily when more "realistic" parameter values were trialled.

The vertical hydraulic conductivity (Kz) through the fractured zone (Figure 10.3) is a reasonable interpolation between undisturbed host values and geotechnical model estimates for fracture permeability. The enhanced Kz is represented by a ramp function covering about 1.5 orders of magnitude up to a constant height of 260 m above the mined coal seam. However, the absolute values are higher than the reviewer has experienced at other active mines, but they are consistent with database values presented by Tammetta (2015[4]; Figure 2) for geometric mean K (equal to $[Kx\ Kz]^{1/2}$). For a greenfield project there is no control on what values are appropriate.

The ramp value Kz multipliers (applied to the undisturbed values), from layer 8 to layer 1 respectively, are 500, 500, 100, 50000, 25, 1 and 5. Tammetta (2015; Figure 3) found that the geometric mean enhancement in a fractured zone is only a factor (R) of 100 or so (±1 order of magnitude).

From the values in Table 10.2, it is possible to calculate the relative contributions of open fractures and rock blocks to mine inflow. For example in layer 7 (interburden), weighted averaging requires a fracture ratio of 0.009%. For 210 m longwall width, the total aperture widths of all fractures reaching the goaf would be 18 mm for a fracture Kz of 864 m/d.

Spoil activation times were raised in Document #2 as a matter of concern at Watermark. It appears that an unchanged procedure has been followed here. The objection relates to the onset of enhanced recharge. Whether or not a delay is imposed depends on the algorithm for migration of recharge water through the unsaturated zone. In theory, no delay is required if SURFACT's variable saturation (van Genuchten) option is used. If a pseudo-soil is used, a manual delay is warranted.

Open cut drains are deactivated at the end of each year, after which spoil is emplaced. Longwall drain deactivation is another matter that affects mine inflow estimates. In this case, longwall panels are actively drained in the model for the entire mine life, as pumping from the network of panels

[4] Tammetta, P., 2015, Estimation of the Change in Hydraulic Conductivity above Mined Longwall Panels. *Groundwater*, vol.53, no.1, Jan-Feb 2015, 122-129.

would be ongoing.

In Section 13.4.1 (water quality triggers), Document #1 states that the "mean and standard deviations could be recalculated" after a lengthy period of "no alarms". As that period could still have mining effects, it is more usual to define a baseline period that cannot conceivably have mining effects and to leave that unchanged (unless the period is too short to be representative). A similar situation could arise with water level triggers (Section 13.4.2) if the trigger levels are reset to the previous 24 month period.

4.3.2 *Model Calibration*

Model calibration is satisfactory for steady-state and transient conditions. The calibration performance (for groundwater levels) is less than 4 %RMS in relative terms and about 10 mRMS in absolute terms. Vertical head gradients are reproduced reasonably well (Figure 9-7). Hydrograph matches in Appendix D are reasonable on the whole.

Baseflow estimates are reasonable, when compared with field estimates, and stream leakage estimates have order-of-magnitude agreement with seepage meter measurements.

Document #1 would have benefitted from inclusion of a spatial residuals map to show whether there are specific areas that are particularly well calibrated or poorly calibrated.

Calibrated formation properties are consistent with field measurements.

The adopted rainfall recharge rates are controlled through a soil moisture model. This is not often attempted.

The dataset from February 2014 to July 2014 was reserved for a verification assessment. This is not often done, and is not a compulsory step in the modelling process (Barnett *et al*., 2012). The verification performance (for groundwater levels) is about 6 %RMS in relative terms and about 7 mRMS in absolute terms.

4.3.3 *Sensitivity Analysis*

In Document #1 (Appendix F), an extensive traditional sensitivity analysis has been conducted for the *prediction* model for variable saturation algorithms and the following model properties:

- host Kx and Kz;
- host Sy and Ss;
- host recharge;
- river stage heights;
- river bed conductance;
- fractured zone Kz;
- van Genuchten (VG) parameters;
- drain activation time;
- spoil and tailings Kx and Kz;
- spoil and tailings Sy and Ss; and
- spoil and tailings recharge.

The multipliers are selected responsibly to investigate practical limits.

As the model could be decalibrated with some extreme perturbations, it is wise to apply each realisation to the *calibration* model so that some sensitivity runs can be excised (e.g. those runs with perturbed RMS > 1.5 x base RMS).

A feature of the sensitivity analysis is the exploration of the effects produced by different desaturation algorithms. In particular, the MODFLOW-SURFACT structured grid base model was converted to a MODFLOW-USG unstructured grid model with a Voronoi (polygonal) mesh with pinched out dummy layers, using an upstream weighting function for the handling of dry cells. The investigation in Document #1 is believed to be the most extensive to date for examination of algorithmic effects.

Of importance is the observation that the pseudo-soil model failed to converge. This matches the reviewer's experience with the pseudo-soil option on many other coal mine models. However, the USG model performed very well, using a similar algorithm to a pseudo-soil. In addition, a calibration-constrained *Monte Carlo* uncertainty analysis was conducted using the USG model, with results retained for about 80% of the 200 model realisations, the remainder being essentially decalibrated.

Document #1 argues the case for a pseudo-soil approach being *less* conservative than a VG approach (depending on the adopted VG parameters), as drawdown effects are prevented from propagating through dry cells, whereas they can be propagated through unsaturated cells.

Figures F-28 to F-32 in Document #1 illustrate the results of the *Monte Carlo* analysis using USG with the upstream weighting algorithm. The mean USG inflow proved to be very similar to VG Scenarios 2 and 3, which have higher alpha and residual saturation than the base model, and lower beta. The VG parameters for Scenarios 2 and 3 are more in keeping with expected values. The base model appears *not* to be conservative for mine inflow.

The alluvial water take is predicted to be much less with the USG model than with any VG model, which can be regarded as conservative for this impact. However, substantial sensitivity to VG parameters is noted during underground mining, at which time the base model is not conservative. During open cut mining, the base model is the most conservative of the trialled models.

Similarly, the Bylong River water take is predicted to be much less with the USG model than with any VG model, which can be regarded as conservative for this impact. Of the various VG models, the base model is the least conservative.

The various VG models differ only marginally in the predicted extent of maximum groundwater drawdown.

The following conclusions can be drawn:

- the choice of VG parameters affects model predictions for water takes but not significantly for drawdown extent;
- a pseudo-soil (or surrogate) approach is likely to give higher mine inflow than some VG model parameterisations;
- a pseudo-soil (or surrogate) approach is likely to predict less environmental impact;
- specifically, a pseudo-soil (or surrogate) approach is likely to give less impact on alluvial water take;
- specifically, a pseudo-soil (or surrogate) approach is likely to give less impact on river water take; and
- it is difficult to affirm consistent conservatism in any adopted approach.

4.3.4 *Unsaturated Zone*

Documents #2 and #4 advocate the use of the pseudo-soil algorithm in preference to full Richards Equation solution based on VG parameters, unless the VG parameters are physically based, spatially variable and layer-dependent. Document #1, however, notes that the pseudo-soil algorithm as implemented in MODFLOW-SURFACT is unstable in this case and such a model does not converge. The VG approach eases model stability problems. However, there will never be sufficient field evidence for the four VG coefficients in a regional groundwater model.

At the reviewer's instigation, a literature review of van Genuchten parameters has been included at Appendix E. Unfortunately, there is an absence of information on these parameters for hard rocks. This reviewer favours the VG approach for the same reasons as AGE - stability and runtime, as he has been invariably disappointed with the stability performance of the pseudo-soil algorithm for large, complex models.

There has never been any expectation that the unsaturated zone was being modelled accurately with the VG approach, given its applicability usually to a single model layer, but there is a claim in Document #2 of significant errors in water table elevation (at Watermark). The reviewer's view is that this was caused by adoption of an extremely low value for alpha (0.01) in the Watermark study. In the present study, the base model has a value of 0.02 m^{-1}, which this reviewer regards also as too low. However, values of 0.3 to 10 m^{-1} for VG Scenarios 1-3 are acceptable. Of the VG Scenarios, Scenarios 2 and 3 have the most reasonable residual saturation. The sensitivity analysis shows that Scenarios 2 and 3 perform very similarly in terms of predicted mine inflow, alluvial water take and river water take. In addition, the mine inflow for these two scenarios is similar to what is predicted by the USG model using upstream weighting (a surrogate for a pseudo-soil). The USG model using upstream weighting, however, is likely to underestimate environmental impacts, as demonstrated by the sensitivity analysis.

We are left with a dilemma. The pseudo-soil approach (if it converges), or equivalently the upstream weighting approach in USG, appears to be more reliable for estimation of mine inflow, but the VG approach appears to be more reliable for prediction of environmental impacts.

This reviewer in the past has recommended against use of a pseudo-soil for a number of reasons: (1) the lack of explanation in the MODFLOW-SURFACT manual as to its mathematical foundation; (2) slower runtime; (3) poor stability, with common spikes in cell heads hugging cell bottom elevations apparently randomly. The authors of Documents #2 and #4 do not share these experiences. Issue (1) is now largely overcome by answers to FAQs on the HydroGeoLogic website and better explanations in some published papers. For example, Schoups et al. (2005)[5] explain pseudo-soil functions as follows:

> *In this approach the nonlinear water retention and conductivity functions at a point are replaced by discrete functions, with degree of saturation and relative hydraulic conductivity equal to zero when the soil water pressure is negative and equal to one when the pressure head is positive. These point values are integrated across the thickness of the grid cell that contains the water table to yield linear soil hydraulic functions. These linear grid-scale representative functions define saturation, S_w, and relative horizontal conductivity, k_{rw}, values that increase linearly from 0, when the water table is at or below the bottom of the grid cell, to 1 when the water table is at or above the top of the cell, or*

$$S_w = k_{rw} = 1, \ \text{for}\ \psi/\Delta z \geq 0.5;$$

$$S_w = k_{rw} = 0.5 + \frac{\psi}{\Delta z}, \ \text{for}\ -0.5 < \psi/\Delta z < 0.5;$$

$$S_w = k_{rw} = 0, \ \text{for}\ \psi/\Delta z \leq -0.5$$

> *where Δz is thickness [L] of the grid cell with the water table and j is the pseudo-pressure head [L] at the node (Huyakorn et al., 1994). In the vertical direction, k_{rw} is always equal to 1.*
>
> *The use of pseudo-soil functions constitutes a computationally attractive compromise between the rigorous variably-saturated flow modeling using the van Genuchten relationships, and the simplified MODFLOW approach for which cells become inactive when the water table drops below the bottom of the cell (McDonald and Harbaugh, 1988). In the approach used here, when the water table drops below the bottom of a grid cell, Eq. (1a) is still solved but with the right-hand side equal to zero, i.e. changes in storage above the water table are neglected. This procedure avoids convergence problems with (in)activation of cells encountered in MODFLOW (Doherty, 2001).*

This explanation gives more confidence in its use in practice. In addition, the authors of MODFLOW-SURFACT give consistent advice on the applicability of pseudo-soil functions in regional models and the inapplicability of the VG algorithm for other than small-scale models with high vertical resolution across the unsaturated zone:

- *"Theoretically, the VG method is scale-independent. However, high vertical resolution is required to describe the vertical variation of moisture. Time steps may have to be very small to track the movement of the front accurately. Because of the high degree of non-linearity, a large number of iterations may be necessary. Therefore it is not practical for regional applications."* (Email from Jeff Fairbanks of HydroGeoLogic to Frans Kalf, 2014)

---

[5] G. Schoups, J.W. Hopmans, C.A. Young, **J.A. Vrugt**, and W.W. Wallender (2005), Multi-objective optimization of a regional spatially-distributed subsurface waterflow model, *Journal of Hydrology*, 20-48, 311(1-4), doi:10.1016/j.jhydrol.2005.01.001

- *"Using Richards equation solution with some default parameters on grids that are as thick as the aquifer layer do not provide a correct solution. I have seen it being done but you have to be careful and explore the correctness of your solution. For regional flow the unconfined flow solutions of MODFLOW (one of the rewetting or the upstream weighted formulation) or Surfact can be used*." (Email from Sorab Panday formerly of HydroGeoLogic to Neil Manewell, 2014)

If the VG approach is to be retained, the adopted soil parameters should be treated as no more than *calibration parameters* "which do not necessarily have a physical meaning" (Kabat *et al*., 1997). There are many papers on upscaling of soil properties to regional scale through dimensional analysis, or by inverse modelling to recover aggregated soil parameters: e.g. Kabat *et al*. (1997)[6]; Vrugt et al. (2004)[7], Hopmans *et al*. (2002)[8]. Inverse modelling is unlikely to be applicable to regional groundwater models as widespread soil moisture measurements are rare. As soil maps, however, are available, dimensional analysis is a feasible approach for near-surface application of the pseudo-soil function. Kabat *et al*. (1997) show that "the heterogeneity within a sub-grid consisting of several textural soil types can be therefore described by a single 'effective' scaling parameter $<D/K_s^2>$", where water diffusivity $D = - K\, dh/d\theta$ (where θ is the volumetric soil moisture content). This potential approach warrants more research by groundwater modellers.

Applicability of the VG approach to desaturation at depth, associated with a coal seam, is another matter. There seems to be no guidance in the literature on what VG properties are appropriate, and hence no opportunity for rigorous upscaling. However, this reviewer expects the drain boundary condition applied to a mined cell, coupled with the adopted vertical hydraulic conductivity in the fractured zone, would overwhelm any sensitivity to VG values.

5. Conclusion

This reviewer finds that the model underpinning the groundwater assessment is "fit for purpose", where the primary purpose of the model is the prediction of environmental impacts in the context of the Aquifer Interference Policy, and estimation of water takes for licensing. A very thorough analysis of the uncertainty in the estimates has been conducted..

Yours sincerely,

Dr Noel Merrick

[6] Kabat, P., Hutjes, R.W.A. and Feddes, R.A. (1997) The scaling characteristics of soil parameters: From plot scale heterogeneity to subgrid parameterization. Journal of Hydrology 190, 363-396.
[7] Vrugt, J. A., G. Schoups, J. W. Hopmans, C. Young, W. W. Wallender, T. Harter, and W. Bouten (2004), Inverse modeling of large-scale spatially distributed vadose zone properties using global optimization, *Water Resour. Res.*, **40**, W06503, doi:10.1029/2003WR002706.
[8] Hopmans, J.W., Nielsen, D.R, and Bristow, K.L. (2002) How useful are small-scale soil hydraulic property measurements for large-scale vadose zone modeling? Washington DC American Geophysical Union Geophysical Monograph Series 01/2002; DOI: 10.1029/129GM20.

Table 1. MODEL APPRAISAL: **BYLONG COAL PROJECT**

Q.	QUESTION	Not Applicable or Unknown	Score 0	Score 1	Score 3	Score 5	Score	Max. Score (0, 3, 5)	COMMENT
1.0	THE REPORT								Good Executive Summary and very good AI Policy summary.
1.1	Is there a clear statement of project objectives in the modelling report?		Missing	Deficient	Adequate	Very Good			Project objectives S3 (SEARs). Modelling objectives S8.1.
1.2	Is the level of model complexity clear or acknowledged?		Missing	No	Yes				Impact Assessment Model, medium complexity. Class 2 confidence.
1.3	Is a water or mass balance reported?		Missing	Deficient	Adequate	Very Good			Steady-state calibration, transient calibration and transient prediction.
1.4	Has the modelling study satisfied project objectives?		Missing	Deficient	Adequate	Very Good			
1.5	Are the model results of any practical use?			No	Maybe	Yes			
2.0	DATA ANALYSIS								
2.1	Has hydrogeology data been collected and analysed?		Missing	Deficient	Adequate	Very Good			Substantial field investigation (App B). Expansive monitoring network.
2.2	Are groundwater contours or flow directions presented?		Missing	Deficient	Adequate	Very Good			Interpolated contours are shown in Fig 7.8. Source data points are posted.
2.3	Have all potential recharge data been collected and analysed? (rainfall, streamflow, irrigation, floods, etc.)		Missing	Deficient	Adequate	Very Good			Bylong River flow duration Fig 7.16. LIDAR DEM. Stream stage from nearest gauge but not interpolated to model cell. Irrigation recharge included. No consideration of flooding as potential recharge source (conservative approach).
2.4	Have all potential discharge data been collected and analysed? (abstraction, evapotranspiration, drainage, springflow, etc.)		Missing	Deficient	Adequate	Very Good			Gw abstraction included (data from FOI request). BoM "actual ET" is quoted. Main streams are included - most are gaining.

2.5	Have the recharge and discharge datasets been analysed for their groundwater response?		Missing	Deficient	Adequate	Very Good			Hydrographs compared with residual mass and streamflow. Comments on vertical head gradients and flow directions. Installed seepage meters to assess gaining/losing streams. Soil moisture transient balance.
2.6	Are groundwater hydrographs used for calibration?			No	Maybe	Yes			
2.7	Have consistent data units and standard geometrical datums been used?			No	Yes				
3.0	CONCEPTUALISATION								
3.1	Is the conceptual model consistent with project objectives and the required model complexity?		Unknown	No	Maybe	Yes			
3.2	Is there a clear description of the conceptual model?		Missing	Deficient	Adequate	Very Good			
3.3	Is there a graphical representation of the modeller's conceptualisation?		Missing	Deficient	Adequate	Very Good			Figure 7.22.
3.4	Is the conceptual model unnecessarily simple or unnecessarily complex?			Yes	No				
4.0	MODEL DESIGN								Class 2. Elements of Class 3.
4.1	Is the spatial extent of the model appropriate?			No	Maybe	Yes			33km x 39km. Extensive all directions. Western extent adjacent to Wilpinjong outer impact zone. 50-500m cell size. 1.16 million cells.
4.2	Are the applied boundary conditions plausible and unrestrictive?		Missing	Deficient	Adequate	Very Good			Appear to be no-flow cells on boundaries, other than GHB on western edge.
4.3	Is the software appropriate for the objectives of the study?			No	Maybe	Yes			Modflow SURFACT, Gw Vistas, PMWIN & custom code. Variable saturation by van Genuchten –pseudo-soil unstable. MODFLOW-USG for uncertainty analysis. Literature review App E.

Q.	QUESTION	Not Applicable or Unknown	Score 0	Score 1	Score 3	Score 5	Score	Max. Score (0, 3, 5)	COMMENT
5.0	CALIBRATION								412 stress periods (2 days length). Nov2011-Jan2014. Very fine stepping.
5.1	Is there sufficient evidence provided for model calibration?		Missing	Deficient	Adequate	Very Good			Scattergram for 94 target water levels. Standard statistics. Table of observed, simulated and residual water levels. Head contours. Vertical head profiles. App D (not C) hydrograph matches - generally good trends. Overly responsive to rain events.
5.2	Is the model sufficiently calibrated against spatial observations?		Missing	Deficient	Adequate	Very Good			3.1%RMS & 9.5mRMS steady state. WL maps Figs 9.4, 9.5. No spatial residual map to see where calibration is good or bad.
5.3	Is the model sufficiently calibrated against temporal observations?	N/A	Missing	Deficient	Adequate	Very Good			94 hydrographs and baseflow. 3.6%RMS & 10.2mRMS. Baseflow Fig 9.10 compared with AWBM model.
5.4	Are calibrated parameter distributions and ranges plausible?		Missing	No	Maybe	Yes			Uniform per layer. Figure 9.1 shows consistency with packer test values for coal Kx. Whisker plot in Figure 9.2 for all lithologies. %Rain recharge from soil moisture model (6% alluvium). 27% irrigation return of pumpage.
5.5	Does the calibration statistic satisfy agreed performance criteria?		Missing	Deficient	Adequate	Very Good			
5.6	Are there good reasons for not meeting agreed performance criteria?	N/A	Missing	Deficient	Adequate	Very Good			Lower baseflow peaks than AWBM. Order of magnitude agreement with seepage meters.
6.0	VERIFICATION								Feb2014-July2014. This is not a compulsory step (Barnett et al., 2012).
6.1	Is there sufficient evidence provided for model verification?	N/A	Missing	Deficient	Adequate	Very Good			Scattergram and statistics.
6.2	Does the reserved dataset include stresses consistent with the prediction scenarios?	N/A	Unknown	No	Maybe	Yes			No mining stress yet.
6.3	Are there good reasons for an unsatisfactory verification?	N/A	Missing	Deficient	Adequate	Very Good			One outlier due to perching. 6.2%RMS, 6.7mRMS.

7.0	PREDICTION								100 stress periods (91.3 days length).
7.1	Have multiple scenarios been run for climate variability?		Missing	Deficient	Adequate	Very Good			Average quarterly rainfall and pumpage. Also above average and below average. Raised ET surface for varying landform. Enforced makeup water requirements from groundwater (9 bores) in Dry Climate scenario.
7.2	Have multiple scenarios been run for operational /management alternatives?	N/A	Missing	Deficient	Adequate	Very Good			One mine plan. Recovery for 1000 years. Three tailings/rejects scenarios. Open cut drains active 1 year. Underground drains permanent to end of mining. No final void. No need for cumulative impacts.
7.3	Is the time horizon for prediction comparable with the length of the calibration / verification period?		Missing	No	Maybe	Yes			25 years projected from 3 years transient calibration and verification (Nov 2011 - July 2014).
7.4	Are the model predictions plausible?			No	Maybe	Yes			Consistent with earlier modelling and experience with neighbouring mines.
8.0	SENSITIVITY ANALYSIS								
8.1	Is the sensitivity analysis sufficiently intensive for key parameters?		Missing	Deficient	Adequate	Very Good			Extensive. Done for Kx, Kz, recharge%, Sy, Ss, RIV conductance, DRN duration, fractured zone Kz, van Genuchten coefficients. Comparison of Richards Equation solution, pseudo-soil and USG upstream weighting.
8.2	Are sensitivity results used to qualify the reliability of model calibration?		Missing	Deficient	Adequate	Very Good			Ranges from 3.4 to 7.0%RMS transient (base 3.6). "PseudoSoil" effectively 5.8 %RMS - uncalibrated. .
8.3	Are sensitivity results used to qualify the accuracy of model prediction?		Missing	Deficient	Adequate	Very Good			For mine inflow, drawdown extent, baseflow and rock-alluvium fluxes.
9.0	UNCERTAINTY ANALYSIS								
9.1	If required by the project brief, is uncertainty quantified in any way?		Missing	No	Maybe	Yes			Calibration-constrained Monte Carlo. Also traditional sensitivity analysis.
	TOTAL SCORE								PERFORMANCE: %