APPENDIX 9 – INDEPENDENT PEER REVIEW OF THE NUMERICAL GROUNDWATER MODEL

NEW LENTON COAL PTY LTD

NEW LENTON COAL PROJECT

NEW LENTON PROJECT

Independent peer review of the New Lenton Project numerical groundwater model

MANEWELL GROUNDWATER

Source: Google earth August 2023

1. Introduction

This report presents an independent technical peer review of the numerical groundwater model prepared for the New Lenton Project (the Project, New Lenton) public environmental report (PER) under the *Environmental Protection and Biodiversity Conservation Act 1999* (EPBC Act) (EPBC 2020/8778). The model incorporates the Project and the adjoining Burton Mine (Burton) and is used to make predictions. The aim of the review is to ensure technical competency and adherence to regulations for the numerical groundwater model.

New Lenton Coal Pty Ltd (NLC), a wholly owned subsidiary of Bowen Coking Coal Ltd (BCC), proposes to construct and operate an open-cut coal mine and associated infrastructure, located 65 kilometres north of Moranbah, Queensland. BCC has engaged SLR to prepare a groundwater impact assessment to support the Project. The Project comprises two opencut pit areas operating over a period of 17 years, after which a final void will remain. The Project will utilise infrastructure at Burton Mine, which is also operated by BCC.

It is noted that the model incorporates both the Project and Burton, which ensures that any cumulative predictions of groundwater impacts are explicitly modelled. Burton Mine targets the Burton/Rider and Leichardt/Vermont seams of the Rangal Coal Measures via open-cut mining. A proposed final landform for Burton Mine contains five residual final voids.

The controlling provisions for matters of national environmental significance (MNES) for the Project include a water resource, in relation to coal seam gas development and large coal mining development ('the water trigger'). The PER has been prepared to determine if there are likely to be significant impacts to water resources.

Pit dewatering is predicted to lower groundwater levels in the adjacent alluvium and Permian strata.. Previous modelling has indicated The Project is unlikely to significantly impact surrounding landholder bores and groundwater dependent ecosystems (GDEs) (SLR, 2020).

The numerical modelling has been developed in support of the PER for submission to DCCEEW under the EPBC Act for the Project. As the model incorporates Burton Mine it will also be used to support a State approvals process for a transitional progressive rehabilitation and closure plan (PRCP) schedule.

The objectives of the New Lenton PER are to predict inflows to mine workings, drawdown, impacts on surface water, changes to alluvial resources, and groundwater levels. The objectives of the Burton PRCP are to simulate post mining groundwater level recovery, long-term drawdown, and flux to residual voids.

2. Documentation

Familiarisation of the Project was undertaken by briefly reviewing the following reports:

- Hydrogeological Study (GHD, 2019)
- Groundwater Modelling Report (GHD, 2019)
- Groundwater Modelling Report (SLR, 2020)
- Conceptual model update report (SLR, 2020)
- Trigger-limits report (SLR, 2020)

This review is based on the following report:

- SLR, New Lenton & Burton. 2023 Groundwater Modelling Technical Report, SLR Project No.: 620.31367.00000, 20 December 2023, Revision: 3.0 + 8 Appendices (Modelling Technical Report)

The Groundwater Modelling Technical Report is an appendix to Project's PER Groundwater Assessment (SLR, 2023)

Manewell Groundwater attended numerous progressive groundwater peer review meetings during the course of model development for SLR 2023, namely the presentation of calibration results, predictive model setup, uncertainty analysis results, post mining setup.

3. Review methodology

The review methodology is in accordance with the principles and guidelines outlined in the Australian Groundwater Modelling Guidelines (Barnett et al, 2012). This technical review incorporates the latest guidance on best practices for uncertainty analysis methods as presented in Uncertainty analysis—Guidance for groundwater modelling within a risk management framework (Middlemis and Peeters, 2023). The technical review is also based on a combination of professional expertise and modelling methodology, together with the model appraisal checklist provided in the Australian Groundwater Modelling guidelines.

The technical review consults the following federal regulatory guidelines to assess the sufficiency of the numerical groundwater model developed for the PER (SLR, 2023) and assess compliance with the following:

- Information guidelines for proponents preparing coal seam gas and large coal mining development proposals, issued by the Independent Expert Scientific Committee of the Commonwealth of Australia (Peeters & Middlemis, 2023).
- Significant impact guidelines 1.3, issued by the Department of the Environment and Energy of the Commonwealth of Australia (DES, 2013).

4. Model review

4.1 Model conceptualisation

Conceptualisation is presented in the SLR, 2023. Sufficient rigor has been applied to the development of this model which represents a culmination a numerous hydrogeological investigations over the Project area.

4.2 Model code/extent and mesh design

The model utilises industry standard MODFLOW-USG with Voronoi cell unstructured model grid. The extent is sufficiently large to overcome boundary condition effects (9 to 14 km from open-pit mining). The grid is refined appropriately around key features, including creeks, rivers, dams, alluvium, faults and pits to reduce structural error. Model cells are ‘pinched’ where hydrostratigraphic units are not present.

4.3 Model layers

A total of 2 layers are used to represent near surface hydrostratigraphy. Layer 1 represents alluvium where present and a nominal 1m thick regolith outside the extent of alluvium. It is likely most of the regolith cells in layer 1 are dry and essentially inactive for the majority of the simulations. Layer 2 represents the underlying regolith and basalt. Thicknesses are based a composite of geological mapping, CSIRO thicknesses (assume 50th percentile estimate), bore data, and geological models. This segregation is appropriate given the expected hydraulic properties of the shallow strata. The Triassic and Permian have been segregated into nine hydrostratigraphic units whose layer elevations are based on geological mapping and models.

Each Triassic and Permian Rangal Coal Measures layer is split laterally into 2 separate layers to represent disconnect across the Burton Range Fault. Potential connection to neighbouring offset strata is appropriately represented through the use of the Lateral Connection Group option in Algomesh.

4.4 Timing

Steady state represents pre-mining (1998) conditions, progressing to a transient historic annual stress periods from 1998 to 2011. Quarterly stress periods commence Jan 2011 to July 2023 to represent seasonal fluctuations and to better represent mining progression.

Predictive simulations employ quarterly stress periods from July 2023 to January 2024. Annual stress periods are simulated during the predictive mining period (Jan 2024 to December 2039). Timing is presented in Section 3.1 of the Modelling Technical Report. Given the absence of significant quarterly variability in groundwater levels this approach is appropriate. Post mining is simulated using 6-monthly to 100 yearly stress periods – a common approach.

4.5 Boundary conditions and stresses

General head boundary (GHB) cells are used to represent regional groundwater flow into and out of the model. GHB cells are appropriately represented in *all* model layers using calculated conductance based on cell dimensions and simulated horizontal hydraulic conductivity. This implies that hydraulic conductance changed during the history matching and predictive uncertainty simulations. This is common practice. The GHB cells are sufficiently far from the mining areas to not compromise the reliability of the predictions.

River cells along the Isaac River are used to appropriately represent primarily ephemeral conditions using zero-stage height at all times. River stage heights are informed by LiDAR elevation data where available. Leakage following high rainfall events is simulated using the recharge package. No baseflow is expected to occur – this assumption has been included as a measurement target during the history matching process. Sufficient rigor has been implemented in the representation of surface water, given site observation data. Dams are represented using measurement stage height information.

Diffuse recharge is appropriately represented using LUMPREM2. Daily rainfall and potential evaporation was simulated using SILO point data. Although simulated temporal recharge is less variable than other soil-moisture bucket methods (i.e., there is positive recharge to the watertable during all quarterly stress periods), LUMPREM2 is an industry standard and commonly accepted approach. The parameters used in LUMPREM2 appear reasonable. It is not apparent if the recharge events from LUMPREM2 align with apparent increases in groundwater levels, and if the parameters were initially calibrated using site information. Regardless, hydrographs presented in Appendix D suggest recharge events align generally successfully to groundwater level signatures. Spatial factors were then applied to the LUMPREM2 generated rates to different hydrostratigraphic units appropriately.

Evapotranspiration cells are used to simulate the removal of groundwater from plant/tree roots that extend to the groundwater table. EVT depths are appropriately history matched using bounds representative of expected rooting depths. Rates are calibrated and based on the BOM EVT estimate map in a reasonable manner.

4.6 Mining

Mine dewatering associated with open-cut mining practices are simulated using the Drain package. Drain elevations are set to the base of the lowest mined elevation at a nominally high conductance rate, and advance at a quarterly or annual rate during the calibration phase. Predictive mine progression advances at an annual basis. Drains are held for a period of 2 years before spoil or mine water storage is emplaced using the TVM package. Open cut mining is represented to industry standards.

Water storages are simulated using the CHD package, and deactivated using the TIB package appropriately. One negative of the CHD package is that injection of groundwater is unrestricted, however, presented results indicate reasonable volumes entering the model. Activation and removal of these storages are provided in the report.

4.7 History matching

The historical model was history-matched using industry standard data assimilation software PESTPP-IES (White, 2018). A sufficiently large ensemble of 250 realisations was used; current guidance is to generate an ensemble of equal size to the number of singular values present in the parameter/observation dataset. A total of 7 optimisation iterations were performed. Groundwater level and drawdown data from a relatively large dataset was used to history match. The spatial coverage of these bores generally extends a little further than the New Lenton Mining Lease. Burton observations are generally close to the pits, although there are a few bores that exist far field.

Sixty-six zero flux baseflow targets are applied to the 3 km section of the New Lenton mine pits. Both head, drawdown, and baseflow is appropriately weighted using a weight balancing script (PWTADJ1).

Horizontal, vertical hydraulic conductivity, specific yield, recharge, river bed vertical hydraulic conductivity, evaporation depths and rates were refined as common practice. Pilot point multipliers in layer 1 were added appropriately to represent heterogeneity. Spatial covariance information generated using PPCOV was added to PESTPP-IES to promote spatial correlation during the optimisation process (i.e., pilot points close to each other within the 'A range' of a semi-variogram will maintain similarity). PPCOV parameters are detailed in the report.

The performance metrics for the best calibrated model is presented in section 2.2.3 of the Modelling Technical Report as scatter plots and residual histograms. The SRMS from the best calibrated realisation is a fairly arbitrary metric for regional models. It is more important that trends and flows are replicated from the ensemble as a whole. Pleasingly, the results from the ensemble are presented and discussed with rigor. Calibrating to historic mining where mine progression and other stressors are uncertain is difficult. Representing complex hydrostratigraphy with uniform parameters that employ a depth dependence function without further spatial variability means that the models can only fit an average head (and drawdown) to the spatial measurement dataset. The ensemble generally spans the range of measured groundwater levels with expected and appropriate precision. This reviewer agrees with the conclusions presented in section 2.2.5.4 of the Modelling Technical Report that explain ongoing mismatch.

Simulated water level and depth to water level maps from the best calibrated realisation reveal levels and saturated extents that are consistent with model conceptualisation. Baseflow rates along Isaac River are low and correspond to LUMPREM2 recharge rates. It is not reported how the baseflow varies within the ensemble, although Figure 2-39 indicates the ensemble is generally well behaved. Importantly however, measured groundwater levels in the Isaac River Alluvium are generally well-matched across the model ensemble.

Whole of model flows appear reasonable and consistent across steady state and transient phases. There are some stress periods where CHD inflows are high as water storage is emplaced, although these appear to be temporary, reducing to more reasonable consistent inflows to the system. EVT is the highest outflow, which suggests that EVT cells act as a surrogate for surface drainage in low lying areas.

Calibrated hydraulic parameters are presented from the best calibrated realisation and appear reasonable and consistent with site measurements and other regional studies. The range of parameters considered in the historic ensemble are presented as Appendix B. Generally, the imposed range is broad and wider than what is typically applied in similar environments. It is noted that specific storage is higher than contentious ranges implied from ongoing research and poroelastic theory. However, given the averages for specific storage was calibrated towards the lower range, it is not anticipated that artificially high values of specific storage would significantly influence the range of predictions. Recharge, evapotransporation and river parameters for the best calibrated parameter set appear reasonable. Recharge to spoil is based on MER (2009) – a common approach.

Despite structural error discussed in section 2.2.5 of the Modelling Technical Report this reviewer agrees the ensemble of models are deemed fit for purpose.

4.8 Predictions

Three predictive models were used to make predictions in cumulative conditions using the standard Cumulative, Approved, Null approach. Mass balances are consistent with the historic phases.

Predictive drain cells are simulated according to an operation window indicated by mine plan surfaces. Enhanced spoil recharge is emplaced immediately with spoil emplacement. In reality, spoil recharge to the watertable occurs several years after emplacement, however, given the progression of the opencut at New Lenton it is likely this potentially buffering capacity is minimised.

Predicted drawdowns in the watertable and alluvium are as expected. Incremental watertable and alluvial drawdown extends up to 2 km from the New Lenton mine pits. Drawdown predictions for the best calibrated realisation appear reasonable.

Absolute inflows to the New Lenton workings appear reasonable. Inflow peaks to over 800 ML/year from East Pit.

Incremental alluvial takes peak at approximately 180 ML/year at 2030, as expected. Negligible take from the Isaac Connors Alluvium is predicted.

No baseflow was encountered in the best calibrated realisation during the predictive phase, therefore there is no anticipated take. If seasonal recharge were applied during the predictive phase, it might be reasonable to assume that the model could predict a maximum of approximately 7 ML/year (from average historic predictions of 0.02 ML/day), although it is important to note that baseflow does not occur, and this rate is more representative of local evapotranspiration from recharge, rather than watertable derived baseflow. Simulating this detail is difficult in models of this scale (even quarterly or monthly stress periods is an oversimplification) and is beyond the capabilities of such a regional scale model.

4.9 Calibration sensitivity analysis

PESTPP-IES does not record parameter sensitivities during the optimisation process. A separate sensitivity analysis was undertaken, assumedly using the best calibrated realisation, by creating a traditional sensitivity or Jacobian Matrix through perturbation. The discussion on identifiability is thorough. Results are as expected given the observation dataset. Values of high identifiability are encountered where a given parameter is estimable to measurement data. If many measurements inform a single parameter, such as Kx-L8-Rewan, identifiability is high. This could be an issue if posterior parameter realisations are generated using this information, however, the sensitivity matrix was not used to generate realisations used for uncertainty analysis.

4.10 Predictive uncertainty analysis

A Type 3 uncertainty analysis was undertaken to assess future impacts. The posterior ensemble from PESTPP-IES was analysed and the posterior standard deviation was adjusted. Then the final posterior was attained via model performance rejection. 162 realisations from the calibration ensemble met this performance criteria, so were included in predictive uncertainty. As such, the 'prior' is not really the 'prior', as it has been history matched already . Regardless, the range of parameters appears wide in all instances, suggesting that the predictive uncertainty steers towards the conservative end.

A total of 362 realisations were tested. Confidence intervals were tested using an increasing number of model realisations to reveal that little would be gained using a higher number of runs. This is expected – generally most information of the pertinent predictions within the 90th percentile are gained using less than 100 realisations.

Inflow predictions into the New Lenton mine are presented according to IESC guidelines. Maximum inflows appear reasonable and likely conservative due to the wide posterior parameter ranges assessed. Uncertainty of take from the alluvium is presented in a table. Results suggest that the calibrated model sits higher than 50th percentile from the uncertainty analysis. Very little impact on the Isaac Connors Alluvium Groundwater sub-area is predicted at the 90th percentile.

Percentiles of maximum drawdown encountered at any time during the predictions are presented. Results suggests that it is unlikely that incremental drawdowns of 1 m or more will extend an additional 1 km from the best calibrated realisation. Given the conservatism of the impacts and the range of parameters assessed in the uncertainty analysis this is likely a conservative estimate. This is because distal drawdown is partially controlled by the drawdown extent of the underlying Triassic/Permian strata.

4.11 Post mining recovery modelling

Following mine closure, the recovery of groundwater is simulated using the industry standard 'high K lake' approach. After an iterative process with a surface water model, CHD cells are used to represent pit lake using predictive transient pit lake stages.

The recovery of pit lake levels within the New Lenton void appear reasonable. Equilibrated groundwater levels appear to remain approximately 20 m below the regional groundwater table, and significantly lower than the base of shallow geological formations, suggesting a perpetual groundwater sink.

Water balances to the void appear reasonable and in line with predictions at the end of predictive mining. The properties of the spoil, evapotranspiration and recharge rates are reasonable.

Post mining drawdown and flow changes from New Lenton mine in the alluvium is consistent with the predictive phase. As expected, drawdown in the Permian strata extends further than the predictive phase.

Particle tracking confirms the voids for both New Lenton and Burton form a long-term sink to groundwater flow.

Hydrographs are presented at key receptors that indicate groundwater receptors do not experience significant impacts, although it important to note this is the best calibrated case, and other realisations are likely to experience higher, yet less likely drawdown at these locations.

Uncertainty analysis on the recovery model was undertaken using the same posterior ensemble used during predictive analysis to predict the likelihood in inflows, indirect flow changes and maximum incremental drawdown. This approach is aligned to guiding principles presented in Tomlin et al (2022). Results suggest all voids remain long-term groundwater sinks, with the exception of the Ellensfield void at Burton, which shows a very low likelihood (10th percentile) of being a long-term source. Indirect flow changes to the alluvium are reasonable and consistent with rates at end of the predictive simulation. It is noted that all realisations use the same transient pit lake recovery curve as calculated from surface water modelling. If predicted equilibrated pit lake levels vary significantly given different hydraulic properties, recharge rates, and evaporation, different pit lake levels could be considered within the ensemble for future post-mining studies.

Predicted incremental post-mining drawdown extents are extensive due to broad parameter ranges assessed. Commendably, lateral connection imposed in the model DISU package allows significant hydraulic connection to adjacent layers. It is likely that strong hydraulic connectivity is impeded by clay gouge along the fault, and layer wide hydraulic connectivity is less than explored in the uncertainty analysis. These factors suggest 90th percentile impact predictions are overstated.

5. Report review

The Modelling Technical Report is a high-quality document of about 190 pages long including 8 appendices that contain information of spatial hydraulic properties, history matching performance, water budgets, drawdowns, hydrographs and parameter distributions. SLR, 2023 contains useful information regarding the implications of drawdown to GDEs. Modelling objectives are clearly stated and adequately addressed in the report.

The report is well-structured and well-written, the majority of the figures and maps are of high quality and easy to understand. An executive summary in the modelling appendix is not provided.

Progressive review comments were provided by the reviewer in the working modelling 'docx' document. Reviewer comments have been considered by SLR and have been accommodated satisfactorily in revisions of the report.

6. Data review

The hydrogeological model for which this model is built upon was derived from four field programs undertaken from 2014 to 2023. Drilling programs and geophysics have been used to build the geological model. Slug testing forms the basis of the hydraulic testing program. This is suitably covered in SLR, 2023. Given slug tests provide a snapshot of average hydraulic properties, literature ranges have been used to form hydraulic parameter assessed under uncertainty. As a result, parameters ranges remain suitably wide. Posterior uncertainty could only be reduced through the acquisition of additional hydraulic testing, such as packer, pumping tests, or laboratory analysis. If there are concerns over the range of predictions presented more hydraulic information is required.

Generally, the quality of groundwater level measurements within the model domain is adequate and captures episodic climate variability and longer-term trends. There are frequent readings from many bores that screen multiple groundwater model layers.

Groundwater recharge and surface water flow has been conceptualised with rigor, considering the arid environment in which the Project exists. Groundwater quality is suitably discussed in SLR, 2023.

The identification of potential GDEs is thorough and implications considering groundwater impacts is discussed. There are very few groundwater users in the model domain.

It is likely that historical groundwater inflows to Burton Mine are uncertain. The model would benefit from measured flows during the history matching process to reduce non-uniqueness.

7. Model level classification

Similar to the Modelling Technical Report, this report applies a new concept for assessing model confidence.

The 2023 IESC Explanatory Note for Uncertainty Analysis defines a model is fit for purpose when results are:

- Useable – Relevant to the decision-making process, provide information on the uncertainty in conceptualisations and model simulations that allows decision-makers understand the effects of uncertainty on project objectives and potential bias.
- Reliable – Demonstrate that the range of model outcomes is consistent with system knowledge. Uncertainties affecting decision-critical predictions are not underestimated.
- Feasible - Trade-offs due to budget, time and technical constraints are reasonable and justifiable within the risk context of the project.

Table 1 presents this reviewers assessment of the four sources of uncertainty (structural, parameterisation, measurement error, predictions).

Table 1 Model confidence assessment

Type	Part	Status	Comment
Structural/ Conceptual	Grid and Model Extent	Fit for purpose	Model extent sufficiently large so that drawdown does not significantly interact with GHB boundaries. MODFLOW-USG Voronoi grid with more than sufficient refinement area key features.
	Layers	Fit for purpose	Top of model uses LiDAR data where available.
		Fit for purpose	Model layer elevation sourced from best available data sources (TEM, CSIRO regolith, drilling, site geological models)
	Conceptualisation – Geological Structure	Fit for purpose, future improvements possible where new data collected	Geological structure based on best available data. Validated from observed and mapped geology. Geological layers displaced by the Burton fault. Connection to different units across the fault automatically calculated using Algomesh.
	Conceptualisation – GDEs	Fit for purpose, future improvements possible	GDE conceptualisation provided in accompanying SLR, 2023.
	Conceptualisation – Surface Water Groundwater Interactions	Fit for purpose, future improvements possible	Surface water-groundwater interaction considered minimal in the New Lenton Project area. Detailed discussion in provided in SLR, 2023.
	Conceptualisation – Saturated Extent of Alluvium and Regolith	Fit for purpose	Saturation extents informed by site monitoring bores.
Parameterisation	Hydraulic Conductivity – Shallow layers	Fit for purpose, future improvements possible	Heterogeneity represented using pilot points. Future analysis could involve pumping tests to key parts of the alluvium to reduce parameter uncertainty and validate variogram used to impose spatial connectivity.
	Hydraulic Conductivity – Depth Dependence	Fit for purpose, future improvements possible	Depth dependence equation used to impose average changes in hydraulic conductivity with depth. Future analysis could involve pumping tests (if possible) Future modelling should consider pilot points/linear structures to validate connectivity between the pit and the Permian/Triassic below the alluvium.
	Storage	Fit for purpose, future improvements possible	Field hydraulic testing of specific storage and specific yield appears limited.
	Vertical Hydraulic Conductivity	Fit for purpose, future improvements possible	Literature and model used to attain vertical hydraulic conductivity estimates. Nested monitoring bores or VWPs could be used to constrain ranges of Kv. Laboratory hydraulic testing could constrain lower bound Kv.
	Spoil Properties	Fit for purpose, future improvements possible	As usual, limited site-specific information is available. The model was used constrain spoil properties. Future work should hydraulically test emplaced saturated spoil.
	Rivers	Fit for purpose	Isaac River is highly ephemeral. Flow occurs for a few days per year. Observed groundwater indicate baseflow is unlikely to exist. Isaac River represented using river cells with zero stage height representing baseflow only. Zero baseflow was targeted during the calibration process.

Type	Part	Status	Comment
	Recharge	Fit for purpose	Recharge calculated using LUMPREM2. Recharge distributed as factors using surface geological zones.
Measurement Error	Observation Data Quality	Fit for purpose, future improvements possible	Bore logs and construction details available for most site bores. Most units have long-term water level data available. Measurement noise implied by the accuracy of the measurement devices/bore construction could be incorporated more accurately into inversion exercise (if not done already).
	Landholder Bore Data Quality	Fit for purpose, future improvements possible	Landholder bore construction subject to more uncertainty. Field census survey data. Uncertainty of screen information could be incorporated into drawdown estimation (e.g., span observation bore level over multiple layers).
	Temporal spread	Fit for purpose	Timeseries water level data gathered for all units over relatively long monitoring period.
Scenario Uncertainties Future Stresses / Conditions	Calibration	Fit for purpose	Steady state/Transient (1998-2023) history-matching to groundwater levels, drawdown and baseflow undertaken using PESTPP-IES. Calibration includes historic mining at Burton Mine.
	Predictive	Fit for purpose	Predictive model simulated approved mining at Buron and proposed open cut at New Lenton.
	Sensitivity and uncertainty	Fit for purpose	Predictive uncertainty analysis undertaken on the 3 scenarios. Recharge and EVT parameters were varied within the ensemble that span the range of potential climate change scenarios. Predictive uncertainty undertaken on post mining recovery scenarios.

Overall, it is this reviewer's opinion that the model is indeed useable, reliable and feasible in accordance with the IESC uncertainty note criteria. There is room for improvement in the next phase of modelling, which would aim to further reduce predictive uncertainty.

8. GMDSI decision-support modelling guiding principles

In a similar vein to the IESC uncertainty note, the Groundwater Modelling Decision Support Initiative released 'Monograph: Decision-Support Modelling Viewed Through the Lens of Model Complexity' (Doherty & Moore, 2021). Manewell Groundwater has distilled the key guiding principles into checklist items to ensure the model is not overly complex and serves as a useful tool that appropriately assesses the likelihood of the "bad thing" from occurring. Table 2 presents pertinent guiding principles as checklist items.

Table 2 Decision-support guiding principles checklist

Category	Guiding Principle Item	Yes/No	Comment
Client/Regulator Concerns	Have the impacts that the client/regulator fears been clearly identified before starting the modelling process?	Yes	Several iterations of this model have been presented to regulators and peer reviewers.
Model Purpose and Uncertainty	Is the model specifically designed to make a particular prediction and quantify the uncertainty associated with that prediction?	Yes	Model designed to predict drawdown, inflow and flow change.
Simplicity Over Complexity	Has the model avoided unnecessary complexity, opting for simpler, effective approximations?	Yes/No	Simpler representations of localised parts of the connection between pit and the receptors possible, however, these are often interpreted to be "too simple", hence a complex simulator is used in isolation.
Conceptual Model Collaboration	Is the conceptual model developed in collaboration with relevant experts, addressing where key information resides?	Yes	Model developed with team of hydrogeologists, groundwater modellers, and surface water modellers.
Exploration of Hydrogeological Representations	Are multiple representations of specific hydrogeological details explored, acknowledging system unknowns?	Yes/No	Pilot points used to represent heterogeneity in the alluvium. Localised variability in the Triassic/Permian not assessed, however it is likely these features exist at a fine scale.
Historical Stress Response Simulation	Can the model simulate the system's response to historical stresses effectively?	Yes	Ensemble of models replicates absolute heads and transient trends appropriately.
Numerical Efficiency and Robustness	Does the model run reasonably fast and maintain numerical robustness for effective hypothesis testing?	Yes/No	Model runs at an appropriate speed to meet the model objectives. Future hypothesis testing would require a model that runs significantly faster, although this is not in the scope of this project.
Model-Partner Software Integration	Is there a non-intrusive interface with model-partner software, capable of running models in parallel?	Yes	History-matching and uncertainty analysis undertaken using PESTPP suite.
History Matching Approaches	Are appropriate numerical strategies, such as Tikhonov regularization and singular value decomposition, used in history matching?	Yes	Pilot point spatial covariance used as regularisation in PESTPP-IES.
Data-Driven Prediction Design	Is the primary model designed for calibration against historical field measurements?	Yes	Quarterly stress periods used to simulate climate response and historic dewatering.
Parameterization in Decision-Support Modelling	Are model parameters plentiful, adjustable, and capable of stochastic representation?	Yes/No	Shallow groundwater system has an appropriate number of pilot points. Deeper units have a simpler parameterisation scheme.

Category	Guiding Principle Item	Yes/No	Comment
Accommodation of Uncertainty in Decision-Modelling	Does decision-modelling accommodate uncertainty, exploring the pessimistic end of the uncertainty bell curve?	Yes	The posterior parameter ranges presents as wide. Given layer wide parameters in the Triassic/Permian are perturbed, it is assumed the predictions are at the pessimistic end of the bell curve.
Reduction of Decision-Critical Uncertainty	Is decision support modelling tuned to reduce the uncertainty of one or a few critical predictions?	Yes	Drawdown, inflow, baseflow and flow change predicted.
Avoiding Type 2 Statistical Error	Does the model avoid overstating the widths of decision-critical predictive uncertainty analysis?	Yes	Range of posterior parameter distribution enforced by expert knowledge, so worse case impacts are within reality.
Predictive Error Accounting	Is predictive error introduced by numerical methodologies accounted for and communicated effectively?	Yes	Posterior parameter distributions presented. Model rejection process adequately discussed. Appropriate range of predictions present as likelihood bands.
System Component Representation	Are essential system components that impact predictions represented in the numerical model?	Yes	Key processes relevant to dewatering and the saturated groundwater phase represented.
Parameter Importance and Hosting of Information	Is the importance of parameters and their role in hosting information acknowledged?	Yes	Pilot points used to expose heterogeneity in response to groundwater level measurements.
Hypothesis Testing for Future Stresses	Is the hypothesis that future stresses will result in negative outcomes only rejected when demonstrably incompatible with known system measurements or behaviours?	Yes/No	Predictive uncertainty used historic groundwater level, drawdown and flow information to reject unrepresentative realisation. Hypothesis testing not undertaken and not required at this stage of the Project.
Prediction Spectrum Awareness	Is there an understanding of the differences between data-driven, expert-knowledge-driven, and mixed predictions?	Yes	Not explicitly addressed, although mixed prediction approach used is appropriate.

Overall, it is this reviewers opinion that the model serves as a useful decision support tool. Generally, the model has the ability to harvest information. If hypothesis testing were required for future studies (e.g., if early dewatering indicates more impacts than simulated), a simpler, highly parameterised version of the model should be developed that has the ability to expose avenues for drawdown that were not previously represented in the complex approvals model.

9. Australian Modelling Guidelines checklists

Table 3 is the NWC Compliance Checklist from the Australian Groundwater Modelling Guidelines (Barnett, 2012). The results conclude that the groundwater model is “fit for purpose”, where the primary purpose is to develop upon an existing model that is reasonably well calibrated to historical groundwater levels, drawdown and baseflow (or lack thereof). The model is a suitable tool for decision support where the purpose is to provide predictions on potential impacts due to mining at New Lenton and Burton Mine.

Table 3 Compliance checklist

No.	Review Questions	Yes/No	Comment
1	Are the model objectives and model confidence level classification clearly stated?	Yes	Model objectives clearly stated, and model level classification discussed in detail.
2	Are the objectives satisfied?	Yes	The objectives of the study of assessment of effects from the mine have been satisfied through predictions of dewatering effects on receptors. Alluvial flow change and baseflow impacts are sufficiently calculated and discussed. The effects of the voids on the post mining groundwater system are presented and discussed.
3	Is the conceptual model consistent with objectives and confidence level classification?	Yes	The conceptual model is consistent to the geological structure and mining induced stresses. The interaction between surface and groundwater stresses is adequately simulated given the conceptualisation, field observations, and scale of the model. Although the model is high in cell count, it proves to be useable, reliable, and feasible. The numerical model is history matched to best practice and uncertainty analysis assessed to industry standards.
4	Is the conceptual model based on all available data, present clearly and reviewed by and appropriate reviewer?	Yes	The model was based on a cumulation of previous modelling designs. Although the conceptual model is not presented in the numerical report, a conceptual model is clearly outlined in SLR, 2023, and there are illustrations that include hydrogeological stresses and interactions. The differences between the previous models is clearly defined.
5	Does the model design conform to best practice?	Yes	An adequate conceptual model has been generated and is reflected in the numerical model. The simplistic surface/groundwater interactions are aligned to the conceptual model. Pilot points are used to represent heterogeneity in the alluvium. Heterogeneity is not represented in the hard rock, which likely results in conservative estimates, given the range of parameters tested remains high at a regional scale.
6	Is the model calibration satisfactory?	Yes	The model correctly replicates absolute groundwater levels and changes to groundwater levels around the areas of importance proximal to New Lenton Mine. Large mismatches remain close to uncertain historic mining/infrastructure. Importantly recharge trends are generally well matched within the ensemble of models. Remaining mismatches are exposed and discussed. The model ensemble is constrained using a baseflow target to reduce non-uniqueness.
7	Are the calibrated parameter values and estimated fluxes plausible?	Yes	History-matched parameter values are within plausible bounds and align to site information (although limited and reliant primarily on slug testing). Recharge rates and hydraulic parameters of the aquifer units are consistent with studies undertaken throughout the region. Fluxes into the mine workings appear plausible. Fluxes to surface water drains are aligned to site information.
8	Do the model predictions conform to best practice?	Yes	Model predictions include time variant pit seepage inflows and drawdown. Composite maximum drawdown maps are presented to show transient maxima. Impacts on baseflow and alluvial flow changes are presented. Predictive hydrographs with layer elevations are presented. Predictions are calculated using the difference between null, approved, cumulative scenarios at a time-step level.
9	Is the uncertainty associated with the predictions reported?	Yes	A comprehensive calibration constrained Monte Carlo predictive parametric uncertainty analysis using 362 realisations spanning a wide range of parameters. Structural uncertainty is discussed and attempts made to reduce error. Measurement error discussed and incorporated into the history matching process (as noise). Different scenarios assessed. Variability in recharge across the uncertainty analysis arguably spans reduction in recharge implied by climate change. Post mining uncertainty analysis was undertaken and presents long-term drawdowns that are likely over-stated.
10	Is the model fit for purpose?	Yes	The model is built and assessed to industry standards. Given site information around New Lenton suggesting surface-groundwater interactions are minimal, and predictive uncertainty presented is likely conservative, the model is considered fit for purpose.

Table 4 Review checklist

No.	Review Questions	Yes/No	Comment
1	**Planning**		
1.1	Are the project objectives stated?	Yes	Designed to comply with submission under the EPBC and EP Acts.
1.2	Are the model objectives stated?	Yes	Designed to comply with submission under the EPBC and EP Acts.
1.3	Is it clear how the model will contribute to meeting the project objectives?	Yes	The report states that the modelling results are to assess the likely impacts of the proposed mining activity.
1.4	Is a groundwater model the best option to address the project and model objectives?	Yes	Measuring the effects of the propagation of drawdowns on piezometric surface through time in this case is best tested through a numerical groundwater model.
1.5	Is the target model confidence-level classification stated and justified?	Yes	Contemporary model level classification not discussed.
1.6	Are the planned limitations and exclusions of the model stated?	Yes	Detailed coverage of model limitations including data gaps and interpolation factors in report.
2	**Conceptualisation**		
2.1	Has the literature review been completed, including examination of prior investigations?	Yes	Yes previous studies and models have been utilised to develop the current model.
2.2	Is the aquifer system adequately described?	Yes	The unconsolidated and consolidated strata have been described with a number of model layers - 9 in total (split in two each side of the Burton range fault). The hydrostratigraphy and vertical gradients have been adequately simulated.
2.2.1	Hydrostratigraphy including aquifer type (porous, fractured rock …)	Yes	The hydrostratigraphy describes the main components of the system related to the desired project objectives - namely unconsolidated alluvial aquifers, weathered overburden and consolidated Permian formations.
2.2.2	Lateral extent, boundaries and significant internal features such as faults and regional folds.	Yes	The extent of the groundwater model is fit for purpose and well thought out. Fault displacement and potential lateral connectivity represented using lateral connection groups. Rivers and creeks are represented as baseflow conditions only, as per site information, although can leak via recharge.
2.2.3	Aquifer geometry including layer elevations and thicknesses.	Yes/No	SLR, 2023 presents layer elevations of base of alluvium and thickness. Elevation and depth of base of Triassic and Leichardt/Vermont seams presented only.
2.2.4	Confined or unconfined flow and the variation of these conditions in space and time?	Yes	Brief discussion on nature of unconfined/confined conditions in numerical report. Saturation thicknesses of the alluvium formation discussed.
2.3	Have data on groundwater stresses been collected and analysed?	Yes	
2.3.1	Recharge from rainfall, irrigation, floods, lakes.	Yes	Recharge calculated using LUMPREM2.
2.3.2	River or lake stage heights.	Yes	River stage heights set to river base to represent ephemeral creeks. Positive river stages used to represent dams and water bodies.
2.3.3	Groundwater usage (pumping, returns etc.)	N/A	Well package not required in the calibration or predictive scenarios.

No.	Review Questions	Yes/No	Comment
2.3.4	Evapotranspiration.	Yes	EVT package used to simulate groundwater take from tree roots. In some areas this is used as a surrogate for surface drainage.
2.3.5	Other?	No	
2.4	Have groundwater level observations been collected and analysed?	Yes	Relatively extensive groundwater level dataset.
2.4.1	Selection of representative bore hydrographs.	Yes	Key hydrographs presented and discussed. All other hydrographs presented in an Appendix.
2.4.2	Comparison of hydrographs.	Yes	Observed and simulated groundwater levels presented on the same chart.
2.4.3	Effect of stresses on hydrographs.	Yes	Groundwater drawdown and short-term recovery replicated.
2.4.4	Watertable maps/piezometric surfaces?	Yes	Historic and simulated contoured heads presented.
2.4.5	If relevant, are density and barometric effects taken in account in the interpretation of groundwater head and flow data?	N/A	Not relevant to the scale of this study.
2.5	Have flow observations been collected and analysed?	Yes/No	Observed baseflow discussed. Estimated inflows to existing workings (at Burton) not discussed in the numerical modelling report as they are not the focus of this Project.
2.5.1	Baseflow in rivers.	Yes	The absence of measured baseflow used as a history-matching target.
2.5.2	Discharge in springs.	N/A	SLR, 2023 suggests no springs were mapped within the vicinity of the project area.
2.5.3	Location of diffuse discharge areas?	N/A	Out of the scope of this study.
2.6	Is the measurement error or data uncertainty reported?	No	Briefly discussed, although measurement error assumed in the inversion exercise not presented.
2.6.1	Measurement error for directly measured quantities (e.g. piezometric level, concentration, flows).	No	Briefly discussed, although measurement error assumed in the inversion exercise not presented.
2.6.2	Spatial variability/heterogeneity of parameters.	Yes/No	Discussion of local heterogeneity discussed in limitations section. Heterogeneity represented in the alluvium via pilot points. Depth dependence applied to Triassic/Permian strata, although local heterogeneity not represented in the model.
2.6.3	Interpolation algorithm(s) and uncertainty of gridded data?	No	Uncertainty of gridded data (e.g. model layer surfaces) not discussed.
2.7	Have consistent data units and geometric datum been used?	Yes	ML/day and ML/year consistently used where appropriate.
2.8	Is there a clear description of the conceptual model?	Yes	Presented in the SLR, 2023 report.
2.8.1	Is there a graphical representation of the conceptual model?	Yes	Presented in the SLR, 2023 report as a hydrostratigraphical cross section with inferred groundwater levels. Groundwater flow and some hydrogeological stresses presented. 3D schematics also presented.
2.8.2	Is the conceptual model based on all available, relevant data?	Yes	Cross section based on geological information from drill-logs and geological models. Key processes are presented. Post mining conditions are presented as a second conceptual model.
2.9	Is the conceptual model consistent with the model objectives and target model confidence level classification?	Yes	The key system components have been integrated to target the project objectives, the objectives being effects of mine pit development on groundwater levels, groundwater users, alluvial aquifer levels and river and stream baseflows. Effects on GDE can be inferred from the section.

No.	Review Questions	Yes/No	Comment
2.9.1	Are the relevant processes identified?	Yes	The main components of the hydrology, hydraulics of the system have been included and the transient stresses from mine pit development. A separate conceptual model in numerical modelling report presents post mining conditions. Key interactions are identified.
2.9.2	Is justification provided for omission or simplification of processes?	Yes	The report limitation section discusses these.
2.10	Have alternative conceptual models been investigated?	No	
3	**Design and construction**		
3.1	Is the design consistent with the conceptual model?	Yes	
3.2	Is the choice of numerical method and software appropriate (Table 4-2)?	Yes	Model is simulated using MODFLOW-USG Transport using the upstream weighting function - details on the nature of layer types and saturation functions not detailed in the report.
3.2.1	Are the numerical and discretisation methods appropriate?	Yes	The cell refinement (60x60m within mining areas and 100-150 around rivers/dams) are appropriate. and the total of number of cells ((291,564) provides adequate spatial resolution and relatively fast model run times.
3.2.2	Is the software reputable?	Yes	MODFLOW-USG is an accepted industry standard for this type of application.
3.2.3	Is the software included in the archive or are references to the software provided?	Yes	MODFLOW-USG is referenced.
3.3	Are the spatial domain and discretisation appropriate?	Yes	Spatial domain is well thought out. The minimum cell size of 60 x 60m replicates the yearly progression of the pit areas more than adequately.
3.3.1	1D/2D/3D.	Yes	3D
3.3.2	Lateral extent.	Yes	These are well thought out and defined in the model. The lateral boundaries are located far enough from area stressed during simulation and prediction.
3.3.3	Layer geometry?	Yes	The model uses nine major model layers (split in half along the Burton range fault), this is thought appropriate for this level of study.
3.3.4	Is the horizontal discretisation appropriate for the objectives, problem setting, conceptual model and target confidence level classification?	Yes	The model could be thought of as finely discretised.
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?	Yes	Aquitards have been included, in some case multiple layers have been used.
3.4	Are the temporal domain and discretisation appropriate?	Yes	
3.4.1	Steady state or transient.	Yes	Steady state and transient calibration undertaken, followed by transient predictive analysis.
3.4.2	Stress periods.	Yes	Stress period information well documented. Appropriately simulated using quarterly periods during historic phase, then annual for predictions. Stress periods grow to 100 years during the recovery phase.

No.	Review Questions	Yes/No	Comment
3.4.3	Time steps?	Yes	Adaptive Time-Stepping (ATS) option used and minimum time step size reported.
3.5	Are the boundary conditions plausible and sufficiently unrestrictive?	Yes	Cumulative drawdown extents do not significantly interference with general head boundaries. River and CHD cells used to represent water storage do not leak unrealistic flows into the groundwater system.
3.5.1	Is the implementation of boundary conditions consistent with the conceptual model?	Yes	GHBs used to simulate regional throughflow. DRN cells used to represent dewatering. CHD cells used to represent post mining pit lake.
3.5.2	Are the boundary conditions chosen to have a minimal impact on key model outcomes? How is this ascertained?	Yes	Predictive impacts from New Lenton do not interact with the GHBs, no flow boundaries and fixed heads. River cells representing drains act as baseflow only. River leakage represented via recharge package and assess via uncertainty analysis.
3.5.3	Is the calculation of diffuse recharge consistent with model objectives and confidence level?	Yes	Recharge has been zoned based on conceptual understanding of the system - e.g., highest recharge over flat permeable sediments and less recharge over steeper slope less permeable sediments. LUMPREM2 used to calculate base diffuse recharge rates.
3.5.4	Are lateral boundaries time-invariant?	Yes	General heads using fixed levels used for the entire simulation.
3.6	Are the initial conditions appropriate?	Yes	Steady state groundwater levels used that override any initial conditions. Pre-mining and steady state representative of 1998 conditions that appear representative of average climate conditions based on hydrographs and flows.
3.6.1	Are the initial heads based on interpolation or on groundwater modelling?	Yes	Steady state levels using ~1998 levels as steady state targets.
3.6.2	Is the effect of initial conditions on key model outcomes assessed?	Yes	Steady state heads reformulated for every parameter combination.
3.6.3	How is the initial concentration of solutes obtained (when relevant)?	N/A	
3.7	Is the numerical solution of the model adequate?	Yes	The numerical model solver settings produce percentage discrepancies well-under the AGMG guidelines (<1%) at 0.02% at any timestep in the simulation.
3.7.1	Solution method / solver.	Yes	Sparse Matrix Solver (SMS) used.
3.7.2	Convergence criteria.	Yes	Appropriate HICLOSE criteria (0.001 m) used.
3.7.3	Numerical precision.	No	No discussion on numerical precision in the report.
4	**Calibration and sensitivity**		
4.1	Are all available types of observations used for calibration?		
4.1.1	Groundwater head data.	Yes	Model history matched using all available head and drawdown measurements.
4.1.2	Flux observations.	Yes/No	Baseflow estimations were qualitatively checked against site information proximal to New Lenton. Flux to existing Burton workings not discussed as they are not available.
4.1.3	Other: environmental tracers, gradients, age, temperature, concentrations etc.	Yes	SLR, 2023 discusses isotopes, concentrations.

No.	Review Questions	Yes/No	Comment
4.2	Does the calibration methodology conform to best practice?	Yes	The reporting of the calibration process is well-documented. History matching was achieved using PESTPP-IES. Statistics, hydrographs, water budgets and adequately presented. Statistics regarding drawdown are not presented (although not essential).
4.2.1	Parameterisation.	Yes	Calibrated to groundwater heads and baseflow while maintaining other model parameters with pre-defined realistic bounds based on site data or literature.
4.2.2	Objective function.	Yes/No	Briefly discussed in the observation weighting scheme. The evolution of phi during the calibration process not discussed or presented.
4.2.3	Identifiability of parameters.	Yes	Identifiability results including tables and maps are thoroughly presented and discussed.
4.2.4	Which methodology is used for model calibration?	Yes	PESTPP-IES using an ensemble smoother approach.
4.3	Is a sensitivity of key model outcomes assessed against?	Yes/No	Not directly assessed in the report, however it is noted that the calibration process revealed the hydraulic parameters and recharge to be the most identifiable (or estimable) parameter.
4.3.1	Parameters.	Yes/No	As above.
4.3.2	Boundary conditions.	Yes/No	Recharge, EVT and river conductance identifiability discussed.
4.3.3	Initial conditions.	Yes	Prescribed groundwater levels at GHBs discussed. Rest of model solves steady state heads at stress period 1
4.3.4	Stresses.	Yes	Recharge, EVT and river conductance identifiability discussed.
4.4	Have the calibration results been adequately reported?	Yes	
4.4.1	Are there graphs showing modelled and observed hydrographs at an appropriate scale?	Yes	Ensemble and best calibrated results presented in report and Appendix.
4.4.2	Is it clear whether observed or assumed vertical head gradients have been replicated by the model?	Yes/No	The absence of nested observations makes vertical gradients difficult to verify. Heads discussed at length, although simulated vertical gradients not specifically discussed.
4.4.3	Are calibration statistics reported and illustrated in a reasonable manner?	Yes	Appropriate statistics, scatter, and histograms presented.
4.5	Are multiple methods of plotting calibration results used to highlight goodness or fit robustly? Is the model sufficiently calibrated?	Yes	Scatter, hydrographs and spatial residual maps presented. The model is sufficiently calibration given the scale of the model.
4.5.1	Spatially.	Yes	Piezometric surfaces and spatial residuals presented.
4.5.2	Temporally.	Yes	Transient calibration hydrographs presented.
4.6	Are the calibrated parameters plausible?	Yes	Within plausible bounds from available data.
4.7	Are the water volumes and fluxes in the water balance realistic?	Yes	Comparison was made to like hydrogeological examples in the regional project area. Water fluxes to alluvial aquifers and streams/ rivers are consistent with the conceptual model.
4.8	Has the model been verified?	No	No verification section in the report.
5	**Prediction**		

No.	Review Questions	Yes/No	Comment
5.1	Are the model predictions designed in a manner that meets the model objectives?	Yes	Progressive drain cells simulate open cut dewatering. Average climate conditions assumed. Predictions calculated as difference from Null and Approved model scenarios.
5.2	Is predictive uncertainty acknowledged and addressed?	Yes	Predictive uncertainty comprehensively addressed and acknowledged through demonstration and summarised in Table 4-1.
5.3	Are the assumed climatic stresses appropriate?	Yes	Diffuse recharge applied using LUMPREM2. Annual climatic stresses during predictive period which are appropriate given the climate. Future modelling exercises may consider applying quarterly stress periods and associated climate stresses (although this would add computational effort).
5.4	Is a null scenario defined?	Yes	All predictive runs are compared to a null scenario (where no mining occurs) and an approved scenario.
5.5	Are the scenarios defined in accordance with the model objectives and confidence level classification?	Yes	Predictive modelling undertaken throughout the entire mining operation and includes a sufficient recovery period.
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?	N/A	There are no pumping stresses simulated in this model.
5.5.2	Are well losses accounted for when estimating maximum pumping rates per well?	N/A	There are no pumping stresses simulated in this model.
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?	Yes	The predictive time period (~25 years) is similar to the transient calibration simulation duration (25 years).
5.5.4	Are the assumed stresses and timescale appropriate for the stated objectives?	Yes	As above. Recovery stress periods are adequate given pit lake estimates are derived from daily stress period models.
5.6	Do the prediction results meet the stated objectives?	Yes	The predictive results measure changes in groundwater level, flux to alluvial aquifers, changes to river baseflows and pit seepage - this meets the stated study objectives.
5.7	Are the components of the predicted mass balance realistic?	Yes	Inflows and discharge rates going into the model are realistic, and adequately reported and presented.
5.7.1	Are the pumping rates assigned in the input files equal to the modelled pumping rates?	N/A	There are no pumping stresses simulated in this model.
5.7.2	Does predicted seepage to or from a river exceed measured or expected river flow?	Yes	The river boundary volumes are realistic.
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)?	No	Fixed head cells representing water storage during predictions and historic phase do not significantly contribute to the groundwater flow budget (aside from a brief moment in 2023).
5.7.4	Is diffuse recharge from rainfall smaller than rainfall?	Yes	Recharge rates range from 0.2 - 4.4% of rainfall.
5.7.5	Are model storage changes dominated by anomalous head increases in isolated cells that receive recharge?	No	

No.	Review Questions	Yes/No	Comment
5.8	Has particle tracking been considered as an alternative to solute transport modelling?	Yes	Particle tracking used to simulate long-term movement of groundwater to the pit lakes.
6	**Uncertainty**		
6.1	Is some qualitative or quantitative measure of uncertainty associated with the prediction reported together with the prediction?	Yes	Calibration constrained Monte Carlo predictive analysis undertaken and presented.
6.2	Is the model with minimum prediction-error variance chosen for each prediction?	Yes/No	An ensemble of models are used in the predictive analysis that meet calibration criteria thresholds. Each model does not necessarily possess a combination of parameters that are statistically minimised through inversion.
6.3	Are the sources of uncertainty discussed?	Yes	Discussed in Table 4-1.
6.3.1	Measurement of uncertainty of observations and parameters.	No	Not specifically discussed. If not specifically provided measurement uncertainty assumed equal to the reciprocal of observation weights.
6.3.2	Structural or model uncertainty.	Yes	Discussed in Table 4-1.
6.4	Is the approach to estimation of uncertainty described and appropriate?	Yes	Parametric uncertainty explored using relatively wide, aquifer wide perturbations. Ensemble of models assessed remain calibrated. Pleasingly, uncertainty applied to the predictive and recovery phases.
6.5	Are there useful depictions of uncertainty?	Yes	Predictions are presented as estimates of likelihood as per the IESC uncertainty guidelines.
8	**Surface water-groundwater interaction**		
8.1	Is the conceptualisation of surface water-groundwater interaction in accordance with the model objectives?	Yes	The conceptual model determined that surface water-groundwater interactions in the New Lenton Project area were minimal. The model validates this assumption and predictions of baseflow are assessed appropriately for the scale of the model. Future iterations may consider simulation of quarterly stress periods, although potential baseflow intercepted would be expected to be very low (<7 ML/year).
8.2	Is the implementation of surface water-groundwater interaction appropriate?	Yes	Given the absence of significant baseflow in the project area the simplistic representation of surface water-groundwater interaction is appropriate.
8.3	Is the groundwater model coupled with a surface water model?	Yes	Although not explicitly coupled, recharge rates are prescribed by LUMPREM2, and the pit lake recovery stages are derived from AWBM simulations.
8.3.1	Is the adopted approach appropriate?	Yes	The simulation of rivers and streams through the application of boundary conditions is appropriate.
8.3.2	Have appropriate time steps and stress periods been adopted?	Yes	Quarterly stress periods used to replicate historic phase. Annual for predictive phase. Given baseflow impacts are not deemed significant these stress periods are appropriate for this stage of assessment.
8.3.3	Are the interface fluxes consistent between the groundwater and surface water models?	Yes/No	Quarterly/annual recharge averaged and applied as a factor from the daily LUMPREM2 model. Similarly, daily pit lake stage estimates are averaged from 2 yearly to 100 yearly representations. This approach is appropriate.

10. Conclusions

It is this reviewer's opinion that the documented groundwater modelling in the Modelling Technical Report meets professional industry standards. The authors should be commended with the rigor applied to the analysis and insights provided in the technical report. The conclusions from the review and checklists indicate that the model is *fit for purpose*, where the purpose is to produce a calibrated groundwater model up to 2023. The ensemble of predictive models are *fit for purpose* for use as a decision support tool to predict future impacts due to open cut mining at the Project.

Stated modelling objectives have been met, the model is reasonably well calibrated, and the likelihood of future predictions appear reasonable. Expected drawdown occurs in the alluvium proximal to the Project mining areas. In the most likely scenario, it is unlikely that groundwater receptors will be significantly affected by dewatering and the post mining void at the Project.

The ensemble of models are well calibrated and simulated baseflow is generally consistent with site observations. Uncertainty analysis is robust and likely overstates expected impacts determined at the higher percentiles. For this reason, the clients and stakeholders can have confidence that the model is a reasonable predictor of future mining impacts.

There are minor improvements that can be made at the onset of the Project to validate the expected impacts, and the range of potential impacts. Future studies should aim to sequentially validate the performance of the model ensemble as more data are available.

Neil Manewell

Neil Manewell
Director | Principal Groundwater Modeller
Manewell Groundwater

11. References

Barnett et al, (2012) Australian groundwater modelling guidelines, Waterlines report, National Water Commission, Canberra.

Doherty, J. and Moore C., (2021). Decision Support Modelling Viewed through the Lens of Model Complexity. A GMDSI Monograph. National Centre for Groundwater Research and Training, Flinders University, South Australia.

Panday, Sorab, Langevin, C.D., Niswonger, R.G., Ibaraki, Motomu, and Hughes, J.D., 2017, MODFLOW-USG version 1.4.00: An unstructured grid version of MODFLOW for simulating groundwater flow and tightly coupled processes using a control volume finite-difference formulation: U.S. Geological Survey Software Release, 27 October 2017, https://dx.doi.org/10.5066/F7R20ZFJ

Peeters LJM and Middlemis H (2023). Information Guidelines Explanatory Note: Uncertainty analysis for groundwater modelling, A report prepared for the Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development through the Department of Climate Change, Energy, the Environment and Water, Commonwealth of Australia, 2023'.

SLR (2020), Environment Authority Amendment Supporting Information Document, New Lenton Mine, SLR Ref: 620.12614, Version No: V 2.0.

SLR (2023), New Lenton Coal Project, PER Groundwater Assessment, 620.31367, Rev: 3.0, incorporating the New Lenton & Burton 2023 Groundwater Modelling Technical Report, 620.31367, Rev: 3.0

Tomlin, J., Phillipson, K., Lenahan, M., Durick, A., Manewell, N., Newton, D., Briody, M., and Dunlop, J. (2022) Review of current approaches to model residual mine voids for rehabilitation planning: Technical paper 1, 2, 3.

White, J., 2018. A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions. Environmental Modelling & Software, v. 109, p. 191-201.