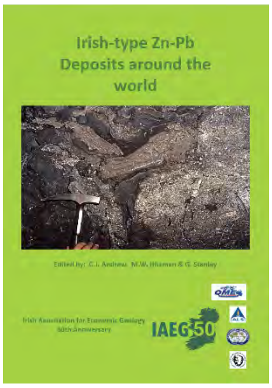

Irish Association for Economic Geology

(founded 1973)

Home Page: https://www.iaeg.ie

A machine learning approach to geological modelling of an Irish-type Zn-Pb deposit

Gary Buchanan[1], Caroline Burden[1], Michael Absolom[1] & Arpit Gole[2]

[1]Maptek Ltd, 3 Darnaway St., Edinburgh, EH3 6DW, United Kingdom

[2]Maptek Pty Ltd, 31 Flemington Street, Glenside, South Australia, 5065, Australia

Corresponding Author: Gary Buchanan gary.buchanan@maptek.co.uk

To cite this article: Buchanan, G., Burden, C., Asbsolom, M., & Gole, A. (2023) A machine learning approach to geological modelling of an Irish-type Zn-Pb deposit. *In:* Andrew, C.J., Hitzman, M.W. & Stanley, G. '*Irish-type Deposits around the world*', Irish Association for Economic Geology, Dublin. 385-392. DOI: https://doi.org/10.61153/UYMP7702

To link to this article: https://doi.org/10.61153/UYMP7702

A machine learning approach to geological modelling of an Irish-type Zn-Pb deposit

Gary Buchanan[1], Caroline Burden[1], Michael Absolom[1] & Arpit Gole[2]

[1]*Maptek Ltd, 3 Darnaway St., Edinburgh, EH3 6DW, United Kingdom*
[2]*Maptek Pty Ltd, 31 Flemington Street, Glenside, South Australia, 5065, Australia*

Abstract: The geological domains of the carbonate hosted zinc-lead deposit forming the Lisheen Mine in County Tipperary was, for many years, 3-dimensionally modelled via explicit methods in the best available computer software technology available at the time. The process required the geologists to load their drilling data onto a 3D CAD system where they made sectional interpretations of the domains, which were then stitched together to form a 3D volumetric model.

The wide availability of detailed geological data following the closure of the mine provides a unique opportunity for technology developers to compare the performance of next-generation software systems against those models meticulously generated by the expert team during the operation of the mine.

Previous studies have investigated the efficacy of a machine learning approach to modelling ore domains in comparison to the manually derived approach employed by the mining operation during its operating life. That study demonstrated a close relationship between the algorithmic and human derived models.

In this paper we will investigate the application of the machine learning approach to changing data, as is the case when new data such as new drilling logs become available. In addition, we investigate the role that the machine learning approach has within the resource modelling team, the impacts on job roles and how it can affect decision making processes.

Keywords: Machine Learning, domain modelling, artificial intelligence, Lisheen Mine

Introduction

A previous presentation demonstrated the efficacy in which a deep learning approach could be used to generate classified domain models directly from spatially referenced pre-coded data (Sullivan *et al*, 2019). This approach was subsequently demonstrated on an Irish-type Zn-Pb deposit, using logged borehole data and the manually derived wireframe models from the Lisheen Mine (Sullivan *et al*, 2020). The Lisheen Zn-Pb Mine, located in County Tipperary is a carbonate hosted Zn-Pb deposit, that for many years utilised the Maptek™ Vulcan™ software for a range of mine planning tasks including the geometric modelling of the ore domains. Following the closure of the operation all mine data has been made publicly available, making it an ideal case study to compare the best-in-class technology applied during the operation of the mine, with modern, data-driven approaches to ore domain modelling.

In the previous study by Sullivan *et.al.*, (2020), the new machine learning-based process, which involved very little human effort aside from preparing the drilling data, produced a model with a resultant volume within 4% of the original hand-drawn models. This result was achieved after around one hour of machine learning processing - a significant improvement in the time-cost compared to conventional methods which although not documented, are typically in the order of several days to weeks of manual effort.

New technologies, including Industry 4.0 artificial intelligence ("AI") software, have the potential to change the processes used by resource modelling teams to generate updated 3D interpretations of their geological data. It will therefore impact the roles of the individual geologists tasked with this work, alleviating demands on certain tasks and reprioritising human effort elsewhere.

These new AI developments coincide with the global shortage of earth science professionals across earth science industries. In resource modelling, these professionals would perform typical geological tasks such as digitising orebody wireframes, which can now be augmented by these new technologies. Correspondingly, resource teams are strained by ongoing demands from miners to shorten the mine planning horizon and therefore generate updated models in ever shortening times. Finally,

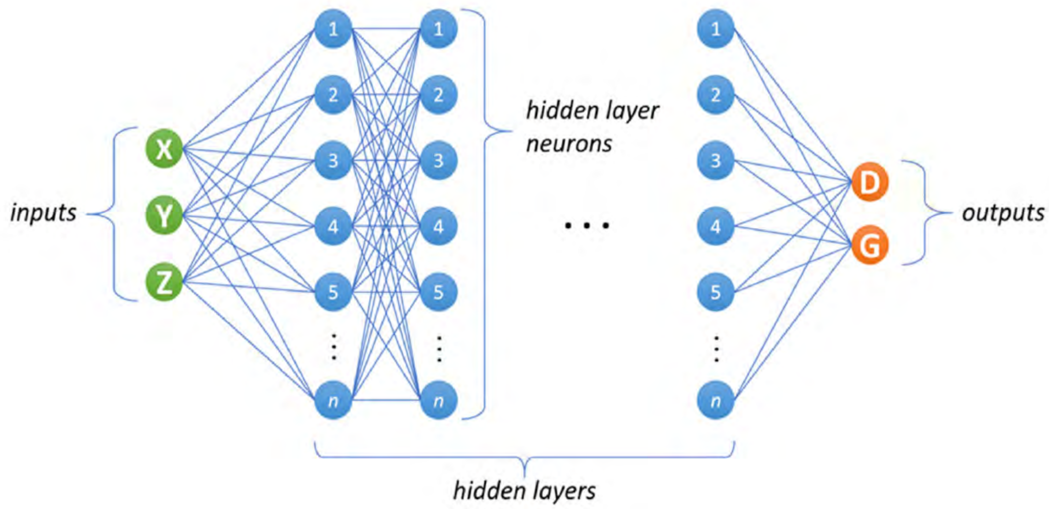

Figure 1: *Architecture of the neural network (after Wu & Zhou, 1993)*

economic and social pressure from key stakeholders to ensure mining operations are executed in the most socially and environmentally responsible manner, drive operators to seek new approaches that change how mining is carried out.

In this paper we will review the process of applying a machine learning application to domain modelling from spatially referenced data, with Maptek's second-generation machine learning application, Domain MCF. We will expand further on the consequences of such algorithmic approaches to domain modelling and the teams with responsibility for generating valid orebody models for mineral extraction.

A Data Driven Approach

Machine learning is a subcategory of artificial intelligence which uses algorithms to automatically learn insights and recognise patterns from data, applying that learning to make increasingly better decisions (*Web link 4).*

There are several merits of using AI for studying data:

- Faster and More Accurate Analysis: AI algorithms can quickly process large amounts of data, much faster than humans, and with greater accuracy. This means that insights and patterns can be identified much more efficiently and effectively.
- Unbiased Analysis: AI algorithms are not influenced by personal biases or preconceived notions, which means that they can analyse data objectively and provide more accurate results.
- Improved Decision Making: By analysing large datasets, AI can provide valuable insights that humans may miss, enabling better decision making. This can be particularly useful in complex decision-making processes, such as in healthcare, finance or, indeed, mining.
- Predictive Analytics: AI algorithms can use data to make predictions about future outcomes, such as customer behaviour or market trends. This can help organisations to prepare for the future and make informed decisions.
- Scalability: AI algorithms can handle large datasets with ease, making it possible to study and analyse data on a large scale. This can be particularly useful for businesses or organisations that need to analyse vast amounts of data.
- Customization: AI can be customised to meet specific needs, making it a valuable tool for a wide range of industries and applications. This means that organisations can tailor their AI systems to their unique needs, ensuring they get the most out of their data.

By way of example, the paragraphs above summarising the benefits of using AI for studying data was generated entirely by an Artificial Intelligence driven online chatbot, named ChatGPT. The author asked the Chatbot a simple question, "What are the merits for using AI for studying data" and was able to copy and paste the answer generated by the algorithm directly into this paper. With this example the reader can see the merits, but perhaps also identify potential consequences (both positive and negative) of technologies that can be applied to tasks that are traditionally performed by a human.

The technology applied to this paper is the Domain MCF software, developed by Maptek, which utilises a neural network, deep learning approach (itself a subset of Machine Learning) to generate classified/domained orebody models, including the estimation of multiple numeric variables and uncertainties — directly from spatially referenced pre-coded sample data. The algorithm generates classified/domained orebody models, including the estimation of multiple numeric variables and uncertainties — directly from spatially referenced pre-coded sample data (Sullivan *et al*, 2019).

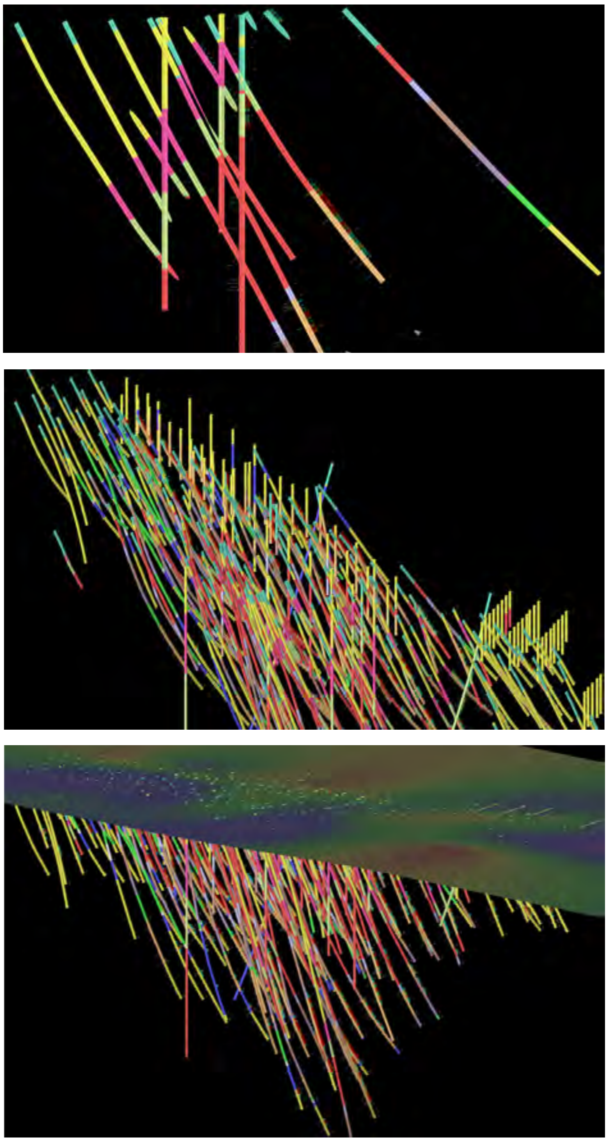

Figure 2: *Example of data requirements. Top and centre: Logged boreholes. Bottom: bounding surfaces such as topography, basement, fault zones, etc. In this case the colour coding is sourced from the domain code legend.*

Methodology

The use of artificial intelligence in orebody modelling is possible by using Domain MCF. An approach developed by Maptek uses raw drill hole data as an input to improve the orebody modelling process. It is designed to capture the spatial distribution in the mining data. Domain MCF builds artificial neural network ("ANN") models to understand spatial distribution of discrete domain values from a set of samples.

ANNs, such as those developed by Domain MCF, typically have an architecture, as shown in Figure 1 . The ANN consists of multiple layers of processing elements (PEs) also known as neurons (McCulloch & Pitts, 1943). There are three types of layers and corresponding PEs—input, hidden and output. PEs from one layer are connected to PEs in the next layer using weighted links known as synapses. PEs transfer the input signal to their outputs using an activation function that differs between the three types of layers. The number of input PEs is controlled by the way samples are presented to the ANN, i.e., the input space configuration. Researchers in the field of ANN applied to grade/resource estimation have used multiple configurations in defining the input space (Burnett, 1995; Clarici *et al*, 1991; Wu & Zhou, 1993; Kapagerides, 1999, 2005; Batchelor, 2019).

The number of hidden layers and PEs per hidden layer can be fixed or controlled by an optimisation process that will find the best configuration according to some performance criteria. Typically, the number of network inputs and outputs and the complexity of the required mapping between them will lead to a different number of hidden layers/PEs. The number of PEs in the output layer is controlled by the number of variables to be modelled.

Learning from examples is the main operation of any ANN. In general terms, learning means the ability of an ANN to improve its performance, defined with some measure, through an iterative process of adjusting its free parameters (weights, number of PEs, etc.). The adjustment of an ANN's free parameters is stimulated by a set of examples. presented to the network during the application of a set of well-defined rules for improving its performance and is called a learning algorithm.

In the case of Domain MCF, sample X, Y, Z coordinates are used as inputs and the sample domain (D) and, optionally, sample grade (G) are used as the required outputs. When both sample domain and grade are used as outputs, the synaptic weights between PEs of successive hidden layers will be affected by both distributions during training, thus leading to some dependency between the learned mappings for each variable. ANN development is data driven and thus largely dependent on the quantity, quality and accuracy of data. Generally, in the case of domain modelling for grade/resource estimation purposes, more samples will be required to produce a representative model in a more geologically complex scenario. A more complex ANN architecture with more PEs and hidden layers, allows a more complicated model to be generated (through development) but also requires more data. After development, the ANN can be used to get output values for any set of X, Y, Z coordinates presented at its input layer (e.g., block centroid coordinates), even outside of the sample coordinates range. However, outputs produced in areas outside of the range of examples introduced to the ANN during development should be treated with caution and examined carefully as to their validity, as in any case of extrapolation by more conventional methods.

Finally, to understand what's happening behind the scenes DomainMCF provides statistics of the process at defined steps during the workflow. These statistics range from the empirical analysis of the data/results on the fly, generating plots like swath plots, distribution curves about the data, confusion matrices and many more. These statistics are valuable to shed some light on the process and can be used in the various tasks.

Process

Sample data, typically generated from core logging (shown in Figure 2) is fed to the software which validates the data and allows the user to specify the data type for each column.

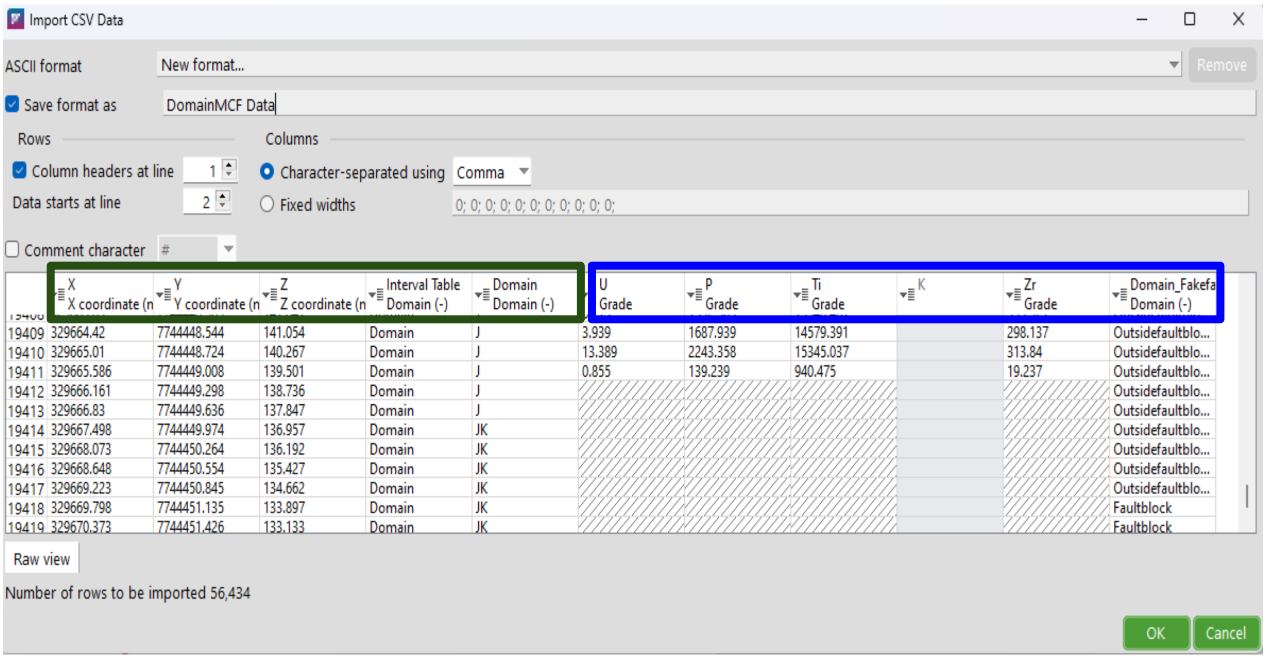

Figure 3: *Borehole and other geological data is converted to sample point format, with each point including the mandatory data (within the green box) for its geospatial location and at least one classification field, as well as other optional data (within the blue box) which can include numerical and categorical data.*

The data, as illustrated in Figure 3, is packaged into a table format with each sample point including the sample location, as well as one or more categorical labels (typically a geological domain code).

Provided the input data is valid, the risk of human error is largely negated. Even if the input data does contain an error, the impact of re-running the entire process is negligible due to the time taken to rebuild a model using this approach. The deep learning algorithm first generates a neural network (NN) from the pre-coded samples, then using the NN constrains the interpolation of numerical attributes. Building a block model is then simply a case of converting neurons to voxels and then aggregating voxels to generate the blocks we are familiar with (*Weblink 5*).

The next step in the process is to define the block model where the result will be output. The block model, which is essentially a 3D grid, is defined by its origin, extent (in X, Y & Z), orientation, dip and plunge, and block size. The block size can be adjusted for each individual axis (X,Y,Z) with a single, proportionate sub-blocking scheme (e.g., '½ of the parent, ¼ of the parent etc.).

Finally, surface limits, such as topography or geological features, can be applied to the final model.

The operator has the option to report the distance to the nearest sample for each block and set a sample search distance for each block.

Once the job is set up the operator selects the 'Preview Job' function, where the software estimates the processing time. This is important from a commercial perspective as this processing time is charged to the customer, on top of the time-cost associated with waiting for the job to process. The operator may wish to make changes to the job, such as reducing the number of samples or increasing the block size.

Once the job is uploaded all processing is handled by the cloud-based Maptek Compute Framework. The operator downloads the result in the format of a block model which has the domain variable populated with the predicted domain code.

Case Study: Lisheen Zn-Pb

Previous research has compared the domain model produced by the Domain MCF against the models generated by site geologists during active operation of the mine (Sullivan *et al*, 2020)

For this paper, the authors examined the impact of new data on the outcome of the domain prediction model, how new data affected the machine learning prediction model, and finally to discuss how this technology might impact the people and processes currently involved in resource modelling.

A selection of data obtained from the Lisheen operation was parcelled for processing by the second generation DomainMCF application, Domain MCF 2.0. The data was split into two groups, the first containing samples from 490 drill holes (80%) of the original data in the sampling area ('control'). The remaining 120 holes (20% of total holes in sampling area) were used to generate 5 groups of 30 holes, selected at random. ('random_1', 'random_2'...'random_5'). With that we have generated one sample group as a control and five additional sample groups containing random data changes.

The first experiment involved running the control data set five times to determine the repeatability we can expect from the

Figure 4: *Resource model update process exploiting machine learning and simulation to determine uncertainty and drill planning software to optimise return on investment for subsequent drilling campaigns.*

process using identical data. The nature of the process involves a degree of randomness, and the merits of reporting volumes and uncertainties has been covered in previous research (Sullivan *et al*, 2020).

Then, a further set of five experimental runs was performed using each of the random datasets in combination with the control data set. This provides some insight into the changes to the model when new data is introduced.

In the real world, this process could fit into a resource modelling update workflow as shown in Figure 4, which the machine learning process for spatial domain prediction informs the resource modelling team of predicted domain boundaries and, when combined with human geological expertise and other statistical approaches such as simulation, can inform the team for the next round of infill drilling.

The generated model can be further interrogated through other geostatistical analysis and visually inspected in appropriate 3D modelling software, as shown below.

Considerations and impacts for the Application of Artificial Intelligence

When new technologies are applied to existing workflows or challenges, there are potential impacts to processes, people and outcomes. Applying an algorithm for predicting domain distribution and contacts is no different. It can have profound effects on the role of the individuals who are typically responsible for building orebody models and opens a question around what role geoscientists will play in the future of resource discovery and modelling, as algorithms take a more prominent role in the construction of orebody models.

Web link 6 a CDC blog on the application of artificial intelligence by the Centre for Disease Control highlighted a number of issues. Some of these are considered below when applying technologies such as Domain MCF:

> *Effective: Ensure AI is the right tool to address the problem/concern. Technology should be used to improve productivity or working conditions and should not be used haphazardly.*

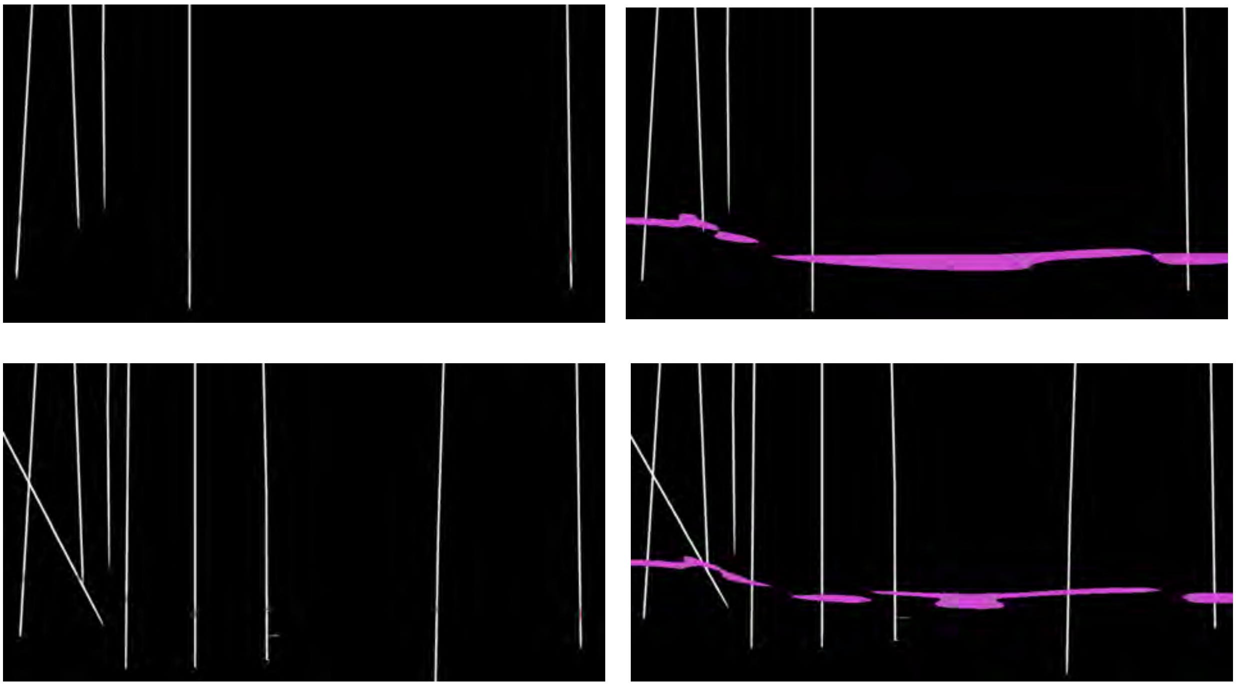

Figure 5: *A visual inspection of the predicted block model against the drilling data provided to the machine learning engine. The two top images show the control data and predicted model. The bottom two images show the predicted model based on the control data including the additional drill holes (bottom left).*

This is particularly important when it comes to applying such technologies such as Domain MCF to orebody knowledge. The task assigned to the software within the overall process of 3D modelling domains is to predict domain boundaries and populate the blocks within the boundaries with the predicted domain value. This task has been selected due to the ability of machine learning algorithms to accurately predict spatial patterns and the time-consuming nature of the human-driven alternative.

> *Explainable: Logic of, and decisions produced by AI should be communicated to stakeholders in a concise and useful manner. This is essential for mitigating risk and assessing impact of unintended, and potentially harmful, consequences.*

The way in which neural networks reach conclusions has long been considered a mysterious black box—that is, a network could not also provide an explanation of how it arrived at the conclusion it did (*Web link 7*).

In the case of Domain MCF, the algorithm is applied complementary to one or more qualified geoscience professionals alongside a robust process for reviewing and accepting the generated model.

> *Accountable: Organizations and individuals should be accountable for the outcomes of the AI systems they develop and implement.*

As with the above, it remains critical that the organisation maintains the systems and expertise required to provide accountability for the results produced by the algorithm.

In considering these key factors related to the application of AI technologies to geoscientific problems, it is clear that the role of the geoscientist remains critical to the task of orebody modelling, with or without AI applications.

The Role of Machine Learning in Resource Modelling

In summary, we recognise that the implementation of machine learning technologies to resource modelling brings a number of changes, challenges and opportunities. The mining industry has evolved extensively throughout history and new technologies associated with the Industry 4.0 revolution. A small survey of mining industry professionals involved in designing resource modelling workflows indicated the belief that AI technologies will have an impact on resource modelling jobs and workflows (Figure 6). A useful future exercise would be to expand this study to a much wider audience.

In general, the survey respondents agreed that AI technologies will play a complementary role alongside existing earth science experts and processes.

For this technology to find an acceptable place in resource modelling there remains the need for robust oversight by appropriately qualified geoscience professionals. Unlike earlier AI models, Deep Neural Networks like those utilised by Domain MCF, with their numerous layers and nodes, can be difficult to interpret (Clarici *et al.,* 1993). Geoscience professionals would remain responsible for data collection and quality, and the interpretation of the resultant model(s) followed by its

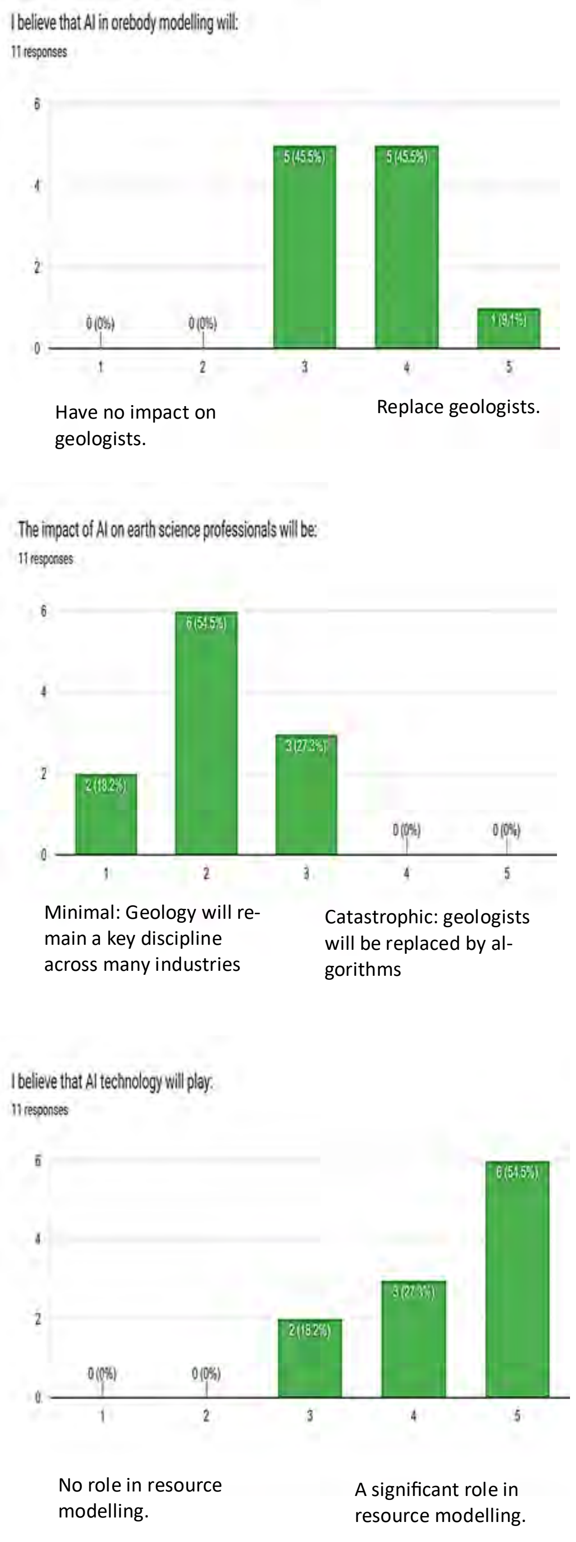

***Figure 6:** Forecast impacts of AI technologies on geologists and resource geology processes.*

acceptance or rejection. They will be challenged in new ways, particularly when it comes to managing and preparing data for analysis and explaining results.

On a macro level, the jobs landscape for earth science professionals may change, as time consuming tasks such as orebody modelling become increasingly automated. There are widely documented shortages of geoscience professionals across the extractive industries. Geology and geoscience related degrees are on the UK Skilled Worker Shortage Occupation list (updated Feb 2022) with geoscience related disciplines facing skill shortages in other advanced economies reliant on the minerals sector, such as Australia (*Web link 7*) and North America (Castelvecchi, 2016). It is perhaps fortuitously timed for a new technology to enter the market and release the geoscience professional from some of the most burdensome tasks.

Finally, new opportunities arise from the introduction of AI technologies such as the ability to generate multiple scenarios within any resource update cycle to provide an understanding of uncertainty of the model. The ability to generate updated resource models quickly, within the mine planning cycle, to provide downstream planners with up-to-date models with which to base their plans.

References

Sullivan, S.; Green, C.; Carter, D.; Sanderson, H.; Batchelor, J. (2019) Deep Learning—A New Paradigm for Orebody Modelling. InProceedings of the AusIMM Mining Geology Conference, Perth, Australia, 25–26 November 2019

Sullivan, S., Green, C., Carter, D., Sanderson, H., Batchelor, J. & Badenhorst, C. (2020) Uncertainty in domain modelling. *Maptek Forge*. September 2020.

Castelvecchi, D. (2016) Can we open the black box of AI? Nature News, v. 538 (7623)

Clarici, E.; Owen, D.; Durucan, S.; Ravencroft, P. (1993) Recoverable reserve estimation using a neural network. In Proceedings of the 24th International Symposium on the Application of Computers and Operations Research in the Minerals Industries (APCOM), Montreal, QC, Canada, 31 October–3 November1993; pp. 145–152.

Wu, X.; & Zhou, Y. (1993) Reserve Estimation Using Neural Network Techniques. *Comput. Geosci*, v. 19, 567–575.

Burnett, C.C.H. (1995) Application of Neural Networks to Mineral Reserve Estimation. Ph.D. Thesis, University of Nottingham, Nottingham, UK, 1995; 254p.

Kapageridis, I. (1999) Application of Artificial Neural Network Systems to Grade Estimation from Exploration Data. Ph.D. Thesis, University of Nottingham, Nottingham, UK, 1999; 260p.

Kapageridis, I. (2005) Input Space Configuration Effects in Neural Network Based Grade Estimation. *Comput. Geosci*. 2005, v. 31, 704–717.

Batchelor, J. (2019) Artificial Intelligence in mining and metallurgical operations, with case studies: Treating rocks and minerals as data objects using AI and immersive visualization. In Proceedings of the Digital Mines: Building Fully Autonomous Mines from Pit to Port, Perth, Australia, 10–12 April 2019.

McCulloch, W.S. & Pitts, W.H. (1943) A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysic*s, v. 5, 115-133.

Web Links

(3) https://ai.engineering.columbia.edu/ai-vs-machine-learning/

(4) www.openai.chat "What are the merits of using AI for studying data".

(5) https://blogs.cdc.gov/niosh-science-blog/2021/05/24/ai-future-of-work/

(6) https://spectrum.ieee.org/black-box-ai

(7) https://asiaminer.com/news/regional-news/10517-greater-asia-home-to-half-of-the-world-s-biggest-mining-companies.htmlhttps://www.americangeosciences.org/geoscience-currents/explanation-predicted-geoscience-workforce-shortage