Mantle to Mine: An integrated Machine Learning, Minerals Systems and Geomechanical Approach to Copper and Gold Exploration

John G McLellan
GMEX; James Cook University
Townsville, QLD
john@gmex.com.au

Paul J Pearson
Newmine Exploration
Lima, Peru
paul@newmine-exploration.com

SUMMARY

Many mineral deposits lie under thick accumulations of post-mineral cover, making exploration expensive and risky. Exploration targeting is thus often focussed mainly on various combinations of gravity and magnetic anomalies. In the case of IOCG style deposits the fluid pathway footprint is generally expressed as district-scale alteration and geochemical haloes. However, it is generally difficult to focus successful exploration due to the extensive geochemical "smoke". In addition, many seemingly attractive accumulations of iron oxide alteration are often barren or at best host low metal grades, and it is rarely obvious which iron oxide masses might host economic Cu and Au deposits.

Our solution for successfully identifying the "fire" in IOCG systems of the Gawler Craton of South Australia is to apply a range of methodologies that combine the strengths of both empirical and conceptual, process-based targeting. Our integrated "**Mantle to Mine**" approach minimizes interpretation bias and reduces the reliance on specific deposit models. Specifically, we have utilized a data-driven methodology (**machine learning**) that systematically tests for the relative importance of a large number of geological factors on known deposits, which is additionally guided by an integrated **mineral systems approach** that analyses the processes operative in the passage of ore fluids and metals to their eventual deposition sites. Finally, we have utilized an analogue-driven methodology to test geological processes (**geomechanical modelling**) that simulates rock failure and ore fluid flow at potential metal trap sites in the upper crust to further hone district-scale targeting. As an independent test of their accuracy, the results of the machine learning and geomechanical modelling were compared with the location of known mineral occurrences and geochemical anomalies. The hit rate was very high in the case of both methodologies and this test has given us confidence in the predictions made for less well-known zones.

Key words: IOCG Deposits, Machine Learning, Geomechanical Modelling, Fluid Pathways, Gawler Province Targeting

INTRODUCTION

The Archaean-Mid Proterozoic Gawler Craton of southern South Australia covers some 200,000 km2, and is demonstrably well endowed with world class copper, gold and uranium deposits. In the several decades since the discovery of Olympic Dam in 1975, intensive exploration has delivered further large iron oxide-associated copper-gold discoveries such as Prominent Hill and Carrapateena, in addition to smaller gold deposits in the Gawler Gold Province further to the west. Recognizing the probable existence of further undiscovered mineral deposits in the Gawler Craton, the South Australian Government recently initiated the 2020 Gawler Craton Challenge, which aimed to leverage the value of their large database to accelerate the discovery process in South Australia. Our response to the challenge is that we have integrated and analysed the majority of the available exploration data through an effective combination of data-driven and conceptually-driven approaches.

THE CHALLENGE OF EXPLORING COVERED AREAS

Exploring extensively covered areas for mineral deposits is becoming increasingly technically challenging and expensive. Regional targeting studies may be empirical or conceptual, or a combination of both (Woodall, 1994). Both approaches have their strengths and weaknesses. Purely empirical, data-driven targeting suffers in areas where data coverage is poor and tends to focus on analogues to known deposits. It tends to emphasize local understanding of ore controls and correlations that may not hold up at larger scales (McCuaig and Sherlock, 2017). On the other hand, purely conceptual targeting analyses the essential elements of mineral systems analysis (Wyborn et al. 1994) and seeks to identify combinations of proxies in the numerous datasets acquired (Hronsky and Groves, 2008). However, this approach is prone to bias, due to the imperfect understanding of mineral systems and the imperfect ability to interpret datasets (McCuaig and Sherlock, 2017).

The solution in a largely covered area like the Gawler Craton, where limited "training data" exist in the form of numerous mineral deposits, is to reduce targeting risk at an early stage and apply methodologies that combine the strengths of both empirical and more conceptual, process-based targeting. This integrated approach can minimize interpretation bias and reduce the reliance on specific deposit models that may or may not apply at larger scales. A data-driven methodology (**machine learning**) that systematically tests for the relative importance of a large number of potential controlling factors in the known deposits represents an important element of the targeting toolbox. In addition, a conceptual, integrated **mineral systems approach** that analyses the processes operative in the transport of ore fluids to their eventual deposition sites serves as an important guide in information extraction. Finally, an analogue-driven methodology (**geomechanical modelling**) that simulates ore fluid flow at potential mineral system sites, as a result of stress partitioning during deformation events, represents a critical tool in identifying key locations of fluid focusing in upper crustal rocks.

UNIQUE APPROACH

Our unique, geologically-grounded approach consists of the following workflow depicted in Figure 1:

- Definition of prospective mineral belts through multivariate analysis of factors controlling heat and fluid migration from the mantle to the mid crust (*data driven, machine learning phase 1*). Because of thick post mineral cover sequences, the overall limits of prospective mineral belts is commonly poorly understood.
- Assuming all processes essentially equal within given mineral belts, prediction of new mineral systems through multivariate analysis of their controlling physiochemical factors in the upper crust (*data driven, machine learning phase 2*);
- Prediction of zones of fossil fluid focussing and ore precipitation through analysis of palaeostress and flow fluid in the upper crust (*forward modelling, geomechanical modelling phase*);
- Identifying potential expressions of mineral systems in the near surface environment, including in hard rock, regolith and groundwaters (*data driven, machine learning phase 3 and expert knowledge-driven*).

The **Mineral Systems Approach** is a logical extension of the traditional petroleum systems approach and was first introduced into mineral science literature by Wyborn et al. (1994). The mineral systems approach has evolved over the past couple of decades, but the basic premises remain the same. Here, the original seven factors of Wyborn et al. (1994) have been further developed, and the idea of a critical time window has been introduced. The critical components leading to formation of a significant ore deposit in the upper crust can be summarized as follows:

- Tectonic events that trigger and define temporal windows for mineralising events;
- Geochemical and tectonic processes that produce a source of ore constituents, include metals, ligands and fluids;
- Energy to drive the mineral system;
- Tectonic processes that cause the activation of favourable crustal and lithospheric architecture that permitted the passage of melts and fluids;
- A focussing mechanism that concentrates (or throttles) the flow of fluids or magmas into depositional 'trap' sites;
- Strong physio-chemical gradients around the site of ore precipitation that cause ore precipitation;
- Post-mineral processes that cause exhumation, preservation and upgrading of mineralisation.

A key concept at the core of understanding mineral systems is that the passage of ore fluids through the crust leaves a much larger footprint than the actual scale of the ore deposits themselves. This is because the passage of potentially ore-forming fluids along the pathways between a deep source and a shallower sink may be marked by modification of the physio-chemical properties of the surrounding rocks which in turn are potentially imaged in the form of geophysical contrasts. Thus, when direct detection of orebodies may be difficult due to thick post-mineral cover and poor geophysical contrast the altered fluid pathway, although perhaps subtle and more widespread, could be the best expression of the mineral system to initially try and detect. A number of examples from the literature document the deeper crustal expression of these probable ore fluid pathways having apparently been directly imaged in geophysical datasets such as seismic and magnetotellurics (Skirrow et al., 2018; Heinson et al., 2018).
It is worth bearing in mind that the next orebody will almost definitely not look the same as those previously discovered, but will likely share common threads in terms of processes, controls and alteration in the surrounding rocks.

MACHINE LEARNING

Machine learning (ML) is a well-documented information technology that is a subset of Artificial Intelligence (AI). ML uses a suite of algorithms to simultaneously seek patterns in massive amounts of multivariate data. Once the algorithm "learns" the patterns or rules relating to these data through bivariate or multivariate statistical means, it can make predictions about other data.

After experimenting on the Gawler data and comparing the results, we settled on the **Random Forest** (**RF**) algorithm as it lends itself to both **classification** and **regression** problems in **supervised learning.** The Random Forest algorithm was found to be particularly efficient in more data poor environments where training data is sparse, commonly outperforming other methods such as Logistic Regression and Artificial Neural Networks (Rodriguez-Galiano et al., 2014; Carranza and Laborte, 2016).

Predictor layers derived from gridding the various geological and geophysical proxies are sampled on a systematic grid to generate a large, 2-dimensional data array with numerous features. Training sets derived from different areas/subsets are used to train the Random Forest model. Unknown test data is then run through the trained Random Forest model: the output from the ML algorithm is either a classification code (in the case of ML Phase 1 a simple binary MIN or NOMIN prediction is returned) or a regression predicted value (in the case of ML2, a "distance to mineralisation" value in metres is returned). The returned values are then imported into the QGIS environment where they are gridded, imaged and contoured. The resultant maps form the basis for validation against known deposits and further integration with the other methodologies in this study, including the geomechanical forward modelling.

GEOMECHANICS

Several methods of geomechanical modelling have been applied within the geosciences in an attempt to better understand geological processes, and in particular the mechanical processes responsible for the migration and focusing of metal laden fluids. The main modelling technique employed here to simulate the response of the structural architecture and stress transfer though the rock packages is a discontinuous method known as Discrete Element Modelling. This type of modelling treats the rock masses as elastic-plastic discrete blocks, and focuses on the deformation along faults and lithological boundaries between such blocks. Areas of low values minimum principal stress (σ_3) and areas requiring lowest values of fluid pressure increases for material failure may indicate dilation and potential sites of fluid focusing, and are of great interest in mineralised hydrothermal systems. The ability to predict areas that have mechanically failed or are more susceptible to failure, and hence focus fluids, can be advantageous in defining sites of increased prospectivity within any region (McLellan et al., 2016).
The data required to undertake this style of modelling can be conceptualised as two main components; 1. The structural and geological architecture of the area or system under investigation, and 2. The geological history of events and understanding of tectonic periods of instability and its relationship to mineralisation. Structural and geological architecture is developed from interpretation of previous mapping and geophysics. Here the geophysical datasets allowed us to carry out a large-scale interpretation and the recently completed Gawler Aeromagnetic Survey was key data to this interpretation. Due to the scale of this project only the main structural features were identified for inclusion in the model geometry.

The application of numerical modelling to mineral exploration problems typically follows a two-stage approach. The two stages are known as the model validation phase and the predictive modelling phase. Initial simulations are designed to test the conceptual geological/mineralisation model(s) for a known deposit in a region of interest: this is the model validation phase. These simulations aim to reproduce, or forward model, the observed distribution of mineralisation, strain or alteration within the known system in order to constrain the critical parameters controlling this distribution. A potential outcome of the model validation phase is the identification of previously unknown or unexpected anomalies, or the identification of preferential structural host orientations, within the ore system. This can yield either a better understanding of the grade distribution within the deposit, or identify new structural orientations or locations requiring a revised detection method (such as a different drill direction) in order to optimise detection and definition. The quantitative results of the model validation phase constrain the input parameters for generic or specific predictive models, which aim to predict the location of anomalies outside the area of known deposits. It is therefore critical that the results must be non-intuitive at some level.

TARGETING

A 2-stage approach using ML was applied, with ML1 used to map out broad permissible belts for copper, gold and silver-lead-zinc mineralisation in the whole of the Gawler Craton, and the second stage ML2 was to predict potential zones of ore precipitation for copper and gold mineralisation in the northern half of the Gawler Craton. The ML used training datasets to predict the regressed distance to ore, and the results are verified against known mineral deposits, with positive results. An example of this is the near perfect prediction of the location of Prominent Hill and other surrounding mineralisation using only the Olympic Dam area training data (Fig. 2) which is over 100 km to the southeast. These predictions provided great confidence in the anomalies and potential targets from the ML analysis.

Geomechanical analyses were then used to further hone and identify the upper crustal locations for rock failure and focused fluid flow. The partitioning of strain and localisation of dilation and fluids showed and incredible correlation with both the Au and Cu Provinces. In the IOCG Province many of the major deposit locations were predicted (Fig. 3), based on values of fluid pressures required for rock failure. Similarly, in the Au Province, the geomechanics predicted corridors of increased prospectivity and correlated exceptionally well with gold distributions for the surface and drillhole geochemical data.

This project defined forty-five target areas, each reflecting a confluence of the various machine learning and geomechanical modelling results. The 45 targets consist of 21 Cu targets for 9,765 km^2 (average 465 km^2 each) and 24 Au targets for 13,580 km^2 (average 565 km^2 each). These targets cover some 2.39% (Cu) and 3.32% (Au) of the total Gawler project area, representing a substantial reduction in the search area. The average area of the target areas defined is a maximum and they effectively define district-scale areas. A further reduction of the search area within each prospective box-shaped target area was achieved by considering in detail the outlines of the prospective zones defined by the individual machine learning and geomechanical modelling studies. The key to ascertaining 'solid' and reliable targeting was to investigate confluences between the two proven and complimentary techniques. Numerous district-scale target zones were defined, which when filtered with the geology (e.g., post mineral processes, geochemistry, depth of cover) to provide robust geologically focused targets (Fig. 4).

DISCUSSION & CONCLUSIONS

Economic Cu-bearing IOCG systems, although extremely valuable, are notoriously difficult to find. The footprint of these systems is moderate to large over the scale of kilometres to tens of kilometres, commonly expressed in the form of district-scale alteration and geochemical haloes. However, it is commonly difficult to focus successful exploration due to the extensive occurrences of geochemical "smoke". Another issue is that large areas of iron oxide alteration, usually as magnetite or hematite, may be geochemically barren or at best of low metal grade and it is rarely obvious which iron oxide accumulation might host economic Cu and Au.

Our dual methodology of mineral systems-guided, data-driven machine learning combined with geomechanical analogue modelling maximizes the probability of locating the "fire" within these systems, recognizing that the Cu and Au are usually precipitated late in the alteration paragenesis. Once we independently define the important controlling factors in the upper crust, localized fluid flow and metal precipitation can be forward modelled as mechanical processes that was focussed in only certain areas. This study has firmly established the relationships between 100 mappable criteria ranging in depths from the upper mantle to the upper crust and the spatial distribution of known mineral deposits/occurrences. By learning these relationships free of human bias, the machine learning has consistently delivered specific zones of high probability where unknown mineral systems may have developed through similar combinations of controlling factors and ore-forming processes.

Since we are targeting mineral systems formed originally in the upper crust under a known ambient regional stress regime, a powerful check and adjunct on the ML predictive process has been the Geomechanical Modelling. This kind of modelling is a proven methodology and has predicted where "exit" zones controlled by variations in rock rheology and geometry subject to a far field stress regime may have failed and focussed fluids and metals within the high probability districts. This strategy has proven its effectiveness in reducing the search area substantially by enabling us to focus on specific zones within the larger target areas. Combined with a detailed appraisal of the "explorability" of each high priority target area and comparing the results with gravity and magnetic datasets, we have further defined internal zones of interest within the broader target areas. Consequently, the potential search area has been even further reduced and this added value is one of the key outcomes of the study. Because of the systematic, process-driven approach taken here to targeting yet undiscovered mineral systems under post-mineral cover, systematic exploration conducted over the prioritized target areas defined here would be expected to have a high probability of success

ACKNOWLEDGMENTS

We would like to acknowledge the South Australian Government and the Unearthed team for the opportunity to take part in the Explore S.A. Gawler Challenge. We would also like to thank the judging panel for their hard work in scrutinising all of the entries in this competition, and off course we are thankful to be deemed winners in such a high calibre field of 2,227 entrants.

REFERENCES

Carranza, E.J.M., Laborte, A.G., 2016, Data-Driven Predictive Modeling of Mineral Prospectivity Using Random Forests: A Case Study in Catanduanes Island (Philippines): Natural Resources Research 25, 35–50, 2016.

Heinson, G., Didana, Y., Soeffky, P., Thiele, S., Wise, T., 2018, The crustal geophysical signature of a world-class magmatic mineral system: Scientific Reports (2018) 8:10608.

Hronsky, J. M. A., and Groves, D.I., 2008, Science of targeting: Definition, strategies, targeting and performance measurement: Australian Journal of Earth Science, 55, 3–12.

McCuaig, T.C. and Sherlock, R.L., 2017, Exploration Targeting: In "Proceedings of Exploration 17: Sixth Decennial International Conference on Mineral Exploration" edited by V. Tschirhart and M.D. Thomas, 2017, p. 75–82.

McLellan, J.G., Oliver, N.H.S., Brown, M., 2016. Advanced Understanding of Structural and Geochemical Controls on Mineralisation in the Eastern Mount Isa Inlier Using Innovative Techniques for Exploration. GSQ Industry Priorities Initiative, Report No. 99142. https://geoscience.data.qld.gov.au/report/cr099142.

Rodriguez-Galiano, V.F., Chica-Olmo, M., Chica-Rivas, M., , 2014: Predictive modelling of gold potential with the integration of multisource information based on random forest: a case study on the Rodalquilar area, Southern Spain: International Journal of Geographical Information Science, Volume 28(7), 1336-1354.

Skirrow, R.G., Van de Wielen, S.E.,Champion, D.C., Czarnota, K., Thiele, S.,2018, Lithospheric Architecture and Mantle Metasomatism Linked to Iron Oxide Cu-Au Ore Formation: Multidisciplinary Evidence from the Olympic Dam Region, South Australia: AGU 100 Advancing Earth and Space Science, Geochemistry, Geophysics, Geosystems, 19, 2673–2705. Research Article 10.1029/2018GC007561.

Wyborn, L.A.I., Heinrich. C.A., and Jaques, A.L., 1994, Australian Proterozoic mineral systems: Essential ingredients and mappable criteria: Australian Institute of Mining and Metallurgy Annual Conference Proceedings, 109–115

Figures

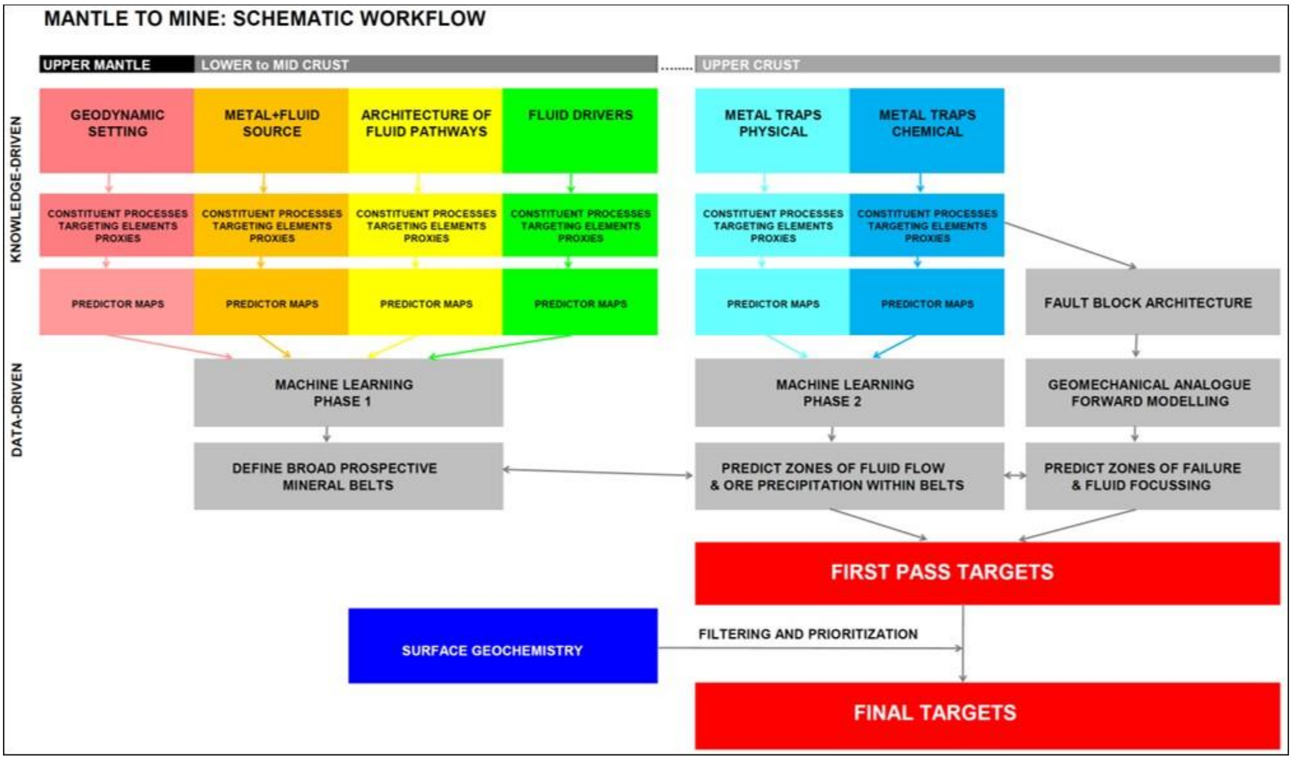

Figure 1. Mantle to Mine Workflow presenting the multiple stages of the knowledge driven and data driven approaches under the Minerals Systems Analysis umbrella. This workflow allows integration of expert geological knowledge in the systematic approach to identifying the 'key' geological parameters in the formation of mineral systems from the Upper Mantle through to the Upper Crust, where the Geomechanical Modelling is the key to identifying the ore localisation as a result of stress partitioning and fluid focusing.

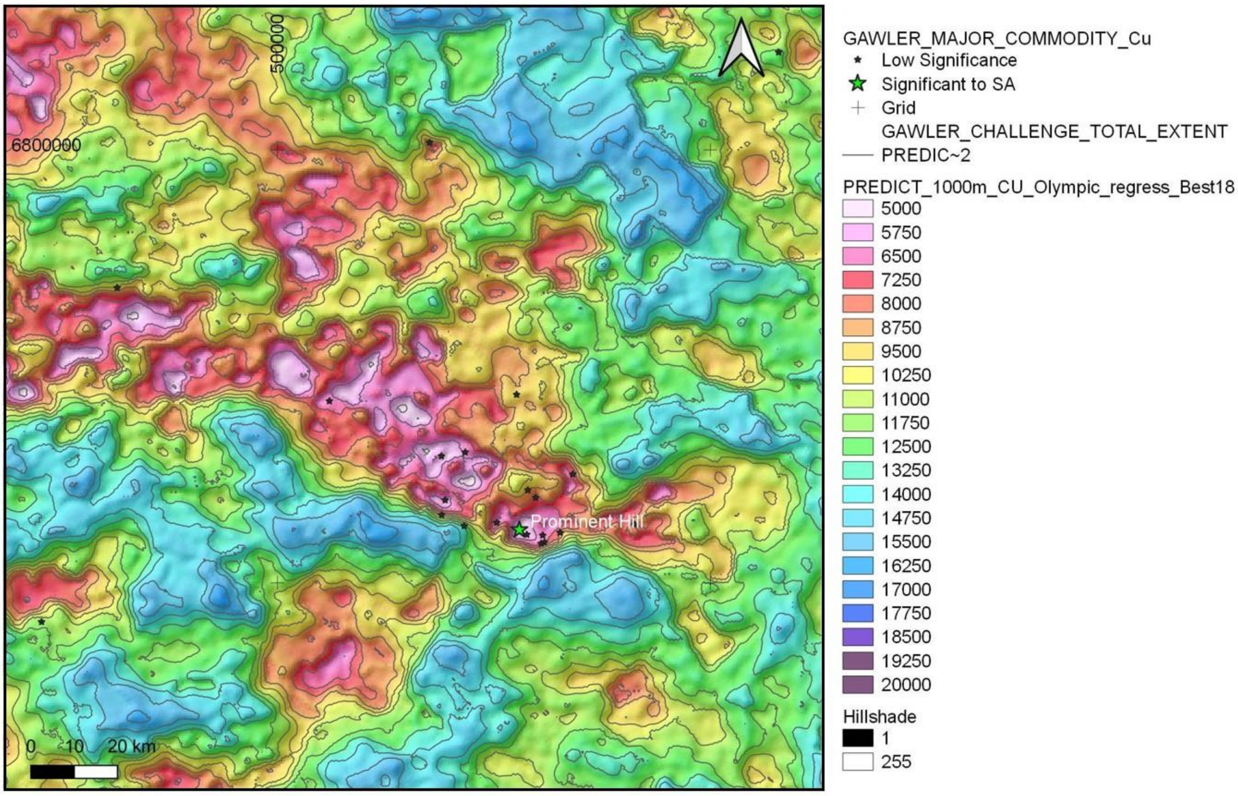

Figure 2. This dataset displays an example of Phase 2 Machine Learning results showing regressed distance to copper, where we see least values correlating exceptionally well and ultimately predicting the location of the Prominent Hill Deposit and other surrounding mineralisation using only the Olympic Dam area training data, which is over 100 km to the southeast.

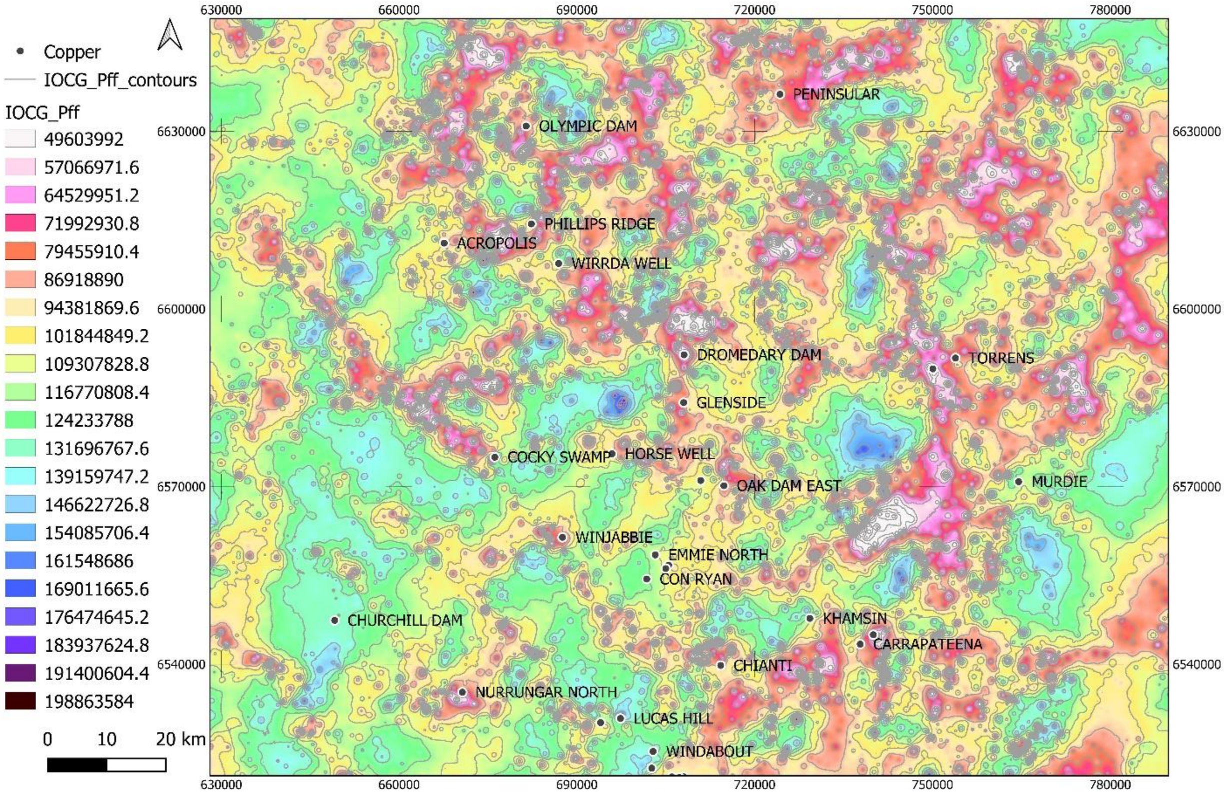

Figure 3. Results of the Geomechanical Modelling of the IOCG Province displaying calculated fluid pressure required for rock failure (Pff) in Pa. Geomechanical predictions of stress localisation during deformation is clearly highlighting areas requiring least additional fluid pressures required for rock failure (whites through to pinks and reds), with an excellent correlation of these lowest values areas with known locations of several IOCG deposits, and in addition highlighting several key structural locations for exploration.

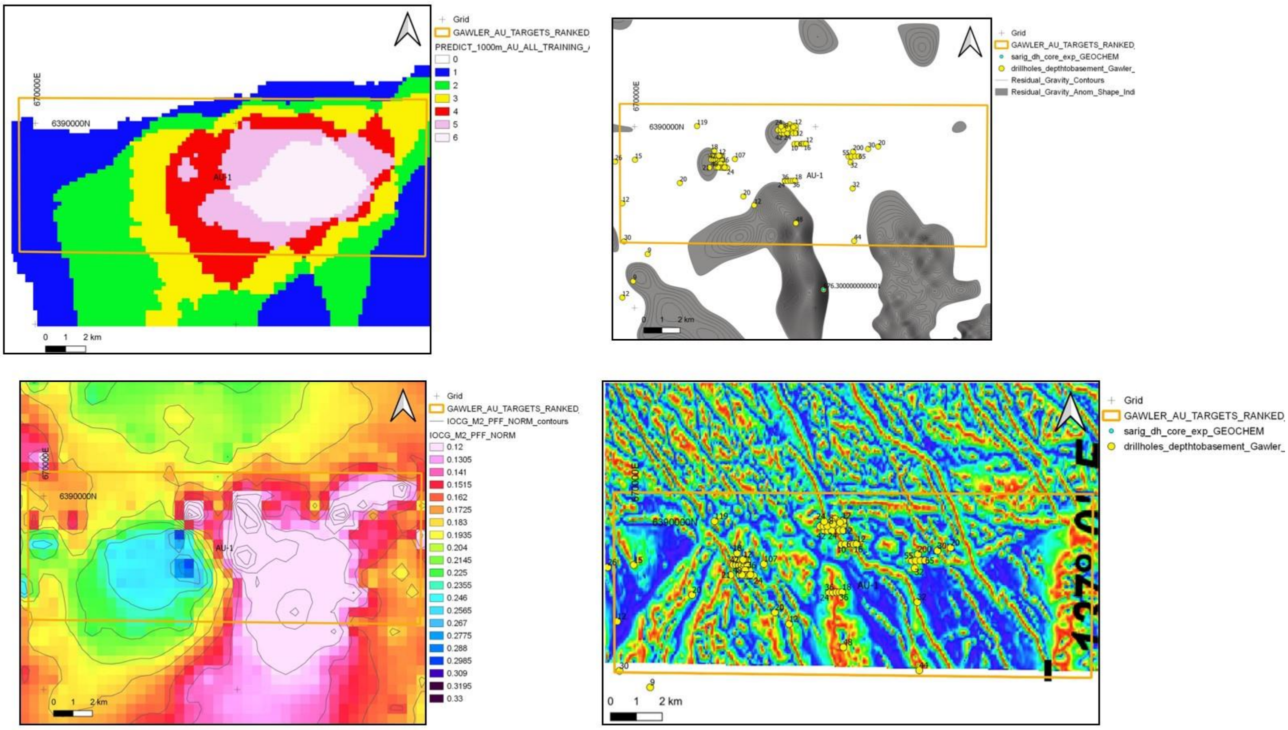

Figure 4. The figure in the top left displays the Machine Learning Phase 2 composite Hit-Scores which were derived from the regressed distance values for gold, these results were then compared with the 'Normalised' Geomechanical outputs which produced a 'failure probability map' and confluences between these two techniques were the primary input in the derived Target Maps. These maps were then filtered by several geological datasets or favourable characteristics e.g., geophysics (top right) and depth to basement criteria (bottom right) in addition to several other datasets to provide the most geologically robust targets possible.