1. Appendix

A. High-resolution model without resistivity

Here we present a model (and submit as a grid) which does not use the resistivity layers in the model (Fig. A14). The AusLAMP (2020) geophysical layer is lower resolution than other layers, so we include this as an example of the impact it makes on the overall model outcomes. All other parameters used in the Cu Model II are the same. We find the major targeting areas remain consistent (Fig. 11a)

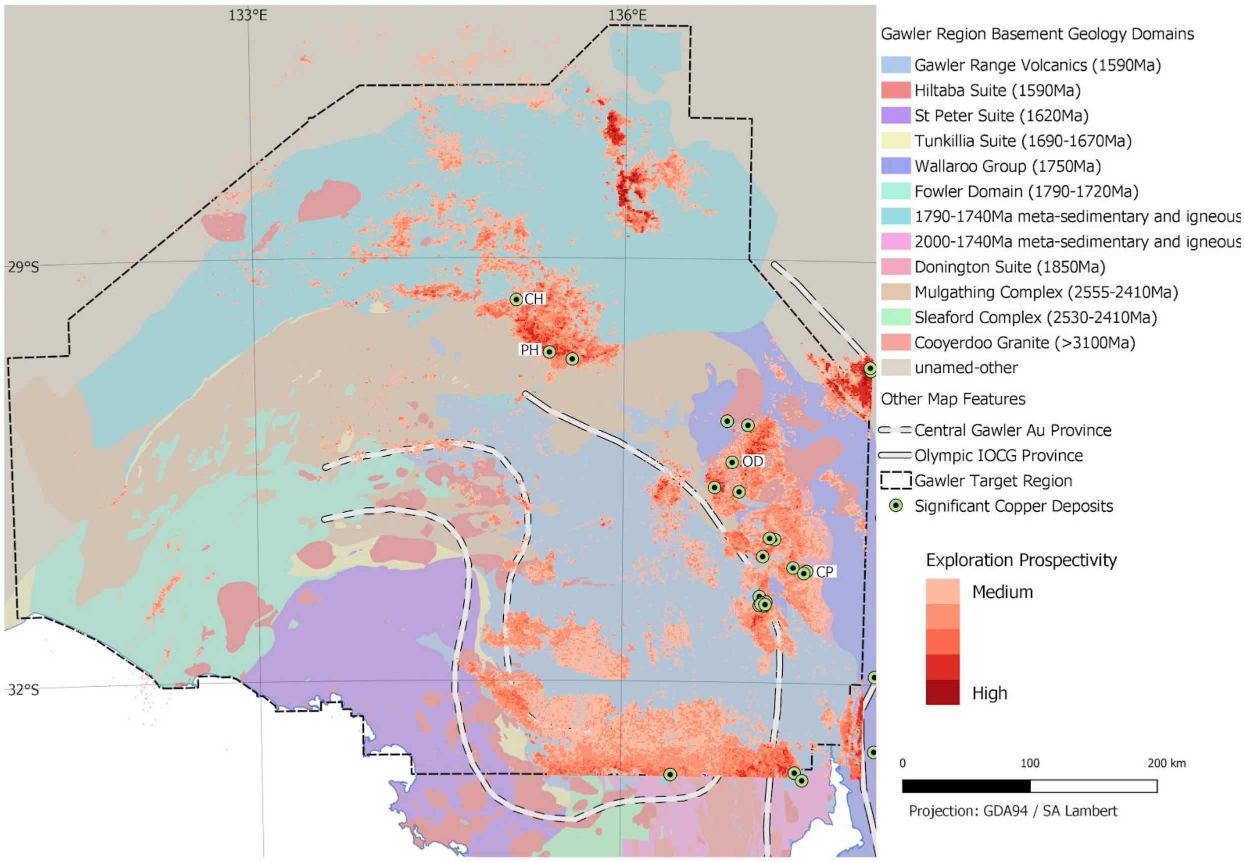

Figure A14: **Cu** Model II with resistivity input layers not included in model, medium (light red) to high (red) prospectivity probability classification plotted over regional geological domains (Fig. 1) and the two major metallogenic provinces after Reid (2019) as black and white dotted lines. Green dots denote locations of recorded “significant” **Cu** ore deposits. PH: Prominent Hill, OD: Olympic Dam, CH: Cairn Hill. CP: Carrapateena.

B. Other Commodities (Ag, Co, DIA, Fe, Mn, Ni, U, Zn)

Here we present results for Silver (Ag), Cobalt (Co), Diamond (DIA), Iron (Fe), Manganese (Mn), Nickel (Ni), Uranium (U), and Zinc (Zn) as subset from *SA Mines and Minerals* (2020) dataset (Table A4, Fig. A15-23). Generally when fewer deposits are available for training the classification variance increases, but when there are too many points used for training the model becomes too restrictive for useful greenfields exploration. None of these examples have been controlled for quality or grade of deposits used for training, and substantial improvements could be made for some of the commodities. All these models are submitted as grids. Figures A16-23 highlight the important features as predicted by the model. The most important layers correlated with the commodity are recommended for future data collection, and fur understanding mineral pathways and processes.

Commodity	Number of deposits	Classification Score	Standard Deviation
Ag	377	0.87	0.05
Co	92	0.83	0.08
DIA	78	0.64	0.19
Fe	251	0.85	0.06
Mn	115	0.92	0.10
Ni	65	0.66	0.17
U	238	0.76	0.12
Zn	265	0.76	0.10

Table A4: 10-fold cross validation classification scores for various commodities from the *SA Mines and Minerals* (2020) dataset. The number of deposits used for training (which is the corresponding number of non-deposits generated and used) is shown. Standard deviation gives an indication of the variance in the classification subsets.

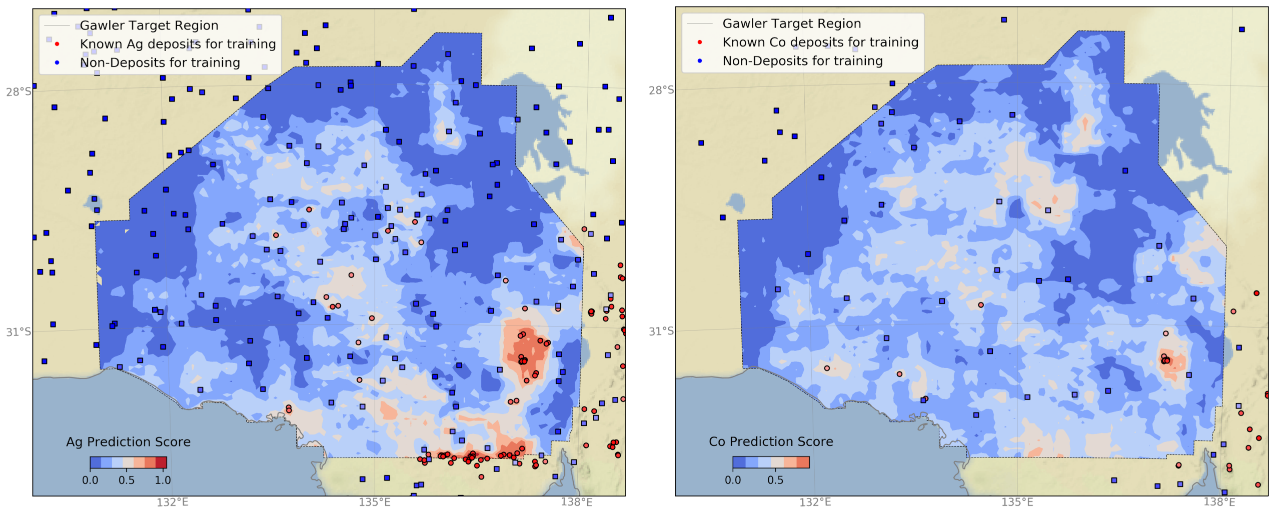

Figure A15. Predicted classification probabilities for additional commodities (as listed on the colorbar for each figure): Ag, Co, DIA, Fe, Mn, Ni, U, Zn. Training points (known deposit locations and random non-deposits) close to the Gawler Region are shown. Note: the predictions made in Gawler Target Region do not use the training points, so the correlation between points and predictions is a robust test for how well the model is working. Figure continued below...

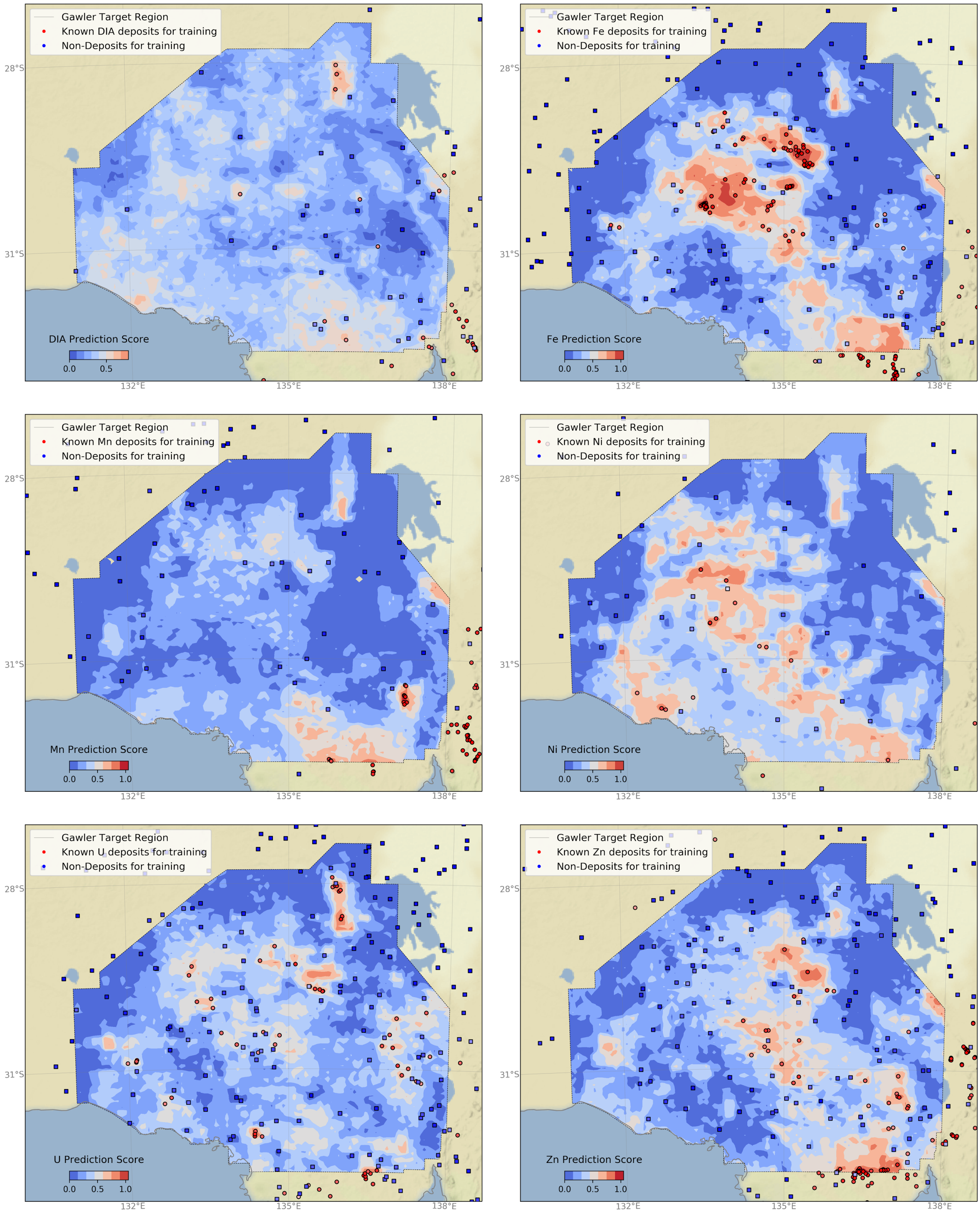

Figure A15. continued.

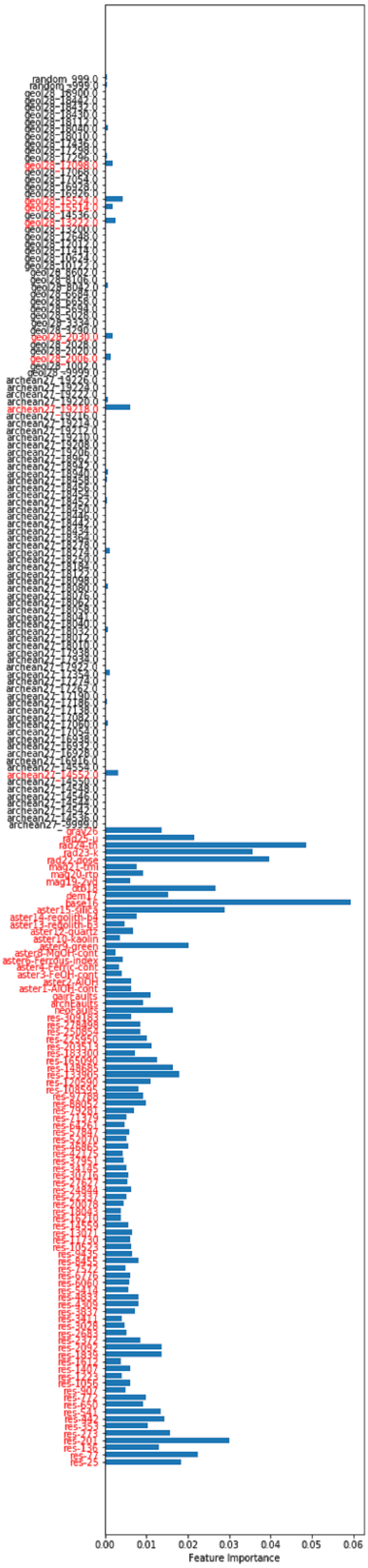

Figure A16. Silver (Ag) Feature Importance.

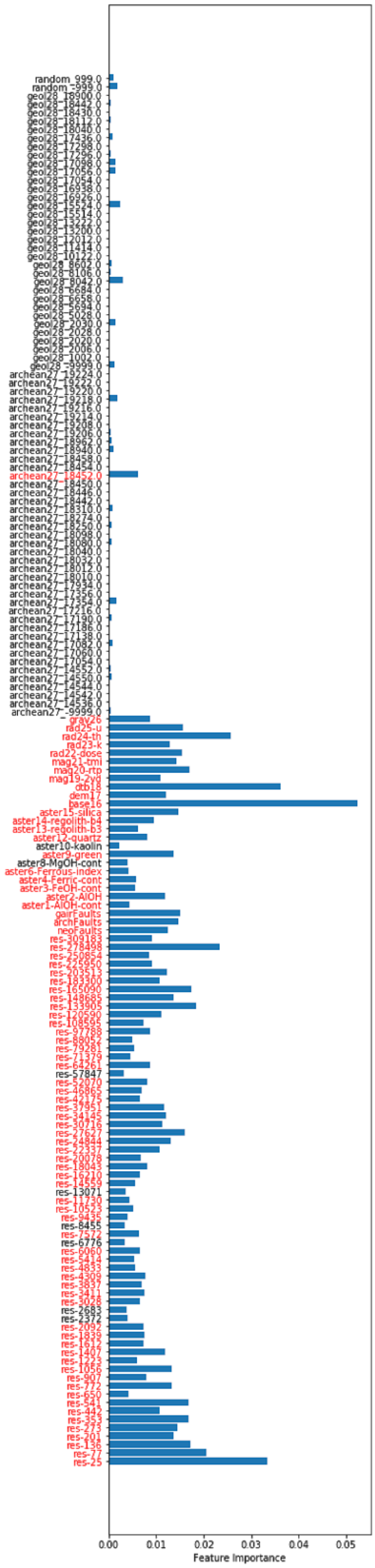

Figure A17. Cobalt (Co) Feature Importance.

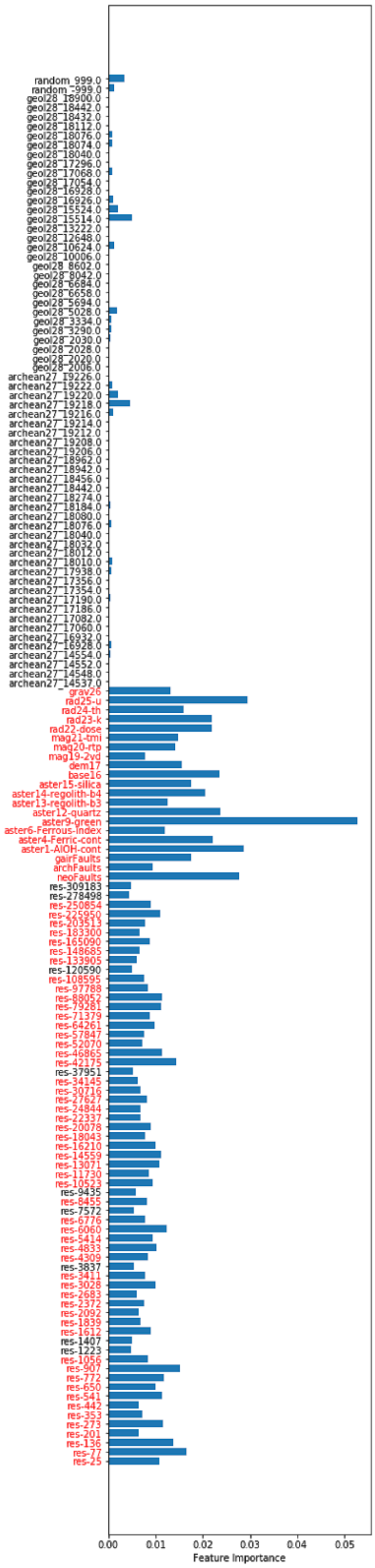

Figure A18. Diamond (DIA) Feature Importance.

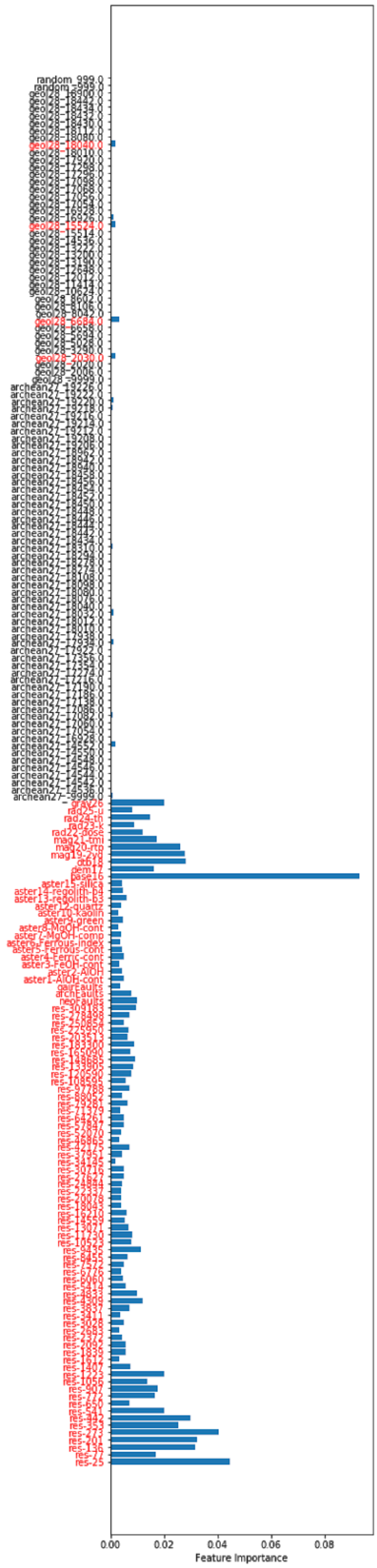

Figure A19. Iron (Fe) Feature Importance.

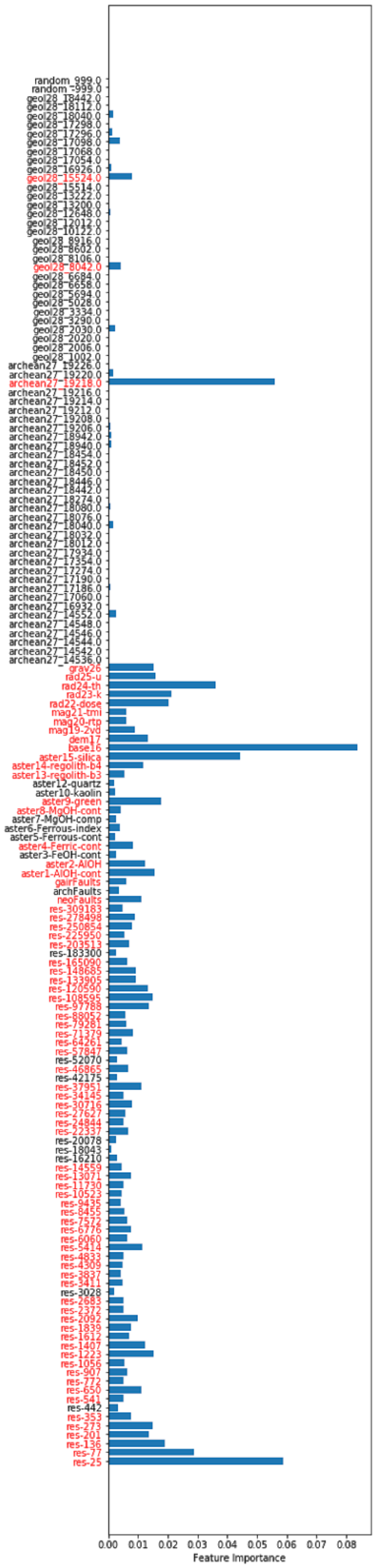

Figure A20. Manganese (Mn) Feature Importance.

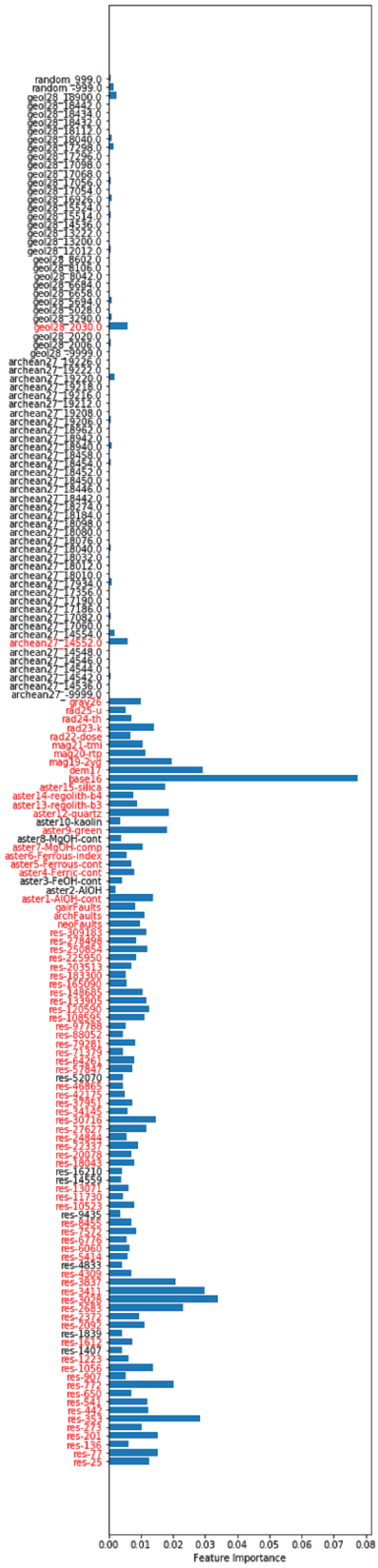

Figure A21. Nickel (Ni) Feature Importance.

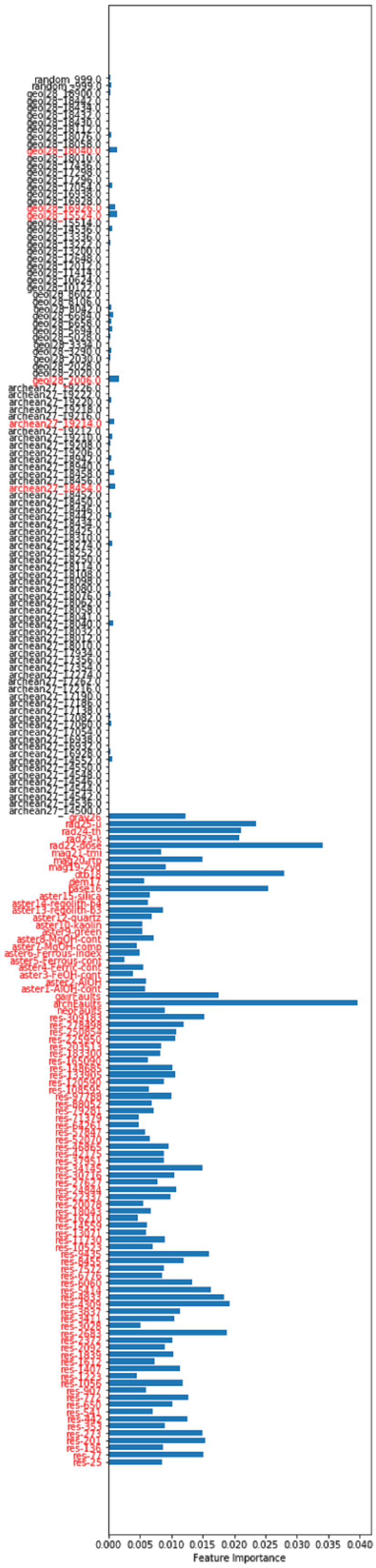

Figure A22. Uranium (U) Feature Importance.

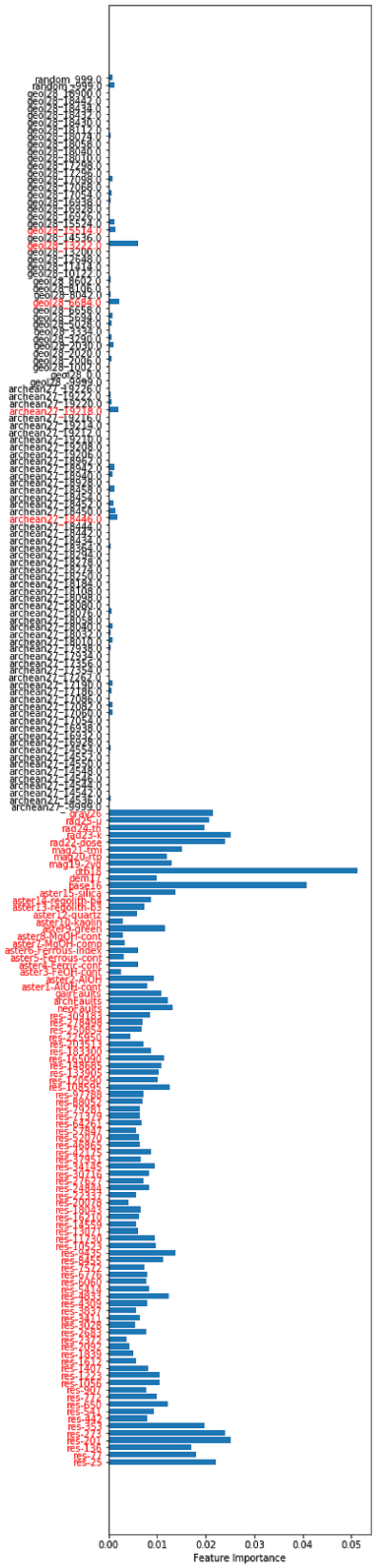

Figure A23. Zinc (Zn) Feature Importance.

C. Sensitivity Analysis

Here we show the results for a model sensitivity analysis we performed on the Cu Method II dataset. Overall the model behaves as expected. Important features absent from the model reduce the classification score, but no single geophysical parameter is significantly important on its own. Starting with all the numerical feature data, we proceed to systematically remove the most "important features" from the model (as determined by the model). As the "important features" are progressively removed the model 10-fold cross-validation classification score reduces to random performance (0.50). We then add the most important features back in systematically, and marvel as the score quickly increases (to a high of 0.90). These results not only highlight the powerful capabilities of the workflow but aides in the knowledge-discovery of understanding the complex interplay between geological features with their relationship to commodities, and subsequently which features we should target to learn more about a particular commodity.

Features	Score	Standard Deviation
All	0.86	0.07
Bottom 41	0.79	0.10
Bottom 20	0.74	0.11
Bottom 10	0.64	0.11
Bottom 4	0.57	0.12
Bottom 1	0.50	0.11
Top 1	0.64	0.10
Top 4	0.80	0.05
Top 10	0.89	0.05
Top 20	0.90	0.06
Top 40	0.90	0.05

Table A3. Sensitivity analysis for different parameter feature combinations in the workflow and their 10-fold cross-validation scores with standard deviation for Cu with Model II deposits. The number of features used begins with "All" 92 numerical data types. We then progressively remove the most important features before running the model again on the new subset of features. After bottoming-out with the least important feature (Bottom 1), we swap this for the most important feature (Top 1), and then progressively introduce the most important training features and continue to re-run the model with each subset of features. Figures of each iteration are shown below (Fig. A24).

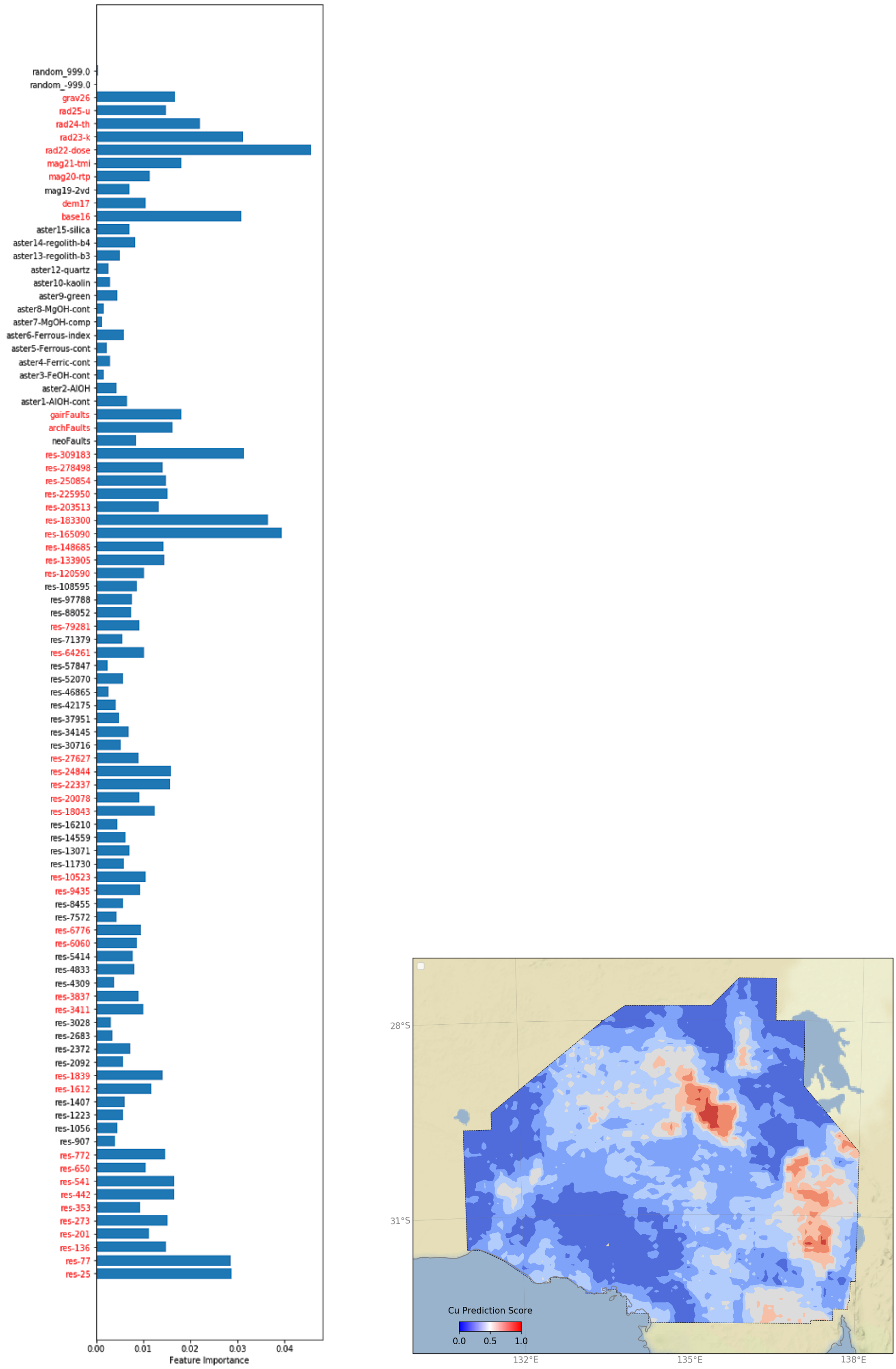

Figure A24. Sensitivity analysis of features used in CU Model II. Here we start with all the numerical features and remove features systematically before adding them back in. Left panel shows all features with corresponding feature importance score. Features highlighted in red are the most important features. We keep the “random” response variables as a benchmark for all tests. The right panel is the predicted probability classification map using this subset of features. Cont...

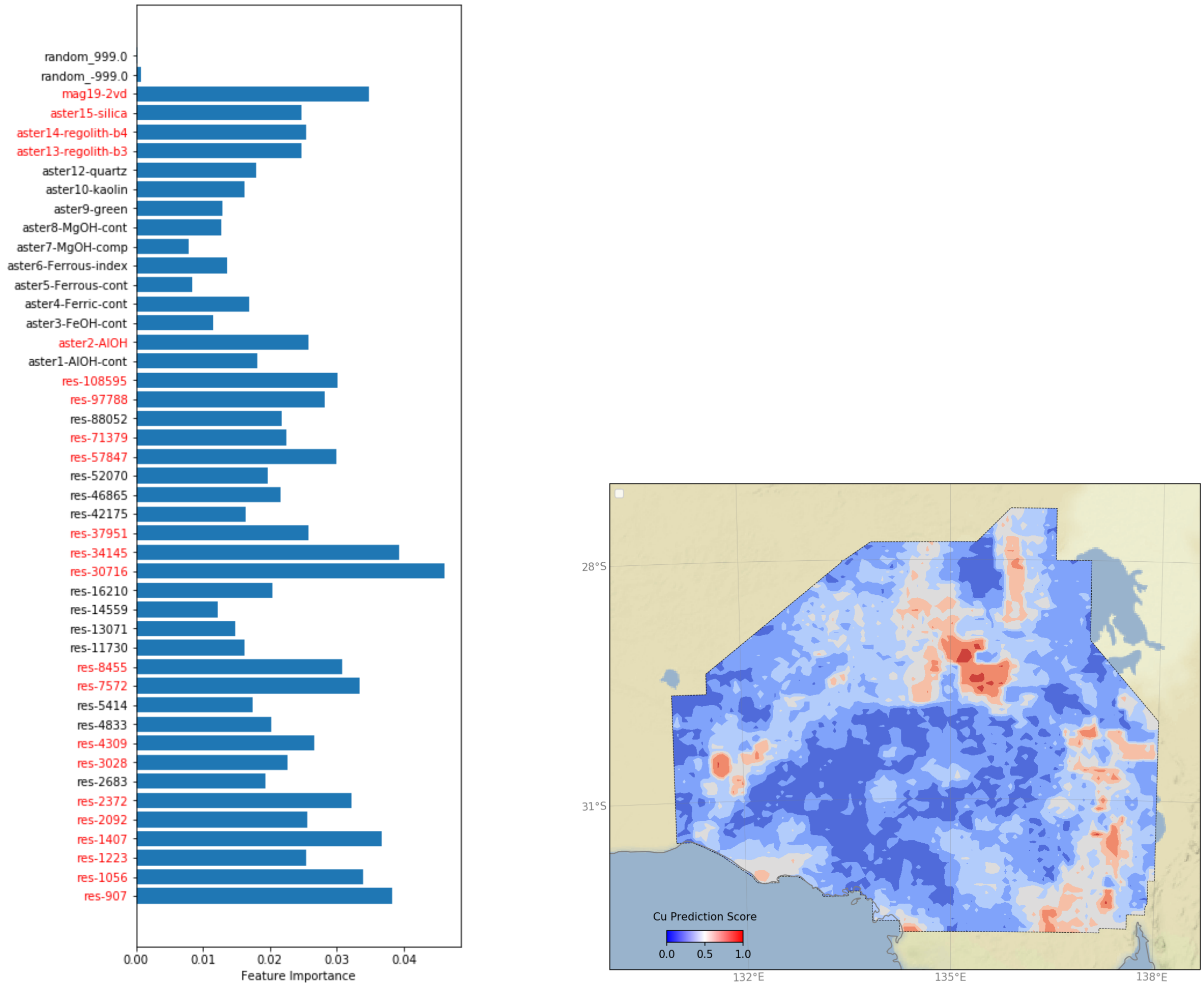

Figure A24 cont. Bottom 41 features. Cont...

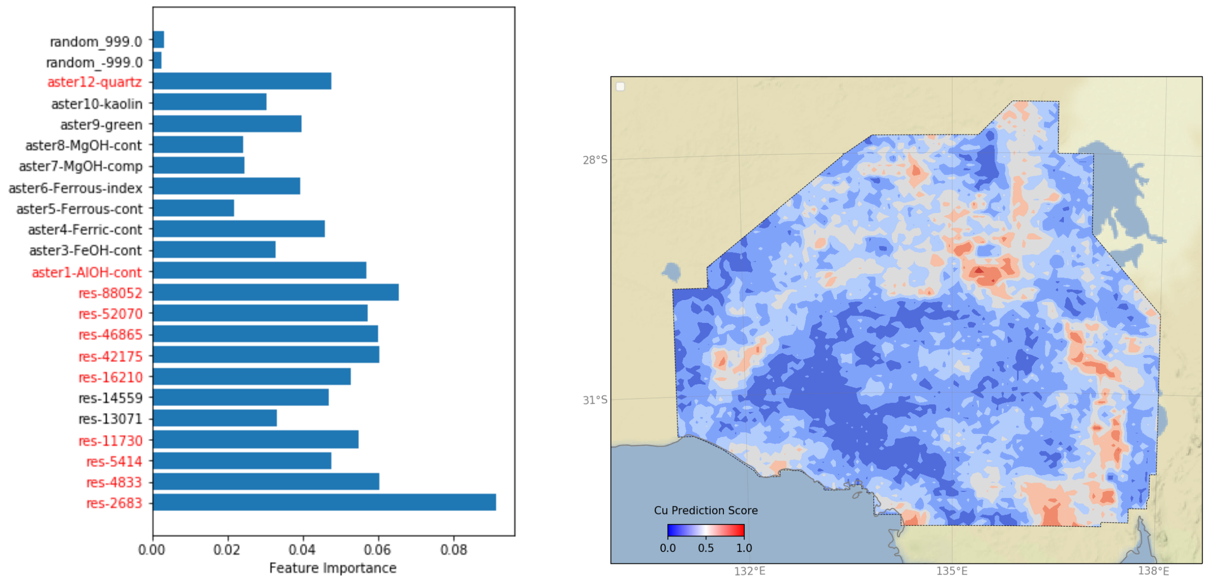

Figure A24 cont. Bottom 20 features. Cont...

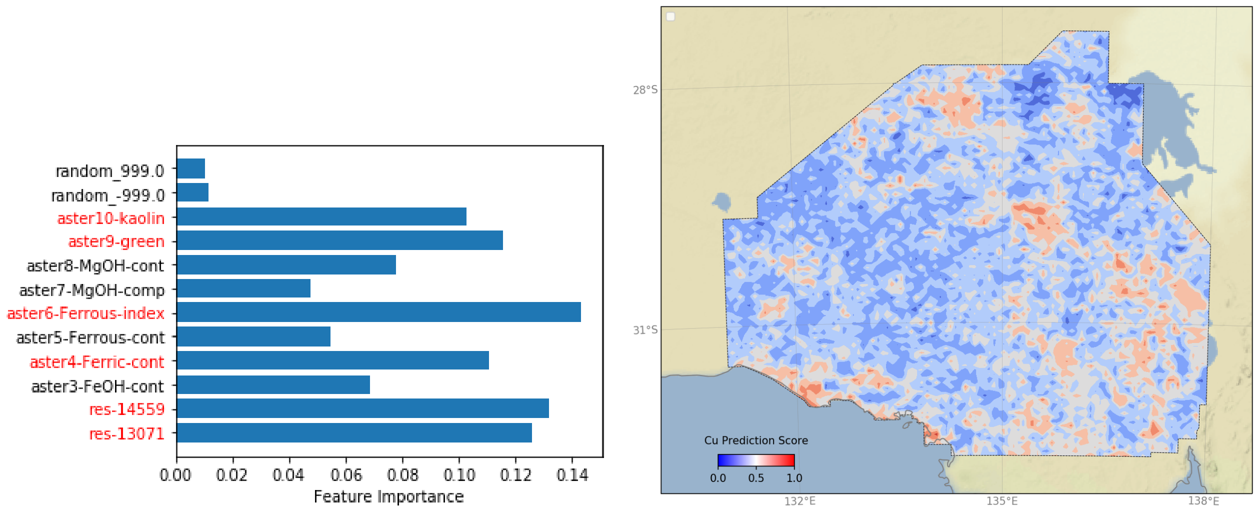

Figure A24 cont. Bottom 10 features. Cont...

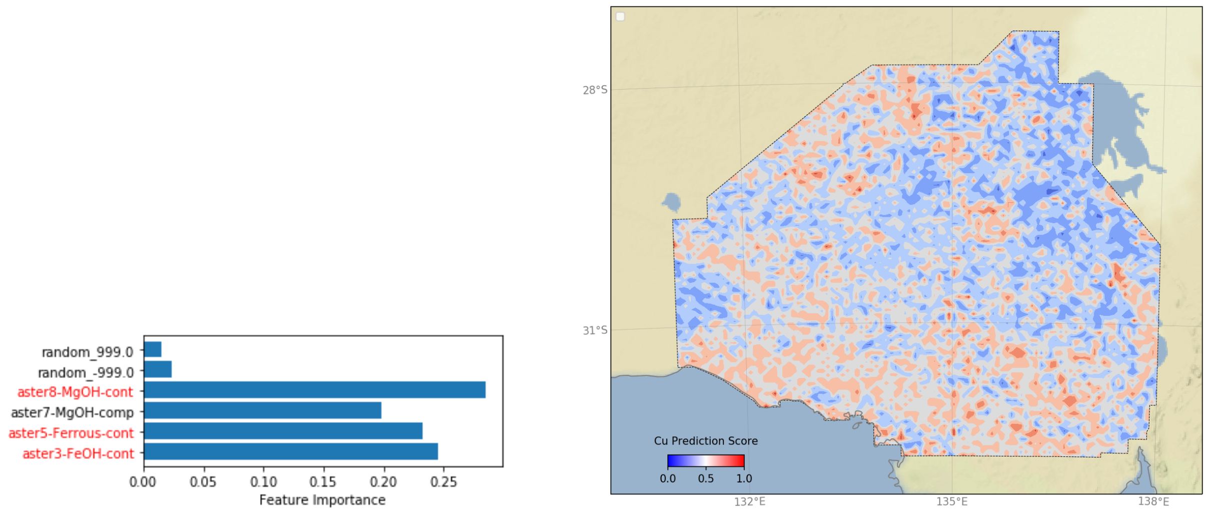

Figure A24 cont. Bottom 4 features. Cont…

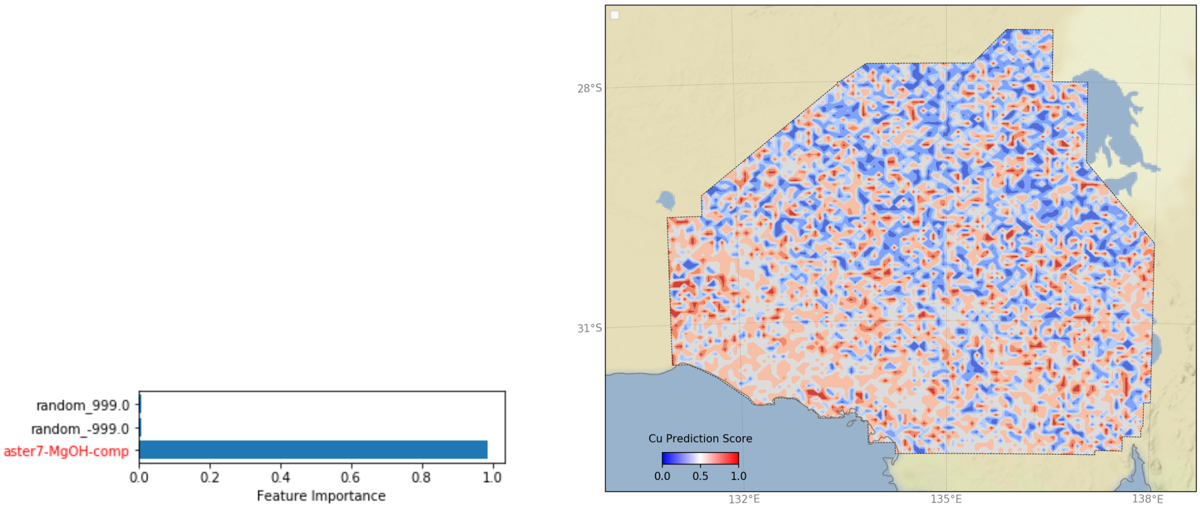

Figure A24 cont. Bottom 1 features. Cont...

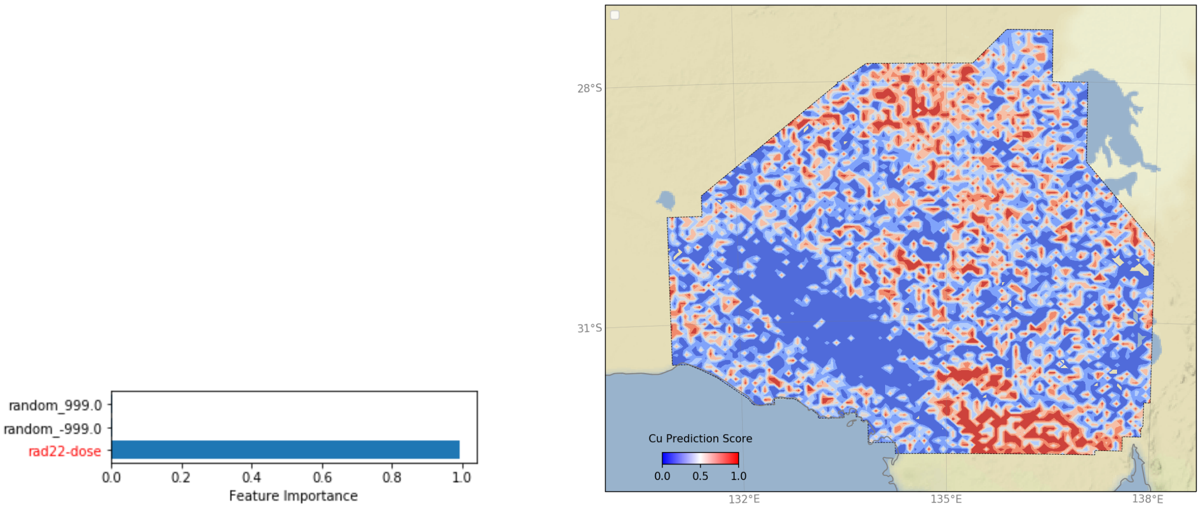

Figure A24 cont. Top 1 features. Cont...

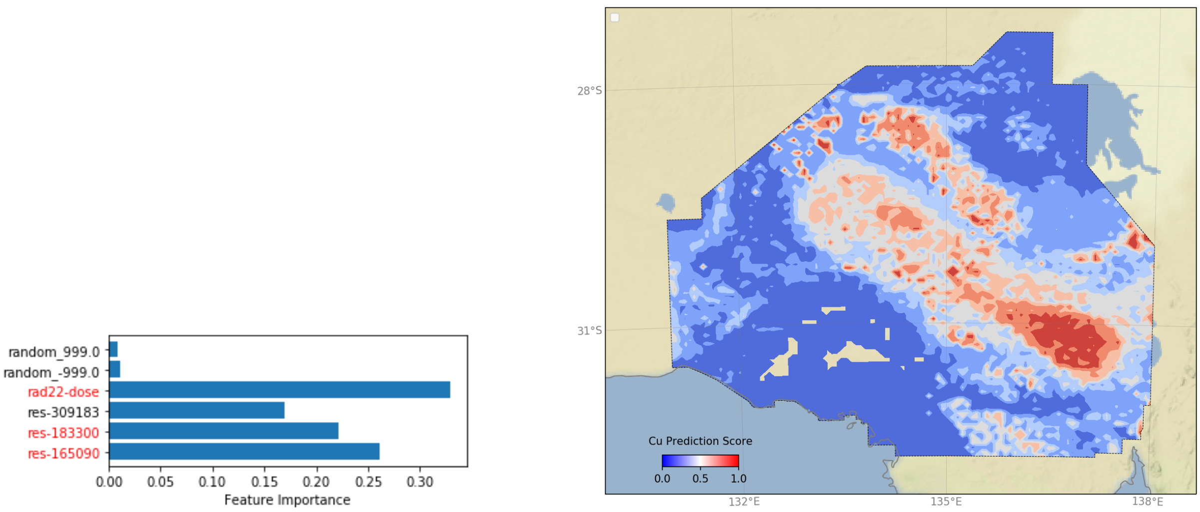

Figure A24 cont. Top 4 features. Cont...

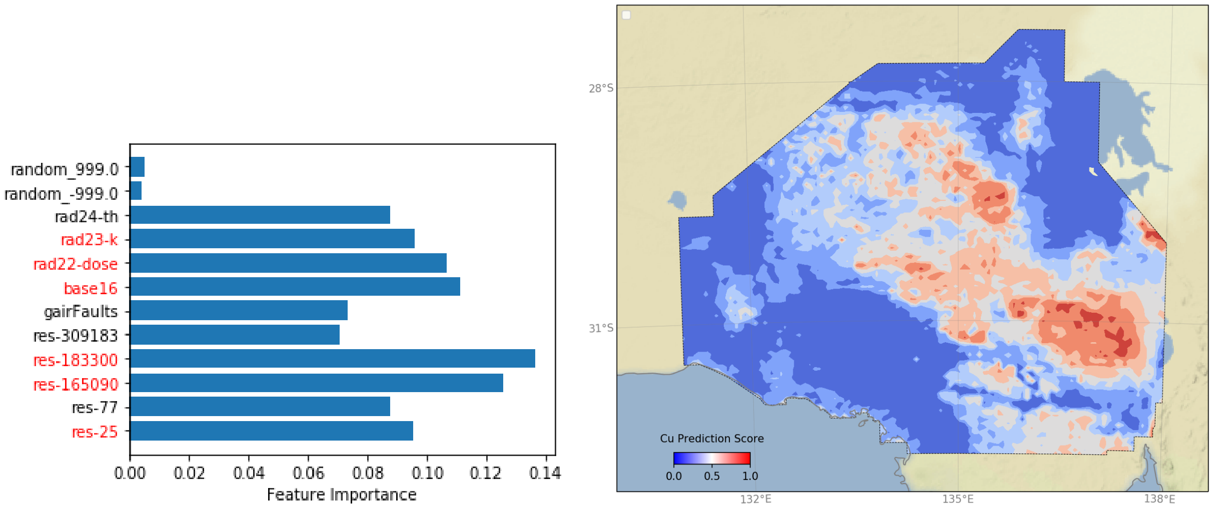

Figure A24 cont. Top 10 features. Cont...

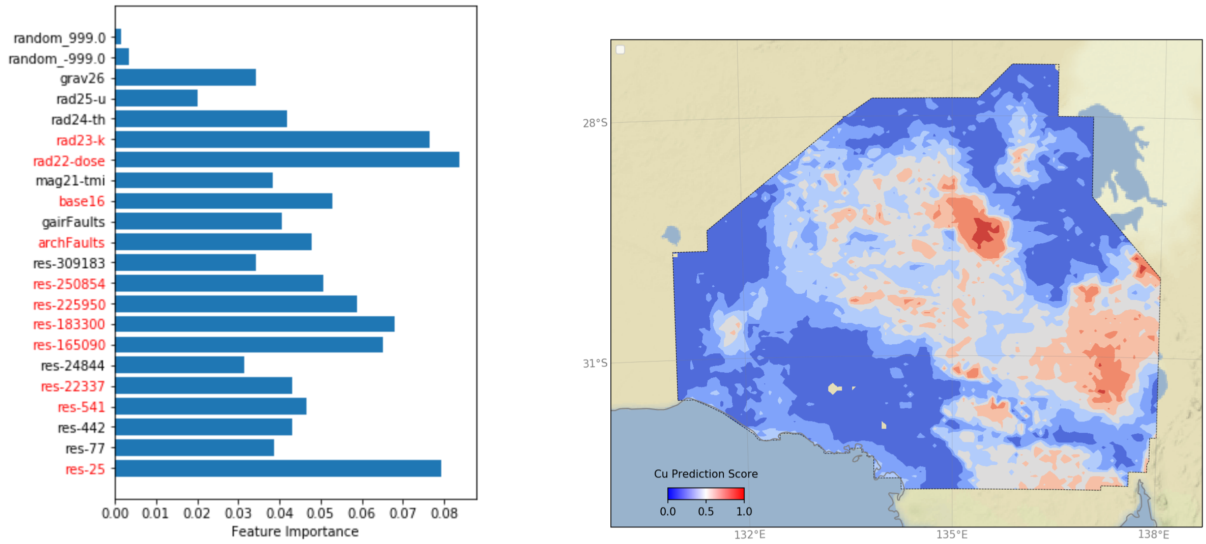

Figure A24 cont. Top 20 features. Cont...

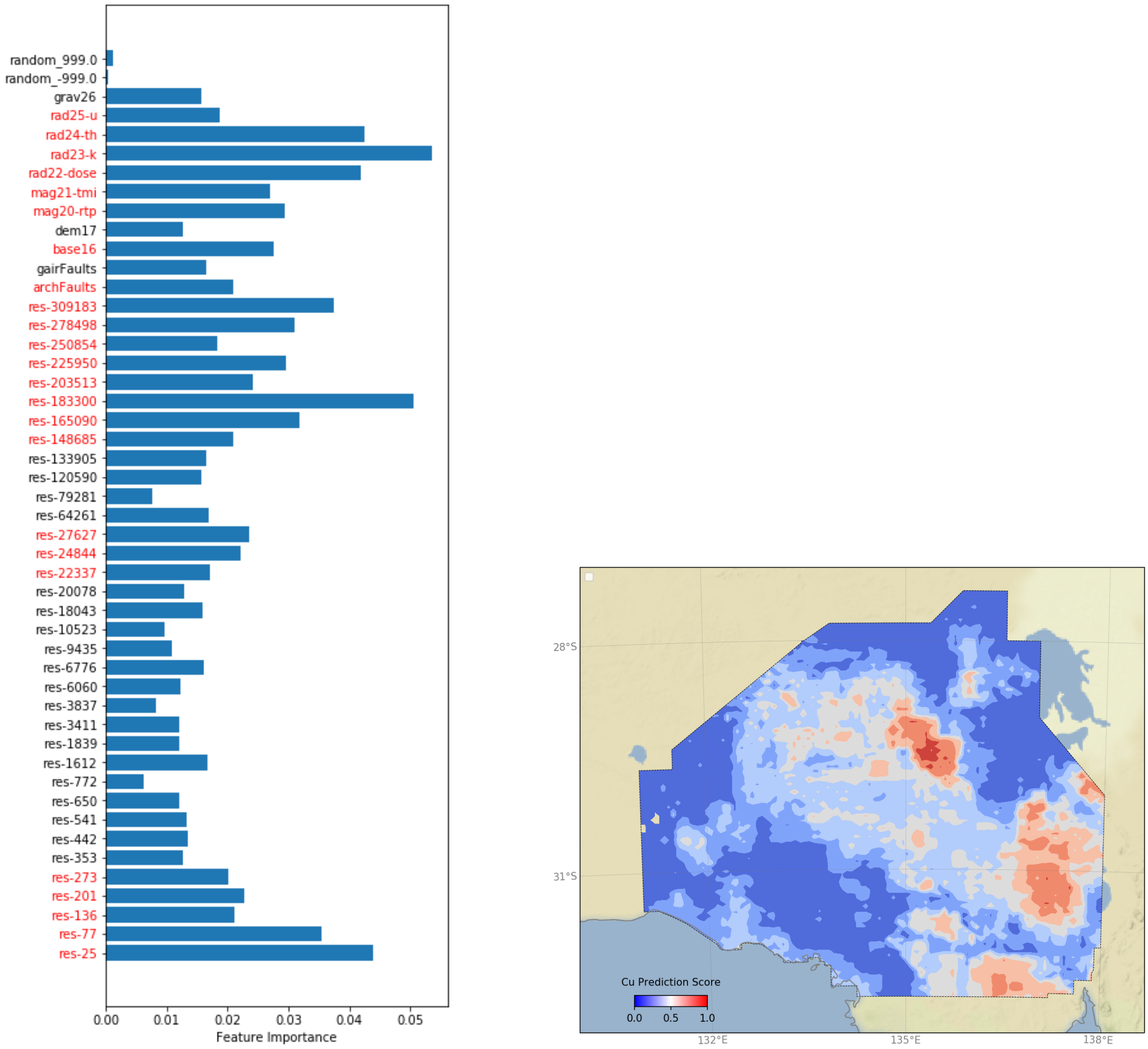

Figure A24 cont. Top 40 features.