Proceedings of ISERME 2025
19-20 September 2025
https://doi.org/10.31705/ISERME.2025.23

Smart Integration of Kriging and Machine Learning for Ash-Based Coal Quality Estimation in Sparse Drillhole Environments

Gebretsadik[1,2] A, Okada [1]N, Itano[1] R, Ohtomo[1] Y, Mathews[3] G, and Kawamura[1] Y*

[1]Division of Sustainable Resources Engineering, Graduate School of Engineering, Hokkaido University, Sapporo, Japan

[2]Department of Mining Engineering, Aksum University, Aksum-7080 (Tigray), Ethiopia

[3]Department of Geological Sciences, School of Mining and Geosciences, Nazarbayev University, Astana, Kazakhstan

*Corresponding author – angstg2007@gmail.com

Abstract

Effective coal quality estimation, particularly ash content prediction, plays a vital role in mine planning, resource evaluation, and environmental management. However, conventional geostatistical methods, such as kriging, often face limitations in sparsely drilled areas due to their linear assumptions and inability to capture complex geological variability. This study implements a residual-based integration of geostatistics and machine learning for ash estimation. Ordinary kriging was first applied using a spherical variogram with strong geometric anisotropy (ranges: 2500 m horizontal, 20 m vertical) and a negligible (≈0) nugget indicated by the experimental variogram. The kriging baseline achieved RMSE = 1.485 (leave-one-out), and kriging variance was retained as an uncertainty indicator. Machine learning algorithms, including support vector machine, XGBoost, random forest, and LightGBM were then trained on the kriging residuals, incorporating the kriging variance as an additional input feature to capture complex non-linear patterns. The models were optimized using advanced hyperparameter tuning methods such as GridSearchCV and genetic algorithm. The hybrid framework substantially improved predictive accuracy, with the optimized random forest model achieving a reduced RMSE of 1.26 on the test set. These results demonstrate that integrating geostatistics with AI-based residual modeling enhances estimation robustness in sparse-data mining environments. The improvement, reflected in a 15% reduction in RMSE compared to kriging alone, provides a statistically significant basis for coal quality prediction, particularly in data-scarce environments.