Exploratory analysis of multivariate drill core depth series measurements

X. Gu[1] A. V. Metcalfe[2] N. J. Cook[3] C. Aldrich[4] L. George[5]

(Received 27 February 2022; revised 11 November 2022)

Abstract

Demand for mineral resources is increasing, necessitating exploitation of lower grade and more heterogeneous orebodies. The high variability inherent in such orebodies leads to an increase in the cost, complexity and environmental footprint associated with mining and mineral processing. Enhanced knowledge of orebody characteristics is thus vital for mining companies to optimize profitability. We present a pilot study to investigate prediction of geometallurgical variables from drill sensor data. A comparison is made of the performance of multilayer perceptron (MLP) and multiple linear regression models (MLR) for predicting a geometallurgical variable. This comparison is based on simulated data that are physically realistic, having been derived from models fitted to the one available drill core. The comparison is made in

DOI:10.21914/anziamj.v63.17192, Published 2023-01-10, as part of the Proceedings of the 15th Biennial Engineering Mathematics and Applications Conference. ISSN 1445-8810. (Print two pages per sheet of paper.)

terms of the mean and standard deviation (over repeated samples from the population) of the mean absolute error, root mean square error, and coefficient of determination. The best performing model depends on the form of the response variable and the sample size. The standard deviation of performance measures tends to be higher for the MLP, and MLR appears to offer a more consistent performance for the test cases considered.

Contents

1 Introduction

The demand for mineral resources is increasing, and companies are mining lower grade and more heterogeneous orebodies. Such high variability leads to higher costs, complexity and environmental impacts. So mining companies need models that can predict processing costs from exploratory drill core data. Geometallurgy is a multi-disciplinary approach that combines geological and mineralogical information to create and predict spatial models aimed at maximising metal recovery, optimising processing efficiency, and minimising associated management risks. Geometallurgical variables are those that contain information on any rock or mineral property that carries economic implications for the business model. Geometallurgical variables are grouped into primary and responsive variables. Primary variables (e.g., grain size, metal grades, mineralogy and other rock properties) are measured directly, whereas responsive variables are predicted from multiple parameters. Responsive variables (e.g., metal recovery, grindability and plant throughput) are key to process efficiency and evaluation of the final cost of mineral exploitation. It is important to initiate a geometallurgical program at the early stage of mine development, as a thorough understanding of geometallurgical variables inherent to the orebody, or parts thereof, can help identify potential processing issues prior to major capital investment.

Since 2009, when the concept of primary-response framework was first established [7], there has been an increasing number of papers using multiple linear regression models to predict geometallurgical variables. Boisvert et al. [3] fitted multiple linear regression (MLR) models to reduced dimension data from the Olympic Dam Mine, South Australia, to predict six sparsely sampled plant performance variables. They reported correlations between true and estimated values ranging from 0.533 to 0.9. Hunt et al. [12] used MLR to predict three comminution variables: the SAG power index, Bond Ball Mill Work Index (BWi), and resistance to abrasion breakage. The results gave average relative errors between 6% and 12%. Both et al. [5] implemented MLR to predict ball mill throughput as a function of rock attributes of blended materials for the Tropicana Gold Mining Complex in Western Australia. They found Pearson correlation coefficients ranging from 0.5 to 0.80. Moreover, Johnson et al. [14] showed that MLR performed the best for gold and copper recovery models, as well as gold grade, among partial least squares (PLS) regression and deep PLS models.

However, non-linear relationships between primary and response geometallurgical variables are well documented [e.g., 16], and MLR, even with quadratic and interaction terms, may not be the best model for predicting geometallurgy. Artificial neural networks (ANN), in particular, can be used to capture non-linearity and there is a considerable number of papers that model geometallurgical variables using different types of ANN models, but there is no consensus about the best approach. Moreover, the extrapolation issue in this field is quite new. Multilayer perceptron (MLP) is a basic form of ANN which is commonly used and has been found to give reliable results in industry and business. In this simulation study, we compare MLR with MLP for predicting a physical plausible geometallurgical variable, especially for extreme values.

2 Analytical methodology

Drill core data were obtained from a Boart Longyear test site at Brukunga, South Australia. We applied standard time series analysis to geological series that are at equally spaced depths (sampled every 10 cm), and are referred to here as depth series. The first principle component (PC1) based on three variables, namely the concentrations of potassium (K), uranium (U), and thorium (Th), is considered. A Generalised Autoregressive Conditional Heteroskedasticity (GARCH) model is fitted to the PC1, and used to construct a study population of 1000 depth series and associated geometallurgical variables. Samples of depth series are taken from this study population. The aim is to predict the geometallurgical variables from features of the sample depth series. These features are either estimates of GARCH parameters or summary measures of a wavelet transform.

2.1 Wavelet analysis

Wavelet analysis is a frequency technique that captures both local spectral and temporal information. There have been many applications of wavelet transforms of geophysical log data to identify borehole layers and boundaries. These studies have shown that the Morlet and Mexican Hat wavelets are efficient for de-noising and blocking geophysical log data for various types of mineral deposits [e.g., 8]. Wavelet analysis is one of the best approaches to detect spatially repeating patterns. For example, the wavelet analysis performed by Webb et al. [17] using density and susceptibility data from the Bellevue and Moordkopjie boreholes, South Africa, has demonstrated the presence of previously unobserved layering on scales.

2.2 GARCH model

The GARCH model, developed by Engle and Bollerslev [10], models changing variability, known as variance volatility [9]. Mathematically, a GARCH(p, q) model has the form

$$y_t = \phi + \epsilon_t\,, \quad \epsilon_t \sim N(0, \sigma_t^2)\,, \tag{1}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2\,, \tag{2}$$

where ϕ, ω, α_i and β_j are constants, and ϵ_t is an error term which follows a normal distribution with mean zero and variance σ_t^2. As variance σ_t^2 is non-negative, Bollerslev [4] imposed the conditions ω, α_i and β_j non-negative, and $\alpha_i + \beta_j < 1$ for $i = 1, \ldots, q$ and $j = 1, \ldots, p$.

2.3 Multilayer perceptron

An MLP model is a class of feedforward neural network and is used in this study with resilient back-propagation and the weight backtracking algorithm. An MLP consists of at least three layers of neurons: an input layer; one or more hidden layers; and an output layer. Each neuron in an MLP is fully connected with neurons in the next layer. To avoid underfitting or overfitting of MLP models, it is reasonable to use a trial-and-error test. We compare one hidden layer with four, three, two, and one neurons, and two hidden layers with $(4, 3)$, $(4, 2)$, $(4, 1)$, $(3, 2)$, $(3, 1)$ and $(2, 1)$ neurons. Sigmoid functions are used as activation functions.

2.4 Settings for simulation studies

2.4.1 Overview

The set up of the simulation study is based on an analysis of a depth series from a single drill core that is considered to be representative of the ore deposits. A GARCH model is fitted to the depth series from this core and plausible distributions for the parameters of the GARCH model are inferred. Then N sets of random parameters are drawn independently from these distributions and assigned to N notional drill cores to make up the study population. For each notional drill core, a depth series of length D is simulated, using the assigned parameters, and a plausible geometallurgical response y is calculated as a function of the assigned parameters. So, we have set up a study population of N drill cores, each with a corresponding depth series and a response value.

As a training set for the simulation, samples of n cores are drawn from the population of N cores, at random without replacement. An MLP model and MLR models for predicting y from features of the depth series are fitted. The models are then tested against the remaining $N-n$ cores. Summary statistics of prediction errors are then calculated. The draw of n from N is repeated M times, allowing for a mean and standard deviation of summary statistics to be calculated. We set $N = 1000$, $M = 100$, and consider five values of n between 800 and 30.

2.4.2 Setting up the population

Depth series along the single drill core are available for concentrations of three elements: K, U and Th. We consider the PC1 rather than choose a single element. In general, there is no reason why the PC1 should be associated with a response [e.g., 11], but in practice it often is, as in applications of multivariate statistical process control [e.g., 15]. Also, use of the PC1 considerably reduces the influence of spikes at zero concentration in the single element depth series. The PC1 calculated using the correlation matrix accounts for 68% of the variability in the original data, which is not unusual for geochemical data. For example, Zuo [18] reported that the PC1 accounts for 52% of the total variability in stream sediment concentration data for Cu, Pb, Zn and Ag. A GARCH(1, 1) was fitted to the PC1 of the K, U, Th depth series and has the form

$$y_t = \phi + \epsilon_t \quad \text{where} \quad \epsilon_t \sim N(0, \sigma_t^2)\,, \tag{3}$$

$$\sigma_t^2 = \omega + \alpha\epsilon_t^2 + \beta\sigma_t^2\,. \tag{4}$$

Lognormal distributions are chosen as plausible distributions for the parameters. This is because of non-negatively constraints and high skewness, which facilitate the prediction performance of the MLP in terms of extrapolation. For each coefficient a lognormal distribution of plausible values is set up. This is equivalent to assuming the logarithm of a coefficient has a normal distribution with mean and variance, respectively,

$$\mu = \log\left(\frac{\mu_X^2}{\sqrt{\mu_X^2 + \sigma_X^2}}\right), \quad \sigma^2 = \log\left(1 + \frac{\sigma_X^2}{\mu_X^2}\right),$$

where μ_X and σ_X are the mean and standard error of the estimates of coefficients returned from the GARCH(1, 1) model fitted to the first principal component. A population of $N = 1000$ drill cores is set up by making 1000 independent random draws from the distributions of ψ, ω, α and β, subject to the constraint that $\alpha + \beta < 1$. For each drill core a depth series is simulated of length $D = 6381$ to match the length of the observed depth series. In addition, values of two plausible geometallurgical response variables are calculated and scaled to the interval $[0, 1]$ to facilitate fitting of the MLP (see Algorithm 1 for details).

2.4.3 Simulation

For a given value of n, a random sample of n from N drill cores was drawn without replacement to form a training set. A GARCH$(1, 1)$ model and a wavelet decomposition was fitted to the depth series for each drill core in the sample. An MLR and MLPs were fitted to both y_1' and y_2' using the four estimated GARCH$(1, 1)$ coefficients as predictor variables. Similarly, an MLR and an MLP was fitted to both y_1' and y_2' using four features extracted from the wavelet decomposition as predictor variables.

Values of n are chosen as 800, 500, 200, 50 and 30, corresponding to sample fractions of 80%, 50%, 20%, 5% and 3%, and are used as training sets. The fits of the models are evaluated using adjusted R^2 on the training sets, defined

Algorithm 1: Setting for plausible populations for simulation.

for i *in* $1 : N$ **do**
  Random draw deviates $(\phi_i, \omega_i, \alpha_i, \beta_i)$ from distributions of the coefficients for the GARCH model;
  Define a plausible metallurgical variable:
  Simulation 1: $y_{1i} = e^{\phi_i + \omega_i + \alpha_i + \beta_i + \phi_i\alpha_i + \phi_i\beta_i}$;
  Simulation 2: $y_{2i} = \sin[6\pi(\phi_i + \omega_i + \alpha_i + \beta_i + \phi_i\alpha_i + \phi_i\beta_i)]$;
  Simulate a depth series of length n to represent observations from the drill core
end
Define $y_1' = \frac{y_1 - \min(y_1)}{\max(y_1) - \min(y_1)}$, $y_2' = \frac{y_2 - \min(y_2)}{\max(y_2) - \min(y_2)}$

as

$$\text{Adjusted } R^2 = 1 - \left(\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-p}\right)\left(\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}\right)^{-1}, \qquad (5)$$

where y_i is an observation in the training set, $\hat{y}_i$ is the fitted value, $\bar{y}$ is the mean value of the training set, and p is the number of parameters fitted. The prediction performances of the models are tested on the remaining $N - n$ drill cores (the test set), using mean absolute error (MAE), root mean square error (RMSE), and $R^2 = 1 - \text{RMSE}^2/\text{var}(y)$ where y is the response variable in the test set. For each value of n the random draw was repeated $M = 100$ times, allowing us to calculate both the mean and standard deviation of the performance measures.

2.4.4 Extrapolation

For each of the four parameters, we define a low value as zero and a high value as the upper 0.001 quantile of its distribution. We then consider all $2^4 = 16$ combinations of high and low parameter values. If $\alpha + \beta \geqslant 1$, α and β are reduced by the same factor so that their sum is reduced to 0.999. These 16 parameter combinations are taken to define outlying drill core values (out-of-ensemble) and corresponding values of y_1 are calculated. The fitted MLP and MLR models are then used to predict y_1 for these drill cores using the parameter values as predictors.

2.4.5 Sample size in practice

In a real situation, 800 drill cores is a very large number, even for a large, complex deposit. Early exploration campaigns will have anywhere from 6 to 20 drill holes, advanced exploration maybe ten times that, and deposits approaching JORC-approved resource reporting and production perhaps a few hundred drill holes. Only giant deposits containing billions of tonnes of ore will be explored by 1000 or more drill holes.

3 Results and discussion

3.1 Exploratory analysis

Summary statistics of the data are given in Table 1. Observations that lie below the minimum detection point of measurement are recorded as zero. The $\text{PC}1 = -0.609\text{K}_s + 0.549\text{U}_s - 0.572\text{Th}_s$, where K_s, U_s, Th_s represent the standardised variables with mean zero and variance of one. The depth series plot of the PC1 is shown in Figure 1, and displays volatility and outlying values (red lines represent mean ± 3 standard deviations) which are typical of geological depth series.

Table 1: Summary statistics of the data (length $D = 6381$ observations).

Variables	Prop. 0s	Mean	Maximum	sd	Skewness	Kurtosis
K	27%	2.825	24.434	2.986	1.250	1.863
U	37%	0.001	0.018	0.001	1.622	5.389
Th	55%	0.002	0.052	0.003	3.110	25.482

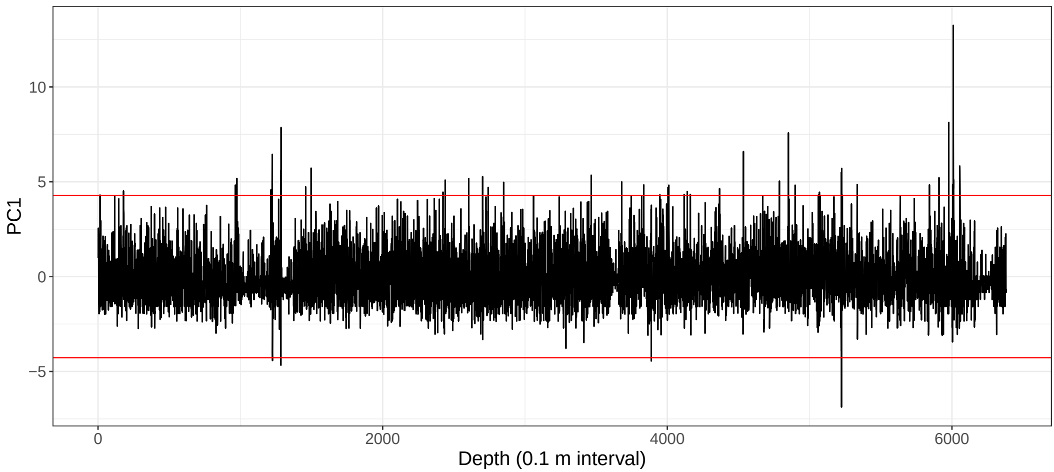


Figure 1: The depth series plot of the PC1.

3.2 Simulation study

The estimates of ϕ, ω, α, β are 0.0014, 0.1382, 0.1025 and 0.8316, respectively, with standard errors 0.013, 0.041, 0.012 and 0.029, respectively (equations (3) and (4)). Two simulation studies are performed (Algorithm 1). For the simulation study using features from the GARCH model, we also examine the performance of MLP and MLR for extrapolation. Figure 2 is a boxplot of the scaled response variables (range $[0, 1]$).

3.2.1 Features from the GARCH model (simulation study 1.1)

The MLR models are fitted with five parameters (one intercept + four covariates) and 15 parameters (one intercept + four linear + six two-way interactions + four quadratic). An MLP with lowest RMSE is chosen which consists of two hidden layers with two and one neurons, respectively, this

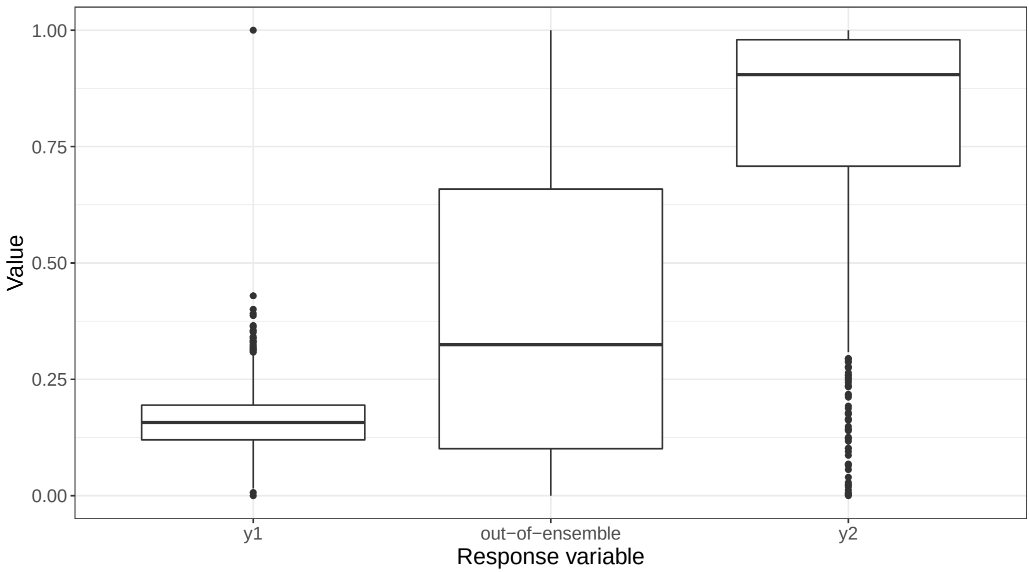


Figure 2: Boxplots of the response variables.

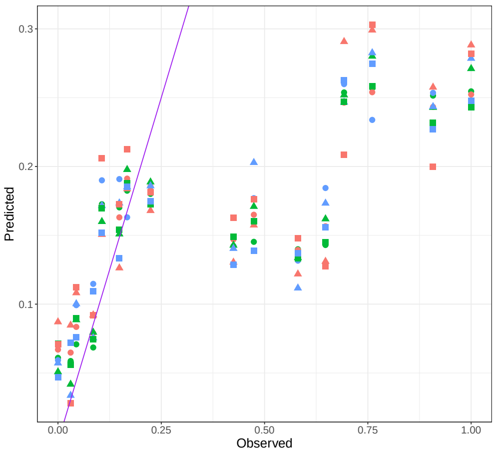
gives an MLP with 15 parameters ($8+2+1$ weights and $2+1+1$ bias). The MLR(5), MLR(15) and MLP(15) are used to represent theses three models.

Interpolation: Table 2 shows negligible differences between the MLP(15), MLR(5) and MLR(5) for all summary statistics for large training sizes. As we decrease the training size, the performance of these models are distinguishable, and the MLP(15) is somewhat more erratic than the MLRs for small training sizes. Both the MLP(15) and the MLR(5) perform well in prediction with features from the GARCH models as predictor variables, as shown by the comparison between RMSE and standard deviation of the response variable ($\text{sd}(y_1') = 0.066$). Even though the MLR(5) gives relatively lower RMSE than the standard deviation of the y_1', and the highest adjusted R^2 (train), it is not suitable in this case. This is because the MLR(5) tends to over-fitting for small training sizes, as R^2 (test) is 0.177 for a training size of 30. Moreover, the adjusted R^2 (train) for the MLP(15) is decreasing, and the standard deviations of the statistics are relatively large compared to the other two models when the training size is smaller than 200. The MLP(15) is not suitable for predicting y_1' when the training size is small.

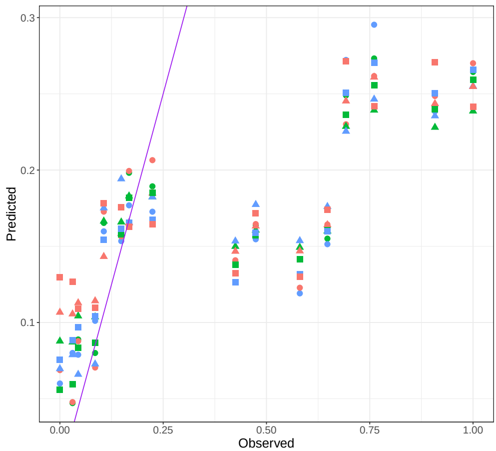
Extrapolation: Figures 3 and 4 show three randomly chosen realisations of the predicted values against the actual values, and suggests that neither the MLP(15) nor the MLR(5) provides useful predictions for values of y_1' above around 0.3 (the purple lines represent the perfect fit). For observations that are greater than 0.3, the predicted values are too low although the correlations between predicted and actual values are quite high (around 0.75). Linear trends tend to be more reliable for extrapolation than non-linear trends, but even so, extrapolation outside the ranges of the fitted predictor variables are unreliable and should be restricted to small proportions of the ranges. Figures 3 and 4 show that the predicted values returned from the MLP(15) are more variable and even less accurate than the MLR(15) values. The limitation of using MLPs for extrapolation was shown by Balabin and Smirnov [1] and Bishop [2].

Table 2: Average MAE, RMSE, adjusted R^2 (train) and R^2 (test), and their associated standard deviations (in parentheses), for the interpolation study using features from the GARCH model (Section 3.2.1). The standard deviation of y_1' is 0.066.

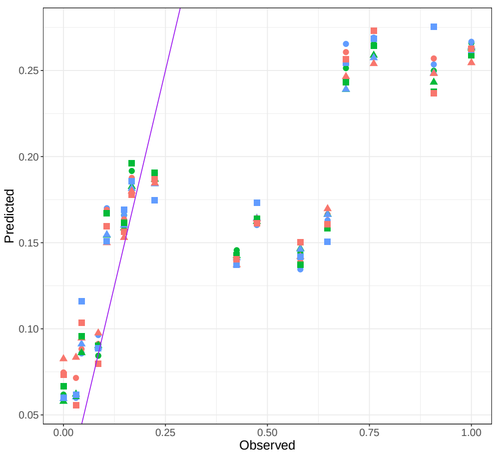
Size of training set		800	500	200	50	30
adjusted	MLP(15)	0.798	0.795	0.857	0.810	0.704
R^2 (train)		(0.052)	(0.086)	(0.084)	(0.136)	(0.354)
	MLR(5)	0.794	0.791	0.849	0.867	0.895
		(0.051)	(0.088)	(0.105)	(0.125)	(0.109)
	MLR(15)	0.798	0.807	0.844	0.903	0.929
		(0.057)	(0.088)	(0.120)	(0.087)	(0.066)
MAE (test)	MLP(15)	0.014	0.014	0.015	0.022	0.023
		(0.002)	(0.001)	(0.003)	(0.009)	(0.009)
	MLR(5)	0.013	0.013	0.014	0.015	0.017
		(0.002)	(0.001)	0.001	(0.004)	(0.010)
	MLR(15)	0.014	0.014	0.015	0.020	0.024
		(0.002)	(0.001)	(0.002)	(0.014)	(0.016)
RMSE (test)	MLP(15)	0.031	0.032	0.034	0.042	0.042
		(0.015)	(0.008)	(0.004)	(0.024)	(0.020)
	MLR(5)	0.031	0.031	0.032	0.033	0.035
		(0.015)	(0.008)	(0.004)	(0.004)	(0.012)
	MLR(15)	0.031	0.032	0.034	0.041	0.048
		(0.016)	(0.008)	(0.002)	(0.026)	(0.033)
R^2 (test)	MLP(15)	0.786	0.775	0.734	0.431	0.479
		(0.142)	(0.077)	(0.085)	(1.381)	(1.194)
	MLR(5)	0.788	0.781	0.770	0.746	0.686
		(0.144)	(0.082)	(0.034)	(0.105)	(0.486)
	MLR(15)	0.782	0.769	0.736	0.411	0.177
		(0.146)	(0.080)	(0.047)	(1.746)	(2.388)



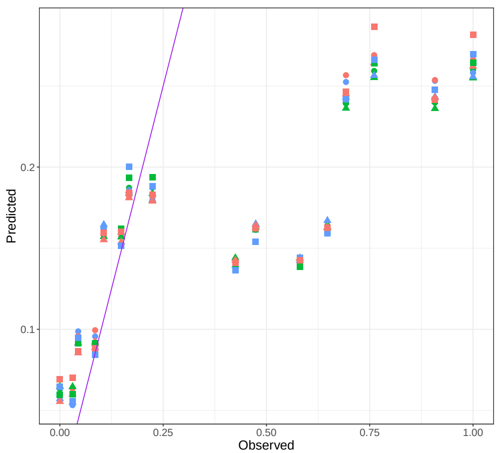
(a) Training size of 30.



(b) Training size of 50.

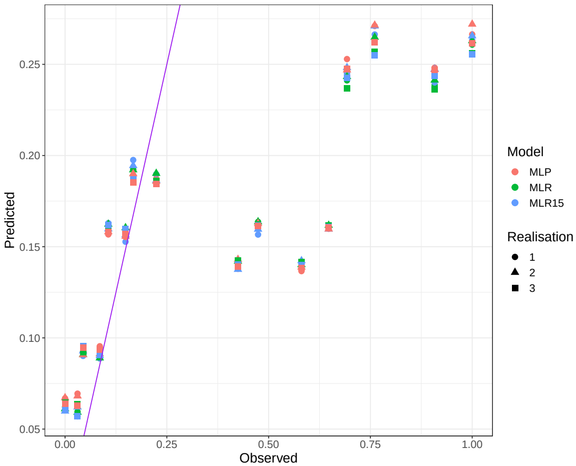


(c) Training size of 200.



(d) Training size of 500.

Figure 3: Three realisations of predicted values against actual values for test sets, by the training set size, for MLR, MLR15, and MLP models. Figure 4 provides the key.



(a) Training size of 800.

Figure 4: Three realisations of predicted values against actual values for test sets, by the training set size, for MLR, MLR15, and MLP models. Figure 3 uses the same key.

3.2.2 Features from wavelet decomposition (simulation study 1.2)

In this section, we aim to predict a geometallurgical variable y_1' using features from wavelet decomposition of simulated depth series. To keep the same structure of models, we calculated four features from Morlet wavelet decomposition. The four features were calculated for each frequency, then means over frequencies were calculated and used as predictor variables. The four features are Tsallis entropy [6], stability (variance of the means), lumpiness (variance of the variances), and the number of observations within a depth series that cross the median line [13]. The MLR models are fitted with five parameters and 11 parameters (one intercept + four linear + six two-way interactions).

Table 3 shows that the MLP(15) outperforms the MLRs in terms of predictions, regardless of training size. However, the standard deviations of the summary statistics of the MLP(15) for training size of 30 are higher than the other two models, which is consistent with the simulation study using features from the GARCH model (Section 3.2.1). The adjusted R^2 (train) for the MLP(15) is

Table 3: Average MAE, RMSE, adjusted R^2 (train) and R^2 (test), and their associated standard deviations, for simulation study using features from wavelet decomposition (Section 3.2.2). The standard deviation of y_1' is 0.066.

Size of training set		800	500	200	50	30
adjusted R^2 (train)	MLP(15)	0.680	0.679	0.691	0.548	0.374
		(0.038)	(0.059)	(0.073)	(0.210)	(0.313)
	MLR(5)	0.416	0.428	0.456	0.512	0.557
		(0.042)	(0.066)	(0.101)	(0.122)	(0.137)
	MLR(11)	0.446	0.466	0.506	0.579	0.628
		(0.050)	(0.075)	(0.111)	(0.134)	(0.167)
MAE (test)	MLP(15)	0.024	0.024	0.026	0.031	0.034
		(0.002)	(0.002)	(0.003)	(0.005)	(0.007)
	MLR(5)	0.033	0.033	0.033	0.036	0.038
		(0.003)	(0.001)	(0.001)	(0.004)	(0.005)
	MLR(11)	0.032	0.032	0.033	0.042	0.056
		(0.003)	(0.001)	(0.002)	(0.009)	(0.019)
RMSE (test)	MLP(15)	0.041	0.043	0.045	0.050	0.053
		(0.014)	(0.009)	(0.009)	(0.009)	(0.015)
	MLR(5)	0.053	0.052	0.053	0.057	0.063
		(0.014)	(0.007)	(0.003)	(0.006)	(0.009)
	MLR(11)	0.053	0.055	0.059	0.083	0.131
		(0.014)	(0.007)	(0.009)	(0.029)	(0.070)
R^2 (test)	MLP(15)	0.622	0.572	0.508	0.405	0.281
		(0.189)	(0.230)	(0.333)	(0.377)	(0.772)
	MLR(5)	0.406	0.385	0.362	0.237	0.081
		(0.117)	(0.070)	(0.051)	(0.184)	(0.318)
	MLR(11)	0.378	0.327	0.192	−0.806	−4.032
		(0.172)	(0.107)	(0.306)	(1.737)	(7.417)

decreasing with decreased training size, whereas the adjusted R^2 (train) is increasing for the MLRs. The MLP(15) is more erratic than the MLRs. Performance using features from the wavelet decompositions as predictors is not as good as using features from the GARCH models. This is what we expect, because we set the physically plausible geometallurgical variable as a non-linear combination of the four features from the GARCH models.

3.2.3 Features from the GARCH model (simulation study 2)

The results from Section 3.2.1 show the accuracy of predictions using the MLP and the MLRs have negligible differences. This may be due to the mild non-linear function that we use to set up y_1, so that the MLRs also capture the relationships. A sine function is used for the second simulation study (Algorithm 1). It performs three cycles over possible combinations of parameter values, but as most parameter values are relatively close to their means and parameters are combined randomly and independently, quadratic terms suffice to predict the majority of y_2'. Nevertheless, outlying values of parameter values will lead to y_2' that are badly predicted by any combination of linear and quadratic terms.

Table 4 shows only the results using features from the GARCH models as predictors for interpolation because the wavelet features do not provide better predictions. The prediction performance of the MLP(15) shows higher accuracy in terms of MAE and RMSE, but is more erratic. The MLRs can only model sinusoidal responses if the predictor variables include sinusoidal functions at the same frequency. The MLPs are better able to model a sinusoidal response without stating that explicit relationship. From these results, we conclude that neither MLP nor MLR is suitable for predicting y_2 with small training size.

Table 4: Average MAE, RMSE, adjusted R^2 (train) and R^2 (test), and their associated standard deviations, for simulation study using features from GARCH model (Section 3.2.3). The standard deviation of y_2' is 0.229.

Size of training set		800	500	200	50	30
adjusted	MLP(15)	0.732	0.664	0.672	0.591	0.470
R^2 (train)		(0.213)	(0.265)	(0.268)	(0.337)	(0.439)
	MLR(5)	0.057	0.060	0.065	0.104	0.155
		(0.009)	(0.018)	(0.034)	(0.121)	(0.193)
	MLR(15)	0.695	0.704	0.729	0.804	0.855
		(0.020)	(0.044)	(0.066)	(0.119)	(0.117)
MAE (test)	MLP(15)	0.071	0.083	0.087	0.128	0.154
		(0.035)	(0.042)	(0.045)	(0.055)	(0.067)
	MLR(5)	0.172	0.172	0.172	0.178	0.189
		(0.008)	(0.004)	(0.004)	(0.011)	(0.019)
	MLR(15)	0.084	0.084	0.087	0.106	0.139
		(0.006)	(0.003)	(0.006)	(0.018)	(0.050)
RMSE (test)	MLP(15)	0.114	0.130	0.135	0.196	0.232
		(0.042)	(0.050)	(0.052)	(0.069)	(0.106)
	MLR(5)	0.223	0.225	0.226	0.240	0.256
		(0.012)	(0.007)	(0.005)	(0.014)	(0.024)
	MLR(15)	0.130	0.135	0.147	0.199	0.258
		(0.019)	(0.013)	(0.014)	(0.040)	(0.099)
R^2 (test)	MLP(15)	0.719	0.631	0.599	0.175	−0.246
		(0.227)	(0.287)	(0.328)	(0.608)	(2.340)
	MLR(5)	0.044	0.039	0.020	−0.104	−0.270
		(0.042)	(0.024)	(0.028)	(0.137)	(0.253)
	MLR(15)	0.672	0.650	0.584	0.215	−0.462
		(0.089)	(0.064)	(0.080)	(0.322)	(1.647)

4 Conclusion

The MLP offers a versatile modelling strategy which can emulate the MLR and so we would expect the MLP to generally outperform the MLR for interpolation. However, geological variables have highly skewed distributions so predictions of geometallurgical variables will sometimes rely on extrapolation. Extrapolation is unreliable with the MLR and arguably more so for the MLP. This study included a comparison of the prediction performance of the MLR and the MLP for outlying values of geological variables used as predictors, which showed the limitations of both models. Nonetheless, the MLP shows its advantages in predictions using wavelet features, but results are more erratic. Additionally, the MLP is preferable in modelling conceptual relationships; that is, without assuming a mathematical formula. In general, the prediction performances of the MLPs are better than the MLRs but are more erratic, although MLRs do perform in some cases. But this simulation study shows the limitations of using MLRs and MLPs for predicting physically plausible geometallurgical variables for small sample sizes. Other methods, such as support vector machine and partial least squares, or other types of artificial neural networks, could also be included for comparison once we have data from multiple drill cores and the associated geometallurgical variables.

In practice, a plausible scenario is that 40 drill cores are taken from a deposit and indicate that mining is viable. Mining begins and values of metallurgical variables associated with drill cores will be available sequentially. Prior models for predicting these geometallurgical variables can be modified by adaptive least squares as the response variable becomes available. We recommend linear MLP and MLR models, and quadratic terms and interactions should be considered at the start of mining operations. As mining continues more weight should be placed on the model with minimum mean squared error.

Acknowledgements Xiaomeng Gu acknowledges a PhD scholarship from the Australian Research Council Training Center for Integrated Operations for Complex Resources.

References

[1] R. M. Balabin and S. V. Smirnov. “Interpolation and extrapolation problems of multivariate regression in analytical chemistry: Benchmarking the robustness on near-infrared (NIR) spectroscopy data”. In: *Analyst* 137.7 (2012), pp. 1604–1610. DOI: 10.1039/c2an15972d (cit. on p. C219).

[2] C. M. Bishop. *Pattern recognition and machine learning*. Springer, 2006. URL: https://link.springer.com/book/9780387310732 (cit. on p. C219).

[3] J. B. Boisvert, M. E. Rossi, K. Ehrig, and C. V. Deutsch. “Geometallurgical modeling at Olympic dam mine, South Australia”. In: *Math. Geosci.* 45 (2013), pp. 901–925. DOI: 10.1007/s11004-013-9462-5 (cit. on p. C210).

[4] T. Bollerslev. “Generalized autoregressive conditional heteroskedasticity”. In: *J. Economet.* 31.3 (1986), pp. 307–327. DOI: 10.1016/0304-4076(86)90063-1 (cit. on p. C212).

[5] C. Both and R. Dimitrakopoulos. “Applied machine learning for geometallurgical throughput prediction—A case study using production data at the Tropicana Gold Mining Complex”. In: *Minerals* 11.11 (2021), p. 1257. DOI: 10.3390/min11111257 (cit. on p. C210).

[6] J. Chen and G. Li. “Tsallis wavelet entropy and its application in power signal analysis”. In: *Entropy* 16.6 (2014), pp. 3009–3025. DOI: 10.3390/e16063009 (cit. on p. C222).

[7] S. Coward, J. Vann, S. Dunham, and M. Stewart. “The primary-response framework for geometallurgical variables”. In: *Seventh international mining geology conference*. 2009, pp. 109–113. URL: https://www.ausimm.com/publications/conference-proceedings/seventh-international-mining-geology-conference-2009/the-primary-response-framework-for-geometallurgical-variables/ (cit. on p. C210).

[8] A. C. Davis and N. B. Christensen. “Derivative analysis for layer selection of geophysical borehole logs”. In: *Comput. Geosci.* 60 (2013), pp. 34–40. DOI: 10.1016/j.cageo.2013.06.015 (cit. on p. C212).

[9] C. Dritsaki. “An empirical evaluation in GARCH volatility modeling: Evidence from the Stockholm stock exchange”. In: *J. Math. Fin.* 7.2 (2017), pp. 366–390. DOI: 10.4236/jmf.2017.72020 (cit. on p. C212).

[10] R. F. Engle and T. Bollerslev. “Modelling the persistence of conditional variances”. In: *Econ. Rev.* 5.1 (1986), pp. 1–50. DOI: 10.1080/07474938608800095 (cit. on p. C212).

[11] A. S. Hadi and R. F. Ling. “Some cautionary notes on the use of principal components regression”. In: *Am. Statistician* 52.4 (1998), pp. 15–19. DOI: 10.2307/2685559 (cit. on p. C214).

[12] J. Hunt, T. Kojovic, and R. Berry. “Estimating comminution indices from ore mineralogy, chemistry and drill core logging”. In: *The Second AusIMM International Geometallurgy Conference (GeoMet) 2013*. 2013, pp. 173–176. URL: http://ecite.utas.edu.au/89773 (cit. on p. C210).

[13] R. Hyndman, Y. Kang, P. Montero-Manso, T. Talagala, E. Wang, Y. Yang, M. O’Hara-Wild, S. Ben Taieb, H. Cao, D. K. Lake, N. Laptev, and J. R. Moorman. *tsfeatures: Time series feature extraction.* R package version 1.0.2. 2020. URL: https://CRAN.R-project.org/package=tsfeatures (cit. on p. C222).

[14] C. L. Johnson, D. A. Browning, and N. E. Pendock. “Hyperspectral imaging applications to geometallurgy: Utilizing blast hole mineralogy to predict Au-Cu recovery and throughput at the Phoenix mine, Nevada”. In: *Econ. Geol.* 114.8 (2019), pp. 1481–1494. DOI: 10.5382/econgeo.4684 (cit. on p. C211).

[15] E. B. Martin and A. J. Morris. "An overview of multivariate statistical process control in continuous and batch process performance monitoring". In: *Trans. Inst. Meas. Control* 18.1 (1996), pp. 51–60. DOI: 10.1177/014233129601800107 (cit. on p. C214).

[16] E. Sepulveda, P. A. Dowd, C. Xu, and E. Addo. "Multivariate modelling of geometallurgical variables by projection pursuit". In: *Math. Geosci.* 49.1 (2017), pp. 121–143. DOI: 10.1007/s11004-016-9660-z (cit. on p. C211).

[17] S. J. Webb, G. R. J. Cooper, and L. D. Ashwal. "Wavelet and statistical investigation of density and susceptibility data from the Bellevue drill core and Moordkopje borehole, Bushveld Complex, South Africa". In: *SEG Technical Program Expanded Abstracts 2008*. Society of Exploration Geophysicists, 2008, pp. 1167–1171. DOI: 10.1190/1.3059129 (cit. on p. C212).

[18] R. Zuo. "Identifying geochemical anomalies associated with Cu and Pb–Zn skarn mineralization using principal component analysis and spectrum–area fractal modeling in the Gangdese Belt, Tibet (China)". In: *J. Geochem. Explor.* 111.1-2 (2011), pp. 13–22. DOI: 10.1016/J.GEXPLO.2011.06.012 (cit. on p. C214).

Author addresses

1. **X. Gu**, School of Civil, Environmental and Mining Engineering, University of Adelaide, South Australia 5005, Australia.
mailto:xiaomeng.gu@adelaide.edu.au
orcid:0000-0001-8930-1552

2. **A. V. Metcalfe**, School of Mathematical Sciences, University of Adelaide, South Australia 5005, Australia.
mailto:andrew.metcalfe@adelaide.edu.au

3. **N. J. Cook**, School of Civil, Environmental and Mining Engineering, University of Adelaide, South Australia 5005, Australia.

http://orcid.org/0000-0002-7470-3935

4. **C. Aldrich**, Western Australian School of Mines: Minerals, Energy and Chemical Engineering, Curtin University, WA 6845, Australia.
http://www.curtin.edu.au/

5. **L. George**, Boart Longyear Asia Pacific, 26 Butler Boulevard, Adelaide Airport, Adelaide South Australia 5950, Australia.
http://www.boartlongyear.com/