

Forecasting mass and metallurgical balance at a gold processing plant using modern multivariate statistics

<http://dx.doi.org/10.1590/0370-44672020730001>

Fernanda Gontijo Fernandes Niquini^{1,2}

<https://orcid.org/0000-0003-1872-1466>

João Felipe Coimbra Leite Costa^{1,3}

<https://orcid.org/0000-0003-4375-370X>

¹Universidade Federal do Rio Grande do Sul - UFRGS, Departamento de Engenharia de Minas, Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais, Porto Alegre - Rio Grande do Sul - Brasil.

E-mails: ²fernanda.gontijo.fn@gmail.com,

³jfelipe@ufrgs.br

Abstract

Knowing the quantity and the quality of products and tailings generated by a beneficiation plant, even before ore processing, can make the mining operations more sustainable, more profitable, and safer. To forecast these values, it is necessary to submit samples to batch tests which mimic the processing workflow used on an industrial scale. Then, the results need to be analysed with the aim of finding a statistical model able to comprehend how *Run of Mine* (ROM) characteristics impact the performance at the beneficiation. After developing a model, it is possible to apply it to blocks where the ROM characteristics are known, but the metallurgical information is not, making it possible to estimate these. With this goal, a geometallurgical model was developed with a neural network technique using 37 samples collected at two Brazilian gold mines. The Au and S grades in ROM, and the mine from where the sample was collected, were used as input variables. The model was able to forecast the following variables with a Pearson correlation coefficient on the cross validation test set equal to the value in parenthesis: mass (0.55) and metallurgical (0.54) recovery in the gravimetric concentrate, mass (0.80) and metallurgical (0.12) recovery in the flotation tailings, mass (0.77) and metallurgical (0.11) recovery in the leaching tailings, mass recovery (0.84) of gas sent to the sulphuric acid plant, and metallurgical recovery (0.65) in the leaching concentrate. The results obtained with neural networks were superior to the ones obtained when three alternative techniques were tested.

Keywords: mass recovery; metallurgical recovery; geometallurgy; neural networks; gold.

1. Introduction

Geometallurgical models are frequently developed focussing on forecasting variables related to saleable products. Geometallurgy applied for tailings grades and mass are rarely explored, either for lack of investment to know variables which will not bring immediate financial return or for ignoring the fact that by knowing the quantity and the quality of tailings, financial benefit can also be added. Knowing the tailings grade beforehand opens up the possibility of deciding where to send it: to the tailings dam, as usual, or to stockpile the tailings to make possible further use of it, if the observed grade is high enough to offset the process costs. Adopting the second option, it would be possible to recover a metal which initially would be lost, so improving the metallurgical recovery. Furthermore, by

knowing the complete plant mass balance, including tailings mass, it would enable one to forecast when the tailings dam will reach its maximum storage capacity, making it possible to better plan its height or building a new one. In a critical scenario, a new solution would be to use dried tailings deposition in waste dumps, aiming at minimising the environmental impact.

The need to predict concentrate geometallurgical characteristics beforehand is mandatory at most mining operations. By doing that, it is possible to classify a block as ore or waste, not only by considering the cut-off grade in ROM, but also by considering the characteristics that the material will present when processed. By adding this information into mine planning, it would avoid feeding the plant with a material

above the cut-off grade but with several contaminants which impact on the flotation recovery, leading to a reduction in the metallurgical recovery and in the concentrate grade. Ultimately, this would avoid producing a concentrate which would not generate a revenue, as it was planned to do, so affecting the cash flow for that specific period of time. Following the same reasoning, it is possible to feed a material below the cut-off, but with low contaminant grades, which can generate a product within the desirable specifications. Basically, ore and waste classification, supported by geometallurgical studies, goes beyond the simple cut-off grade definition. Additionally, geometallurgical models provide the information needed to adjust the plant as the ore characteristics change. Knowing beforehand that a

given ore can lead to a low metallurgical recovery, it would be possible to adjust the milling factors, or flotation circuit, to mitigate the expected loss in the metallurgical recovery.

Considering the benefits in predicting the products and tailings characteristics, this study focussed on building a geometallurgical model able to forecast the mass and the metallurgical balance that will be obtained when the gold ore from two Brazilian mines is processed. The novelty of this work is the capability to forecast, at each part of the flowsheet where there is a material output, either concentrate or tailings, the mass and the metallurgical recoveries and, ultimately, the gold grade in such materials simultaneously respecting mass closure.

Application of machine learning techniques in the mineral industry

This section presents a brief review on the application of machine learning to forecast variables in mining context.

Singh and Rao (2005) used a RBF neural networks to classify particles as manganese enriched, iron enriched or alumina enriched, using as input variables the red, green and blue in the RGB color space and four Harlic's textural parameters (energy, entropy, contrast and homogeneity). The overall accuracy of the model was 88.71%.

Lucay *et al.* (2020) developed a methodology to use a RBF neural network and a multi-layer perceptron to better model the design of the experiment's (DoE) data. First, the authors created a DoE using the frother concentration, the collector concentration and the water conductivity as input variables to forecast the froth stability. After the results of the experiment were processed, they decided to estimate (using variography and kriging) the froth stability at unsampled points which were after sampled using the Monte Carlo method. Then, the created samples using kriging were used to train a RBF neural networks and a multi-layer perceptron and the real values from the DoE were used as test set samples. Other applications of this methodology were presented by the authors to forecast the ash recovery in coal flotation and gold recovery at a plant, using other input variables. In all cases, the methodology proposed overcame the classical response surface methodology.

A challenge related to dealing with geometallurgical data is that most are non-additive (Richmond and Shaw, 2009). Carrasco *et al.*, (2008) proved that mass recovery is an additive variable, but metallurgical recovery is not, and because of this, it needs to be studied with prudence. According to Coward *et al.*, (2009), additivity is the "property that allows the mean of some variables to be calculated by a simple linear average". The non-additivity of an attribute restricts the use of linear averages, since it would generate biased or inconsistent results according to the studied phenomenon.

One of the most applied techniques to forecast geometallurgical variables is linear regression analysis, uni- or multivariate, which uses, as independent variables, information about *in*

situ ore characteristics including grade, grain size distribution, or lithology, to predict variables such as ore recovery at the processing plant. Examples in geometallurgical modelling are found in MacMillan *et al.*, (2011), Montoya *et al.*, (2011), Fernandes and Cabral (2016), and Vieira and Costa (2016), among others. Since regression analysis provides a linear estimation, its use to forecast non-additive variables can generate results that do not ideally represent the studied phenomenon.

In view of the increasing application of machine learning techniques in several fields, and their capacity to deal with non-additive data, it was decided to compare the results obtained by using regression analysis with the ones obtained using machine learning techniques.

Lishchuk, Lund and Ghorbani (2019) tested random forest, SMO (a variation of SVR), linear regression, M5 and M5P (variation of decision tree), among other machine learning techniques to forecast six geometallurgical variables at the Leveäniemi iron ore mine. Included among them the iron metallurgical recovery. The authors concluded that the best forecast was found when techniques which used decision trees were applied to forecast non-additive variables.

Prades and Deutsch (2016) evaluated the results obtained to forecast the acid consumption, the recovery and the impurity in a copper ore processing plant, using linear and quadratic regression, ACE (alternating conditional expectations), ridge regression, random forest and gradient boosted models (GBM). According to the authors, random forest and GBM presented the best results based on R^2 values.

Drumond (2019) applied logistic regression, discriminant analysis, support vector machine, random forest, neural networks, k-nearest neighbors, decision tree and naïve bayes techniques to forecast mineral particles of ore and waste in an ore-sorting equipment, using X-ray sensor readings as independent variables. The author considered that the support vector machine, k-nearest neighbors, random forest and the neural networks showed the best results.

Panda and Tripathy (2014) adjusted a neural network to forecast the

grade and the metallurgical recovery of chromite in a processing plant with a gravity concentrator using the wash water flow rate, the deck tilt angle and the slurry feed rate as input variables. The network built used only one hidden layer with eleven hidden nodes and the two output variables were forecasted simultaneously. The R^2 obtained in the test set was 0.96 to forecast the grade and 0.99 to forecast the recovery.

Other studies related to machine learning applied to problems in ore processing, excluding geometallurgy articles, can be read in McCoy and Auret (2018). Examples of the neural networks application to forecast variables related to transportation and blasting can be seen in Franco-Sepúlveda, Del Rio-Cuervo and Pachón-Hernández (2018).

According to the authors knowledge, no geometallurgical model has been elaborated so far to forecast simultaneously the grades and the mass and metallurgical recoveries of all outputs in a plant, either concentrates or tailings. To achieve this goal, it was necessary to build several models using different machine learning techniques aiming at identifying the one which would lead to the most precise and accurate forecast and to recognize the one able to better honor the mass and the metallurgical closure of the predictions. Among all techniques described in this section, three were chosen as the most promising to be evaluated: neural networks (NN), support vector regression (SVR), and random forest (RF).

2. Material and methods

The software used to build all models was developed in Python®. The scikit-learn library was used to build the SVR and the random forest models, to standardize the variables in the preprocessing step and to calculate the mean absolute error and the mean squared error. The keras library was used to build the neural network model and the statsmodels package was used to build the linear regression model. To calculate the correlations between the true values and the forecasts made, the scipy library

was used.

To better comprehend the variables involved in this problem, a brief explanation of the process flowsheet will be presented. The ROM is processed initially at Plant 1, which has crushing and milling circuits, followed by a gravimetric separation and a flotation circuit. Next, the flotation concentrate is sent to Plant 2, in which the material is submitted to a calcination followed by a leaching circuit. Figure 1 illustrates the simplified flowsheet with all mass

and metal output points. The variables studied are the ones represented in the green and blue boxes. The gas metallurgical recovery box is shaded as no metal is lost in the gas generated by the calcination process, so its value is zero. The leaching concentrate mass recovery is also shaded as the mass related to this part of the process is irrelevant compared to the others, and consequently, this mass recovery is rounded to zero. Finally, the shaded variables will not be modelled, leaving the remaining eight to be studied.

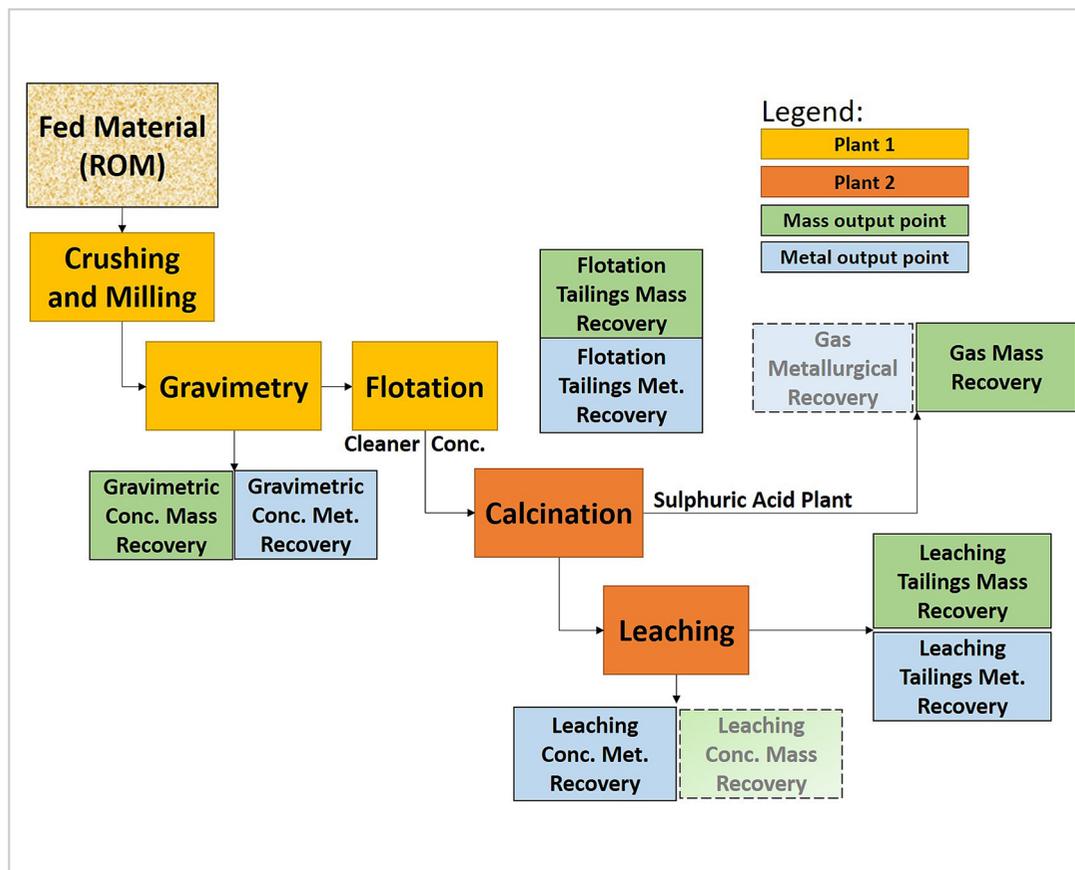


Figure 1 – Simplified flowsheet from Plant 1 and 2.

Planning to mimic the performance of the processing plant, bench scale tests were run using samples with different chemical and geological features. Each sample passed through each of the concentrator steps: first, the samples entered a gravimetric concentrator and its tailings were fed to the flotation. Next, the flotation concentrate was fed into the calcination process and the material remaining was sent to the leaching. In total, 37 lab tests were made allowing building a statistical model able to relate the ROM grades to the quantity and quality of the concentrate and tailings generated at the plant.

Table 1 presents the descriptive statistics for the 37 samples regarding Au (g/t) and S (%) grades at ROM, and the eight metallurgical variables, divided between the two mines under study. Note that the mean gold grade at Mine A is higher than that at Mine B (11.42 against 2.38 g/t). The same applies for the S mean grade, the Mine A mean is higher (8.55 against 7.53%).

The metal flow for Mine A ore shows that (on average) 60.49% of the gold fed to the plant goes to the gravimetric concentrate, 7.36% to the flotation tailings, 30.42% to the leaching concentrate, and 1.73% to the leaching

tailings. For the Mine B ore, the average numbers are: 70.67% of the gold goes to the gravimetric concentrate, 4.95% lost to the flotation tailings, 22.51% to the leaching concentrate, and 1.88% lost to the leaching tailings. Analyzing the mass fed from Mine A, on average, 5.57% goes to the gravimetric concentrate, 78.70% to the flotation tailings, 11.01% to the leaching tailings, and 2.75% to the sulphuric acid gas. For Mine B, these numbers are: 5.16% of all mass fed goes to the gravimetric concentrate, 87.32% to the flotation tailings, 6.06% to the leaching tailings, and 1.53% goes to the sulphuric acid gas.

Table 1 – Descriptive statistic by mine.

Mine	Variable	N	Mean	Std. Deviation	Minimum	Q1	Median	Q3	Maximum
Mine A	Au (g/t)	32	11.42	10.58	1.07	5.32	8.21	13.75	51.33
	S (%)	32	8.55	7.58	0.38	3.62	6.33	10.41	31.64
	Gravimetric Conc. Mass Recovery (%)	32	5.57	1.09	2.82	5.10	5.83	6.14	7.99
	Gravimetric Conc. Met. Recovery (%)	32	60.49	14.43	25.60	51.00	59.12	71.44	87.17
	Flotation Tailings Mass Recovey (%)	32	78.70	9.92	50.08	73.11	80.89	87.80	92.94
	Flotation Tailings Met. Recovey (%)	32	7.36	6.33	0.52	3.10	5.10	9.53	21.61
	Sulphuric Acid Gas Mass Recovery (%)	19	2.75	1.70	0.57	1.53	2.21	3.77	6.53
	Leaching Conc. Met. Recovery (%)	32	30.42	13.28	10.36	21.39	27.27	39.45	66.11
	Leaching Tailings Met. Recovery (%)	32	1.73	1.17	0.07	0.71	1.64	2.31	4.35
	Leaching Tailings Mass Recovery (%)	19	11.01	6.30	2.39	6.31	9.83	12.97	27.52
Mine B	Au (g/t)	5	2.38	2.31	0.46	0.83	2.12	2.21	6.27
	S (%)	5	7.53	11.67	0.85	1.48	3.46	3.56	28.30
	Gravimetric Conc. Mass Recovery (%)	5	5.16	0.85	4.21	4.38	5.42	5.55	6.25
	Gravimetric Conc. Met. Recovery (%)	5	70.67	10.13	56.53	66.69	69.61	78.08	82.45
	Flotation Tailings Mass Recovey (%)	5	87.32	3.33	83.87	85.58	86.73	87.73	92.71
	Flotation Tailings Met. Recovey (%)	5	4.95	1.58	3.34	3.36	5.04	6.33	6.67
	Sulphuric Acid Gas Mass Recovery (%)	3	1.53	0.21	1.33	1.42	1.51	1.63	1.75
	Leaching Conc. Met. Recovery (%)	5	22.51	8.07	12.07	17.78	24.58	24.64	33.48
	Leaching Tailings Met. Recovery (%)	5	1.88	1.09	0.78	0.81	2.13	2.34	3.33
	Leaching Tailings Mass Recovery (%)	3	6.06	1.33	4.70	5.40	6.11	6.74	7.37

From Table 1, it can be seen that each mine has different ROM grades with metallurgical variables showing different be-

haviors. Thus, a categorical variable (D1) was created representing the mine from where the sample was collected, making it

possible to introduce this information into the models. Equation 1 illustrates how this variable was built:

$$D1 = \begin{cases} 0, & \text{if the sample belongs to Mine B} \\ 1, & \text{if the sample belongs to Mine A} \end{cases} \quad (1)$$

Thus, Au (g/t) and S (%) grades, and the variable D1 were used as independent (input) variables in this study.

To build a forecast model an isotopic database is necessary, which is not the case, since only 22 of the 37 samples have information about the sample mass after the calcination process. The lack of this information prevents the calculation of the gas mass recovery and the leaching tailings mass recovery. To complete the database, it was decided to use the maximum likelihood technique (Enders, 2010), which is the imputation state of the art for missing at random (Rubin, 1976) information.

After the imputation, the next step was to transform the Au and S grades into values within the 0 to 1 range (the same used for the D1 variable), to ensure that, in the models created, a variable does not impact more than the others due to its magnitude.

Finished the preprocessing step, the imputed and standardized dataset was used to create the geometallurgical models. Four different techniques were investigated: regression analysis (Hair

et al., 2009), random forest (Hastie, Tibshirani and Friedman, 2009), support vector regression (Izenman, 2008), and artificial neural networks (Haykin, 1999). Basics on these methods are presented herein.

The models made with regression analysis were evaluated according to the significance t-tests (Triola, 2008) and the residual analysis (Hair *et al.*, 2009) was run. The models developed for the flotation and leaching tailings metallurgical recovery presented coefficients statistically equal to zero, indicating that their use for prediction purposes is discouraged. The leaching tailings mass recovery and the flotation tailings metallurgical recovery models violated the normality assumptions of the residuals, which can affect the t-test results to such a point as to invalidate them. Thus, three of the eight models developed using regression analysis are not statistically valid, and the results obtained by using this technique will not be further discussed.

The random forest algorithm using 1000 trees was used to model the eight metallurgical variables under

study. To determine the minimum number of samples by leaves in the trees, values were tested ranging from 2 to 15. The correlation coefficient obtained in the out-of-bag (Hastie, Tibshirani and Friedman, 2009) was calculated for each minimum number of samples in the leaves tested, and the value which presented the higher correlation value was chosen.

The models using support vector regression were divided into two parts: first, the models were adjusted using a linear kernel and then, adjusted using a non-linear kernel (Radial Basis Function). For both cases an epsilon (ϵ) value equal to 0.0001 was used. The determination of the best C parameter value, in the linear case, and of the best C and gamma (Izenman, 2008) values, for the non-linear case, was made using a grid search with values ranging from 0.001 and 100 for both parameters. For each dependent variable the parameter (linear SVR), or the combination of parameters (non-linear SVR), which returned the lower mean absolute error (MAE) in the test set was chosen.

Table 2 - 5-fold cross validation results.

Metric	Set	Variables	RF	Linear SVR	Non-Linear SVR	NN
Correlation	Training	Gravimetric Conc. Mass Recovery	0.7686 ± 0.0153	0.6292 ± 0.0235	0.8108 ± 0.0203	0.5457 ± 0.0109
		Flotation Tailings Mass Recovery	0.9407 ± 0.0046	0.8394 ± 0.02	0.8993 ± 0.0142	0.8449 ± 0.0211
		Leaching Tailings Mass Recovery	0.9239 ± 0.0053	0.8028 ± 0.0234	0.8664 ± 0.0204	0.8029 ± 0.0262
		Gas Mass Recovery	0.9396 ± 0.0083	0.868 ± 0.0134	0.9349 ± 0.0093	0.8803 ± 0.0158
		Gravimetric Conc. Met. Recovery	0.9148 ± 0.0066	0.5954 ± 0.0126	0.8644 ± 0.0103	0.5955 ± 0.0117
		Flotation Tailings Met. Recovery	0 ± 0	0.2589 ± 0.0228	0.7014 ± 0.0268	0.128 ± 0.013
		Leaching Conc. Met. Recovery	0.918 ± 0.0084	0.6556 ± 0.0214	0.8548 ± 0.0199	0.6824 ± 0.0179
		Leaching Tailings Met. Recovery	0.445 ± 0.0304	0.3376 ± 0.0164	0.5713 ± 0.0476	0.3003 ± 0.0449
		Mean	0.731	0.623	0.813	0.598
	Test	Gravimetric Conc. Mass Recovery	0.6176 ± 0.0891	0.4788 ± 0.1866	0.6639 ± 0.1076	0.5528 ± 0.1183
		Flotation Tailings Mass Recovery	0.8128 ± 0.0509	0.7754 ± 0.0813	0.8102 ± 0.081	0.8039 ± 0.0707
		Leaching Tailings Mass Recovery	0.7651 ± 0.056	0.7053 ± 0.0925	0.5438 ± 0.2041	0.7708 ± 0.0743
		Gas Mass Recovery	0.8417 ± 0.0438	0.8265 ± 0.053	0.742 ± 0.1368	0.8366 ± 0.0556
		Gravimetric Conc. Met. Recovery	0.4126 ± 0.1816	0.3117 ± 0.1417	0.1554 ± 0.2025	0.5372 ± 0.1072
		Flotation Tailings Met. Recovery	0 ± 0	0.1709 ± 0.1024	0.4619 ± 0.175	0.1176 ± 0.1471
		Leaching Conc. Met. Recovery	0.5276 ± 0.1667	0.6268 ± 0.1102	0.5216 ± 0.1849	0.6495 ± 0.1144
		Leaching Tailings Met. Recovery	0.4127 ± 0.1265	0.1494 ± 0.0838	0.1817 ± 0.2	0.1119 ± 0.208
		Mean	0.549	0.506	0.510	0.548
MAE	Training	Gravimetric Conc. Mass Recovery	0.0052 ± 0.0001	0.0059 ± 0.0001	0.0038 ± 0.0001	0.0065 ± 0.0002
		Flotation Tailings Mass Recovery	0.0239 ± 0.0011	0.036 ± 0.001	0.0235 ± 0.0013	0.0379 ± 0.0014
		Leaching Tailings Mass Recovery	0.0198 ± 0.0011	0.027 ± 0.0002	0.018 ± 0.001	0.0297 ± 0.0008
		Gas Mass Recovery	0.0051 ± 0.0004	0.0083 ± 0.0002	0.0048 ± 0.0004	0.0079 ± 0.0004
		Gravimetric Conc. Met. Recovery	0.0543 ± 0.0017	0.0884 ± 0.0038	0.0417 ± 0.0025	0.0942 ± 0.0042
		Flotation Tailings Met. Recovery	0.0452 ± 0.0031	0.04 ± 0.0028	0.0192 ± 0.0017	0.0451 ± 0.003
		Leaching Conc. Met. Recovery	0.0427 ± 0.0024	0.0783 ± 0.005	0.0398 ± 0.0016	0.0748 ± 0.0033
		Leaching Tailings Met. Recovery	0.0093 ± 0.0002	0.0081 ± 0.0001	0.0068 ± 0.0001	0.009 ± 0.0001
		Mean	0.026	0.037	0.020	0.038
	Test	Gravimetric Conc. Mass Recovery	0.0068 ± 0.0002	0.0085 ± 0.0004	0.0073 ± 0.0022	0.0067 ± 0.001
		Flotation Tailings Mass Recovery	0.044 ± 0.0045	0.0397 ± 0.0039	0.0461 ± 0.005	0.0398 ± 0.0061
		Leaching Tailings Mass Recovery	0.0377 ± 0.0029	0.0321 ± 0.0016	0.0403 ± 0.0063	0.0302 ± 0.004
		Gas Mass Recovery	0.0084 ± 0.0017	0.0083 ± 0.001	0.0097 ± 0.0024	0.0084 ± 0.0017
		Gravimetric Conc. Met. Recovery	0.1015 ± 0.0158	0.1239 ± 0.0194	0.1244 ± 0.0189	0.096 ± 0.0154
		Flotation Tailings Met. Recovery	0.0486 ± 0.0068	0.0452 ± 0.0117	0.0473 ± 0.0144	0.0501 ± 0.0061
		Leaching Conc. Met. Recovery	0.0807 ± 0.0171	0.0862 ± 0.008	0.0869 ± 0.0091	0.0793 ± 0.0095
		Leaching Tailings Met. Recovery	0.0094 ± 0.0009	0.0112 ± 0.0008	0.0102 ± 0.0012	0.0094 ± 0.0005
		Mean	0.042	0.044	0.047	0.040
RMSE	Training	Gravimetric Conc. Mass Recovery	0.0067 ± 0.0001	0.0083 ± 0.0001	0.0061 ± 0.0002	0.0088 ± 0.0003
		Flotation Tailings Mass Recovery	0.0336 ± 0.0019	0.0525 ± 0.0014	0.0414 ± 0.0013	0.0504 ± 0.0017
		Leaching Tailings Mass Recovery	0.0279 ± 0.0014	0.0421 ± 0.0012	0.0348 ± 0.0014	0.0411 ± 0.0013
		Gas Mass Recovery	0.0076 ± 0.0007	0.0109 ± 0.0002	0.0075 ± 0.0004	0.01 ± 0.0004
		Gravimetric Conc. Met. Recovery	0.0664 ± 0.0023	0.1166 ± 0.0069	0.0709 ± 0.0038	0.1126 ± 0.0056
		Flotation Tailings Met. Recovery	0.0581 ± 0.0023	0.0604 ± 0.0032	0.0429 ± 0.0028	0.0577 ± 0.0023
		Leaching Conc. Met. Recovery	0.0553 ± 0.0032	0.0997 ± 0.0065	0.0651 ± 0.0023	0.093 ± 0.0048
		Leaching Tailings Met. Recovery	0.0111 ± 0.0002	0.0114 ± 0.0003	0.0092 ± 0.0002	0.0107 ± 0.0002
		Mean	0.033	0.050	0.035	0.048
	Test	Gravimetric Conc. Mass Recovery	0.0084 ± 0.0004	0.0111 ± 0.0004	0.0107 ± 0.0037	0.0086 ± 0.0012
		Flotation Tailings Mass Recovery	0.0611 ± 0.0044	0.0554 ± 0.0056	0.0596 ± 0.0067	0.0503 ± 0.0073
		Leaching Tailings Mass Recovery	0.0502 ± 0.0027	0.0473 ± 0.0036	0.0544 ± 0.0071	0.0406 ± 0.0054
		Gas Mass Recovery	0.0111 ± 0.0021	0.011 ± 0.0011	0.0112 ± 0.0025	0.0101 ± 0.0017
		Gravimetric Conc. Met. Recovery	0.1168 ± 0.0143	0.1515 ± 0.0233	0.1715 ± 0.0296	0.1127 ± 0.0194
		Flotation Tailings Met. Recovery	0.0592 ± 0.0097	0.0601 ± 0.0165	0.0614 ± 0.0192	0.0602 ± 0.0092
		Leaching Conc. Met. Recovery	0.0964 ± 0.0197	0.1021 ± 0.0148	0.1146 ± 0.0134	0.095 ± 0.0165
		Leaching Tailings Met. Recovery	0.0112 ± 0.0007	0.0135 ± 0.001	0.0123 ± 0.0017	0.0111 ± 0.0006
		Mean	0.052	0.057	0.062	0.049

The first step to create a neural network is to determine the optimisation algorithm used in its construction. The gradient descent (Bishop, 1995) and the Adam (Kingma and Ba, 2015) algorithms were tested, with the latter producing the best results. Next, the learning rate (Haykin, 1999) needs to be determined. Values between 0.00001 and 0.5 were tested and the best results were obtained with a learning rate of 0.00003. Then, the number of hidden layers and hidden nodes needed to be defined. Bearing in mind the reduced number of samples available, and the possibility of overfitting, it was decided to use no more than three hidden layers and six hidden nodes in each hidden layer. Considering all possible combinations of hidden layers and hidden nodes, the best results were found when two hidden layers with four nodes in the first layer and five nodes in the second layer were used. Finally, it was necessary to determine which activation function will be used in each hidden layer and in the output layer. The following functions were tested: linear, sigmoid, ReLU, hyperbolic tangent, and soft-plus. The best results were obtained when the linear function was chosen in all the layers.

The results obtained using the techniques described earlier were

compared. The k-fold cross validation technique (Hastie, Tibshirani and Friedman, 2009), using $k = 5$, was used to check each model's forecast capacity, in data not used to build them. The techniques were evaluated considering the mean absolute error (MAE), the square root of the mean squared error (RMSE), and the Pearson correlation coefficient between the predicted and the real values. Table 2 consolidates the results presenting the mean and the standard error of the built 5-folds. To make the comparison easier, the mean of the 5-fold mean for the eight dependent variables was calculated and presented.

Analysing the mean values for each metric calculated for all variables in the test set, leads to the result that the best correlation value was obtained when random forest was used, and the best MAE and RMSE values were obtained when the neural networks technique was applied.

The good correlation values found for the gas mass recovery and leaching tailings mass recovery were not impacted by the imputation method used, once when correlations were recalculated considering only the real samples, the test set mean \pm standard error found was 0.8113 ± 0.038 for the gas mass recovery and 0.7594 ± 0.1568 for leaching tailings mass recovery,

values close to the ones obtained when all data (real + imputed) were used.

Considering the neural networks results and thinking that the flotation and leaching tailings, and the gas mass recoveries, together, represent on average 93% of all mass from Mine A, allows one to conclude that correlations above 0.80 for these variables indicates that the models are good to forecast the mass balance. For Mine B, this percent is even higher, 95%. Considering that, on average for Mine A, 90% of all metal goes to the gravimetric and the leaching concentrates, it is possible to consider that the variables with the highest impact on the metallurgical balance, presented models with good results. For Mine B this percentage is 94%. Thus, the leaching and flotation tailings metallurgical recoveries, which are related to models with a poor predictive capacity, are the variables with the smallest metal outputs.

For each of the built 5-folds, the percentage of samples in which the closure error was below $\pm 1\%$ was also evaluated. Next, the average percentage for all 5 folds was calculated. The closure results are in Table 3. Note that the neural networks presented the highest percentage for mass and metallurgical closures. All other techniques presented poor metallurgical closure results.

Table 3 – Closure analysis for all techniques applied.

Set	Variable	RF	Linear SVR	Non-Linear SVR	NN
Training	Mass closure	0.940	0.878	0.778	0.973
	Metallurgical closure	0.243	0.108	0.480	0.946
	Mean	0.591	0.493	0.629	0.960
Test	Mass closure	0.943	0.864	0.679	0.971
	Metallurgical closure	0.114	0.082	0.411	0.836
	Mean	0.529	0.473	0.545	0.904

Using the developed models, it is possible to predict not only the mass and

metallurgical recoveries of each product or tailings, but also their grades, according to:

$$Grade_{(Concentrate\ or\ Tailings)} = \frac{Metallurgical\ Recovery_{(Concentrate\ or\ Tailings)}}{Mass\ Recovery_{(Concentrate\ or\ Tailings)}} * Feed\ Grade \quad (2)$$

Consequently, the gravimetric concentrate, flotation tailings, and leaching tailings grades were calculated using the mass and metallurgical recoveries related to them. However, it

is not possible to calculate the leaching concentrate grade, since it was assumed that the leaching concentrate mass recovery is zero. It was possible to estimate the mentioned grades and to

compare them to the real ones. Table 4 summarises the results obtained when the 5-fold cross validation was applied to the data along with the 5-fold mean for the three studied variables.

Table 4 – 5-fold cross validation results for grades forecast.

Metrics	Set	Variables	RF	Linear SVR	Non-Linear SVR	NN
Correlation	Training	Flotation Tailings Grade	0.6687 ± 0.0522	0.6742 ± 0.0487	0.9422 ± 0.0063	0.6317 ± 0.0526
		Gravimetric Concentrate Grade	0.5256 ± 0.0277	0.4922 ± 0.0158	0.5813 ± 0.024	0.499 ± 0.0202
		Leaching Tailings Grade	0.4499 ± 0.0532	0.5192 ± 0.037	0.6187 ± 0.0375	0.5472 ± 0.045
		Mean	0.548	0.562	0.714	0.559
	Test	Flotation Tailings Grade	0.7663 ± 0.1615	0.7635 ± 0.1647	0.4155 ± 0.3077	0.7548 ± 0.1625
		Gravimetric Concentrate Grade	0.5556 ± 0.0921	0.5168 ± 0.1233	0.5759 ± 0.1042	0.5768 ± 0.1012
		Leaching Tailings Grade	0.291 ± 0.2533	0.3873 ± 0.219	0.3046 ± 0.2601	0.4021 ± 0.2232
		Mean	0.538	0.556	0.432	0.578
MAE	Training	Flotation Tailings Grade	0.5838 ± 0.0626	0.483 ± 0.0504	0.2554 ± 0.0168	0.5416 ± 0.0489
		Gravimetric Concentrate Grade	42.7776 ± 2.9545	43.5283 ± 2.1743	40.3763 ± 3.077	43.254 ± 2.8522
		Leaching Tailings Grade	0.9754 ± 0.0572	0.9792 ± 0.0395	0.8619 ± 0.0394	0.9389 ± 0.0447
		Mean	14.779	14.997	13.831	14.912
	Test	Flotation Tailings Grade	0.6124 ± 0.1598	0.5046 ± 0.1911	0.9951 ± 0.4035	0.596 ± 0.1723
		Gravimetric Concentrate Grade	48.5453 ± 10.4583	43.9584 ± 9.856	57.972 ± 19.4493	43.3076 ± 9.8478
		Leaching Tailings Grade	0.9483 ± 0.1683	1.0151 ± 0.1862	0.9674 ± 0.2019	0.9672 ± 0.166
		Mean	16.702	15.159	19.978	14.957
RMSE	Training	Flotation Tailings Grade	1.0065 ± 0.1826	0.9156 ± 0.1764	0.1869 ± 0.0208	0.9759 ± 0.1785
		Gravimetric Concentrate Grade	8380.52 ± 1350.66	8932.44 ± 1396.28	7565.10 ± 1276.59	8579.40 ± 1369.65
		Leaching Tailings Grade	2.3259 ± 0.2765	2.2284 ± 0.2606	1.803 ± 0.2092	1.9966 ± 0.2233
		Mean	2794.616	2978.528	2522.362	2860.791
	Test	Flotation Tailings Grade	1.1242 ± 0.5827	1.0057 ± 0.7465	4.895 ± 2.7864	1.2006 ± 0.7421
		Gravimetric Concentrate Grade	9201.84 ± 5215.35	8856.74 ± 5510.86	11655.58 ± 7639.38	8311.21 ± 5153.48
		Leaching Tailings Grade	2.3553 ± 1.0408	2.2103 ± 0.7987	2.3743 ± 1.0466	2.0798 ± 0.8861
		Mean	3068.439	2953.320	3887.616	2771.496

Note that all metrics analysed presented the best results for the neural networks prediction models.

3. Discussion

The forecasts obtained by the geometallurgical models presented herein can be used in several ways. The grade information of the flotation or leaching tailings makes it possible for the plant operators to decide the tailings destination. If the tailings grade is above the cut-off grade adopted to feed the plant, it is possible, for example, to store it in piles ready to re-feed it, opening up the possibility to better use the metal content and increase the operation profit. If the grade is below the cut-off, the material must be sent to the dam with the conviction that it will not contain a metal quantity that should belong to the saleable product.

Additionally, this study also provides the means to more precisely sizing the tailings mass sent to the pond (dam) by using the flotation and leaching tailings mass recoveries information. The environmental impact generated by the tailings can

be determined even before the tailings are generated, making it possible to better plan its disposal. Moreover, it is possible to know the percentage of all metal feeds such that their destination to the flotation and leaching tailings, and to the concentrate, make it possible to study alternative processing routes and different dosing reagents, aimed at increasing the metallurgical recovery.

Another gain is related to the gas mass that is sent to the sulphuric acid plant. Knowing that this gas is sold to other companies, it is important to know beforehand the profit that will be generated by selling this material. When the gravimetric concentrate grade and mass recovery are known, it is also possible to predict the operating profit, making the predicted cashflow more adherent to the reality.

After presenting all model benefits, it is necessary to discuss their limitations.

The first one is related to the model boundary conditions. It is important to apply the model in blocks (or piles) which kept the same grade range and the same lithologies used to build the neural networks. The second limitation is associated to the relation between the bench lab tests and the plant results. The bench tests led to higher metallurgical recovery than the plant, a fact related to the higher selectivity of the equipment used at laboratory scale. But the company was not able to quantify this difference at the moment. It is known the need to find an up-scaling factor to make the lab results comparable to the plant actual response. The last limitation is related to the number of samples used to build the model. Collecting and inserting more samples in the study will reduce the MAE and RMSE values found, making the model's predictions more adherent to the reality.

4. Conclusions

The results indicated that, in terms of MAE and RMSE, the neural networks technique stood out when compared with the other techniques applied. The mass and metallurgical closure results also indicate that this technique presents adequate re-

sults to be used in actual situations. Even though, it is important to note that this superiority is related to the specific data studied, being necessary to check its validity when data from other deposits are used.

The developed neural network model

can be used to build 3D geometallurgical block models using the Au and S predicted at each block, and the mine which the ore came from. The same applies to piles built to feed the plant, making it possible to forecast the metallurgical ore behaviour before

the material actually feeds the process, bringing several benefits to the operation.

Analysing all the presented infor-

mation, it is possible to conclude that the development of geometallurgical models capable of predicting products and tail-

ings characteristics can bring financial and environmental benefits, becoming a good practice to be used at mines.

Acknowledgment

This study was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) [grant number 141594/2017-9]. The

authors would like to thank the geologist Mariana Lemos Gazire for all the assistance provided during this study and the company who provided the data

and permitted the publication of this article. The authors are also grateful for the support provided by Fundação Luiz Englert.

References

- BISHOP, C. M. *Neural networks for pattern recognition*. Oxford: Clarendon Press, 1995.
- CARRASCO, P.; CHILÈS, J. P.; SÉGURET, S. Additivity, metallurgical recovery and grade. *In: INTERNATIONAL GEOSTATISTICS CONGRESS*, 8., 2008, Santiago, Chile. *Proceedings* [...]. Santiago, Chile: Geostats, 2008. p. 237-246.
- COWARD, S.; VANN, J.; DUNHAM, S.; STEWART, M. The primary-response framework for geometallurgical variables. *In: INTERNATIONAL MINING GEOLOGY CONFERENCE*, 7., 2009, Perth, WA. *Proceedings* [...]. Melbourne: Australasian Institute of Mining and Metallurgy, 2009. p. 109-113.
- DRUMOND, D. A. *Estimativa e classificação de variáveis geometalúrgicas a partir de técnicas de aprendizado de máquinas*. 2019. 205 f. Tese (Doutorado em Engenharia de Minas, Metalurgia e de Materiais) – Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019.
- ENDERS, C. K. *Applied missing data analysis*. New York: The Guilford Press, 2010.
- FERNANDES, F. G.; CABRAL, I. E. Regression model utilization to estimate the mass recovery of a phosphate mine in Brazil. *REM – Rev. Esc. Minas*, Ouro Preto, v. 69, n. 1, p. 75-77, 2016.
- FRANCO-SEPÚLVEDA, G.; DEL RIO-CUERVO, J. C.; PACHÓN-HERNÁNDEZ, M. A. State of the art about metaheuristics and artificial neural networks applied to open pit mining. *Resources Policy*, v. 60, p. 125-133, 2018.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The elements of statistical learning*. New York: Springer-Verlag, 2009.
- HAIR, J. F.; ANDERSON, R. E.; TATHM, R. L.; BLACK, W. C.; BABIN, B. J. *Análise multivariada de dados*. 6. ed. New Jersey: Prentice Hall, 2009.
- HAYKIN, S. *Neural networks: a comprehensive foundation*, 2nd. ed. New Jersey: Prentice Hall, 1999. 823p.
- IZENMAN, A. J. *Modern multivariate statistical techniques*. New York: Springer, 2008.
- KINGMA, D. P.; BA, J. L. Adam: a method for stochastic optimization. *In: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS*, 3., 2015, San Diego. *Proceedings* [...]. [S. l.]: ICLR, 2015. p. 1-13.
- LISHCHUK, V.; LUND, C.; GHORBANI, Y. Evaluation and comparison of different machine-learning methods to integrate sparse process data into a spatial model in geometallurgy. *Minerals Engineering*, v. 134, p. 156-165, 2019.
- LUCAY, F. A.; SALES-CRUZ, M.; GÁLVEZ, E. D.; CISTERNAS, L. A. Modeling of the complex behavior through an improved response surface methodology. *Mineral Processing and Extractive Metallurgy Review*, p. 1-27, Feb. 2020.
- MACMILLAN, E.; EHRIG, K.; LIEBEZEIT, V.; KITTLER, P.; LOWER, C. Use of geometallurgy to predict tailings leach acid consumption at olympic dam. *In: AUSIMM INTERNATIONAL GEOMETALLURGY CONFERENCE*, 1., 2011, Brisbane, Queensland. *Proceedings* [...]. Melbourne: Australasian Institute of Mining and Metallurgy, 2011. p. 93-102.
- MCCOY, J. T.; AURET, L. Machine learning applications in minerals processing: a review. *Minerals Engineering*, 132, p. 95-109, 2019.
- MONTOYA, P. A.; KEENEY, L.; JAHODA, R.; HUNT, J.; BERRY, R.; DREWS, U.; CHAMBERLAIN, V.; LEICHLITER, S. Techniques applicable to prefeasibility projects: La Colosa case study. *In: AUSIMM INTERNATIONAL GEOMETALLURGY CONFERENCE*, 1., 2011, Brisbane, Queensland. *Proceedings* [...]. Melbourne: Australasian Institute of Mining and Metallurgy, 2011. p. 103-111.
- PANDA, L.; TRIPATHY, S. K. Performance prediction of gravity concentrator by using artificial neural network - a case study. *International Journal of Mining Science and Technology*, v. 24, p. 461-465, 2014.
- PRADES, C. F.; DEUTSCH, C. V. Comparison of machine learning techniques for predicting and learning from geometallurgical multivariate databases. *In: CENTRE for Computational Geostatistics Annual Report 18*. Alberta, Canada: CCG, 2016. 20 p. (Paper 308).
- RICHMOND, A.; SHAW, W. J. Geometallurgical modelling – Quo Vadis? *In: INTERNATIONAL MINING GEOLOGY CONFERENCE*, 7., 2009, Perth, WA. *Proceedings* [...]. Melbourne: Australasian Institute of Mining and Metallurgy, 2009. p. 115-118.
- RUBIN, D. B. Inference and missing data. *Biometrika*, v. 63, n. 3, p. 51-592, 1976.
- SINGH, V.; RAO, M. Application of image processing and radial basis neural network techniques for ore sorting and ore classification. *Minerals Engineering*, v. 18, p. 1412-1420, 2005.
- TRIOLA, M. F. *Introdução à estatística*. 10. ed. Rio de Janeiro: LTC, 2008.
- VIEIRA, M. C. A.; COSTA, J. F. C. L. Geometallurgical modelling to help in predicting zinc metallurgical recovery. *In: WORLD MINING CONGRESS*, 24., 2016, Rio de Janeiro. *Proceedings* [...]. Rio de Janeiro: IBRAM, 2016. Theme: Mineral Exploration. p. 44-54.

Received: 1 January 2020 - Accepted: 22 June 2020.



All content of the journal, except where identified, is licensed under a Creative Commons attribution-type BY.