Physicochem. Probl. Miner. Process., 60(4), 2024, 191709

http://www.journalssystem.com/ppmp

Physicochemical Problems of Mineral Processing

ISSN 1643-1049

Exploring the impact of thiol collectors system on copper sulfide flotation through machine learning-driven modeling

Mustafa K. Guner, Ozge Akyildiz, Hakan Basarir, Pshem Kowalczuk

Department of Geoscience and Petroleum, Faculty of Engineering, Norwegian University of Science and Technology, Trondheim, 7031, Norway

Corresponding author: mustafa.k.guner@ntnu.no (Mustafa K. Guner)

Abstract: Collector selection is a critical step in flotation, as it has a direct impact on product quality, flotation recovery, and selectivity. Collectors can consist of different components, and their effectiveness can vary depending on the type of ore being processed. The general practice in both literature and in industry is to use a mixture of collectors rather than a single collector. However, the use of a collector mixture introduces several complex issues. It is challenging to determine the specific effects of each collector on different minerals, as well as to understand the synergistic effects of mixed collectors in flotation. This study presents a novel investigation focusing on the impact of blends of NAX, AEROPHINE® 3422, and AERO® MX 5149, in varying dosages and combinations, on the flotation performance of Kupferschiefer copper ore. Kinetics flotation tests were conducted using a mechanical flotation cell with various combinations and dosages of listed collectors. For this investigation, different predictive models such as machine-learning (ML) and conventional regression analyses were developed. For model construction, a database including the results of comprehensive experimental results was constructed. The best performing model was selected considering statistical performance indicators and their performance on unseen data. A sensitivity analysis was conducted on the model to justify contributions of collectors on the copper recovery and grade. The results showed that the ML-based models provide compatible results with the expert opinions and have higher statistical performance than conventional modelling tools. According to the experimental results and models' findings, it has shown that AEROPHINE® 3422 (a blend of isopropyl ethyl thionocarbamate and dithiophosphinate) was the most influential collector for the copper recovery. In addition, two ternary graphs were generated from the modeled data to formulate mixtures for different grades and recovery priorities.

Keywords: machine learning, thiol collector, copper flotation, random forest, dosage optimization

1. Introduction

Various factors affect a flotation system, such as mineralogy (i.e., ore type, mineralogical composition, liberation degree), operational (i.e., flow rates of feed, gas tailings, froth level, slurry density, circuit design), and chemical (i.e., types and dosages of flotation reagents) factors (Klimpel and Hansen, 1988; Drzymala, 2007; Kawatra, 2011; Wills, 2016). It is important to control all these variables during a flotation process, and metallurgists must deal with the uncertainty created by these variables. Chemistry plays a central role in flotation systems (Klimpel and Hansen, 1988). Flotation reagents fall into three main categories: collectors, frothers, and modifiers. Collectors, in particular, include a wide range of organic chemical compounds with different chemical functionalities (Bulatovic, 2007). The primary objective of collectors is to create a water-repellent surface on targeted mineral particles within the slurry, enabling proper interaction between hydrophobic particles and air bubbles to transfer these particles to a froth phase (Drzymala, 2007; Kowalczuk et al., 2015). The most widely used collectors in copper sulfide flotation are xanthates, thionocarbamates, and dithiophosphates, (Bulatovic, 2007; Fuerstenau et al., 2007; Drzymala, 2007; Wills, 2016; Nagaraj and Farinato, 2016). All these collector groups have different strength and selectivity properties on different mineral surfaces, as well as

DOI: 10.37190/ppmp/191709

different prices and environmental impacts. Many studies in the literature acknowledge that using different combinations of these collectors improves the flotation efficiency, even though the mechanisms between collectors are not fully explained in real-ore flotation apart from single collector-single mineral systems (Rao and Forssberg, 1997; Sheridan et al., 2002; Lotter and Bradshaw, 2010; Guner et al 2024). Selection of a type of collector and optimization of a dosage are crucial for flotation outcomes in both industrial and batch scales (Nagaraj, 2005). In batch scales, determining the collector with the highest performance can serve as both an early study for a resource project and an improvement for current flotation plants, leading to the maximization of recovery, grade, kinetics, and selectivity of the process (Greet, 2010; Thompson, 2016).

The general methodology for selecting the appropriate collector involves characterizing the feed material based on its mineralogical, chemical, and physical properties. Lab-scale flotation tests should be designed and conducted using different types of selected reagents based on the mineralogical composition of a feed material (Nagaraj and Ravishankar, 2007). Different dosages of collectors should be tested to achieve the highest recovery and grade without overdosing the system (Nagaraj, 2005). Statistical tools can be employed to identify trends and investigate the relationship between collector types, dosages, and recovery. Given the complex nature of the flotation process, additional statistical methods and advanced modeling approaches have been utilized to assist metallurgists in decision-making (Gholami et al., 2022). Design of an Experiment (DoE) is one of the most common tools for reagent optimization (Vazifeh et al., 2010; Napier-Munn, 2014; Corpas-Martínez et al., 2019; Azizi et al., 2020). Arancibia-Bravo et al. (2022) have compiled a comprehensive list of 24 Response Surface Methodology (RSM)-based DoE studies involving various experimental methods and ores. The main limitation of DoE is its lack of flexibility to accommodate external interventions. Initial levels of parameters must be predetermined, meaning that all parameter levels must be established before conducting the experiments, but it significantly reduces number of experiments. Machine learning (ML) models offer more applicability due to their flexibility in data collection. The construction of a consistent database is crucial for model development. Typically, these data are gathered during industrial activities (Saravani et al., 2014; Allahkarami et al., 2017), and some experimental databases are also suitable for running ML models for flotation process (Cook et al., 2020; Çilek, 2002). Furthermore, ML-based models allow the addition of new data to modify and upgrade an existing model in order to extend the data range and increase the accuracy. Table 1 presents the examples of modeling studies using ML-based models for copper flotation. It is important to emphasize that neither DoE nor ML models aim to explain mechanisms at a fundamental level, even though they are powerful tools for prediction and optimization.

Table 1. Literature summary of machine learning-based models used in copper sulfide flotation.

Algorithm	Inputs	Outputs	Scale	References
Artificial Neural Networks (ANN)	pH, pulp density, liberation degree and particle size distribution of feed and middling, grade of feed, flotation times	Grade and mass pull of concentrate and tailing	Batch	(Çilek, 2002)
Artificial Neural Networks (ANN) and Multivariate Non-Linear Regression (MNLR);	Froth height, gas hold-up, frother dosage, collector dosage, wash water flow rate, air flow rate, and grades of products.	Grade and recovery	Pilot	(Nakhaei, et al., 2012)
Fuzzy model	Gas velocity, pulp density, frother dosage, and type	Concentrate grade and recovery	Industrial	(Saravani et al., 2014)
Neural Networks	Feed grade, collector dosage, frother dosage, particle size distribution, pulp density	Recovery and enrichment ratio	Industrial	(Massinaei et al., 2014
Neural Networks	Gas flow rate, slurry density,	Recovery and	Batch	(Jahedsaravani

	frother dosage, collector dosage, pH	grade, mass pull, mean bubble size		et al., 2014)
Hybrid genetic algorithm – artificial neural network (GA-ANN)	pH, collector and frother dosage, fuel oil dosage, slurry density, moisture percent, particle size distribution, feed grades	Recovery and grade	Industrial	(Allahkarami et al., 2017)
Multilayer perceptron ANN; support vector machine (SVM); M5Prime model tree algorithm; RF and the firefly algorithm.	Collector dosage, frother dosage, depressant types and dosage, pH, impeller speed, air flowrate	Recovery and grade	Batch	(Cook et al., 2020)

The primary aim of this study was to evaluate how the flotation efficiency of Kupferschiefer copper ore was influenced by the type and amount of collectors used, specifically NAX (a mixture of sodium isobutyl xanthate and sodium ethyl xanthate), AEROPHINE® 3422 (a mixture based on isopropyl ethyl thionocarbamate and dithiophosphinate) and AERO® MX 5149 (a mixture based on reaction mass of S-allyl O-(2-methyl butyl) dithiocarbonate and S-allyl O-pentyl dithiocarbonate, and n-butoxycarbonyl-O-n-butyl thionocarbamate. Additionally, the study sought to understand effects of mixtures of these collectors. This study deliberately limited the operational and equipment-based parameters of a flotation system, focusing solely on collectors. Other chemical parameters, such as pH, Eh, frother type and dosage, were not included in the study to remain consistent with the processing conditions at the Polkowice Copper Concentrator (KGHM Polska Miedz S.A.), where the material came from. A secondary objective was to explore the application of machine learning modeling to flotation studies and compare its efficacy with conventional regression analyses, within the context of a narrowly defined scope of parameters (specifically, the collector system for this study). It is worth mentioning that this study did not aim to compare DoE with ML approaches. All models were developed using existing data.

2. Experimental studies

2.1. Material and reagents

This research utilized a copper sulfide ore sourced from the Kupferschiefer deposit, obtained from the Polkowice-Sieroszowice Copper Plant operated by KGHM Polska Miedz S.A, Poland. Elemental analyses of feed, concentrate, and tailings were performed using Inductively Coupled Plasma Mass Spectrometry (ICP-MS) at both ALS Scandinavia and the Chemical/Mineralogical Laboratory (IGP, Department of Geoscience and Petroleum, NTNU), Portable X-ray Fluorescence Analyzers (pXRF, Thermal Science, XL3t) at NTNU Mineral Processing Laboratory and LECO - Infrared Spectroscopy (IR) at ALS Scandinavia. The elemental composition of the feed is presented in Table 2.

Table 2. Elemental analyses of feed material with analytical tools

Element	Content	Unit	Method
Cu	1.8	%	HNO_3 digestion followed by ICP-MS
Ag	40	ppm	HNO_3 digestion followed by ICP-MS
Fe	1.1	%	HNO_3 digestion followed by ICP-MS
Pb	1455	ppm	Four acid digestion followed by ICP-AES
Zn	400	ppm	Four acid digestion followed by ICP-AES measurement
Ni	43	ppm	Four acid digestion followed by ICP-AES measurement
As	119	ppm	Aqua regia digestion followed by ICP-MS measurement
Si	11.1	%	HNO_3 digestion followed by ICP-MS
S	1.3	%	Total sulfur (IR Spectroscopy) by LECO
C	6.1	%	Total carbon (IR Spectroscopy) by LECO

A laboratory-scale ball mill was used for wet grinding, with a solid ratio of 55% by weight, to reduce the particle size. The particle size distribution of the ground material was then obtained using the Malvern Panalytical Mastersizer 3000, as shown in Fig. 1. The size of the feed material decreased from a d_{80} of 1.3 mm to a d_{80} of 45 µm after grinding. Additionally, as shown in Fig. 1, the d_{50} and d_{30} sizes of the ground material were approximately 15 µm and 8 µm, respectively.

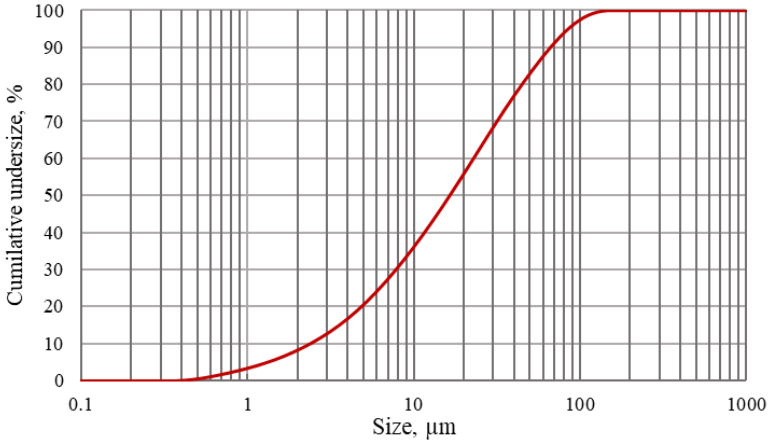

Fig. 1. Particle size distribution of ground material for flotation test

The following chemicals were used as collectors: AEROPHINE® 3422 PROMOTER (a blend of isopropyl ethyl thionocarbamate and dithiophosphinate), AERO® MX-5149 PROMOTER (reaction mass of S-allyl O-(2-methyl butyl) dithiocarbonate, S-allyl O-pentyl dithiocarbonate, and n-Butoxycarbonyl-O-n-butyl thionocarbamate), and NAX (a blend of sodium isobutyl xanthate and sodium ethyl xanthate). AEROPHINE® 3422 PROMOTER and AERO® MX-5149 PROMOTER were provided by Solvay S.A., and NAX (Minova-Ksante, Sp. z o.o,) was provided by KGHM Polska Miedz S.A. Additionally, Nasforth 245 (a type of polyethylene glycol butyl ether with the formula $C_4H_9(C_2H_5O)_nOH$, where n ranges from 2 to 5), and a frother from NASACO (provided by KGHM Polska Miedz S.A.), were also used in the flotation experiments.

2.2. Flotation tests

Kinetic flotation tests were carried out using an MMS (Maelgwyn Mineral Services) mechanical laboratory cell (1 dm^3) over specific intervals, with concentrates being collected at the 1st, 3rd, 7th, 15th, and 30th min. A pulp level was maintained consistently using wash water that had the same concentration of frother by volume with the initial slurry. The weights of the products were recorded, and the copper (Cu) content was analyzed using portable X-ray fluorescence (pXRF). The recovery for each time interval was calculated using Eq. 1:

$$R = \frac{Cc}{Ff}\% \tag{1}$$

where 'C' represents a concentrate weight, 'c' the Cu content in the concentrate, 'F' the final weight of all products, and 'f' the calculated head grade of Cu content in all products. The initial solid ratio of the slurry was fixed at 30% and prepared using tap water. The material's own pH (between 7.9 and 8.1) was consistently maintained for each test. The pH levels of the slurry were continually monitored using a WTW ProfiLine pH 3110 Portable Meter. The initial dosage of the frother for both the slurry and wash water was set at 15 mg/dm^3. The constant conditions applied across all tests are detailed in Table 3. The collectors used were labeled as NAX (R1), AEROPHINE® 3422 (R2), and AERO® MX-5149 (R3). Details about the combinations and dosages of these collector mixtures are provided in Table 4. The initial dosages of each reagent were determined based on a literature review and prior experimental work conducted at the NTNU Mineral Processing Lab. The first 20 tests of the experimental work were taken from a previous study that included a Central Composite Design (CCD), an example of a Design of Experiment (DoE) application for mixtures (D1-D20 in Table 4). This included six repetitions at the central levels of the factors (D1, D3, D4, D8, D14, D17). Subsequently, additional tests were added with

varying dosages and combinations (D21-42). This study does not to focus on optimization through CCD, and thus discussion on either superiority or limitations of CCD was not included.

Table 3. Constant conditions

Flotation time	1, 3, 7, 15 and 30 minutes
Conditioning time	3 minutes for collector, 1 minute for frother
Frother dosage	15 mg/dm^3
Agitator speed	1200 rpm
Solid ratio of pulp	30%
Cell volume	1 dm^3
Gas flowrate	4 dm^3/min
pH	7.9-8.1

Table 4. Collector recipes

Experiment code	Collector dosage, g/Mg			Experiment code	Collector dosage, g/Mg		
	R1	R2	R3		R1	R2	R3
D1	20	10	10	**D22**	10	0	0
D2	30	5	15	**D23**	20	0	0
D3	20	10	10	**D24**	40	0	0
D4	20	10	10	**D25**	120	0	0
D5	30	15	5	**D26**	0	10	0
D6	10	5	5	**D27**	0	20	0
D7	30	5	5	**D28**	0	40	0
D8	20	10	10	**D29**	0	120	0
D9	10	15	15	**D30**	0	0	10
D10	30	15	15	**D31**	0	0	20
D11	10	15	5	**D32**	0	0	40
D12	10	5	15	**D33**	0	0	120
D13	20	10	0	**D34**	20	20	0
D14	20	10	10	**D35**	20	0	20
D15	20	10	20	**D36**	0	20	20
D16	20	20	10	**D37**	10	0	30
D17	20	10	10	**D38**	30	0	10
D18	40	10	10	**D39**	0	10	30
D19	0	10	10	**D40**	0	30	10
D20	20	0	10	**D41**	10	30	0
D21	0	0	0	**D42**	30	10	0

2.3. Flotation results

The results of single collector tests (D22 to D33) are shown in the grade-recovery and kinetics curves for three reagents: R1: NAX (Fig. 2), R2: AEROPHINE® 3422 (Fig. 3), and R3: AERO® MX-5149 (Fig. 4). D21 presents the results of the flotation test using only frother without a collector (collectorless flotation). The maximum recovery of Cu achieved was 44% with a Cu grade of 2.8% at the end of the 30-minute flotation period without a collector. The highest Cu content in this experiment was 5.8%, which was reached within 3 minutes. Thus, it can be concluded that flotation was not successful without a collector, but the floated material, which can be classified as natural hydrophobic, still carried low amounts of copper content. It could be a result of either entrainment or flotation of copper rich carbonaceous matter

As shown in Fig 2., the use of 10 and 20 g/Mg of R1 (a mixture of sodium isobutyl xanthate and sodium ethyl xanthate resulted in relatively low flotation performance presented in the grade-recovery and kinetics curves. An increase in the collector's dosage led to faster kinetics as well as higher recovery and grade. The highest results were achieved with a dosage of 120 g/Mg, resulting in 92% Cu recovery after 30-minute of flotation. Furthermore, there was a significant difference between Cu grades in the first concentrates at 40 g/Mg and 120 g/Mg, which were 15% and 22%, respectively, although the kinetics curves were similar.

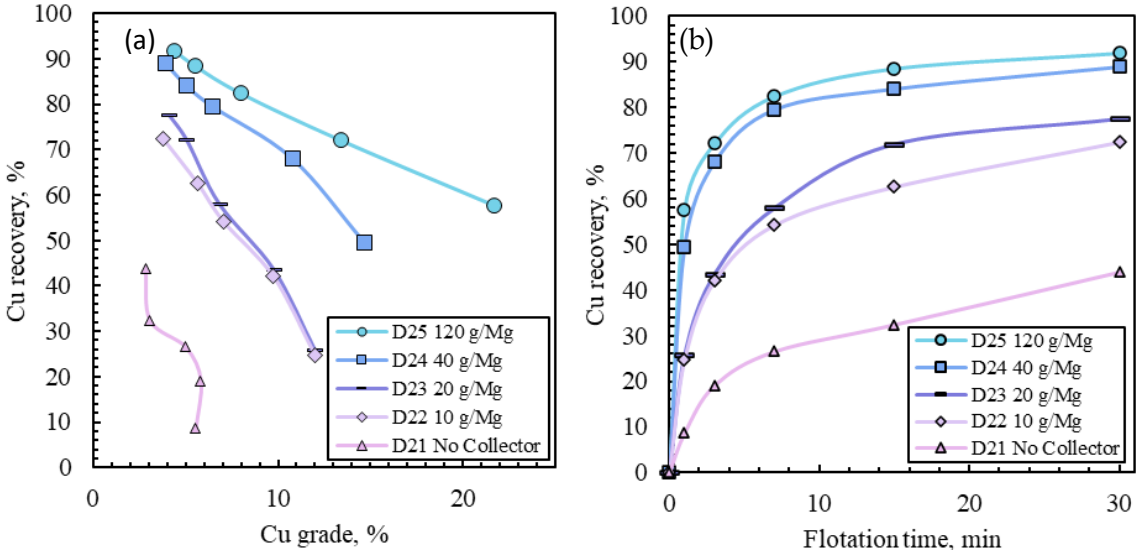

Fig. 2. (a) Grade-recovery and (b) kinetics curves of flotation tests with using NAX (R1)

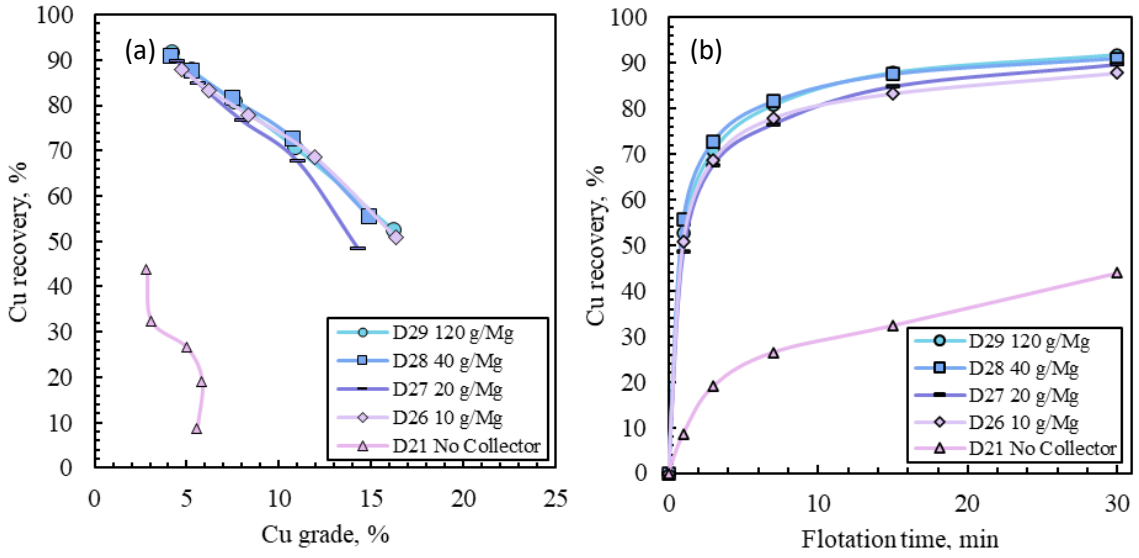

Fig. 3. (a) Grade-recovery and (b) kinetics curves of flotation tests with using AEROPHINE® 3422 (R2)

In the single collector flotation system, R2 (AEROPHINE® 3422, thionocarbamate and sodium dithiophosphinate mixture) showed better performance than the other two reagents in low dosages. According to Fig. 3, the result of the test using 10 g/Mg R2 resulted in the concentrate containing almost the same amount of Cu (ca. 16%) in the 1st minute of flotation as the test using 120 g/Mg R2. All the grade-recovery curves of different dosages of R2 are very close to each other. As expected, higher dosages (40 and 120 g/Mg) resulted in slightly faster kinetics.

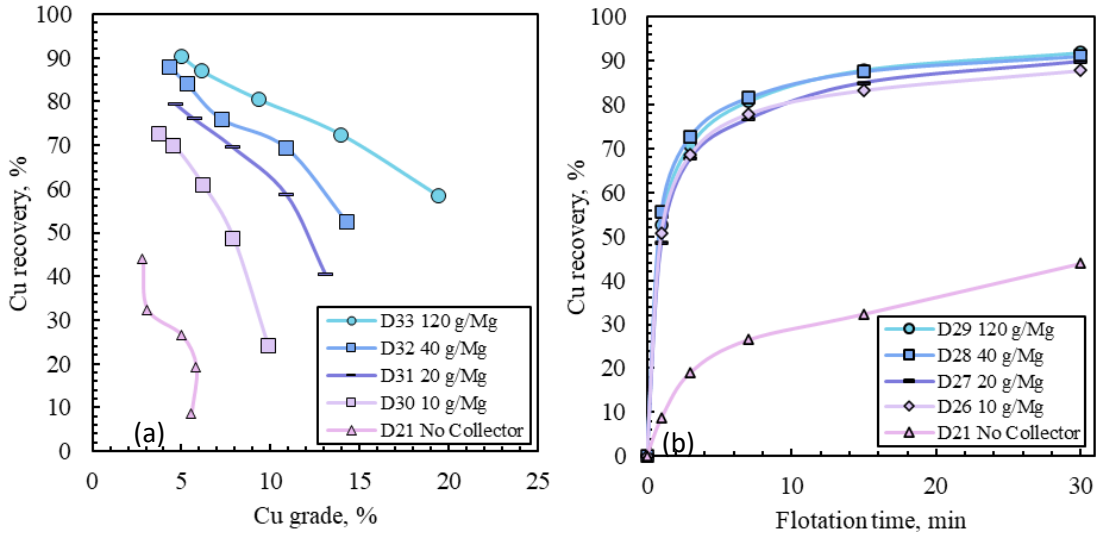

Fig. 4. (a) Grade-recovery and (b) kinetics curves of flotation tests with using AERO® MX-5149 PROMOTER (R3)

AERO® MX-5149 PROMOTER (R3) had quite similar results with R1. 10 and 20 g/Mg were not enough dosages to perform an average flotation performance. As the dosage increased, grade, recovery, and kinetics all increased significantly. The performances of D32 (40 g/Mg) and D33 (120 g/Mg) were much higher than D30 and D31. It was also observed that kinetics curves reached a plateau after 15 mins. The grade-recovery curve shifted from a vertical to a more horizontal position depending the on increase in dosages.

In addition to single reagents, a binary collector system was also tested using proportions of 100:0, 75:25, 50:50, 25:75, and 0:100 for blends of R1-R2, R1-R3, and R2-R3. The results of these tests are shown in Fig. 5. This comparison was performed at 40 g/Mg because of the large number of data for this dosage. According to the results, the binary collector system generally resulted in better outcomes than the single collector system with the same dosage. In terms of the mixture of R1 and R2, the mass ratio that appeared to provide a good balance of high grade and high recovery was at 75:25. At this ratio, both grade and recovery were relatively high after 30 min of flotation. However, it is important to note that the results were very close to each other, indicating that experimental error played a significant role in these tests, especially over longer flotation times. The most optimal combination for R2 and R3 was 50:50 in terms of both grade and recovery, but the results became more similar for this blend over extended flotation times, similar to R1:R2. An opposite trend between grade and recovery was observed for the mixture of R1 and R3. The highest recovery was obtained with a 25:75 ratio of R1:R3, while the situation was completely reversed for grade, with 75:25 obtaining the highest grade. The highest recoveries for the 50:50 mixture of R2 and R3 were 57%, 72%, 83%, 88%, and 92% for 1, 3, 7, 15, and 30-minute flotation intervals, respectively. The corresponding copper grades were 24%, 16%, 8%, 6%, and 5%, respectively.

Similar to the flotation experiments with binary collector mixtures, using mixtures of three reagents yielded more favorable outcomes than utilizing single collectors. Overall, the highest recoveries and grades for the first three concentrates were achieved in tests with a ternary collector mixture, specifically in D10 (30 g/Mg R1, 15 g/Mg R2, and 15 g/Mg R3) and D16 (20 g/Mg R1, 20 g/Mg R2, and 10 g/Mg R3). Since 15 and 30 minutes represent very long flotation times under the conditions introduced in Table 3, it was more straightforward to analyze the differences in results caused by the collectors using the first three concentrates. The trend of higher dosages correlating with higher recovery and grade was

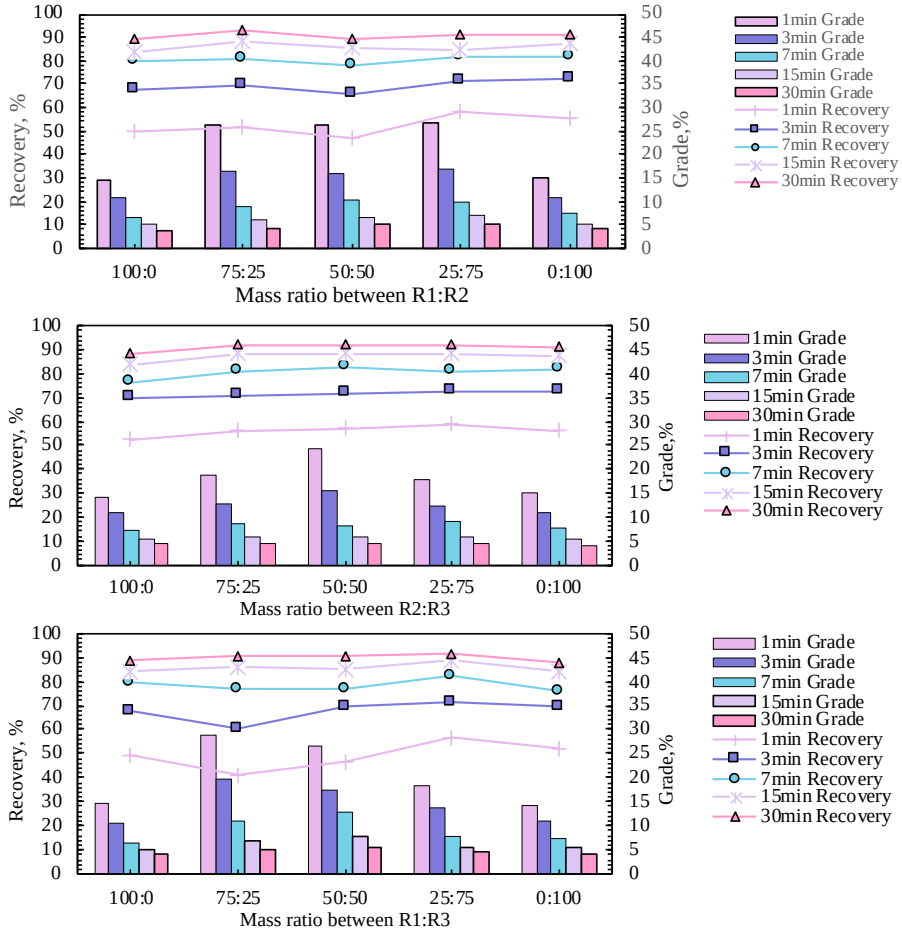

Fig. 5. Cu recovery and grade of binary collector mixture by time (R1: NAX, R2: AEROPHINE® 3422, R3: AERO® MX-5149 PROMOTER, the total dosage of each blend is 40 g/Mg)

Table 5. Results of experiments conducted with ternary mixtures (for recipes see Table 4) (Rec.: Cu-recovery, %)

Experiment code	1 min		3 mins		7 mins		15 mins		30 mins	
	Rec. %	Grade %	Rec. %	Grade %	Rec. %	Grade %	Rec. %	Grade %	Rec. %	Grade %
D1	54.9	20.6	71.8	13.4	79.8	9.2	85.9	6.4	90.2	4.8
D2	57.5	21.5	70.7	15.3	79.8	10.1	85.7	6.9	89.8	5.2
D3	53.6	20.7	68.4	14.2	77.9	9.9	83.5	7.2	88.8	5.6
D4	52.8	14.7	67.7	14.7	79.4	9.1	85.9	5.0	89.7	5.0
D5	57.3	22.0	72.5	15.6	79.7	11.3	86.2	7.5	90.6	5.5
D6	45.4	22.2	62.5	15.3	75.7	9.0	83.5	6.3	87.9	4.9
D7	51.1	23.7	66.0	16.0	76.1	11.2	84.0	7.0	88.7	5.4
D8	53.2	21.0	70.0	15.3	78.6	11.0	85.9	7.0	89.9	5.2
D9	58.8	21.2	71.0	15.1	81.4	9.2	87.3	6.7	90.5	5.5
D10	59.2	22.9	76.4	15.5	82.8	10.4	87.6	7.2	91.4	5.5
D11	53.1	21.4	71.6	14.4	79.6	9.5	86.1	6.5	90.2	5.0
D12	57.0	19.2	70.5	14.8	79.2	10.1	86.4	6.5	89.8	5.3
D14	53.0	20.0	70.6	12.9	80.8	9.1	85.7	6.5	90.7	4.8
D15	57.2	22.0	71.1	14.9	78.0	10.4	85.6	6.9	89.6	5.4
D16	60.0	21.9	73.0	15.5	82.6	9.4	87.9	6.9	91.5	5.3
D17	56.2	21.6	70.6	13.7	81.0	8.4	86.7	5.9	90.9	4.5
D18	58.2	23.3	74.8	15.9	82.3	10.2	88.0	7.0	91.5	5.5

consistent for the ternary mixture, similar to the singular and binary mixture systems. However, it was observed that achieving the same or higher grade and recovery was possible with a lower total dosage compared to the binary collector systems. For example, D17 (20 g/Mg R1, 10 g/Mg R2, and 10 g/Mg R3, totaling 40 g/Mg) achieved approximately 56% Cu recovery with a grade of about 22% in the first concentrate, whereas D30 (40 g/Mg R3) had a Cu content of approximately 15% and a recovery of about 53%. It was also noted that the recovery increased for the ternary mixture, even though Cu grades were higher than the single collector system.

Overall, the results showed that binary and ternary collector systems were generally better than the single collector resulting in improved grade, recovery. It is also necessary to highlight that the long flotation times maximized the recovery and made the evaluation of the results very difficult. Therefore, grade and recovery in the 1st, 3rd, and 7th minutes, which is for the first three concentrates, gave more meaningful results in order to evaluate the effects of the collector.

3. Modeling

3.1. Data preparation and model selection

Inputs are defined as the amounts of the reagents (R1, R2, R3) and time (T), while the outputs of the models are defined as Cu grade and recovery. The statistical descriptions of input variables are presented in Table 6. Due to the lack of amount of data at certain levels, such as 120 g/Mg for experiments D25, D29, and D33, these experiments were excluded during the data preparation stage before modeling. The Cu-recovery % (Eq. 1) and Cu-grade % were assigned as outputs, and the structure of dataset is shown in Table 7.

Table 6. Descriptive statistic of inputs

Variable	R1	R2	R3	T
Total Count	190	190	190	190
St. Dev	11.9	9.3	9.3	10.6
Variance	141.3	87.2	87.2	111.9
Min	0.0	0.0	0.0	1.0
Q1	0.0	0.0	0.0	3.0
Median	20.0	10.0	10.0	7.0
Q3	20.0	15.0	15.0	15.0
Max	40.0	40.0	40.0	30.0

Table 7. Dataset structure

	Inputs				Outputs	
	R1 (n=1)	R2 (n=2)	R3 (n=3)	T (n=4)	Recovery (Cu)	Grade (Cu)
1	↓	↓	↓	↓	↓	↓
2						
3						
.						
.						
.						
m						

m=190, n=4
mxn = [190,4]

3.1.1. Linear (LR) and nonlinear (NLR) regressions

Linear regression is a technique used to model the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the observed data. When the relationship between variables is too complex to be captured by a linear equation, non-linear regression is employed to model these intricate relationships (Bingham et al., 2010). Linear regression is widely used in various

fields such as economics, biology, engineering, and social sciences to understand and predict relationships between variables and it serves as a foundational tool for more complex statistical models (Gareth et al., 2021). However, its effectiveness relies on a critical assumption stating that the relationship between the predictor(s) (independent variables) and the response (dependent variable) must be approximately linear (Bingham et al., 2010).

3.1.2. Random Forest (RF)

The RF algorithm is based on the concept of 'ensemble learning,' which involves combining multiple decision trees to improve the overall performance. Introduced by Breiman (2001), the RF algorithm is versatile and can be used for both classification and regression tasks. It handles both categorical and continuous input features (independent variables). While the mathematical foundation of a single decision tree is straightforward, the RF algorithm is often considered a 'black box' model due to its complexity, involving a large number of trees (Hastie et al., 2009). The RF algorithm constructs a function $f(x)$ from a set of 'base learners' $h1(x), \ldots, hj(x)$ which are then combined to form the 'ensemble predictor' $f(x)$. For regression problems, the final output is obtained by averaging the predictions from each tree in the model. The model's error can be evaluated using various criteria, but the mean squared error (MSE) is most commonly used. The model is trained until the MSE reaches its minimum value. In classification problems, predictions are determined by a majority vote. The RF model is trained until the splitting criterion, such as misclassification error, Gini index, or entropy, reaches its minimum (Cutler et al., 2012). Several hyperparameters can be adjusted to build a successful RF model. In this study, no changes were made to the hyperparameters of the RF model for several reasons:

Firstly, the model's default configuration has already produced useful and satisfactory results, indicating that the current hyperparameter settings are effective for this task. Secondly, optimizing hyperparameters can be a lengthy process, involving extensive searching and tuning, often with numerous iterations needed to find the optimal parameters. Given the satisfactory performance of the current model, further optimization would not justify the additional time required. Additionally, hyperparameter optimization is computationally expensive, requiring significant resources that might not be readily available or could be better used elsewhere. By not tuning the hyperparameters, we save these resources while still maintaining a high-performing model.

RF can be applied as both regression trees and classification trees in mining activities (Akyildiz et al., 2023). In this study, the dataset in Table 6 was split randomly, with 80% used for training and 20% for testing. Further, the training data was divided into five sets through K-fold, containing four for training and one for testing in each fold—a process known as "cross-validation". The model with the highest accuracy from cross-validation was chosen. After testing with the reserved data and achieving satisfactory results, the final model was assessed using input values not present in the original dataset. The development process of the RF model is depicted in Fig. 6.

3.1.3. Genetic Programming (GP)

The GP is of the evolutionary algorithms that encompass various topics including genetic algorithms, evolutionary programming, evolution strategies and classifier systems. Evolutionary algorithms operate iteratively, with each iteration commonly referred to as a "generation." The basic evolutionary algorithm process starts with a population of randomly selected individuals.

Specifically, the GP is an automated programming method that evolves computer programs to solve or approximately solve problems. It begins with thousands of randomly generated programs, which are then gradually evolved over numerous generations using principles such as the Darwinian concept of survival of the fittest. In GP, these programs are referred to as parse trees rather than lines of code. Starting with thousands of randomly created parse trees, a population of trees progressively evolves over many generations using principles such as the Darwinian concept of survival of the fittest.

The programs in the population consist of elements from the function and terminal sets, which are predefined sets of symbols chosen to be suitable for solving problems within the specific domain of interest. In GP, the crossover operation involves taking randomly selected subtrees from individuals (chosen based on their fitness) and exchanging them. Typically, GP does not use mutation as a genetic operator (Asthana, 2000).

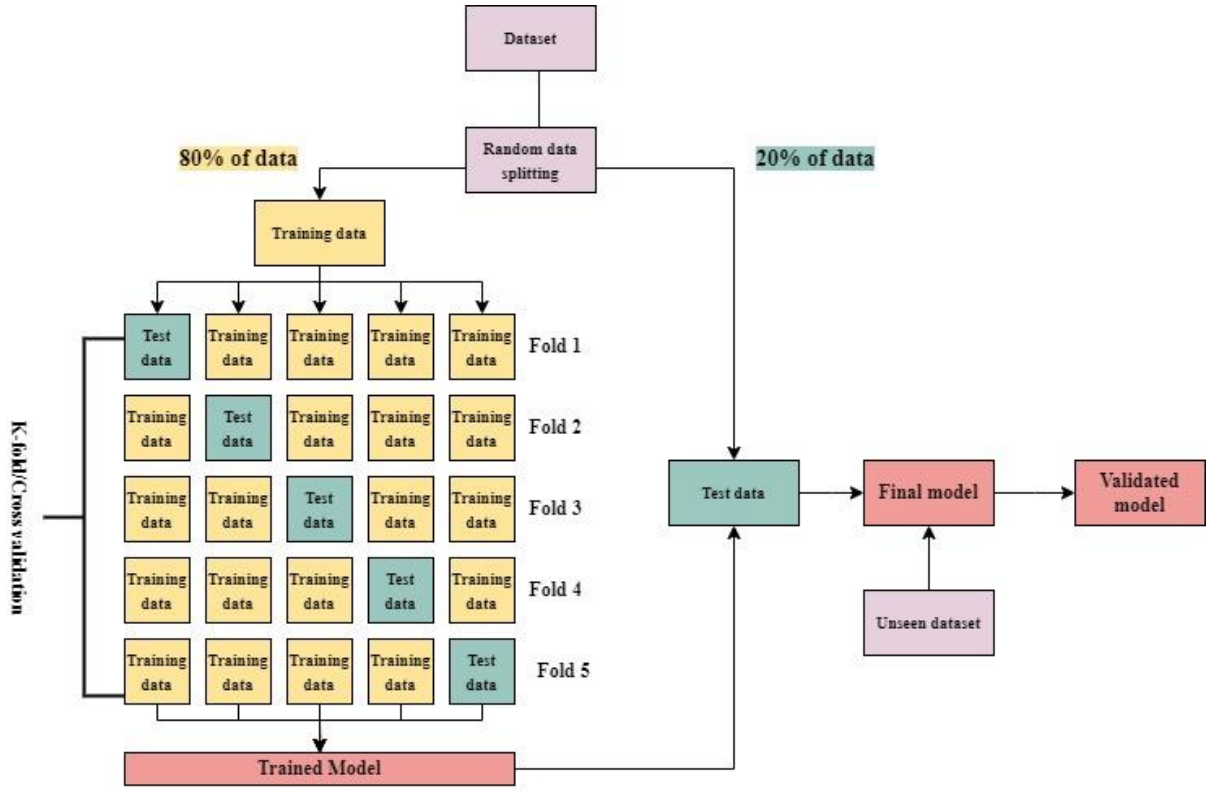

Fig. 6. Steps of training, testing and validation of RF model

3.2. Modeling results

ML model results were evaluated using mean absolute error (MAE), mean squared error (MSE), and R^2 values, which were calculated by comparing model predictions with experimental data. Data on actual outcomes versus model predictions were also plotted in a 'Predicted vs. Actual' graph for all models. All models were evaluated in Table 8, with training and test scores provided separately for the RF and GP models.

Table 8. Statistical evaluation of models (R^2 - coefficients of determination , MAE - mean absolute error, MSE - mean squared error)

		Recovery				Grade			
		MAE	MSE	R^2	RMSE	MAE	MSE	R^2	RMSE
Linear Regression		7.3	76.1	0.66	8.72	3.0	13.3	0.63	3.65
Nonlinear Regression		4.1	26.1	0.88	5.11	1.81	6.09	0.83	2.47
Random Forest	Training	1.3	3.51	0.98	1.87	0.7	1.36	0.96	1.17
	Test	4.5	61.6	0.74	7.85	2.0	9.81	0.63	3.13
	Overall	1.9	15.1	0.94	3.89	1.0	3.05	0.92	1.75
Genetic Programing	Training	18.8	3.2	0.92	1.79	1.38	3.63	0.90	1.91
	Test	18.1	3.0	0.92	1.73	1.37	5.73	0.82	2.39
	Overall	18.6	3.1	0.92	1.76	1.42	4.35	0.88	2.09

3.2.1. Linear (LR) and nonlinear (NLR) regressions

The assessment of linear regression's suitability for analyzing a dataset on kinetic flotation processes, despite their inherent non-linearity (Napier-Munn, 2012), initially focused on Eq. 2 for recovery:

$$R_{Cu,\%} = 47.62 + 0.31R1 + 0.55R2 + 0.41R3 + 1.05T \quad (2)$$

and Eq. 3 for grade:

$$c_{Cu,\%} = 12.46 + 0.11R1 + 0.10R2 + 0.08R3 - 0.44T \quad (3)$$

Notably, the collector R2 exhibited the most significant impact on the Cu recovery with an intercept value of 0.55, a finding visually supported by comparisons across Figs. 2, 3, and 4 single collector systems. However, the models' coefficients of determination (R^2 values) were 0.66 for recovery, and 0.63 for grade, as further detailed in Table 8, suggesting a low fit.

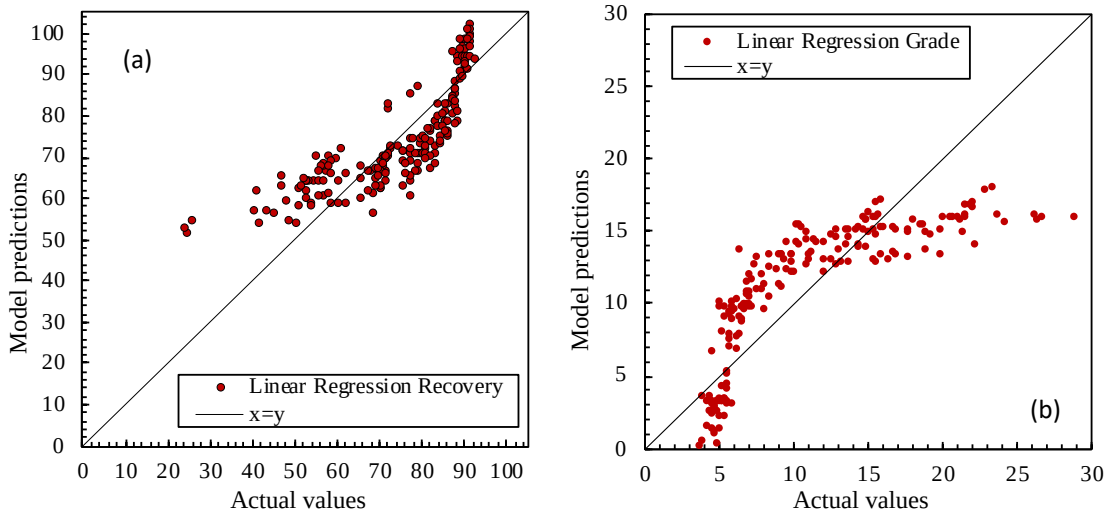

Fig. 7. Predicted data by model versus actual data of linear regression for Cu recovery (a) and grade (b) (R^2= 0.66 for recovery, R^2 = 0.63 for grade)

The linear regression models revealed high intercepts for recovery (48%) and grade (12%). This configuration led to significant overestimations of recovery at low and high dosages (e.g., a predicted recovery of 52% versus an actual of 24% in the D30 test with 0 g/Mg R1, 0 g/Mg R2, and 10 g/Mg R3). Conversely, the model tended to underestimate recovery at moderate dosages, highlighting a critical limitation in its predictive accuracy for this dataset.

Given these observations, it became clear that linear regression, while providing some insights, was ultimately not the best fit for capturing the dataset's dynamics. This realization prompted the exploration of nonlinear regression techniques as a more appropriate analytical method to account for the dataset's complex behavior and better match the actual recovery and grade outcomes.

The nonlinear distribution observed in the predicted vs. actual data plots for linear models, as shown in Fig. 8, suggested that a nonlinear approach might result in better accuracy using the same dataset Consequently, a quadratic nonlinear regression model with four independent variables was developed, resulting in the following equations for recovery:

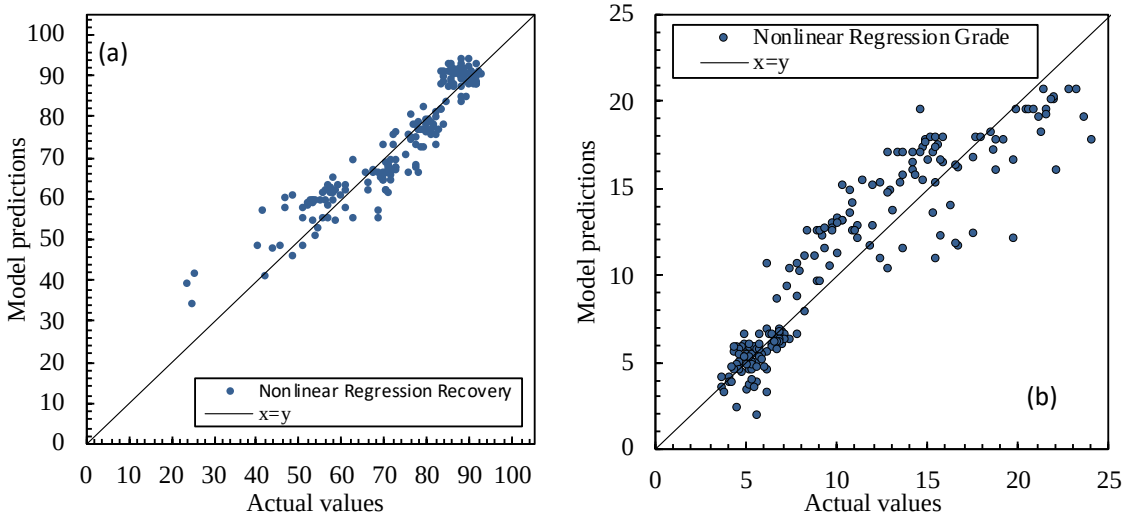

Fig. 8. Predicted data by model versus actual data of nonlinear regression for Cu recovery (a) and grade (b) (R^2= 0.88 for recovery, R^2 = 0.83 for grade)

$$R_{Cu,\%} = 23.79516 + 0.6737657R_1 + 2.4745793R_2 + 1.2751012R_3 + 3.7580174T + 0{,}0006715R_1^2 - 0.0358944R_2^2 - 0.0097735R_3^2 - 0.0051108T^2 - 0.0404622R_1R_2 - 0.0443650R_1R_3 - 0.0117114R_1T - 0.0121262R_2R_3 - 0.0099031R_2R_4 - 0.0767678R_3R_4 \quad (4)$$

and grade:

$$c_{Cu,\%} = 11.06091346 + 0.3263932R_1 + 0.47578347R_2 + 0.35184401R_3 - 1.19667845T - 0.00311324R_1^2 - 0.00598659R_2^2 - 0.0018382R_3^2 - 0.00594995T^2 - 0.00710672R_1R_2 - 0.00928089R_1R_3 - 0.00518962R_1T - 0.00492281R_2R_3 - 0.00398046R_2T + 0.03015359R_3T \quad (5)$$

The intercept values for the nonlinear model were approximately 24% for the recovery, and 11% for the grade. According to the statistical analysis detailed in Table 8, the nonlinear regression demonstrated a more satisfactory fit compared to the linear regression, with R^2 values of 0.88 for the recovery model, and 0.83 for the grade model. Despite this improvement, the analysis indicated a need for more advanced models, particularly to enhance accuracy at both low and high dosages for recovery and grade.

3.2.2. Random Forest (RF)

Reflecting on the limitations identified in both linear and nonlinear regression models, as discussed in the previous chapter, the exploration of more sophisticated techniques became imperative. This led to the testing of advanced modeling approaches, among which the RF model emerged as significantly more accurate in correlating actual with predicted data than both the linear and nonlinear models. The RF model demonstrated a marked reduction in both overestimation and underestimation errors for recovery predictions, as visually represented in Fig. 9.

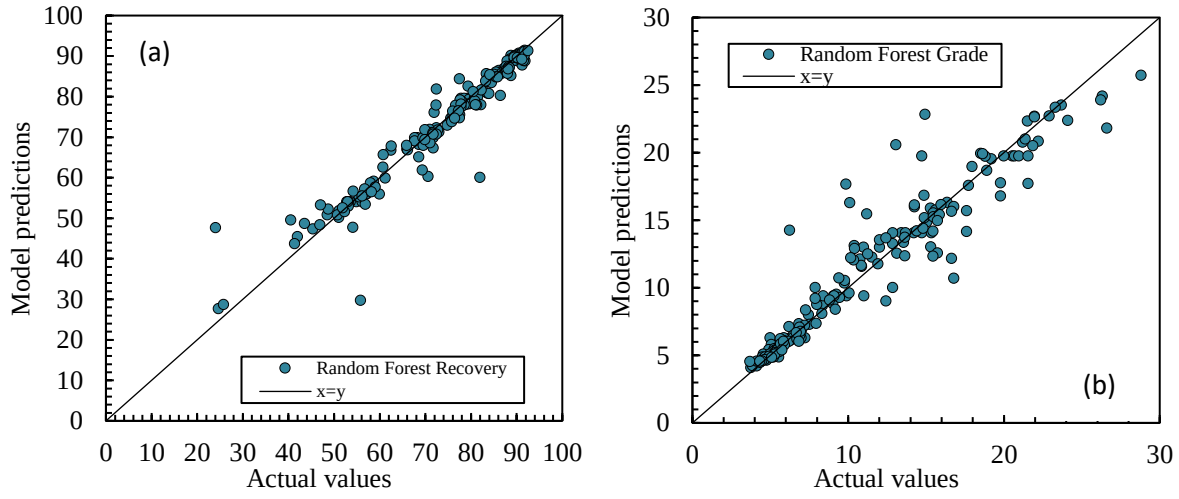

Fig 7. Predicted data by model versus actual data of RF model for Cu recovery (a) and grade (b) (R^2= 0.94 for recovery, R^2 = 0.92 for grade)

Notably, the RF model achieved the highest R^2 scores among all models evaluated, with an R^2 of 0.94 for recovery and 0.92 for the grade (Table 8). These results highlighted the RF model's robustness and its capability to offer a more precise prediction of outcomes, thereby addressing some of the critical limitations faced by the earlier linear and nonlinear regression models.

3.2.3. Genetic Programming (GP)

The GP model exhibited a commendable performance in predicting both recovery and grade, as evidenced by its consistent and high R^2 values across both training and testing phases, with values of 0.92 for the recovery and 0.90 for grade in training, maintaining at 0.92 for recovery and slightly dipping to 0.82 for grade in testing. The overall performance metrics further solidify its predictive capability, with R^2 values of 0.92 for recovery and 0.88 for grade, showcasing the model's robustness and reliability. The GP was more successful than linear and nonlinear regression, but its fitness between actual and

predicted values was lower than RF. Although the overall scores of the GP were lower than those of the RF, there was a higher consistency between the models on the test and training.

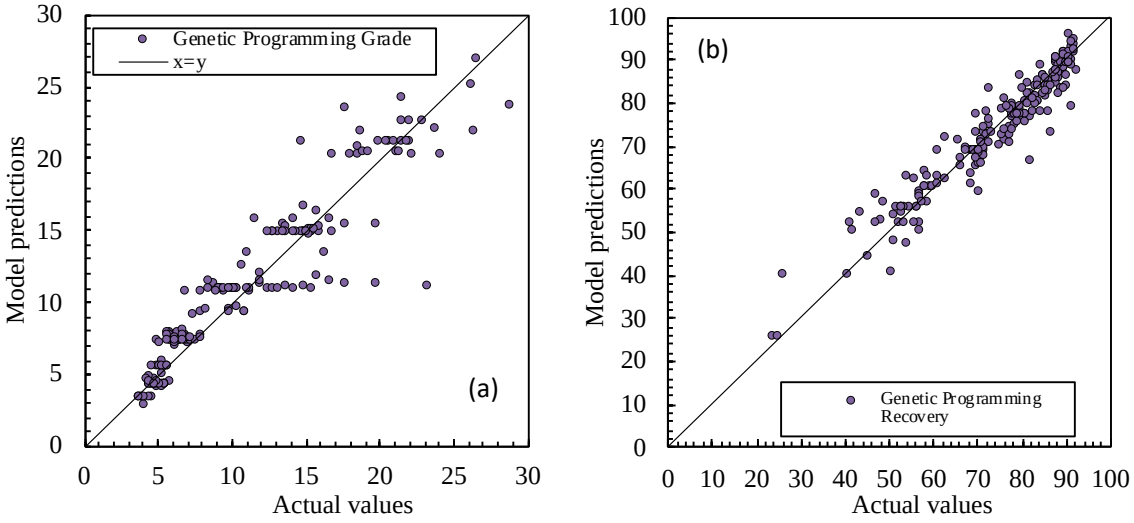

Fig. 8. Predicted data by model versus actual data of GP for Cu recovery (a) and grade (b) (R^2= 0.92 for recovery, R^2 = 0.88 for grade)

3.3. Model justification and selection

In this section, new data (which were not used for model development) were utilized for the model performance evaluation. Two additional experiments at different dosages were conducted for this validation. The selection of collector dosages was based on the data ranges used to train the model. Initially, an experiment was conducted using 10 g/Mg from each of the three collectors, followed by a second experiment with 20 g/Mg for each collector. Recovery and grade were then calculated as functions of time.

Table 9. Prediction of models using unseen data and scores of the models (LR: Linear Regression, NLR: Nonlinear Regression, RF: Random Forest and GP: Genetic Programing)

Condition	Time	Actual recovery	LR recovery	NLR recovery	RF recovery	GP recovery	Actual grade	NR grade	NLR grade	RF grade	GP grade
Mixture 1	1	53.3	61.3	57.2	53.2	52.6	21.9	14.8	18.1	20.1	20.5
10g/Mg R1,	3	70.3	63.4	63.6	69.0	68.1	14.6	13.9	15.6	13.9	14.9
10g/Mg R2,	7	79.1	67.6	74.5	79.4	77.3	9.8	13.9	11.4	9.3	10.9
10g/Mg R3	15	87.3	75.9	88.9	85.0	84.3	6.3	8.7	5.9	6.3	7.4
	30	91.1	91.6	89.5	89.2	89.8	5.0	2.2	6.1	4.8	4.2
Mixture 2	1	57.1	73.9	58.6	57.6	62.5	19.1	17.6	19.8	23.2	21.1
20g/Mg R1,	3	71.6	76.0	64.5	72.2	74.6	14.5	16.8	17.0	15.5	14.9
20g/Mg R2,	7	80.1	80.2	74.3	81.2	82.6	10.3	15.0	12.2	13.2	10.9
20g/Mg R3	15	87.4	88.6	86.6	87.2	88.9	6.5	11.5	5.5	7.1	7.4
	30	91.0	104.3	83.1	90.1	93.9	5.2	5.0	3.4	5.5	4.3
Model's score			**0.56**	**0.90**	**0.99**	**0.96**		**0.59**	**0.90**	**0.92**	**0.97**

According to the predictions made by the models using new data (Table 9), there was a noticeable shift in performance metrics. The accuracy of linear regression (LR) decreased for both recovery and grade. Conversely, in the case of nonlinear regression (NLR), the model's score for recovery improved from 0.88 to 0.90, and for grade, it increased from 0.83 to 0.90. The GP algorithm stood out with a prediction accuracy of 0.97 for grade, outperforming all other models in this aspect. Given the GP's training and test scores of 0.92 and 0.90, respectively, its predictions for grade with the new unseen data demonstrated remarkable success. The RF model exhibited exceptional prediction accuracy, achieving

a score of 0.99 for recovery. Its score for grade reached 0.92. The overall performance of both RF and GP with validation data was quite satisfactory. However, the sensitivity analysis was conducted using RF due to its superior performance in recovery estimation.

4. Discussion

4.1. Contribution of collectors (sensitivity analysis)

Since time (T) was identified as the most influential factor affecting recovery and grade across all models, as expected, and given the primary focus on collectors in this study, it was used as a controlled input variable. The sensitivity analysis was then applied based on the type of collector. It was possible to determine the most significant contributor from the model equations provided by linear and nonlinear regressions. Linear regression, which had the lowest performance, indicated that the most effective collector for recovery was R2 (AEROPHINE® 3422), and for grade R1 (NAX), based on the coefficient values in the objective function. Similarly, nonlinear regression also identified R2 as the most dominant collector for both recovery and grade.

Since RF demonstrated higher accuracy on Cu-recovery than other tested models (Table 9), the sensitivity analysis was performed only on the RF model. Considering that RF is part of a black-box model family, interpreting its behavior through coefficient values is not feasible. Although it is possible to assess the impact of factors on outputs, with some stock commands depending on the software used, this study followed a basic approach to conduct the sensitivity analysis by physically observed changes in results based on variable levels. Table 6 outlines the average value and standard deviation of the dosages, and a new dataset was generated by adjusting one input variable to its average value and ± one standard deviation from it, while keeping the other two input variables' mean values constant. Running the final RF models with this new dataset allowed for a comparison of the variable effects based on changes in the results. In summary, the difference between one standard deviation above and below the mean value of the reagent's dosage, while keeping the other parameters at their mean values, was considered an indicator of the collector's contribution to grade and recovery.

Fig. 11 demonstrates the effect on recovery and grade of modifying the dosage of each collector by plus or minus one standard deviation within the RF. The variations in recovery ranged from an increase of 4.1% to 2.0% for R1, a substantial increase of 10.4% to 4.4% for R2, and a modest increase of 5.1% to 0.5% for R3 (Fig. 11b). It was noted that the impact of these changes diminished as the flotation time extended. However, as shown in Fig. 11, R2 consistently produced the largest change in recovery, underscoring its significant influence on flotation recovery. The sensitivity analysis conducted for grade revealed distinct patterns for variables: RF identified R2 as the least effective collector, and R1 as the most effective one (Fig. 11a).

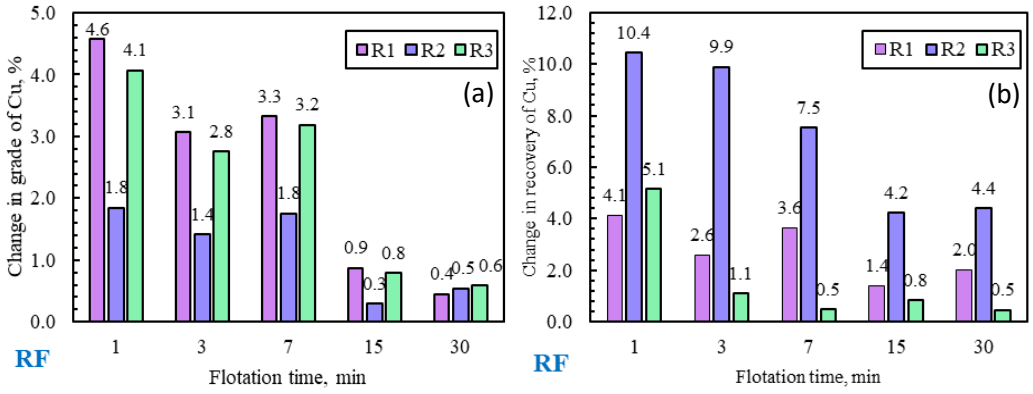

Fig. 11. Sensitivity analysis of the dosages of collectors on recovery (a) and grade (b) based on RF: Changes in recovery and grade at differences of ± standard deviations

4.2. Correlation between reagent chemistry and model's outputs

The components of the reagents used in this study are as follows; -R1- a mixture of sodium isobutyl xanthate and sodium ethyl xanthate, -R2- a mixture based on isopropyl ethyl thionocarbamate and

dithiophosphinate and -R3- a mixture based on reaction mass of S-allyl O-(2-methyl butyl) dithiocarbonate and S-allyl O-pentyl dithiocarbonate, and n-butoxycarbonyl-O-n-butyl thionocarbamate. Each listed components have different adsorption properties on different minerals. It was proven that it is not possible to apply a uniform single adsorption theory on specific minerals extracted from various ores (Fuerstenau, 1979). For instance, either the absence or the presence of oxygen and also pH can switch the adsorption mechanism as chemisorption or physisorption for the same mineral with xanthate (Bulatovic, 2006). Moreover, the observations on adsorption properties of specific components on a single mineral don't really work as it is in real ore systems, i.e. galvanic reaction of sulfide minerals (Rao and Finch, 1988). Even though synergism between thiol collector leads to increased grade and recovery, the background mechanism in this synergistic effect is still unclear in real ore flotation due to variability and uncertainty of the nature of the flotation itself. This study has similar drawback. Both experimental results and modeled output showed that synergism between different substances led to enhanced flotation, but it was impossible to determine the mechanism behind. The potential reason for this improvement can be caused by larger surface coverage of weak/strong site by strong/selective collectors as explained elsewhere (Bradshaw et al., 1998; Lotter and Bradshaw, 2010; McFadzean et al., 2013) or the multilayer formation between metal thiolate (chemisorption between mineral surface and dithiocarbamate) and dixanthogen (physisorption between metal-thiolate and xanthate) (Bagci et al., 2007; McFadzean et al, 2012; Taguta et al., 2018).

Using mixtures in the flotation system introduces variability in functional groups and hydrocarbon chain lengths. For example, when R1 and R2 are used together, the different functional groups (-OCS_2 for R1 and -NCS_2, -PS_2 for R2) are expected to provide a higher adsorption rate due to varied collector strengths and different surface adsorption tendencies. This assumption is supported by experimental findings. Additionally, the RF method was used to generate data for creating ternary plots with the three reagents in different combinations (with a total dosage of 60 g/Mg). Long flotation time made it difficult to observe the collector's effect, as grade and recovery accumulated within a narrow range. For copper, it was recommended to conduct flotation tests at a bench scale for 6-8 minutes (Kawatra, 2011); consequently, 7 minutes of flotation was chosen for creating ternary plots graphs for grade and recovery (Fig. 12).

Dithiocarbamates are known to have higher collector strength than xanthates due to nitrogen being less electronegative than oxygen, which makes it more inclined to donate electrons. This results in a collector that is less selective (Lotter and Bradshaw, 2010; Bagci et al., 2007). Since R2 and R3 contained types of dithiocarbamates, the xanthate mixture in R1 became more selective within the ternary mixture system. As shown in Fig. 12, as the dosage of R2 and R3 increased in the mixture, the flotation recovery was higher, while the grade was a function of the dosage of R1 in the mixture. An antagonistic behavior regarding the grade was observed when R2 and R3 were used together at high dosages.

The ternary graphs created by RF (Fig. 12), enabled the identification of optimum levels of reagents for achieving targeted grades and recovery. For a higher grade, the optimum dosage should be between

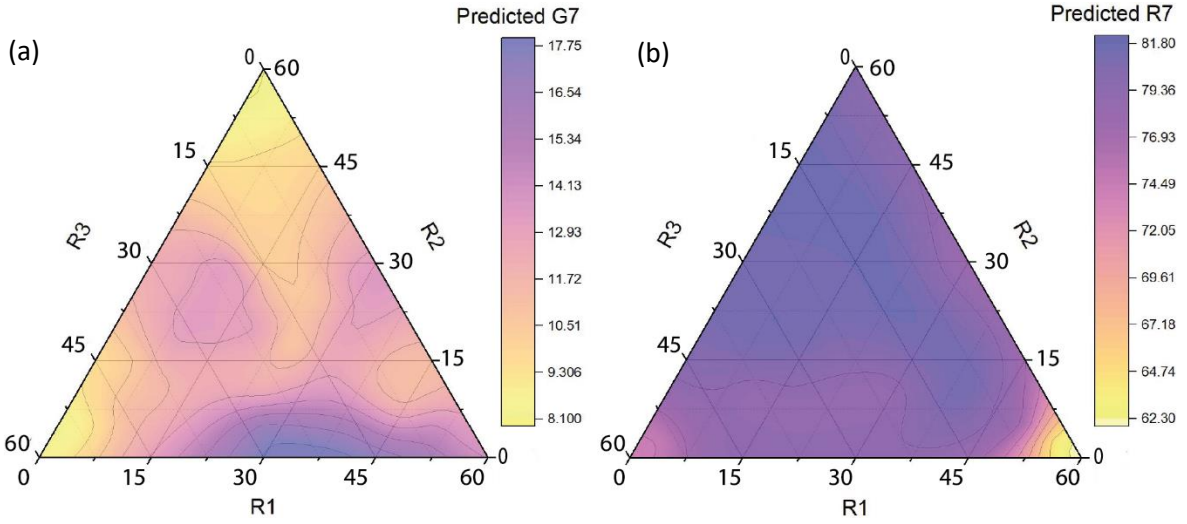

Fig. 12. Predicted grade (a) and recovery (b) at 7-minute flotation of the ternary collector mixtures

30-40 g/Mg for R1, 5-10 g/Mg for R2, and 10-20 g/Mg for R3. For higher recovery, the recommended dosages are 0-15 g/Mg for R1, 35-50 g/Mg for R2, and 10-20 g/Mg for R3. While it is not possible to maximize both grade and recovery simultaneously, it is feasible to determine an optimum range based on the process's priority. For achieving the Cu grade greater than 13% and the Cu recovery exceeding 80%, recommended recipes are as follows: recipe 1: 10 g/Mg R1, 30 g/Mg R2, and 20 g/Mg R3; or recipe 2: 40 g/Mg R1, 10 g/Mg R2, and 10 g/Mg R3. It should be noted that these recipes prepared based on the mixture at total of 60 g/Mg, but RF can be easily utilized to make ternary graphs for different total dosages for different flotation time using the same method and optimum recipes can be determined in different conditions.

5. Conclusions

This research aimed not only at predicting the copper grade and recovery using machine learning (ML) models but also at interpreting the impact of reagents on grade and recovery as well as at assessing the improvement achieved when using collectors alone or in combinations (i.e., binary and ternary mixtures). Rather than using variables as inputs such as pH, gas flowrate, and particle size distribution from different components that constitute the flotation system, it focused on three collectors that directly interact with each other. The conclusions of the study are summarized as follows:

- Machine learning tools can be used to model the results of batch-scale flotation tests, just like simple statistical methods. However, these tools do better than linear and multilinear regressions in being accurate and efficient. The Genetic Programming (GP) model had more consistent model accuracy between training and testing stages, but the overall score of the Random Forest (RF) model was slightly higher. Importantly, GP was more accurate with new data when predicting grade, whereas RF was better at predicting recovery.
- Considering the complexity of the flotation process, it is important to thoroughly review the model outcomes, instead of relying solely on statistical assessment, to prevent any discrepancies between model results and experiential understanding.
- Experimental findings revealed that using AEROPHINE® 3422 led to higher copper recovery and grade in a single collector system, particularly at low doses. However, a combination of three collectors, i.e., NAX, AEROPHINE® 3422, and AERO® MX-5149, generally resulted in better flotation performance than using one collector.
- According to the sensitivity analysis, which is one of the most critical stages of this study, AEROPHINE® 3422 was the most dominant collector in terms of recovery, but its effect on grade was relatively lesser compared to AERO® MX-5149 and NAX. This supports the existing literature discussed in this study: dithiocarbamates have higher collector strength than xanthates, dithiophosphates, and dithiophosphinates.
- The RF model can be successfully used to prepare mixtures in different dosages by creating a ternary graph for desired recovery and grade values. This application is recommended as a useful tool to examine interactions in all flotation applications using a ternary system, not just for collectors.

For future studies, it is essential to conduct an optimization study including the financial and environmental indexes of these three reagents. The optimization process will help to identify the most cost-effective and environmentally friendly combination of reagents while maximizing the desired outcomes. By taking into account the cost and environmental factors, the study can lead to improved efficiency and sustainability in the flotation process.

Acknowledgments

This activity was co-funded by EIT Raw Materials, an entity of the European Union, within the framework of Horizon 2020, the EU's Framework Programme for Research and Innovation (project RFC-Upscaling: New Reflux Flotation Cell Technology Upscaling for Ore Flotation, project number: 999407100). The authors would like to thank Dr. Tarun Bhambhani (Solvay) for the providing the reagents and Mr. Marcin Czekajło (KGHM Polska Miedz S.A.) for supplying copper ore used in testing. The first author would like to express appreciation to Prof. Steinar Løve Ellefmo (NTNU), and Dr. Ozan Kokkilic (McGill University) for discussions on experimental design, and to Marie-Anaïs Chabal for laboratory assistance. We are profoundly grateful to Camilo Mena Silva, Kornel Mateusz Tobiczyk,

Gustav Ward, Jesper Hansen Levinsen, Evy Eilin Stamnes Valås, Tobias Holen Kokkin, and Laurentius Tijhuis (from NTNU) for their unwavering support.

References

AKYILDIZ, O., BASARIR, H., VEZHAPPARAMBU, V. S., ELLEFMO, S., 2023. *MWD Data-Based Marble Quality Class Prediction Models Using ML Algorithms*. Mathematical Geosciences, 1-16.

ALLAHKARAMI, E., SALMANI NURI, O., ABDOLLAHZADEH, A., REZAI, B., MAGHSOUDI, B., 2017. *Improving estimation accuracy of metallurgical performance of industrial flotation process by using hybrid genetic algorithm–artificial neural network (GA-ANN*). Physicochemical Problems of Mineral Processing, 53.

ARANCIBIA-BRAVO, M. P., LUCAY, F. A., SEPÚLVEDA, F. D., CORTÉS, L., CISTERNAS, L. A., 2022. *Response Surface Methodology for Copper Flotation Optimization in Saline Systems*. Minerals, 12(9), 1131.

ASTHANA, R. G. S., 2000. *Evolutionary algorithms and neural networks. In: Soft Computing and Intelligent Systems,* Academic Press, pp. 111-136.

AZIZI, A., MASDARIAN, M., HASSANZADEH, A., BAHRI, Z., NIEDOBA, T., SUROWIAK, A., 2020. *Parametric optimization in rougher flotation performance of a sulfidized mixed copper ore*. Minerals, 10(8), 660.

BAGCI, E., EKMEKCI, Z., BRADSHAW, D., 2007. *Adsorption behaviour of xanthate and dithiophosphinate from their mixtures on chalcopyrite.* Minerals Engineering, 20(10), 1047-1053.

BASTRZYK, A., POLOWCZYK, I., SADOWSKI, Z., SIKORA, A., 2011. *Relationship between properties of oil/water emulsion and agglomeration of carbonate minerals.* Sep. Purif. Technol. 77, 325-330.

BREIMAN, L., 2001. Random forests. Machine learning, 45, 5-32.

BRADSHAW, D.J., O'CONNOR, C.T., 1994. *The flotation of pyrite using mixtures of dithiocarbamates and other thiol collectors.* Minerals Engineering, 7(5-6), 681-690.

BRADSHAW, D.J, HARRIS, PJ, O'CONNOR, C.T. (1998). *Synergistic interactions between reagents in sulphide flotation.* Journal of the Southern African Institute of Mining and Metallurgy, *98*(4), 189-193.

BULATOVIC, S. M., 2007. *Handbook of flotation reagents: chemistry, theory and practice: Volume 1: flotation of sulfide ores*. Elsevier.

CILEK, E. C., 2002. *Application of neural networks to predict locked cycle flotation test results*. Minerals Engineering, 15(12), 1095-1104.

COOK, R., MONYAKE, K. C., HAYAT, M. B., KUMAR, A., ALAGHA, L., 2020. *Prediction of flotation efficiency of metal sulfides using an original hybrid machine learning model*. Engineering Reports, 2(6), e12167.

CORPAS-MARTÍNEZ, J. R., PÉREZ, A., AMOR-CASTILLO, C., NAVARRO-DOMÍNGUEZ, R., MARTÍN-LARA, M. A., CALERO, M., 2019. *Optimal depressants and collector dosage in fluorite flotation process based on DoE methodology*. Applied Sciences, 9(3), 366.

CUTLER, A., CUTLER, D. R., STEVENS, J. R., 2012. *Random Forest. In: Ensemble Machine Learning, Methods and Applications,* Zhang, C. & Ma, Y. (Eds.), Springer, New York, NY, pp. 157-176.

EB, K., 2011. *Multi-objective optimisation using evolutionary algorithms: an introduction*. In Multi-objective evolutionary optimisation for product design and manufacturing. Springer London, London, pp. 3-34.

DRZYMALA, J., 2007. Mineral processing. Foundations of theory and practice of minerallurgy. Ofic. Wyd. PWr, Wroclaw, Poland.

DRZYMALA, J., KOWALCZUK, P. B., OTENG-PEPRAH, M., FOSZCZ, D., MUSZER, A., HENC, T., LUSZCZKIEWICZ, A., 2013. *Application of the grade-recovery curve in the batch flotation of Polish copper ore*. Minerals Engineering, 49, 17-23.

FUERSTENAU, M. C., JAMESON, G. J., YOON, R. H. (Eds.), 2007. *Froth flotation: a century of innovation*. SME.

GARETH, J., WITTEN, D., HASTIE, T., TIBSHIRANI, R. (2013). *An introduction to statistical learning: with applications* in R. Springer.

GHOLAMI, A., ASGARI, K., KHOSHDAST, H., HASSANZADEH, A., 2022. *A hybrid geometallurgical study using coupled Historical Data (HD) and Deep Learning (DL) techniques on a copper ore mine*. Physicochemical Problems of Mineral Processing, 58(3).

GOLDBERG, D., 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, MA, USA.

GREET, C.J. (Ed.), 2010. *Flotation plant optimisation: A metallurgical guide to identifying and solving problems in flotation plants.* Australasian Institute of Mining and Metallurgy.

GUNER, M. K., LODE, S., MALAFEEVSKIY, N., AASLY, K., KOWALCZUK, P., 2024. Mixed thiol collector system in the flotation of copper ore using the REFLUX Flotation Cell (RFC). Minerals Engineering, 216, 108843.

GURIA, C., VERMA, M., GUPTA, S.K., MEHROTRA, S.P., 2005. *Simultaneous optimization of the performance of flotation circuits and their simplification using the jumping gene adaptations of genetic algorithm*. International Journal of Mineral Processing, 77(3), 165-185.

HASTIE, T., TIBSHIRANI, R., FRIEDMAN, J., 2009. T*he Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd Edition). Springer.

HU, W., HADLER, K., NEETHLING, S.J., CILLIERS, J.J., 2013. *Determining flotation circuit layout using genetic algorithms with pulp and froth models.* Chemical Engineering Science, 102, 32-41.

JAHEDSARAVANI, A., MARHABAN, M.H., MASSINAEI, M., 2014. *Prediction of the metallurgical performances of a batch flotation system by image analysis and neural networks.* Minerals Engineering, 69, 137-145.

KLIMPEL, R.R., HANSEN, R.D., 1988. *Frothers*. In: Reagents in mineral technology. Routledge, 385-409.

KOWALCZUK, P.B., ZALESKA, E., DANCZAK, O., 2015. *Flotation of carbonaceous copper shale–quartz mixture with poly (ethylene glycol) alkyl ethers*. Transactions of Nonferrous Metals Society of China, 25(1), 314-318.

KOWALCZUK, P.B., DRZYMALA, J., 2017. *Selectivity and power of frothers in copper ore flotation*. Physicochem. Probl. Miner. Process, 53(1), 515-523.

LASKOWSKI, J.S., RALSTON, J. (Eds.), 2015. *Colloid chemistry in mineral processing*. Elsevier.

LOTTER, N.O., BRADSHAW, D.J., 2010. *The formulation and use of mixed collectors in sulphide flotation*. Minerals Engineering, 23(11-13), 945-951.

MASSINAEI, M., SEDAGHATI, M.R., REZVANI, R., MOHAMMADZADEH, A.A., 2014. *Using data mining to assess and model the metallurgical efficiency of a copper concentrator.* Chemical Engineering Communications, 201(10), 1314-1326.

McFADZEAN, B., CASTELYN, D.G., O'CONNOR, C.T., 2012. T*he effect of mixed thiol collectors on the flotation of galena.* Minerals Engineering, 36, 211-218.

McFADZEAN, B., MHLANGA, S.S., O'CONNOR, C.T., 2013. *The effect of thiol collector mixtures on the flotation of pyrite and galena*. Minerals Engineering, 50, 121-129.

AGARAJ, D.R., 2005. *Reagent selection and optimization – the case for a holistic approach*. Minerals Engineering, 18(2), 151-158.

NAGARAJ, D.R., RAVISHANKAR, S.A., 2007. *Flotation reagents – A critical overview from an industry perspective*. In: Froth flotation: A century of innovation, 375-424.

NAGARAJ, D.R., FARINATO, R.S., 2016. *Evolution of flotation chemistry and chemicals: A century of innovations and the lingering challenges.* Minerals Engineering, 96, 2-14.

NAKHAEI, F., MOSAVI, M.R., SAM, A., VAGHEI, Y., 2012. *Recovery and grade accurate prediction of pilot plant flotation column concentrate: Neural network and statistical techniques*. International Journal of Mineral Processing, 110, 140-154.

NAPIER-MUNN, T.J., 2012. *Statistical methods to compare batch flotation grade-recovery curves and rate constants.* Minerals Engineering, 34, 70-77.

NAPIER-MUNN, T.J., 2014. *Statistical Methods for Mineral Engineers*. JKMRC, Australia.

SARAVANI, A.J., MEHRSHAD, N., MASSINAEI, M., 2014. *Fuzzy-based modeling and control of an industrial flotation column.* Chemical Engineering Communications, 201(7), 896-908.,

RAO, K. Hanumantha, FORSSBERG, K.S.E., 1997. *Mixed collector systems in flotation*. International Journal of Mineral Processing, 51(1-4), 67-79

SADROSSADAT, E., BASARIR, H., 2019. *An evolutionary-based prediction model of the 28-day compressive strength of high-performance concrete containing cementitious materials*. Advances in Civil Engineering Materials, 8(3), 484-497.

SHERIDAN, M.S., NAGARAJ, D.R., FORNASIERO, D., RALSTON, J., 2002. *The use of a factorial experimental design to study collector properties of N-allyl-O-alkyl thionocarbamate collector in the flotation of a copper ore*. Minerals Engineering, 15(5), 333-340.

TAGUTA, J., O'CONNOR, C.T., McFADZEAN, B., 2018. *Investigating the interaction of thiol collectors and collector mixtures with sulphide minerals using thermochemistry and microflotation*. Minerals Engineering, 119, 99-104.

THOMPSON, P., 2016. *Laboratory testing for sulfide flotation process development*. Minerals & Metallurgical Processing, 33, 200-213.

WILLS, B.A., FINCH, J., 2015. *Wills' mineral processing technology: an introduction to the practical aspects of ore treatment and mineral recovery*. Butterworth-Heinemann.

VAZIFEH, Y., JORJANI, E., BAGHERIAN, A., 2010. *Optimization of reagent dosages for copper flotation using statistical technique.* Transactions of Nonferrous Metals Society of China, 20(12), 2371-2378.