

A. Rustemov^{1*} , B. Daulet¹ ,
B. Kenzhaliyev¹ , S. Aibagarov² 

¹Institute of Metallurgy and Ore Beneficiation JSC, Satbayev University, Almaty, Kazakhstan

²LLP «DigitAlem», Almaty, Kazakhstan

*e-mail: adilbek.rustemov@mail.mcgill.ca

MICROFLOTATION OF GOLD-BEARING ORE: WITHIN-BATCH EFFECTS AND GROUP-AWARE CLASSIFICATION OF REAGENT REGIMES

Abstract. Fine particles below approximately 20 μm are recovered poorly in conventional flotation, and microbubble-assisted methods are among the options proposed to address this. This paper reports a statistical assessment of multicomponent microflotation (MMF) and an auxiliary classification of reagent regimes, based on 132 laboratory flotation tests on gold-bearing feed from an operating concentrator. The tests formed 14 daily batches, each from a separate feed sample, comprising 101 microflotation and 31 baseline tests. Because feed properties varied between days, the effect was assessed within batches rather than across the pooled sample. Median Au recovery was higher under microflotation in 10 of 14 batches, with a median within-batch increase of 2.275 percentage points (cluster bootstrap 95% CI 0.025–3.650; Wilcoxon signed-rank $p=0.0166$). The corresponding change in enrichment efficiency did not reach significance at $\alpha=0.05$ and is reported only as a tendency. Textual reagent regimes were converted into tabular features, and regimes with Au recovery of at least 84% were classified under five-fold GroupKFold validation by daily batch. A nested weighted ensemble reached ROC-AUC 0.717 (95% CI 0.632–0.792), indicating moderate discrimination rather than accurate prediction.

Keywords: microflotation, gold-bearing ore, reagent regime, grouped cross-validation, machine learning classification.

1. Introduction

During grinding at flotation concentrators, a considerable quantity of poorly floatable fine particles is produced, particularly in the size range below approximately 20 μm . The reduced flotation efficiency of such particles is commonly attributed to their low probability of collision and attachment to conventional flotation bubbles, which can result in significant losses of valuable minerals to tailings [1], [2]. Particle size is therefore one of the fundamental factors governing flotation performance.

One of the approaches proposed to improve the recovery of fine particles is the use of smaller bubbles. Reduced bubble size can increase the bubble surface area available for attachment and improve the probability of bubble–particle collision [3], [4]. Various approaches involving micro- and nano-bubbles, hydrodynamic effects, and fine-particle aggregation have consequently been investigated as methods for improving flotation of difficult fine fractions [7–9]. Bubble size and stability remain important practical considerations, since generated

fine bubbles may undergo coalescence and form larger bubbles during flotation [5], [6].

The recovery of fine particles is especially important for gold-bearing ores. Fine and partially liberated native gold and electrum can exhibit poor recovery under conventional flotation conditions, depending on particle size, mineral association, surface condition, and reagent environment [10], [11]. These factors make the development of flotation approaches capable of improving fine-gold recovery of particular practical interest.

The multicomponent microflotation (MMF) technology investigated in this work uses micro-suspensions containing microbubbles and micro-emulsions generated with flotation reagents. The approach is intended to reduce microbubble coalescence by exploiting the chemical properties of reagents already used in flotation practice. In the investigated configuration, the micro-suspension is generated outside the treated pulp and subsequently mixed with the pulp before passing through a tubular static flotoreactor under turbulent-flow conditions. From a practical perspective, the technol-

ogy is designed to be integrated without substantial modifications to the existing process flowsheet or prolonged interruption of plant operation.

Evaluation of flotation technologies under laboratory conditions is complicated by the strong dependence of flotation performance on feed and operating conditions. Variables such as mineralogy, particle-size distribution, head grade, reagent regime, and other process parameters can substantially influence metallurgical response [12]. Consequently, comparisons between flotation tests conducted on different feed samples may reflect feed variability rather than the effect of the investigated treatment itself. Data-driven methods have increasingly been considered in mineral processing and flotation as tools for analysing complex relationships between operating variables and metallurgical outcomes [13], [14]. However, such analyses require careful validation because observations originating from the same experimental batch or feed sample cannot necessarily be regarded as statistically independent [16], [17].

This issue is particularly relevant to the present experimental programme because feed samples collected on different days differed in mineralogical and metallurgical characteristics. Pooling all tests without accounting for daily feed batches could therefore confound the effect of the reagent regime with between-day feed variation. A second analytical difficulty is that reagent regimes were originally recorded as free-text descriptions, making direct quantitative comparison and modelling inconvenient.

Accordingly, the contribution of this study is threefold. First, MMF performance is evaluated using within-batch comparisons that hold the daily feed sample constant, thereby reducing confounding from between-day feed variability. Second, free-text reagent descriptions are converted into structured dosage features suitable for quantitative analysis. Third, an auxiliary classifier is evaluated using batch-grouped cross-validation, preventing tests originating from the same feed sample from appearing simultaneously in training and validation data. This approach combines a controlled metallurgical comparison with leakage-aware data analysis rather than treating the pooled laboratory tests as independent observations.

The primary aim of the study is therefore to quantify the effect of microflotation on Au recovery through comparisons performed within individual daily batches. In addition, the structured reagent-regime data are used in an auxiliary classification

analysis to examine whether regimes associated with Au recovery of at least 84% can be distinguished under group-aware validation. The classification analysis is intended as an assessment of the strength and reproducibility of the available signal rather than as a model for accurate recovery prediction or direct prescription of reagent regimes.

2. Materials and methods

2.1. Feed samples and test organisation

Pulp samples of the rougher flotation feed at the Pustynnoye gold recovery plant were taken on a daily basis from hydrocyclone overflows by quality control technologists, in volumes varying according to requirements. A total of 132 flotation tests were performed.

All flotation tests were carried out and compared within groups of tests conducted during a single day; samples were not held over to the following day. This organisation was adopted specifically to exclude the variability of mineralogy, size characteristics, and head grade between days. In the present analysis these daily groups are treated as 14 batches, each corresponding to a separate feed sample. Of the 132 tests, 101 were microflotation tests and 31 were baseline tests. Batch sizes ranged from 4 to 12 tests. Laboratory head grade of Au ranged from 0.57 to 1.29 g/t across batches.

2.2. Equipment and test procedure

Flotation tests were performed on a standard Metso laboratory flotation machine (Figure 1) with a 3-litre cell for rougher and scavenger flotation and a 0.5-litre cell for cleaner flotation.

Tests followed the standard plant flotation flowsheet, in which the rougher concentrate reports to cleaner flotation and the rougher tailings report to scavenger flotation. Under laboratory conditions, the scavenger concentrate is analysed as a separate product. The scavenger tailings constitute the final tailings.

Flotation concentrate and tailings were filtered on separate press filters, the concentrate on a Laarmann vacuum press filter and the tailings on a standard FLSmidth press filter, in order to exclude the risk of sample cross-contamination. Samples were dried in an ESSA FLSmidth DO2 0010 laboratory drying oven. Sample preparation for chemical analysis was performed by senior technicians of the metallurgical laboratory. All samples were submitted for chemical analysis to the ALS laboratory in accordance with internal procedures.



Figure 1. Metso laboratory flotation machines: 3 litre and 0.5 litre cells

The microsuspension (microbubbles and microemulsions) was generated by laboratory modified equipment (LME). The concentration of the reagents used to generate the microsuspension varied within the range 0.00125-0.1%, with recycled process water. The microsuspension was fed directly into the laboratory flotation cell through a PVC tube of 2 mm diameter. Depending on the ratio of collector microemulsions to frother activity, the LME is able to generate microbubbles, microemulsions, and emulsion microbubbles of complex structure (EMB).

Five principal microsuspension configurations were investigated, distinguished by the reagent used as the basis of the microsuspension: X-133, PAX, copper sulphate, diesel fuel, and the authors' MP-series reagents. These configurations represent alternative reagent formulations evaluated within the same laboratory microflotation framework; the present dataset does not support assigning separate causal effects to the individual components.

2.3. Calculated indicators

In addition to recovery and concentrate grade, the enrichment efficiency of each flotation test was calculated. The enrichment efficiency of a mineral separated into two products is determined by the Hancock-Luyken formula:

$$E = (\varepsilon - \gamma) \cdot \beta / (\beta - \alpha) \quad (1)$$

where α is the head grade, %; ε is the recovery of the valuable element into the concentrate, %; β is the grade of the valuable element in the concentrate, %; and γ is the yield of the valuable element into the concentrate, %.

2.4. Conversion of reagent regimes into features

The reagent regime of each test was recorded as free text specifying the reagents, their dosages in g/t, and whether each dosage belonged to the microsuspension or to the base addition. These records were parsed into structured numerical columns. For each reagent, three quantities were derived: the dosage in the microsuspension, the dosage in the base addition, and the total dosage. The reagents represented in the dataset are PAX (present in 125 tests, 100-300 g/t), X-133 (111 tests, 6-100 g/t), CuSO₄ (14 tests, 20-40 g/t), and the MP-series reagents MP-1 (36 tests, 10-100 g/t), MP-1000 (13 tests), MP-9445 (10 tests), MP-102 (5 tests), MP-4650 (5 tests), MP-X2 (5 tests), together with BTF (20 tests, 50-100 g/t). Aggregate features were also derived: total microsuspension dosage, total base dosage, and the count of active reagents (2-4 per test). All dosages were parsed without missing values.

2.5. Feature selection and leakage control

Model inputs were restricted to quantities that are known before a test is run: the laboratory head grade of Au, a binary microflotation indicator, the structured reagent dosages, and the aggregate dosage features. The calculated head grade, which is a product of the material balance, was excluded, as were all measured flotation product outcomes (weights, grades, and metal contents of the final concentrate, final tailings, cleaner tailings, and scavenger concentrate) and the final concentrate yield. Retaining any of these would introduce target leakage, since they are determined jointly with recovery.

The effect of microflotation was assessed within batches. For each of the 14 batches, the median Au recovery of the microflotation tests was compared with the median Au recovery of the baseline tests from the same batch, giving one paired difference per batch. This design keeps the feed sample constant within each comparison. The same procedure was applied to enrichment efficiency.

The paired differences were tested with the Wilcoxon signed-rank test. Confidence intervals for the median effect were obtained by cluster bootstrap resampling at the level of the batch, so that the resampling unit matches the unit of independence in the design [17].

2.6. Classification framework and group-aware validation

The classification target was defined as Au recovery of at least 84%, which is satisfied by 72 of the 132 tests. A regression formulation for continuous recovery was attempted first but performed unsatisfactorily under grouped validation, motivating the use of a binary screening formulation. The 84% value is treated here as an operational, dataset-specific screening threshold rather than as a universal metallurgical recovery criterion; conclusions from the classifier therefore concern separation of regimes around this cutoff only.

Four approaches were compared using scikit-learn [15]: regularized logistic regression, an RBF-kernel support vector machine on a compact feature subset, Extra Trees on total-dosage features, and a nested weighted ensemble of the three.

The primary validation was five-fold GroupK-Fold with the daily batch as the grouping variable, so that no feed sample contributes to both the training and the test side of any fold [16, 17]. Given only 14 independent batches, this is the essential precaution: a random split would place tests from the same feed sample on both sides and would inflate

the apparent performance. Ensemble weights and the decision threshold were selected only inside the training folds, using an inner loop, and were never exposed to the held-out batches [18]. The selections made per fold varied (weights between (1,1,1) and (2,3,1); thresholds between 0.280 and 0.685), which is itself an indication of limited stability.

Performance was reported as ROC-AUC, PR-AUC, balanced accuracy, precision, recall, specificity, and F1, computed on out-of-fold predictions. The confidence interval for ROC-AUC was obtained by cluster bootstrap at the batch level.

3. Results

3.1. Dataset and within-batch effects

The composition of the dataset and the estimated within-batch effects are summarised in Table 1. Median Au recovery was higher under microflotation in 10 of the 14 batches. The mean within-batch increase was 1.829 percentage points and the median within-batch increase was 2.275 percentage points. The cluster bootstrap 95% confidence interval for the median effect was 0.025-3.650 percentage points, and the Wilcoxon signed-rank test gave $p=0.0166$.

Table 1. Dataset and within-batch effects of microflotation

Item	Value
Laboratory flotation tests	132
Daily batches (separate feed samples)	14
Microflotation / baseline tests	101 / 31
Laboratory Au head grade range, g/t	0.57-1.29
Tests with Au recovery $\geq 84\%$	72 of 132
Au recovery	
Batches with higher median under microflotation	10 of 14
Mean within-batch effect, percentage points	+1.829
Median within-batch effect, percentage points	+2.275
Cluster bootstrap 95% CI for the median effect	0.025-3.650
Wilcoxon signed-rank p	0.0166
Enrichment efficiency (Hancock-Luyken)	
Median within-batch effect, percentage points	+1.750
Cluster bootstrap 95% CI for the median effect	-0.350-2.750
Wilcoxon signed-rank p	0.0735

Two aspects of this result should be stated plainly. The confidence interval, although positive, has a lower bound very close to zero, so the data are compatible with an effect that is only marginally above zero. The effect is also not uniform: in 4 of the 14 batches the

median recovery under microflotation was lower than baseline, with the largest negative difference at -2.6 percentage points in batch 10, while the largest positive difference was +5.0 percentage points in batch 11. The batch-level pattern is shown in Figure 2.

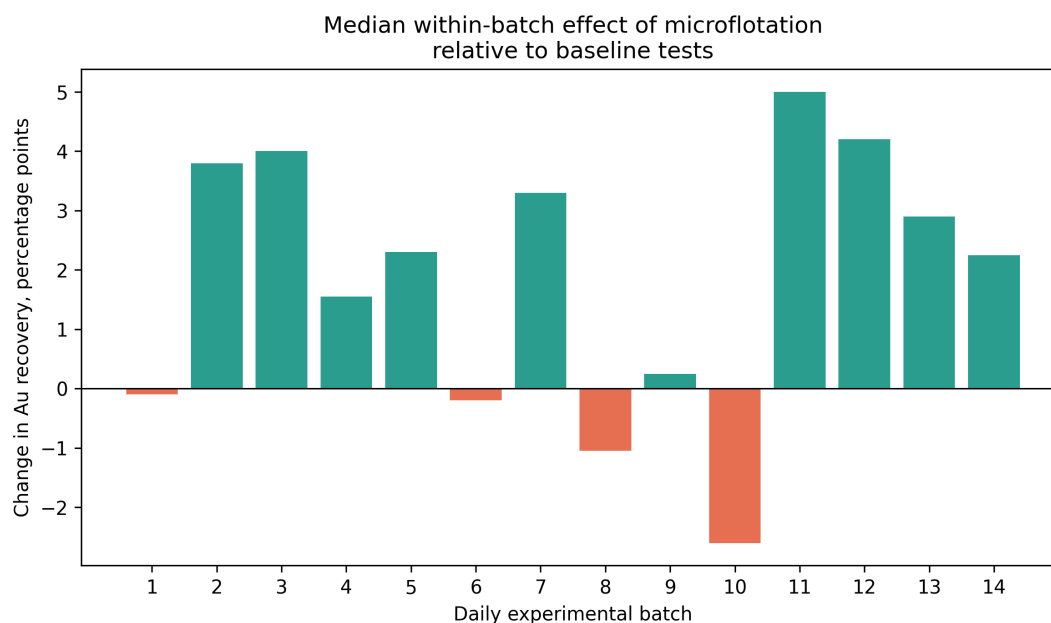


Figure 2. Within-batch change in median Au recovery, percentage points. Each bar represents the difference between the median recovery of microflotation tests and baseline tests within the same daily batch; positive values favour microflotation

For enrichment efficiency, the median within-batch increase was 1.750 percentage points, with a 95% confidence interval of -0.350 to 2.750 percentage points and $p=0.0735$. Since the interval includes zero and the p -value does not reach the conventional $\alpha=0.05$ level, this result is described as a statistical tendency only, and not as a confirmed effect. It is worth noting that the recovery effect and the efficiency effect need not coincide, as efficiency also reflects concentrate grade and yield.

3.2. Classification of reagent regimes

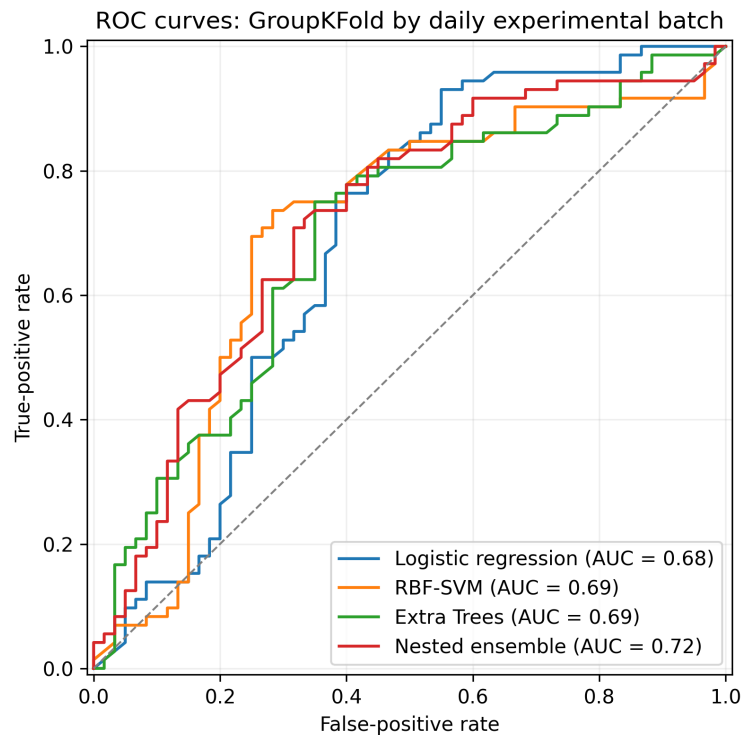
Classification results under five-fold GroupK-Fold validation by daily batch are given in Table 2. This analysis is auxiliary to the primary within-batch experimental comparison and is intended to assess whether the recorded reagent-regime features contain a reproducible screening signal. The nested weighted ensemble achieved the highest ROC-AUC

at 0.717, with a cluster bootstrap 95% confidence interval of 0.632-0.792, and the highest PR-AUC at 0.722. A ROC-AUC of 0.717 corresponds to moderate discrimination. It should not be read as high predictive accuracy, and it is not sufficient to support the use of the model for prescribing reagent regimes. The lower bound of the interval, 0.632, is the more informative figure for practical purposes: the data are consistent with performance that is only modestly better than the individual models.

The differences between the four approaches are small relative to the width of the confidence intervals, which overlap substantially. The ensemble's advantage in ROC-AUC over the compact RBF-SVM (0.717 against 0.694) should therefore be treated as provisional rather than as a demonstrated ranking. The ROC curves for the grouped cross-validation folds are shown in Figure 3, where the variation between folds is visible.

Table 2. Classification of regimes with Au recovery $\geq 84\%$, GroupKFold by batch (n=132; 72 positive)

Model	ROC-AUC	PR-AUC	Bal. acc.	Precision	Recall	Specificity	F1
Logistic regression (total features)	0.684	0.645	0.665	0.663	0.847	0.483	0.744
RBF-SVM (compact features)	0.694	0.676	0.683	0.701	0.750	0.617	0.725
Extra Trees (total features)	0.692	0.696	0.653	0.659	0.806	0.500	0.725
Nested weighted ensemble	0.717	0.722	0.642	0.645	0.833	0.450	0.727

**Figure 3.** ROC curves under GroupKFold validation by daily batch. The spread reflects the small number of independent batches

The ensemble's operating characteristics are asymmetric. Recall is comparatively high at 0.833, meaning that most regimes achieving recovery of at least 84% were identified, but specificity is low at 0.450, meaning that fewer than half of the regimes below the threshold were correctly rejected. Precision

is 0.645, so roughly a third of the regimes flagged as promising did not in fact reach the threshold. Balanced accuracy is 0.642, which is lower than that of the RBF-SVM (0.683), reflecting the fact that the thresholds selected in the inner folds favoured sensitivity. The confusion matrix is shown in Figure 4.

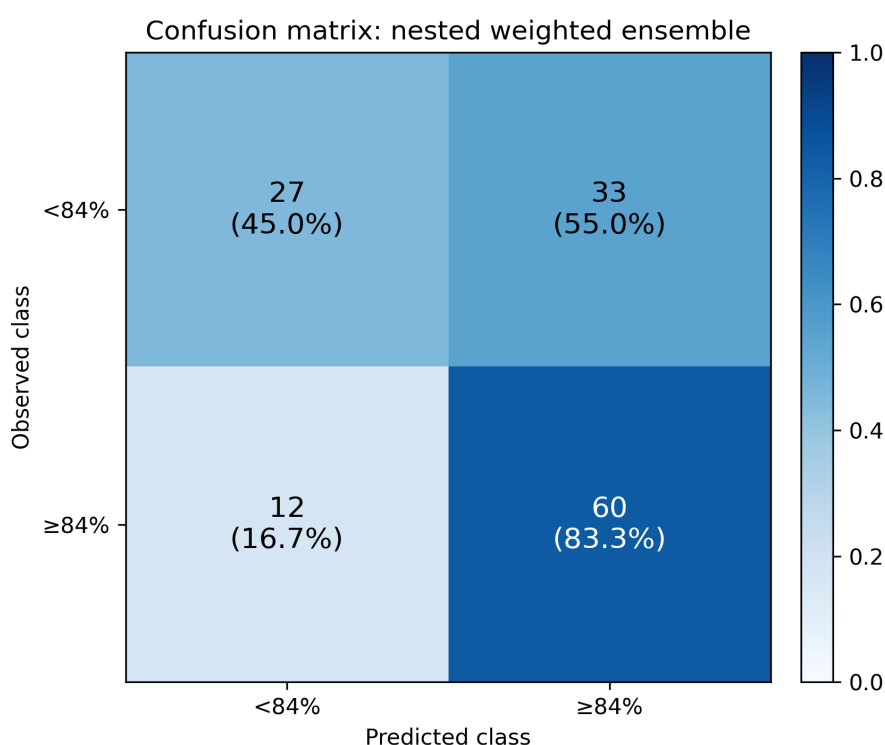


Figure 4. Confusion matrix of the nested weighted ensemble. Of 60 negative tests, 33 were misclassified as positive (specificity 0.450)

For a screening application in which the cost of experimentally testing an unpromising regime is modest but the cost of overlooking a potentially promising regime is higher, the high-recall operating point may be useful for prioritising candidates for follow-up laboratory testing. However, the low specificity means that the model should not be used to reject regimes or prescribe reagent settings without additional experimental confirmation.

4. Discussion

The within-batch analysis provides moderate evidence that microflotation increased Au recovery under the conditions tested. Restricting comparisons to tests performed on the same daily feed sample removes the between-day variation in mineralogy, size distribution, and head grade, which in this test programme is the most obvious confounder. The resulting estimate, a median increase of 2.275 percentage points, is of a magnitude that would be of practical interest for a gold concentrator if it were reproduced at scale. But the confidence interval reaching 0.025

percentage points at its lower bound means the present data cannot exclude an effect close to zero, and the estimate rests on only 14 independent observations at the batch level. The result should be read as encouraging rather than conclusive.

The variability of the effect across batches deserves attention in its own right. The fact that the difference was negative in 4 of 14 batches suggests that the response to microflotation depends on feed characteristics, on the specific reagent regime, or on both. This heterogeneity is therefore not merely statistical noise but an important practical result: the technology may be beneficial only under a subset of feed and operating conditions. The dataset does not contain the variables needed to resolve this dependence, because pH, particle size distribution, mineralogy, aeration rate, and stage durations were not recorded in a form suitable for analysis. Future experiments should record these variables systematically so that conditions associated with positive and negative MMF responses can be identified directly.

The weaker result for enrichment efficiency is consistent with the recovery effect being accompa-

nied by some change in concentrate grade or yield, since the Hancock-Luyken formula incorporates both. A recovery gain that is partly offset in the efficiency metric is a familiar pattern in flotation and is not by itself surprising. Given $p=0.0735$ and an interval spanning zero, no claim is made here beyond the existence of a tendency in the same direction.

The classification results should be interpreted with corresponding restraint. A ROC-AUC of 0.717 indicates that the structured reagent features carry some information about whether a regime reaches 84% recovery, but not enough for reliable prediction. Several factors plausibly contribute. The number of independent batches is small, and grouped validation, while necessary for an honest estimate, leaves few groups per fold. Several MP-series reagents appear in only 5 to 13 tests each, which is too few for their individual contributions to be estimated with any confidence. And the omitted process variables are likely to account for a substantial part of the variance in recovery [12, 13].

The instability of the inner-fold selections reinforces this reading. Ensemble weights and decision thresholds varied considerably between folds, with thresholds ranging from 0.280 to 0.685. A model whose optimal operating point shifts this much across folds is not one whose behaviour has been pinned down by the data [18].

It bears emphasis that the associations found by the classifier do not establish causal effects of individual reagents. The reagent regimes were not assigned by randomisation, and dosages of different reagents are correlated with one another and with the microflotation indicator by virtue of how the test programme was designed. A feature that appears informative may be acting as a proxy for the configuration in which it was tested. Attributing recovery gains to a particular reagent on the basis of these models would not be warranted.

Finally, the scope of the conclusions is confined to the ore investigated. The tests were laboratory tests on samples of industrial feed, not industrial trials, and there is no independent industrial test set against which the models could be checked. The practical claims made for MMF in the introduction, particularly regarding kinetics and throughput, are not addressed by the present data and remain to be evaluated. A larger prospective campaign should therefore combine more independent feed batches with systematic recording of pH, particle size distribution, mineralogy, aeration, and stage duration, followed by external validation before any prescriptive use of the classifier is considered.

5. Conclusions

Several constraints bound what can reasonably be drawn from this work, and they are best stated before the findings themselves. The most fundamental is the number of independent observations: although 132 tests were performed, the tests within a single day share one feed sample, so the effective sample size for the within-batch analysis is the number of batches, and 14 is small. The dataset contains no pH, particle size distribution, mineralogy, aeration, or stage-duration features, because these were not recorded in the source materials in analysable form and were not reconstructed for this work; their absence limits both the interpretation of the microflotation effect and the classification performance that could be achieved. All of the tests were laboratory tests, and although the feed samples were taken from an operating plant, the results therefore do not constitute industrial validation, nor was an independent industrial test set available against which the models could be checked. The machine learning associations do not establish causal effects of individual reagents, for the reasons set out in Section 4, and no claim of applicability beyond the investigated ore is made. It should also be recorded that a quantitative regression formulation for Au recovery was attempted first and produced an unacceptable coefficient of determination, with R^2 close to zero under grouped validation, which is what motivated the classification formulation reported here; that outcome is itself informative, since it indicates that the available features do not support point prediction of recovery on this dataset.

Read in that light, the results are as follows. Within daily experimental batches, median Au recovery was higher under multicomponent microflotation than in baseline tests in 10 of 14 batches, with a median within-batch increase of 2.275 percentage points (cluster bootstrap 95% CI 0.025-3.650; Wilcoxon signed-rank $p=0.0166$); since the lower bound of the interval lies close to zero, this evidence is moderate rather than conclusive, and it is further qualified by the fact that the effect was not uniform, ranging from -2.6 to +5.0 percentage points across batches and so suggesting a dependence on feed conditions that the present dataset cannot resolve. The within-batch change in enrichment efficiency (median +1.750 percentage points; 95% CI -0.350-2.750; $p=0.0735$) points in the same direction but is reported only as a statistical tendency, not as a confirmed effect. On the methodological side, free-text reagent regimes were converted into tabular features

covering microsuspension dosage, base dosage, and total dosage for each reagent, with all 132 records parsed without missing values; under five-fold GroupKFold validation by daily batch, which prevents the same feed sample from appearing in both training and testing, a nested weighted ensemble classified regimes with Au recovery of at least 84% at ROC-AUC 0.717 (95% CI 0.632–0.792) and PR-AUC 0.722. This represents moderate discrimination and is not sufficient for prescriptive use, and the model is markedly more sensitive (recall 0.833) than specific (specificity 0.450), which may suit a screening role but limits its value as a filter for rejecting unpromising regimes. The principal directions for further work follow directly from the limits set out above: the accumulation of a larger number of independent feed batches, and the systematic recording of pH, size distribution, mineralogy, aeration, and stage durations, without which neither the conditions favouring microflotation nor the contributions of individual reagents can be established.

Funding

This research was funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR24992757).

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, A.R., B.D. and B.K.; *Methodology*, A.R., S.A. and B.D.; *Software*, A.R.; *Validation*, B.D. and B.K.; *Formal Analysis*, A.R.; *Investigation*, A.R., S.A. and B.D.; *Resources*, B.D. and B.K.; *Data Curation*, A.R. and B.D.; *Writing—Original Draft Preparation*, A.R., S.A.; *Writing—Review and Editing*, A.R., S.A., B.D. and B.K.; *Visualization*, A.R.; *Supervision*, B.K.; *Project Administration*, B.K.; *Funding Acquisition*, B.K. All authors have read and agreed to the published version of the manuscript.

References

1. T. Miettinen, J. Ralston, and D. Fornasiero, “The limits of fine particle flotation,” *Miner. Eng.*, vol. 23, no. 5, pp. 420–437, 2010, doi: 10.1016/j.mineng.2009.12.006.
2. W. J. Trahar, “A rational interpretation of the role of particle size in flotation,” *Int. J. Miner. Process.*, vol. 8, no. 4, pp. 289–327, 1981, doi: 10.1016/0301-7516(81)90019-3.
3. R.-H. Yoon, “The role of hydrodynamic and surface forces in bubble-particle interaction,” *Int. J. Miner. Process.*, vol. 58, no. 1–4, pp. 129–143, 2000, doi: 10.1016/S0301-7516(99)00071-X.
4. D. Tao, “Role of bubble size in flotation of coarse and fine particles — A review,” *Sep. Sci. Technol.*, vol. 39, no. 4, pp. 741–760, 2005, doi: 10.1081/SS-120028444.
5. R. T. Rodrigues and J. Rubio, “New basis for measuring the size distribution of bubbles,” *Miner. Eng.*, vol. 16, no. 8, pp. 757–765, 2003, doi: 10.1016/S0892-6875(03)00181-X.
6. J. Rubio, M. L. Souza, and R. W. Smith, “Overview of flotation as a wastewater treatment technique,” *Miner. Eng.*, vol. 15, no. 3, pp. 139–155, 2002, doi: 10.1016/S0892-6875(01)00216-3.
7. S. Calgaroto, K. Q. Wilberg, and J. Rubio, “On the nanobubbles interfacial properties and future applications in flotation,” *Miner. Eng.*, vol. 60, pp. 33–40, 2014, doi: 10.1016/j.mineng.2014.02.002.
8. W. Zhou, H. Chen, L. Ou, and Q. Shi, “Aggregation of ultra-fine scheelite particles induced by hydrodynamic cavitation,” *Int. J. Miner. Process.*, vol. 157, pp. 236–240, 2016, doi: 10.1016/j.minpro.2016.11.003.
9. N. Ahmed and G. J. Jameson, “The effect of bubble size on the rate of flotation of fine particles,” *Int. J. Miner. Process.*, vol. 14, no. 3, pp. 195–215, 1985, doi: 10.1016/0301-7516(85)90003-1.
10. R. Dunne, “Flotation of gold and gold-bearing ores,” in *Gold Ore Processing: Project Development and Operations*, 2nd ed. Amsterdam, Netherlands: Elsevier, 2016, pp. 315–338, doi: 10.1016/B978-0-444-63658-4.00018-9.
11. G. C. Allan and J. T. Woodcock, “A review of the flotation of native gold and electrum,” *Miner. Eng.*, vol. 14, no. 9, pp. 931–962, 2001, doi: 10.1016/S0892-6875(01)00103-0.
12. I. Jovanović and I. Miljanović, “Contemporary advanced control techniques for flotation plants with mechanical flotation cells — A review,” *Miner. Eng.*, vol. 70, pp. 228–249, 2015, doi: 10.1016/j.mineng.2014.09.022.
13. J. T. McCoy and L. Auret, “Machine learning applications in minerals processing: A review,” *Miner. Eng.*, vol. 132, pp. 95–109, 2019, doi: 10.1016/j.mineng.2018.12.004.
14. Y. Pu, A. Szmigiel, and D. B. Apel, “Purities prediction in a manufacturing froth flotation plant: The deep learning techniques,” *Neural Comput. Appl.*, vol. 32, pp. 13639–13649, 2020, doi: 10.1007/s00521-020-04773-2.
15. F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
16. G. Varoquaux, P. R. Raamana, D. A. Engemann, A. Hoyos-Idrobo, Y. Schwartz, and B. Thirion, “Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines,” *NeuroImage*, vol. 145, pp. 166–179, 2017, doi: 10.1016/j.neuroimage.2016.10.038.

17. D. R. Roberts et al., “Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure,” *Ecography*, vol. 40, no. 8, pp. 913–929, 2017, doi: 10.1111/ecog.02881.
18. G. C. Cawley and N. L. C. Talbot, “On over-fitting in model selection and subsequent selection bias in performance evaluation,” *J. Mach. Learn. Res.*, vol. 11, pp. 2079–2107, 2010.

Information about the authors:

Adilbek Rustemov – PhD student at the Kazakh-British Technical University (Almaty, Kazakhstan, adilbek.rustemov@mail.mcgill.ca, ORCID ID: 0009-0009-2866-9295).

Bakyt Daulet – Lead Engineer at JSC «Institute of Metallurgy and Ore Beneficiation». (Almaty, Kazakhstan, Daulet.B@stud.satbayev.university, ORCID ID: 0009-0007-0518-9526).

Bagdaulet Kenzhaliyev – General Director of JSC «Institute of Metallurgy and Ore Beneficiation». (Almaty, Kazakhstan, bagdaulet_k@satbayev.university, ORCID ID: 0000-0003-1474-8354).

Serik Aibagarov – Junior Researcher, LLP «DigitAlem». (Almaty, Kazakhstan, awer1307dot@gmail.com, ORCID ID: 0009-0009-4946-4926).

Submission received: 28 July, 2026

Revised: 20 August, 2026

Accepted: 20 August, 2026