

Statistical Tools for Experimental Design, Inference and Process Optimisation in Mineral Processing: Review of ANOVA, Response Surface Methods and Data-Driven Models

I. Isa ¹, M. Ali ², M.N.S Udaini ³

^{1,2,3} Department of Mineral and Petroleum Resources Engineering, Kaduna Polytechnic, Kaduna, Nigeria
Email: isaibrahim@kadunapolytechnic.edu.ng

ARTICLE INFO

Keywords:

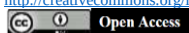
analysis of variance,
design of experiments;
mineral processing,
response surface
methodology,
regression,
process optimisation;
flotation,
critical minerals,
machine learning.

Vol.01, No.04, 2026

<https://doi.org/10.5281/zenodo.22635428>

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CCBY4.0).

<http://creativecommons.org/licenses/by/4.0/>



ABSTRACT

Mineral processing research concerns heterogeneous materials and strongly interacting operating variables, so experimental conclusions depend as much on design quality and statistical reasoning as on laboratory technique. This critical review examines the connected roles of descriptive analysis, analysis of variance (ANOVA), design of experiments (DOE), regression, response surface methodology (RSM), multivariate analysis, mixed-effects modelling and data-driven surrogate models in mineral processing. The discussion covers comminution, classification, gravity and dense-medium separation, magnetic separation, froth flotation, hydrometallurgy, dewatering and the beneficiation of lithium-bearing and other critical-mineral ores. Particular attention is given to experimental units, randomisation, replication, blocking, interaction effects, assumption diagnostics, effect sizes, model uncertainty, multiple-response optimisation and confirmation testing. ANOVA remains valuable for partitioning variation and testing designed effects, but p-values alone provide an incomplete basis for engineering decisions. Regression and RSM support prediction and optimisation when model form is adequate, while mixed and multivariate methods address hierarchical sampling structures and correlated responses. For nonlinear or stochastic systems, Gaussian-process, artificial-neural-network and ensemble models may improve prediction, provided that cross-validation, sensitivity analysis, interpretability and independent experimental verification accompany their use. The article proposes a sequential framework that links ore characterisation, DOE, statistical inference, predictive modelling, constrained optimisation and scale-up validation. It also presents a reporting checklist intended to improve reproducibility and comparability. The central conclusion is that mineral-processing statistics should function as an integrated research architecture rather than as a final software step performed after testwork. This approach can reduce unnecessary experiments, expose factor interactions, quantify uncertainty and produce more transferable process knowledge.

1. Introduction

Mineral processing converts run-of-mine material into products that satisfy metallurgical, chemical or physical specifications. The field encompasses comminution, sizing and classification, gravity and dense-medium separation, magnetic and electrostatic separation, froth flotation, leaching, solid-liquid separation and product handling. Each operation responds to a combination of material characteristics and operating conditions. Mineralogical composition, texture, liberation, particle shape and size distribution interact with reagent chemistry, pulp conditions, residence time, energy input and equipment configuration. As a result, two tests conducted under nominally identical settings may produce different grades, recoveries or yields because the feed and the measurement system vary as well as the controlled factors (Wills and Finch, 2016).

This variability creates a scientific problem that cannot be solved through arithmetic comparison of treatment means alone. A higher recovery under one condition may reflect a genuine process effect, random sampling error, a systematic difference between ore batches or an interaction that was not represented in the experimental plan. Statistical reasoning supplies the language and tools needed to separate these possibilities. It guides the definition of experimental units, the allocation of treatments,

the estimation of residual variation, the testing of effects and the assessment of model uncertainty. In this sense, statistics supports both scientific explanation and engineering decision-making.

ANOVA has long occupied a central position in mineral-engineering research because it partitions observed variability into components associated with specified factors and experimental error. Applications include flotation, coal-characterisation and multi-factor beneficiation studies (Niedoba and Pięta, 2016). However, ANOVA represents one element of a larger analytical system. DOE determines whether effects and interactions can be estimated efficiently; regression translates factor effects into predictive relationships; RSM describes local curvature and searches for favourable operating regions; multivariate methods evaluate correlated responses; and mixed-effects models represent blocks, ore domains, operators or repeated measurements. More recent work has combined DOE with artificial neural networks, stochastic simulation and global sensitivity analysis when polynomial response surfaces describe flotation behaviour inadequately (Arancibia-Bravo et al., 2022; Sepúlveda et al., 2024).

The increasing attention given to critical minerals strengthens the need for this integrated perspective. Lithium ores, tailings and brines present variable mineral associations and often require multi-stage beneficiation or extraction. The experimental objective may involve several competing responses, such as lithium recovery, concentrate grade, impurity rejection, acid consumption and energy demand. A method that maximises one response without considering the others can produce an operating condition that has little technical or economic value. Statistical optimisation must therefore incorporate multiple responses, constraints, uncertainty and confirmation testing.

This article critically reviews statistical tools used in mineral-processing research and organises them around the research cycle. It first defines the review scope and the foundations of statistical thinking. It then discusses DOE, ANOVA, regression, RSM, multivariate and advanced data-driven methods, followed by applications across major unit operations. The paper concludes with an integrated framework, reporting checklist and research agenda. The aim is not to promote a single method, but to show how method choice should follow the experimental question, data structure and engineering decision.

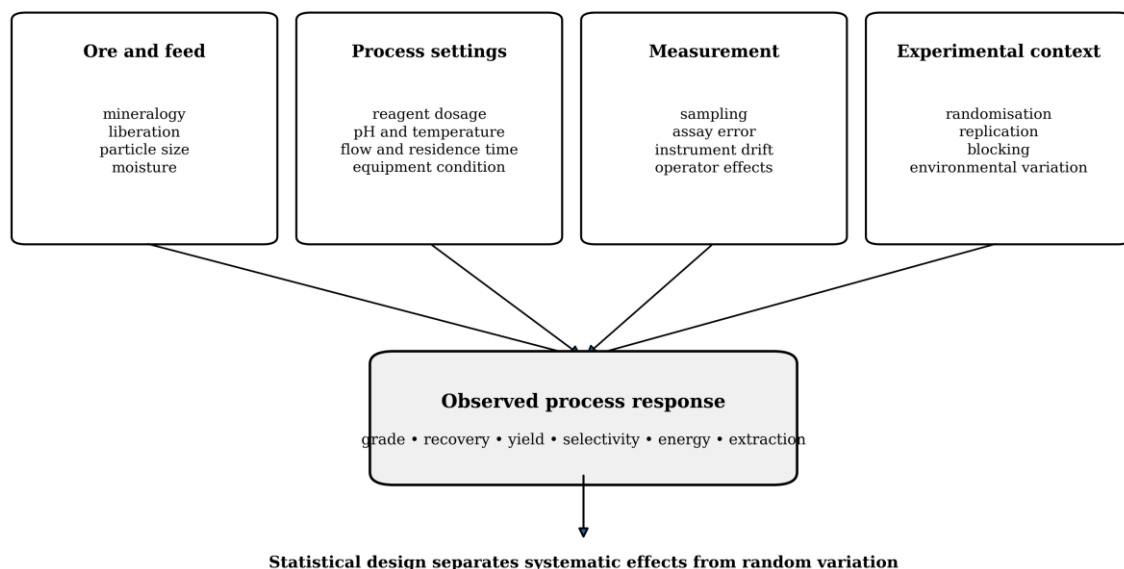


Figure 1. Principal sources of variation in mineral-processing experiments and the role of statistical control. Source: Author's conceptualisation.

2. Review Scope and Analytical Approach

This paper adopts a critical narrative review design. It does not claim the exhaustive coverage or formal risk-of-bias assessment associated with a systematic review or meta-analysis. The literature was selected to represent the principal statistical concepts used in mineral-processing experimentation, foundational texts in experimental design and response-surface methods, and illustrative applications in

flotation, hydrometallurgy, coal and critical-mineral processing. Recent publications were included where they demonstrate the movement from conventional ANOVA and polynomial models towards multivariate, stochastic and machine-learning approaches.

The analytical organisation follows the sequence in which statistical decisions arise during research. The review begins with material and process characterisation, then considers factor selection, experimental design, exploratory analysis, statistical inference, predictive modelling, optimisation and validation. This structure differs from a catalogue of statistical tests because it treats design and analysis as connected decisions. A statistically elaborate model cannot repair an experiment that lacks randomisation, replication or an identifiable experimental unit. Similarly, a well-designed experiment can still produce weak conclusions when authors report only p-values and omit effect magnitudes, uncertainty and diagnostic evidence.

The review uses examples from several mineral-processing operations to distinguish general principles from process-specific considerations. Flotation receives substantial attention because it involves many interacting chemical, hydrodynamic and mineralogical variables. However, the same design principles apply to grinding, gravity separation, magnetic separation, leaching and dewatering. The proposed framework is therefore intended for laboratory, pilot and plant-based investigations, although the sources of variation and feasible designs differ across these scales.

Table 1. Representative factors, sources of uncontrolled variation and responses in mineral-processing studies.

Unit operation	Typical controllable factors	Important material or nuisance variables	Common responses
Comminution and classification	mill speed, grinding time, media size, solids concentration, classifier settings	ore hardness, mineral texture, feed size, liner condition	P80, specific energy, throughput, circulating load, liberation
Gravity and dense-medium separation	particle-size range, fluid flow, deck slope, stroke, medium density	particle shape, density distribution, near-density material, feed rate	grade, recovery, yield, separation efficiency, probable error
Magnetic/electrostatic separation	field intensity, rotor speed, gap, feed rate, voltage	magnetic susceptibility, conductivity, moisture, particle size	grade, recovery, mass pull, impurity rejection
Froth flotation	pH, collector, frother, depressant, air rate, pulp density, conditioning and flotation time	mineralogy, oxidation, liberation, water chemistry, clay content	grade, recovery, yield, selectivity, kinetics, water recovery
Hydrometallurgy	temperature, time, reagent concentration, liquid–solid ratio, agitation	mineral association, passivation, porosity, impurity chemistry	extraction, selectivity, reagent consumption, kinetics, residue quality
Dewatering and filtration	flocculant dosage, pressure, vacuum, cycle time, bed depth	particle-size distribution, surface chemistry, clay content	moisture, settling rate, filtrate clarity, throughput

3. Statistical Foundations for Mineral-Processing Experiments

3.1 Factors, responses and experimental error

A mineral-processing experiment normally contains controlled factors, measured responses and sources of error. Factors may be quantitative, such as pH, temperature, grinding time or reagent dosage,

or qualitative, such as collector type, ore domain or equipment configuration. Responses may include recovery, concentrate grade, mass yield, selectivity index, energy consumption, moisture content or extraction. Several responses are often recorded from the same run, and they are usually correlated. Increasing mass pull, for example, can raise recovery while lowering concentrate grade.

The general relationship between a response and its determinants can be represented as follows:

$$y = f(x_1, x_2, \dots, x_k) + \varepsilon \quad (1)$$

In Equation (1), y denotes the measured response, x_1 to x_k denote controllable or recorded factors, f represents the systematic process relationship and ε represents unexplained variation. The error term includes sampling variation, assay uncertainty, unrecorded process disturbances and natural differences among experimental units. Statistical analysis estimates the systematic component without pretending that the error component disappears.

3.2 Experimental units, replication and pseudoreplication

The experimental unit is the smallest unit that can independently receive a treatment. In batch flotation, a separately prepared and conditioned pulp may constitute an experimental unit. Repeated assays from the same concentrate do not create independent process replicates; they quantify analytical repeatability. Treating subsamples or duplicate assays as independent experimental runs inflates the apparent sample size, underestimates error and produces pseudoreplication. Researchers should distinguish process replication, sampling replication and analytical replication in the method section and statistical model.

Replication estimates experimental error and improves the precision of factor effects. However, replication should reflect the variation relevant to the intended conclusion. Three repeated measurements from one homogeneous pulp do not represent day-to-day feed variation. When the research question concerns robustness to ore variability, replication should include independently collected or prepared ore lots, and ore lot may enter the model as a block or random effect.

3.3 Randomisation, blocking and nuisance variables

Randomisation protects the experiment against systematic ordering effects. Laboratory temperature, instrument drift, reagent ageing, operator fatigue and progressive changes in recycled process water can coincide with the run order. Random allocation spreads these influences across treatments, allowing the residual error to represent them more fairly. Restricted randomisation may be required when factor changes are expensive or when a split-plot design reflects the way equipment can be operated.

Blocking groups experimental units that share a nuisance condition. Ore batch, day, operator, instrument, flotation cell or leaching reactor can serve as blocks. A randomised complete block design removes block-to-block variation from the error term and can substantially increase statistical power. When block levels represent a sample from a larger population, mixed-effects models offer a suitable analysis because they estimate fixed treatment effects and random variation among blocks.

3.4 Descriptive statistics and exploratory analysis

Exploratory analysis should precede formal testing. Means and standard deviations summarise central tendency and repeatability, while medians and interquartile ranges can be informative when distributions are skewed. The coefficient of variation permits relative comparison among responses that use different units, although it can become unstable when the mean approaches zero. Minimum and maximum values help expose factor-setting or transcription errors.

Graphical analysis provides information that a table of means cannot. Box plots compare treatment distributions; scatter plots reveal curvature and heteroscedasticity; interaction plots show whether factor effects change across other factor levels; and run-order plots reveal drift. Residual plots, normal quantile plots and influence diagnostics become important after a model is fitted. Outliers should prompt investigation of sampling, assay and process records. Automatic deletion solely because an observation appears unusual can remove genuine process behaviour and produce biased results.

4. Design of Experiments

DOE structures the allocation of experimental effort so that factor effects and interactions can be estimated with controlled uncertainty. The one-factor-at-a-time approach changes one variable while

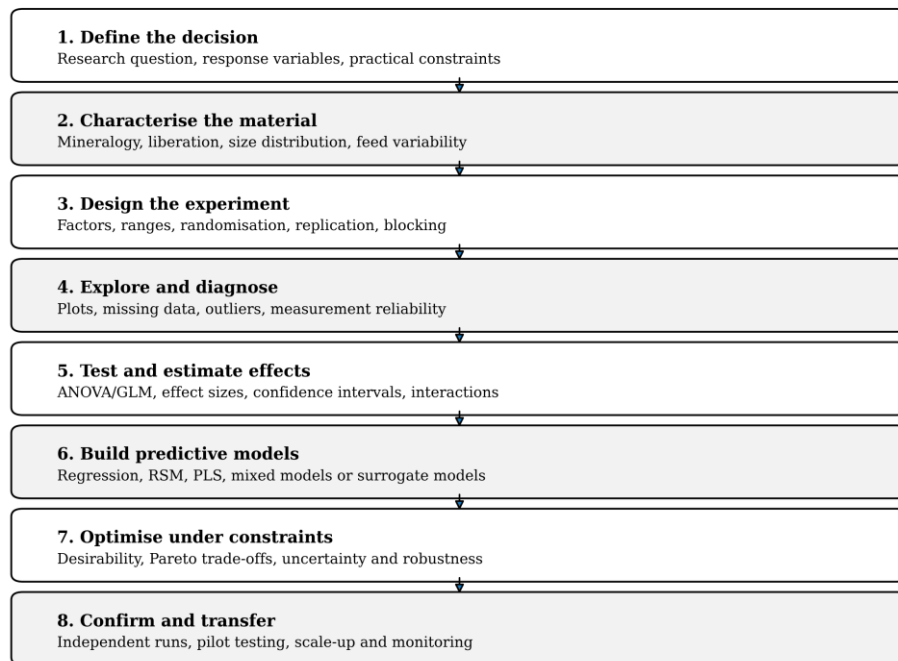
holding the others constant. Although easy to understand, it cannot efficiently estimate interactions and often requires more runs to explore the same factor space. Factorial designs vary factors together and make interactions estimable, which is particularly important in flotation and leaching where chemical effects depend on pH, temperature, particle size or mineralogy (Montgomery, 2020).

Experimental design should follow the research objective. Screening designs identify a small set of influential factors from a larger candidate list. Full or fractional factorial designs estimate main effects and selected interactions over a defined region. Response-surface designs add centre and axial or intermediate points so that curvature can be estimated. Mixture designs apply when proportions sum to a constant, as in reagent blends or composite media. Split-plot designs apply when some factors are difficult to change and others can be varied for each run.

Factor ranges require engineering judgement. A statistically convenient range may cross unsafe, chemically meaningless or industrially impossible conditions. Conversely, a range that is too narrow can make important effects difficult to detect and can conceal curvature. Preliminary tests, mineralogical evidence and plant constraints should therefore inform the coded levels. Centre points help estimate pure error and reveal lack of fit, while confirmation points test whether the fitted model remains accurate away from the original design points.

Table 2. Selection guide for experimental designs used in mineral-processing research.

Design	Best use	Strengths	Limitations and cautions
Completely randomised one-factor design	comparison of one factor at several levels under homogeneous conditions	simple analysis; direct treatment comparison; supports post-hoc tests	cannot estimate interactions; sensitive to hidden gradients when blocking is needed
Randomised block design	treatment comparison across ore lots, days, operators or machines	removes known nuisance variation; improves precision	requires each block to contain suitable treatment combinations; missing runs complicate analysis
Full factorial design	estimation of main effects and interactions for a modest number of factors	complete information on included interactions; efficient use of runs	run count grows rapidly as factors and levels increase
Fractional factorial or Plackett–Burman design	screening many variables	substantial run reduction; identifies candidates for later study	aliasing limits interpretation; strong interactions can distort main effects
Central composite design	quadratic modelling and local optimisation	estimates linear, interaction and quadratic terms; flexible rotatability	axial points may fall outside practical limits; sensitive to model misspecification
Box–Behnken design	quadratic modelling without extreme corner combinations	fewer extreme settings; efficient for three or more factors	does not include factorial corners; unsuitable when boundary behaviour is central
Mixture design	optimisation of component proportions	respects the constant-sum constraint; estimates blending effects	ordinary factorial analysis is inappropriate; interpretation depends on mixture geometry
Split-plot design	experiments with hard-to-change and easy-to-change factors	matches operational reality; reduces costly changeovers	requires correct whole-plot and subplot error terms; ordinary ANOVA is invalid



Feedback from validation can require revised factor ranges, new blocks or a different model class.

Figure 2. Sequential statistical workflow for mineral-processing research. Source: Author's conceptualisation.

5. Analysis of Variance and Related Inference

5.1 Partitioning variation and testing factor effects

ANOVA compares the variation explained by a factor or model with the unexplained variation among observations. In a one-way design, the total sum of squares is partitioned into between-treatment and residual components. In factorial designs, the model component is further separated into main effects and interactions. The basic decomposition can be written as:

$$SS_{total} = SS_{model} + SS_{error} \quad (2)$$

Each sum of squares is divided by its degrees of freedom to obtain a mean square. The F statistic compares the mean square associated with a factor against an appropriate error mean square:

$$F = MS_{factor} / MS_{error} \quad (3)$$

A large F ratio indicates that variation among fitted factor levels is large relative to residual variation. The p-value expresses how unusual such an F ratio would be under the specified null hypothesis. It does not measure the size of the effect, the probability that the null hypothesis is true or the probability that a result will replicate. Authors should therefore report estimated differences, confidence intervals and effect sizes alongside p-values.

For a factorial experiment, interaction terms deserve substantive interpretation. A significant pH-by-collector interaction means that the collector effect cannot be summarised without stating the pH level. Main effects averaged over an interaction can mislead. Interaction plots, estimated marginal means and simple-effect comparisons often communicate the result more effectively than a list of p-values.

5.2 One-way, factorial, nested and mixed ANOVA

One-way ANOVA applies when one categorical factor has three or more levels and independent replicates are available. A significant omnibus test shows that the group means are not all equal but does not identify which pairs differ. Tukey-type comparisons or planned contrasts may follow, with adjustment for multiple testing. Planned contrasts are preferable when the hypotheses were defined before the experiment, such as a linear dosage trend or a comparison between a control and the average of several treatments.

Factorial ANOVA analyses two or more factors and their interactions. Nested designs arise when lower-level units belong exclusively to higher-level units, such as subsamples nested in ore lots. Repeated measurements on the same experimental unit violate independence and require repeated-measures or mixed-effects analysis. Mixed models are especially valuable in pilot and plant studies because they can represent random variation among days, shifts, cells, feed campaigns or sampling locations while estimating fixed process effects.

5.3 Assumptions, diagnostics and robust alternatives

Classical ANOVA assumes independent errors, an appropriate mean structure, approximately constant residual variance and residual distributions that support the inferential procedure. Normality concerns the residuals, not the raw response values. With balanced designs and moderate replication, ANOVA can tolerate modest departures from normality, but strong skewness, heavy tails, unequal variances or influential observations can affect results. Independence violations are more serious because they alter the effective sample size.

Diagnostics should combine graphical and formal evidence. Residual-versus-fitted plots reveal changes in spread and unmodelled curvature. Quantile–quantile plots reveal heavy tails or skewness. Levene-type tests assess equality of variance, while the Shapiro–Wilk test can assess normality, although formal tests can reject trivial deviations in large datasets and have low power in very small datasets. Transformations, Welch ANOVA, robust standard errors, generalised linear models or non-parametric methods may be more suitable when assumptions do not hold. The remedy should follow the measurement scale and scientific question, not a mechanical sequence of tests.

Table 3. Common inferential problems, diagnostic evidence and possible remedies.

Issue	Diagnostic evidence	Possible response	Mineral-processing example
Non-independence	repeated observations from the same batch, run-order structure, residual autocorrelation	mixed model, repeated-measures model, time-series or spatial structure; redefine experimental unit	several timed samples taken from one flotation batch
Unequal variance	funnel-shaped residual plot; group spreads increase with mean	variance-stabilising transformation, Welch ANOVA, heteroscedastic model, robust standard errors	recovery variability increases at high mass pull
Skewed or bounded response	asymmetric residuals; values near 0 or 100 per cent	appropriate transformation, beta or binomial model, bootstrap interval	percentage extraction close to the upper boundary
Unmodelled curvature	systematic curved residual pattern; significant lack-of-fit test	add polynomial or mechanistic terms; use RSM or nonlinear regression	grinding response changes rapidly near an optimum time
Influential observations	large studentised residual, leverage or Cook's distance	check records, repeat assay if justified, report sensitivity analysis	one concentrate assay affected by sample contamination
Multiple testing	many factor, response and pairwise tests	pre-specify hypotheses; adjust family-wise error or false discovery rate	screening ten factors against four response variables

5.4 Effect sizes, confidence intervals and practical significance

Engineering decisions depend on magnitude. Eta-squared and omega-squared describe the proportion of variation associated with an effect, while standardised differences and regression coefficients quantify changes on interpretable scales. For process research, an unstandardised effect often has the greatest value: a 0.5 percentage-point increase in concentrate grade, a 4 percentage-point increase in recovery or a 2 kWh/t reduction in energy consumption. Confidence intervals show the range of effects compatible with the data and help distinguish precise small effects from uncertain large effects.

Practical significance should be evaluated against technical, economic and environmental thresholds. A statistically significant improvement may not justify additional reagent cost, energy consumption or impurity penalties. Conversely, a meaningful industrial improvement can fail to reach conventional statistical significance when experimental variability is high and replication is limited. Power analysis and precision-based sample-size planning should therefore occur before testwork, using the smallest effect that would alter an engineering decision.

6. Correlation, Regression and Predictive Modelling

Correlation quantifies association but does not by itself establish process causation. Pearson correlation measures linear association, while Spearman correlation captures monotonic ordering. Strong correlations can arise because two responses share a common denominator, because factor levels were deliberately co-varied or because an unmeasured variable influences both. In flotation, grade and recovery often show a structural trade-off; interpreting their correlation without considering mass balance and operating conditions can be misleading.

Regression estimates the relationship between a response and one or more predictors. Multiple regression permits adjustment for several variables and can estimate interactions, curvature and covariate effects. Categorical ore domains can enter through indicator variables, while continuous mineralogical or particle-size variables can enter as covariates. The interpretation of each coefficient depends on the model specification and should be stated in engineering units.

Model quality should not be judged by R^2 alone. A high R^2 can result from overfitting, a narrow response range or leakage between training and validation data. Adjusted R^2 penalises unnecessary terms but does not replace prediction assessment. Residual diagnostics, prediction intervals, lack-of-fit tests and independent validation are required. When the dataset is small, leave-one-out cross-validation can have high variance; repeated k-fold or grouped cross-validation may be preferable. Grouping is essential when multiple rows come from the same ore lot or experiment because random row-wise splitting can place closely related observations in both training and test sets.

7. Response Surface Methodology and Process Optimisation

RSM combines designed experiments, regression and optimisation to describe a response over a local factor region. A second-order model commonly takes the following form:

$$\hat{y} = \beta_0 + \sum \beta_i x_i + \sum \beta_{ii} x_i^2 + \sum \sum \beta_{ij} x_i x_j \quad (4)$$

The linear terms describe directional change, the squared terms describe curvature and the interaction terms describe conditional factor effects. Central composite and Box–Behnken designs are widely used because they estimate these terms with fewer runs than a full three-level factorial design. ANOVA tests model terms and lack of fit, while contour and three-dimensional response plots communicate the fitted surface.

RSM has supported optimisation of copper flotation in saline systems and oxidised copper flotation under mineralogical constraints (Arancibia-Bravo et al., 2022; Oumesaoud et al., 2024). It has also been applied to lithium extraction and leaching conditions (Yessengaziyev et al., 2025). These applications demonstrate the capacity of RSM to combine factor screening, interaction assessment and local optimisation. However, the fitted optimum is conditional on the chosen model, factor ranges, material tested and response definition.

Multiple-response optimisation requires explicit trade-offs. Desirability functions transform each response onto a common scale and combine them, but the selected weights and target limits can determine the recommended setting. Pareto-front analysis presents non-dominated alternatives without concealing trade-offs in a single score. Robust optimisation goes further by seeking conditions that perform satisfactorily when feed properties or operating settings vary. In mineral processing, a slightly

lower predicted maximum can be preferable when it is less sensitive to ore hardness, water chemistry or reagent delivery error.

Confirmation experiments distinguish mathematical optimisation from process evidence. Replicated tests should be performed at the proposed optimum and, where possible, at nearby settings. The observed responses should be compared with prediction intervals. Validation on a second ore lot or pilot campaign tests transportability. Extrapolation beyond the design region should be avoided unless mechanistic evidence and new experiments support it.

8. Multivariate, Hierarchical and High-Dimensional Methods

8.1 Multivariate responses and latent-variable methods

Mineral-processing datasets often contain many correlated measurements. Particle-size fractions, elemental assays, mineral abundance, image features and time-resolved process signals can create high-dimensional data matrices. Principal component analysis reduces correlated variables to orthogonal directions that describe dominant variation. It supports ore-domain exploration, process monitoring and visualisation, but it does not automatically identify causal process effects.

Partial least squares regression relates a large set of correlated predictors to one or more responses and can be useful when ordinary regression suffers from multicollinearity. Cluster analysis identifies groups of similar samples or operating states, although the result depends on scaling, distance measure and clustering algorithm. MANOVA tests designed effects on a vector of responses, while ANOVA simultaneous component analysis extends the ANOVA idea to structured multivariate data. Recent tutorial work presents ASCA as a bridge between DOE and chemometric interpretation for high-dimensional experimental data (Camacho et al., 2026).

8.2 Hierarchical data and mixed-effects models

Laboratory and industrial datasets frequently have nested or crossed structures. Assays are nested in samples; samples are nested in runs; runs are nested in days or ore campaigns; and several cells may operate in parallel. Mixed-effects models estimate average process effects while acknowledging random variation at these levels. They prevent the false precision that results when all rows are treated as independent.

Random slopes allow factor effects to vary among ore lots or operating campaigns. This can answer a question that fixed-effect ANOVA cannot: whether a treatment works consistently across heterogeneous materials. Variance components also quantify where uncertainty enters the process. If between-lot variation dominates analytical variation, improved sampling and ore-domaining may offer greater benefit than more duplicate assays.

8.3 Compositional and mass-balance constraints

Mineralogical percentages, size fractions and elemental compositions sum to a constant. Standard correlation and regression can produce spurious relationships when applied directly to such compositional data. Log-ratio transformations and compositional-data methods may be necessary, particularly when the objective is to relate mineral proportions to process response. Zero values and detection limits require transparent handling.

Mass-balance reconciliation presents another structured statistical problem. Feed, concentrate and tailings measurements contain error but must satisfy conservation constraints. Data reconciliation adjusts measured values according to their uncertainty so that balances close. The procedure should use realistic measurement-error models; otherwise, highly uncertain variables can exert disproportionate influence on the reconciled solution.

9. Advanced Data-Driven and Hybrid Models

Polynomial RSM models offer interpretability and efficient estimation in designed regions, but some mineral-processing systems show discontinuities, saturation, strong stochasticity or high-order interactions. Artificial neural networks, random forests, gradient-boosted trees and Gaussian-process models can represent nonlinear relationships without specifying a low-order polynomial. Their flexibility can improve prediction, but it also increases the risk of overfitting and obscures the relationship between factors and responses.

Arancibia-Bravo et al. (2022) showed that ANN models and global sensitivity analysis could supplement conventional RSM in copper flotation when polynomial fits were inadequate. More recent flotation research has combined DOE with stochastic simulation to validate locked-cycle predictions (Sepúlveda et al., 2024). Bayesian surrogate models and global sensitivity measures offer another route to uncertainty-aware process understanding, especially when physical experiments are costly. However, advanced models require enough independent data to support their flexibility. A neural network trained on a small, highly replicated dataset may memorise experimental noise.

Interpretability methods should connect prediction to engineering understanding. Partial dependence, accumulated local effects, permutation importance, Sobol indices and SHAP values can indicate influential variables, but each method has assumptions and can behave poorly when predictors are strongly correlated. Sensitivity analysis should therefore accompany domain checks and controlled experiments. A variable that predicts recovery because it acts as a proxy for ore type should not be described as a controllable causal driver.

Hybrid modelling combines mechanistic equations with statistical or machine-learning components. A mass-balance or kinetic model can impose physical structure, while a data-driven component represents unknown sub-processes. Physics-informed models may generalise better than purely empirical models, especially under changing feed conditions. The value of a hybrid model should be demonstrated through out-of-sample prediction, uncertainty calibration and comparison with simpler baselines.

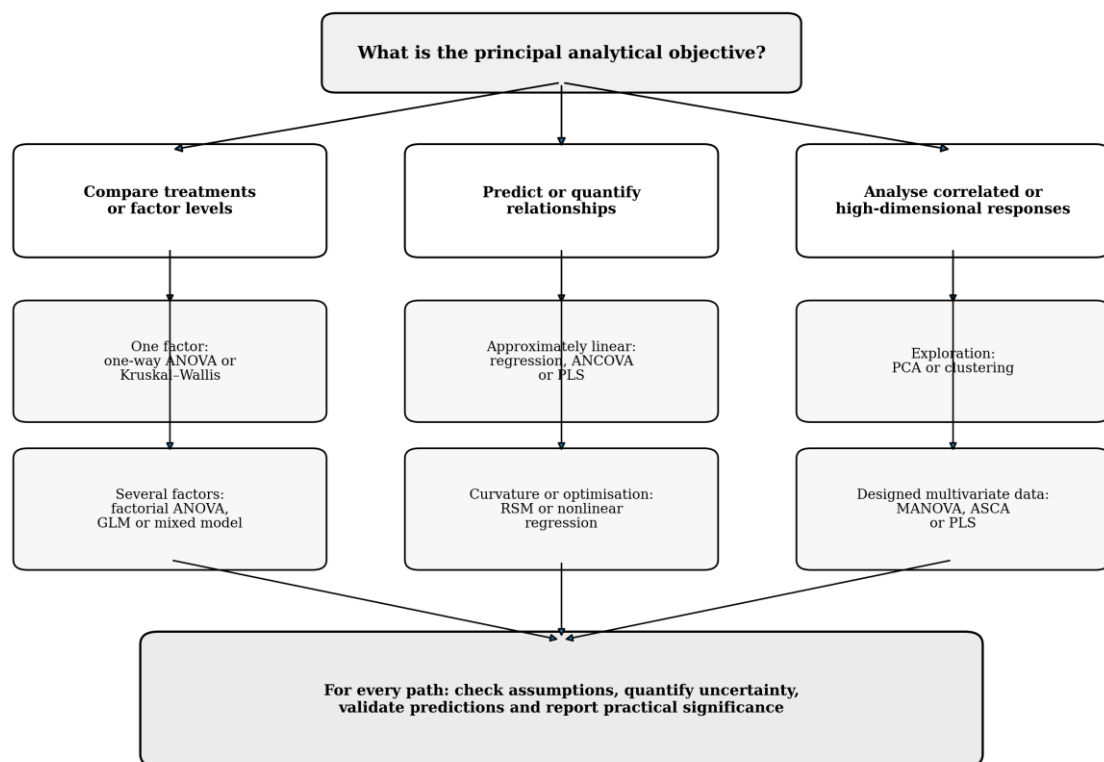


Figure 3. Decision pathway for selecting a statistical method according to the analytical objective and data structure. Source: Author's conceptualisation.

Table 4. Roles, expected outputs and cautions for major statistical method families.

Method	Primary purpose	Recommended outputs	Main caution
Descriptive and graphical analysis	summarise distributions, repeatability, trends and anomalies	mean/median, spread, run-order and response plots, measurement uncertainty	summary statistics cannot separate treatment effects from confounding
ANOVA/GLM	test designed factor effects and interactions	estimated effects, F tests, p-values, confidence intervals, effect sizes	validity depends on design, error term and residual structure
Mixed-effects models	analyse blocks, repeated measurements and hierarchical data	fixed effects, variance components, random-effect estimates, prediction intervals	random structure must reflect sampling and treatment allocation
Regression/RSM	predict responses and optimise factor settings	coefficients, diagnostics, prediction intervals, contour plots, confirmation results	polynomial models can fit poorly outside a local region
PCA/PLS/ASCA	analyse correlated or high-dimensional variables and responses	scores, loadings, explained variation, validated prediction or effect components	scaling and component selection influence interpretation
Machine learning and surrogates	represent nonlinear relationships and complex interactions	cross-validated error, calibration, uncertainty, feature importance, sensitivity indices	high flexibility can overfit and obscure causal interpretation
Bayesian models	combine prior knowledge with data and propagate uncertainty	posterior distributions, credible intervals, posterior predictive checks	results depend on prior specification and model adequacy

10. Applications Across Mineral-Processing Operations

10.1 Comminution and classification

Comminution studies often examine energy input, grinding time, mill speed, media characteristics, solids concentration and feed size. Responses include product P80, specific energy, throughput, circulating load and liberation. Ore hardness and texture can dominate treatment effects, so blocking by ore domain or including hardness as a covariate can improve interpretation. Repeated size analyses from one ground product represent analytical or subsampling replication, not independent grinding runs.

Regression models can relate specific energy to product size, while factorial designs can reveal interactions between feed size and operating conditions. Nonlinear models may be required near grind limits or when breakage regimes change. Scale-up claims should include validation because laboratory mills and industrial circuits differ in transport, classification and energy distribution.

10.2 Gravity and dense-medium separation

Gravity separation responds to particle size, density, shape, wash-water rate, table inclination, stroke and feed rate. Dense-medium performance also depends on medium stability and the proportion of near-density material. DOE can estimate operating-factor interactions, while partition curves and probable error provide process-specific performance measures. Statistical analysis should account for repeated size or density fractions from the same test because they are connected by the partitioning process.

When several product fractions are measured, multivariate or compositional approaches may be more appropriate than separate ANOVAs for each fraction. Robustness analysis should examine how the recommended setting changes as feed density distribution varies.

10.3 Magnetic and electrostatic separation

Field intensity, rotor speed, feed rate, particle size, splitter position and moisture influence magnetic or electrostatic separation. A factorial screening design can identify the dominant settings, followed by RSM over a smaller region. Mineralogical composition and magnetic susceptibility should be recorded because an apparent equipment effect can reflect different feed lots. Grade and recovery should be analysed together, and confirmation tests should include material from a second lot when the intended application covers a variable deposit.

10.4 Froth flotation

Froth flotation represents one of the richest applications of statistical design because chemical, hydrodynamic and mineralogical variables interact. pH alters mineral-surface chemistry and collector adsorption; particle size affects collision, attachment and entrainment; air rate and froth depth affect recovery and selectivity; and water chemistry can change reagent response. A one-factor experiment can therefore produce a result that fails when another condition changes.

Factorial designs, D-optimal designs and RSM have been used to evaluate reagent systems and flotation conditions. Niedoba and Pięta (2016) reviewed ANOVA applications in mineral engineering, while Nakhaei et al. (2016) used D-optimal design for copper–molybdenum sulphide flotation. Arancibia-Bravo et al. (2022) combined RSM concepts with ANN and global sensitivity analysis for copper flotation in saline systems, and Oumesaoud et al. (2024) integrated Box–Behnken optimisation with mineralogical evidence for oxidised copper ore. These studies show that process modelling gains strength when mineralogy accompanies statistical effects.

Flotation datasets also contain kinetic and image information. Time-dependent recovery curves require nonlinear or repeated-measures models. Froth images generate high-dimensional features and temporal dependence. Models for control must distinguish prediction under observed operating conditions from causal response to deliberate changes. Reviews of flotation-control modelling emphasise the challenge of representing nonlinear dynamics and transferring models from simulation to operating systems (Quintanilla et al., 2021).

10.5 Hydrometallurgy

Leaching and extraction studies commonly vary temperature, time, reagent concentration, liquid–solid ratio, agitation and particle size. These variables influence reaction kinetics, mass transfer, passivation and impurity dissolution. Factorial and response-surface designs can estimate interactions and curvature, but kinetic interpretation should not be replaced by polynomial fit alone. Temperature and time effects may have mechanistic forms that support nonlinear kinetic models.

Percentage extraction is bounded and can cluster near zero or 100 per cent, producing non-constant variance. Transformations or beta-type models may be suitable. Multiple-response optimisation should consider target-metal recovery, impurity extraction, reagent consumption and residue stability. Recent RSM research on lithium-bearing aluminosilicate tailings illustrates the use of designed thermochemical and leaching experiments to identify favourable operating conditions (Yessengaziyev et al., 2025).

10.6 Dewatering and process-water studies

Thickening and filtration studies examine flocculant dosage, solids concentration, pressure, vacuum and cycle conditions. Responses include settling velocity, underflow density, cake moisture, filtrate clarity and throughput. Clay content and particle-size distribution can cause strong between-sample variation. Blocked designs or random slopes across feed types can show whether a reagent performs consistently. Recycled water introduces time-dependent chemistry, so run order and water age should be recorded and randomised where possible.

11. Lithium and Other Critical-Mineral Systems

Lithium beneficiation and extraction illustrate the need to connect statistical design with mineralogical context. Hard-rock lithium resources may contain spodumene, lepidolite, petalite or other lithium-bearing phases associated with quartz, feldspar, mica, iron minerals and alteration products.

Beneficiation routes can include crushing, grinding, dense-medium separation, gravity separation, magnetic separation and flotation, followed by roasting and leaching. The response to each stage depends on liberation, mineral association and prior treatment (Dessemond et al., 2019; Meng et al., 2021).

A statistically efficient programme should begin by defining the material domains to which the conclusion will apply. If tests use a single composite sample, the fitted optimum describes that composite and does not establish deposit-wide robustness. Ore-domain blocks, stratified sampling and geometallurgical covariates can extend inference. Multi-stage experiments may require sequential designs: screening for beneficiation factors, local optimisation of the selected route, and confirmatory testing of combined stages.

Critical-mineral research also faces small sample quantities and expensive mineralogical analyses. Fractional factorial designs, Bayesian sequential design and informative priors can reduce testwork, but they should not conceal uncertainty. Transparent reporting of sample provenance, preparation, mineralogy, analytical detection limits and data exclusions is essential. Nigerian lithium-bearing pegmatites and other emerging resources present an opportunity to embed these practices early, avoiding the fragmented one-factor studies that are difficult to compare or scale.

Table 5. Representative studies illustrating the development of statistical practice in mineral processing and related experimental systems.

Study	Process and design	Main statistical contribution	Lesson for future research
Niedoba and Pięta (2016)	flotation and coal-property examples; ANOVA review	demonstrated ANOVA's usefulness and limitations in mineral engineering	link the ANOVA model to the multidimensional structure of mineral data
Nakhaei et al. (2016)	copper–molybdenum flotation; D-optimal design	efficient estimation of reagent and operating effects	use designed factor combinations instead of isolated one-factor tests
Arancibia-Bravo et al. (2022)	copper flotation in saline systems; RSM, ANN and global sensitivity analysis	compared conventional polynomial modelling with nonlinear alternatives	validate model class and use sensitivity analysis when interactions are nonlinear
Oumesaoud et al. (2024)	oxidised copper flotation; Box–Behnken RSM and automated mineralogy	connected statistical optimisation with mineral associations and industrial validation	interpret statistical effects through mineralogical evidence
Sepúlveda et al. (2024)	flotation circuits; DOE and stochastic simulation for locked-cycle validation	extended optimisation from batch tests to cycle-test prediction	test the transfer from designed batch experiments to circuit behaviour
Yessengazyev et al. (2025)	lithium recovery from aluminosilicate tailings; RSM	optimised thermochemical and leaching variables	treat extraction, impurity behaviour and operating constraints as multiple responses

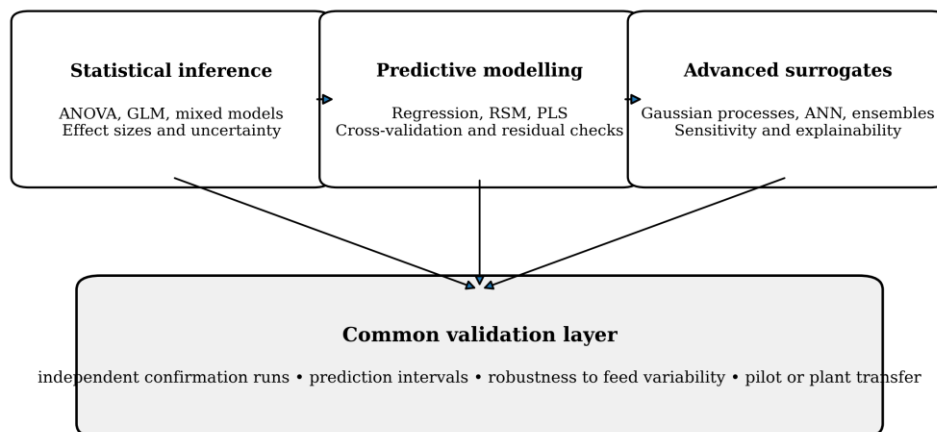
Study	Process and design	Main statistical contribution	Lesson for future research
Camacho et al. (2026)	high-dimensional designed data; ASCA tutorial review	linked DOE and multivariate ANOVA decomposition	use multivariate effect models when many correlated measurements are collected

12. Integrated Framework for Statistical Research and Optimisation

The proposed framework treats statistical analysis as a sequence of connected decisions. Stage one defines the engineering decision and the smallest effect that would matter. Stage two characterises the material and identifies the population of ores, feed campaigns or operating conditions to which the conclusion should apply. Stage three identifies controllable factors, nuisance variables, responses and constraints. Stage four selects a design that permits the required effects and interactions to be estimated.

Stage five executes the experiment with randomisation, calibration, traceable sampling and planned replication. Stage six performs exploratory analysis and verifies data integrity. Stage seven fits an inferential model that reflects the treatment allocation, blocks and repeated measurements. Stage eight develops predictive models only after the principal effects and data structure are understood. Stage nine optimises the relevant responses under practical constraints and uncertainty. Stage ten conducts independent confirmation experiments. Stage eleven evaluates scale-up and establishes monitoring rules for future operation.

The framework is iterative. Confirmation failure may indicate that the factor region was too narrow, the feed material changed, the model omitted an interaction or the measurement system was unstable. The appropriate response is not to select a more elaborate algorithm automatically. Researchers should revisit the experimental design and physical assumptions, then collect targeted data that can distinguish competing explanations.



Model sophistication should increase only when the data volume, experimental design and validation evidence support it.

Figure 4. Hierarchy of model classes supported by a common validation layer. Source: Author's conceptualisation.

13. Reporting and Reproducibility Requirements

A journal article should make the chain from material to conclusion auditable. Sample provenance, preparation, storage and compositing procedures determine the population represented by the testwork. Experimental units, treatment allocation, randomisation and replication should be stated explicitly. Factor levels should appear in actual engineering units as well as coded units. Authors should describe instrument calibration, assay quality control, missing data, outlier decisions and software versions.

Statistical reporting should provide model equations or sufficient coefficient tables, not only statements that a factor was significant. ANOVA tables should identify the error terms and degrees of freedom. Effect estimates, confidence intervals and diagnostic evidence should accompany p-values. For optimisation, authors should state the objective function, response weights, constraints and predicted uncertainty. Confirmation experiments should report both predicted and observed values.

Data and code sharing improve reproducibility, although confidentiality and commercial restrictions may apply. At minimum, authors can share anonymised design matrices, factor definitions, response calculations and analysis scripts. When raw plant data cannot be released, a synthetic example or detailed computational appendix can make the analytical procedure reviewable. Reproducibility also requires consistent nomenclature and units across text, tables and figures.

Table 6. Reporting checklist for statistically defensible mineral-processing research.

Manuscript component	Minimum information to report	Preferred evidence
Material and sampling	deposit or source, sampling plan, compositing, preparation, mineralogy and particle-size basis	representativeness argument; ore-domain or lot identifiers; analytical QA/QC
Experimental design	experimental unit, factors, levels, blocks, replication, randomisation and run order	design matrix or supplementary file; rationale for ranges and centre points
Measurement system	response definitions, instruments, calibration, duplicate or reference samples	measurement uncertainty or repeatability study
Statistical model	model formula, fixed and random effects, transformations, software and version	effect estimates, confidence intervals, ANOVA table and effect sizes
Diagnostics	residual behaviour, variance structure, influential observations and missing data decisions	plots and sensitivity analyses in supplementary material
Optimisation	objective, constraints, response weights and search region	prediction intervals, robustness assessment and alternative trade-off settings
Validation and scale-up	number of confirmation runs, new material or campaign, observed versus predicted response	independent validation, pilot or plant evidence and limitations
Reproducibility	data availability, code availability and nomenclature	machine-readable data, analysis scripts and documented workflow

14. Challenges, Research Gaps and Future Directions

Ore heterogeneity remains the most persistent challenge. Many studies optimise a process on one composite sample and generalise the result to a deposit or ore type. Future work should combine DOE with geometallurgical sampling, hierarchical models and transportability testing. Reporting the mineralogical envelope in which a model was developed will make its limits clearer.

Small experimental datasets create a second challenge. Expensive assays and limited sample mass encourage models with more terms than the data can support. Sequential experimentation offers a better strategy: begin with a defensible screening design, update factor ranges from the results, and collect new confirmation data. Bayesian design can formalise this sequence when prior knowledge is available, but prior assumptions must be reported and tested.

Multiple responses and sustainability metrics require stronger integration. Grade and recovery should be analysed alongside water use, energy consumption, reagent intensity, tailings characteristics and economic value. Multi-objective optimisation and uncertainty analysis can identify operating regions that balance these outcomes. A single predicted maximum is seldom sufficient for industrial decision-making.

Model validation remains underdeveloped. Random train–test splits do not constitute independent validation when observations share ore lots or experimental runs. Future studies should use grouped validation, temporal validation or entirely new material. Prediction intervals and calibration should accompany point predictions. Comparative studies should include simple baselines so that the additional value of a complex model can be judged.

High-dimensional mineralogical and sensor data will increase. ASCA, PLS, Gaussian processes, deep learning and physics-informed models offer promising tools, but future research should preserve interpretability and mass-balance consistency. Data-driven models should support, not replace, process knowledge. Global sensitivity analysis, uncertainty propagation and explainable modelling can connect flexible prediction with engineering understanding.

Reproducible research practices also require attention. Standardised reporting templates, open design matrices, code repositories and shared benchmark datasets would permit meaningful comparison across methods. For emerging critical-mineral regions, including Nigeria, coordinated datasets could reduce duplication and accelerate the development of ore-specific beneficiation knowledge.

15. Conclusions

Statistical tools provide the structure through which mineral-processing experiments become defensible engineering evidence. ANOVA remains valuable for partitioning variation and testing effects, but it should operate inside a broader framework that begins with the definition of experimental units and ends with confirmation testing. Randomisation, replication and blocking determine the credibility of inference before software is opened.

DOE improves experimental efficiency and exposes interactions that one-factor methods cannot estimate. Regression and RSM translate designed effects into predictive and optimising models, provided that factor ranges, residual diagnostics and lack of fit are considered. Mixed-effects and multivariate methods address the hierarchical and correlated structure of modern mineral-processing data. Advanced surrogate models can represent nonlinear behaviour, but their value depends on honest validation, uncertainty analysis and interpretable links to process knowledge.

Across comminution, gravity concentration, magnetic separation, flotation, hydrometallurgy and dewatering, the recommended sequence can be summarised as material characterisation, DOE, exploratory analysis, inference, predictive modelling, constrained optimisation, confirmation and scale-up. Applying this sequence can reduce unnecessary testwork, strengthen reproducibility and produce process recommendations that remain useful beyond the particular dataset from which they were derived.

References

1. Arancibia-Bravo, M. P., Lucay, F. A., Sepúlveda, F. D., Cortés, L., and Cisternas, L. A. (2022). Response surface methodology for copper flotation optimization in saline systems. *Minerals*, 12(9), 1131. <https://doi.org/10.3390/min12091131>
2. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1), 289–300.
3. Box, G. E. P., Hunter, J. S., and Hunter, W. G. (2005). *Statistics for Experimenters: Design, Innovation, and Discovery* (2nd ed.). Wiley.

4. Box, G. E. P., and Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society: Series B*, 13(1), 1–45.
5. Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
6. Camacho, J., Ezenarro, J., Schorn-García, D., and Westerhuis, J. A. (2026). From design of experiments to analysis of variance of multivariate data: A tutorial review on ANOVA simultaneous component analysis. *Journal of Chemometrics*, 40(6), e70151. <https://doi.org/10.1002/cem.70151>
7. Cochran, W. G., and Cox, G. M. (1957). *Experimental Designs* (2nd ed.). Wiley.
8. Dessemond, C., Lajoie-Leroux, F., Soucy, G., Laroche, N., and Magnan, J.-F. (2019). Spodumene: The lithium market, resources and processes. *Minerals*, 9(6), 334. <https://doi.org/10.3390/min9060334>
9. Fisher, R. A. (1935). *The Design of Experiments*. Oliver and Boyd.
10. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd ed.). CRC Press.
11. Iooss, B., and Lemaître, P. (2015). A review on global sensitivity analysis methods. In G. Dellino and C. Meloni (Eds.), *Uncertainty Management in Simulation-Optimization of Complex Systems* (pp. 101–122). Springer. https://doi.org/10.1007/978-1-4899-7547-8_5
12. Kruskal, W. H., and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583–621. <https://doi.org/10.1080/01621459.1952.10483441>
13. Levene, H. (1960). Robust tests for equality of variances. In I. Olkin et al. (Eds.), *Contributions to Probability and Statistics* (pp. 278–292). Stanford University Press.
14. Meng, F., McNeice, J., Zadeh, S. S., and Ghahreman, A. (2021). Review of lithium production and recovery from minerals, brines, and lithium-ion batteries. *Mineral Processing and Extractive Metallurgy Review*, 42(2), 123–141. <https://doi.org/10.1080/08827508.2019.1668387>
15. Montgomery, D. C. (2020). *Design and Analysis of Experiments* (10th ed.). Wiley.
16. Myers, R. H., Montgomery, D. C., and Anderson-Cook, C. M. (2016). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments* (4th ed.). Wiley.
17. Nakhaei, F., Irannajad, M., Sam, A., and Jamalzadeh, A. (2016). Application of D-optimal design for optimizing copper–molybdenum sulphides flotation. *Physicochemical Problems of Mineral Processing*, 52(1), 252–267.
18. Niedoba, T., and Pięta, P. K. (2016). Applications of ANOVA in mineral processing. *Mining Science*, 23, 43–54. <https://doi.org/10.5277/msc162304>
19. Oumesaoud, H., Faouzi, R., Aboulhassan, M. A., Naji, K., Benzakour, I., Faqir, H., Oukhrib, R., and Elboughdiri, N. (2024). Iron oxide–copper mineral associations in supergene zones: Insights into flotation challenges and optimization using response surface methodology. *ACS Omega*, 9(23), 24438–24452. <https://doi.org/10.1021/acsomega.4c00105>
20. Quintanilla, P., Neethling, S. J., and Brito-Parada, P. R. (2021). Modelling for froth flotation control: A review. *Minerals Engineering*, 162, 106718. <https://doi.org/10.1016/j.mineng.2020.106718>
21. Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical Science*, 4(4), 409–423.
22. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Global Sensitivity Analysis: The Primer*. Wiley.
23. Sepúlveda, F. D., Cortés, L. A., Arancibia-Bravo, M. P., Delgado, J., Lucay, F. A., Chacana, C., Galleguillos, F., and Castellón, C. I. (2024). Optimizing flotation circuits: A comprehensive approach using design of experiments and stochastic simulation in cycle test validation. *Minerals Engineering*, 219, 108978. <https://doi.org/10.1016/j.mineng.2024.108978>
24. Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591–611. <https://doi.org/10.2307/2333709>

25. Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1–3), 271–280. [https://doi.org/10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6)
26. Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, 5(2), 99–114. <https://doi.org/10.2307/3001913>
27. Wasserstein, R. L., and Lazar, N. A. (2016). The ASA statement on p-values: Context, process, and purpose. *The American Statistician*, 70(2), 129–133. <https://doi.org/10.1080/00031305.2016.1154108>
28. Wills, B. A., and Finch, J. A. (2016). *Wills' Mineral Processing Technology: An Introduction to the Practical Aspects of Ore Treatment and Mineral Recovery* (8th ed.). Elsevier.
29. Yessengazyev, A., Karshyga, Z., Yersaiynova, A., Tastanova, A., Smailov, K., Mukangaliyeva, A., and Orynbayev, B. (2025). Optimization of lithium recovery from aluminosilicate tailings via sulfation roasting and leaching: Experimental study and RSM modeling. *Metals*, 15(10), 1133. <https://doi.org/10.3390/met15101133>
30. Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., and Smith, G. M. (2009). *Mixed Effects Models and Extensions in Ecology with R*. Springer.