applied sciences

MDPI

Article

Classification of Four Igneous Host-Rock Types Using Apatite Trace-Element Geochemistry: Source-Aware Evaluation of Deep Learning and Geochemical Representations

Yangyang Zhong [1,2,*], Haowei Huang [1,2], Chengbiao Leng [3], Fengge Han [4] and Wenyao Fan [1]

[1] School of Software, East China University of Technology, Nanchang 330013, China; harvey1250451024@163.com (H.H.)
[2] Jiangxi Engineering Technology Research Center of Nuclear Geoscience Data Science and System, East China University of Technology, Nanchang 330013, China
[3] National Key Laboratory of Uranium Resources Exploration-Mining and Nuclear Remote Sensing, East China University of Technology, Nanchang 330013, China
[4] School of Sciences, East China University of Technology, Nanchang 330013, China
* Correspondence: zhyy168@ecut.edu.cn

Abstract

Apatite rare-earth-element (REE) compositions can discriminate igneous host-rock types, but systematic comparisons of deep-learning models and alternative REE representations remain limited, while analyses from shared publications or geological bodies complicate estimates of generalization. We evaluated four-class classification in a cohort of 4148 apatite analyses from 76 connected source groups. Six primary models (SVM, RF, XGBoost, MLP, ResNet, and FT-Transformer) were compared across seven representations derived from the same 14 REEs. Logistic regression and linear discriminant analysis served as baselines, and ILR features were examined separately. Source-group-aware nested cross-validation used five outer and three inner folds, with candidate selection and tuning confined to outer training groups. The selected procedure achieved a five-seed pooled out-of-fold Macro-F1 of 0.556 (95% source-group bootstrap CI, 0.398–0.652). With Raw inputs, logistic regression achieved 0.573 (0.436–0.672). Confidence intervals overlap, and paired comparisons after multiplicity adjustment do not support a reliable ranking of the candidates or a consistent advantage for deeper architectures or combined representations. Raw, Geo, CLR, and ILR are alternative representations of the same REE measurements rather than independent chemical information. Logistic regression using Raw inputs remained competitive, while transfer to unseen geological sources remains uncertain, especially for classes represented by few independent sources.

Keywords: apatite; trace elements; host-rock discrimination; machine learning; deep learning; geochemical representation; source-aware validation

Check for updates

Academic Editors: Zifu Wang and Cheng-Yu Ku

Received: 11 September 2026
Revised: 16 October 2026
Accepted: 19 October 2026
Published: 21 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Apatite occurs widely in igneous, metamorphic, and sedimentary rocks and is a common accessory mineral. Owing to its crystal structure, Sr, Y, Mn, Th, U, and the rare earth elements (REEs) can enter the apatite lattice in substantial amounts, making its trace-element composition sensitive to the magmatic conditions under which it crystallized [1]. Apatite is therefore regarded as an effective recorder of host-rock properties and magmatic evolution, and has been applied in provenance tracing, the analysis of magmatic processes, and the assessment of metallogenic settings.

Apatite compositions vary in REE patterns, as well as in other trace elements such as Sr, Mn, Y and Th. Discrimination diagrams, including Sr–Mn, Sr–Y and Th–Ce plots, illustrate geochemical approaches that use combinations of REE and non-REEs to distinguish host-rock types [1,2]. The present study focuses on a common 14-element La–Lu REE panel; Sr, Mn, Y and Th are contextual examples only and were not model predictors. Although visually intuitive, binary and ternary diagrams represent only a few variables at once and may give ambiguous classifications when compositions overlap [3]. Machine-learning methods can instead evaluate multivariate compositional patterns jointly.

In the geosciences, machine learning has been applied to the classification of various mineral and geochemical datasets, for example, discriminating Pb–Zn deposit types from sphalerite trace elements [4], assessing the fertility of porphyry Cu systems from zircon compositions [5], discriminating deposit types from chlorite compositions [6], and distinguishing IOCG from IOA deposits from magnetite compositions [7]. For apatite, machine learning has been used for the discrimination of ore-deposit types [3], the genetic classification of ore sources [8], and the evaluation of mineralization potential [9]. In the direction of host-rock discrimination, Han et al. [10] used random forest and artificial neural networks to discriminate ultramafic, mafic, intermediate, and felsic igneous rocks from apatite trace elements, establishing the methodological basis for this task.

Deep learning has achieved substantial gains in image recognition [11] and natural language processing [12], aided in part by residual-network architectures [13]. Geoscience applications include lithology identification from drill-core and borehole images [14,15] and three-dimensional mineral prospectivity prediction [16]. These applications motivate an evaluation of deep models for apatite geochemistry, but do not establish their suitability for small- to medium-sized tabular datasets.

Apatite trace-element measurements form a table of continuous variables [17], without the spatial adjacency or natural sequence that characterizes images and time series [18–20]. Consequently, performance gains observed for deep models on structured image or spatial data may not transfer to this classification setting.

On small- to medium-sized tabular classification tasks, deep models do not consistently outperform random forests or gradient-boosted trees [19,20]. Their relative performance depends on dataset and feature structure [19,20]. An apatite-specific comparison therefore requires common data partitions, tuning procedures and evaluation criteria across model families.

More broadly, data-driven modeling of complex multivariate measurements uses nonlinear feature extraction and learned or engineered representations [18–20]. Recent industrial-process studies illustrate manifold-learning and dynamic data-driven methods for nonlinear multivariate systems [21,22]. These dynamic engineering applications provide methodological context rather than direct evidence for static geological classification. Deep architectures can model interactions among predictors, whereas alternative input representations change feature geometry and inductive bias. Their potential value depends on data structure, source dependence and validation design. Apatite REE measurements provide a tabular setting for assessing these choices under source-group-aware evaluation.

Existing apatite studies demonstrate the utility of multivariate trace-element data for supervised classification of deposit types [3], ore sources [8], mineralization potential [9] and igneous host rocks [10]. However, the geological provenance of analyses also affects evaluation: records from one publication, sample or geological body can share characteristics that are unavailable for a genuinely new source.

For four-class host-rock discrimination, a unified source-group-held-out comparison of conventional and deep tabular models remains limited [10]. Such a comparison is needed

to assess whether model complexity improves generalization to unseen geological sources, rather than performance on related records within a source.

Input representation may also influence classification. The 14 measured REE concentrations can be used directly (Raw), transformed into geochemical ratios, anomalies and abundance summaries (Geo) [1,10,23], or expressed as centered log-ratio coordinates (CLR) [24,25]. Each representation emphasizes different aspects of the same measurements; their combination does not add independently measured geochemical variables.

Previous apatite studies have generally selected particular elements or feature sets [3,8–10]. The effect of Raw, Geo and CLR representations and their combinations across both conventional and deep models has not been established under a common source-group-held-out evaluation protocol.

This study evaluates four-class igneous host-rock classification with source-group-aware nested cross-validation. Connected source groups were kept intact within each training or evaluation partition, so records linked by a publication or identified geological body did not cross the corresponding split. The primary comparison includes SVM, RF, XGBoost, MLP, ResNet and FT-Transformer; logistic regression and LDA provide simpler baselines. Candidate selection and hyperparameter tuning use only the training groups of each outer fold. We quantify performance and uncertainty for unseen sources and examine whether model architecture or REE representation yields a consistent improvement.

The seven primary feature schemes comprise the non-empty combinations of Raw, nine geochemically derived features (Geo) and 14 CLR coordinates. A 13-coordinate isometric log-ratio (ILR) representation and Raw+ILR are evaluated separately as sensitivity analyses. Any observed effect of combining representations is interpreted as a change in model input geometry or inductive bias, not as evidence of additional chemical measurements. We also describe source-balanced REE patterns and class-specific errors to relate the numerical comparisons to the geochemical variation represented in the dataset.

2. Materials and Methods

2.1. Data Provenance, Labels and Independent Groups

Dataset assembly started from 3246 records and was expanded through auditable frozen stages of 3364, 3460, 3704, 3745 and 4180 records. Inclusion required an apatite analysis with complete, strictly positive La–Lu concentrations, a traceable source record, an operational host-rock label resolved from source-reported lithology and no duplicate 14-REE vector under the stated rule. These decisions were completed before outer-fold construction. Two records with conflicting legacy labels were moved to the pending ledger, yielding the frozen dataset of 4178 analyses. Before primary fold assignment, 30 mafic analyses associated with the James Ross and Daye cases were conservatively excluded because their publication/source identifiers overlapped with development records. They were removed from the evaluation cohort to avoid any shared-source overlap and were not treated as an independent external test. The resulting development and evaluation cohort contains 4148 analyses: 1591 felsic (coded acidic in the data), 1514 intermediate, 658 mafic and 385 ultramafic. Each record is uniquely identified by its analysis_id and retains its assigned source citation/reference and source_group in the provenance crosswalk. Forty-one archived source-file/worksheet/row locator tuples were duplicated across 82 records. A post-archive verification uniquely matched each affected analysis to a record in the original GEOROC 2024-12 CSV using its complete 14-REE vector; all 82 sample names also agreed. The verified CSV record numbers and raw-file checksum are provided in Supplementary Methods and Code S1. The historical locator fields remain unchanged, and citation and source-group assignments remain many-to-one.

Operational host-rock labels were assigned before modeling from source-reported lithological names through the frozen lookup in data/final_frozen_rock_label_mapping.csv; REE predictor values and retrospective whole-rock compositional thresholds were not used to derive the labels. The lookup groups dacite, granite and rhyolite as acidic; andesite, diorite, granodiorite, tonalite and related names as intermediate; basalt, dolerite and gabbroic names as mafic; and clinopyroxenite, hornblendite, peridotite, pyroxenite, kimberlite and the source-reported pyroxenite–carbonatite entry within the study's operational ultramafic class. Thus, granodiorite and tonalite follow the intermediate lookup, whereas pyroxenite–carbonatite follows the operational ultramafic lookup. These labels reproduce the mapping actually used in the frozen analysis and should not be interpreted as a new whole-rock classification scheme. Retaining the compound pyroxenite–carbonatite entry preserves the target definition used in the frozen four-class evaluation; it does not establish carbonatites as a separately validated lithological class. Ambiguous, non-target and ore-related descriptions were excluded or retained in the pending ledger; identical REE vectors assigned to different classes were resolved before dataset freezing. Additional-record deduplication used the complete 14-REE vector rounded to six decimal places, both against the existing records and within added records; sample metadata were not required to identify a vector collision. Analysis IDs were checked separately for uniqueness. Distinct analyses sharing a sample name were retained but assigned to the same evaluation group. Grain identifiers were not consistently reported in the source publications and were not inferred from analysis counts. The processed dataset, original rock names, mapping ledger, exclusions and deduplication scripts accompany the study. The operative rock-name-to-class mapping is data/final_frozen_rock_label_mapping.csv in Supplementary Methods and Code S1; the legacy inventories document earlier collection stages and do not define the final cohort.

The validation unit was the connected source_group. All records from one publication were kept together, and publications linked to the same identified geological body were merged into a connected group. More specific body and sample keys were retained and audited. This procedure produced 76 non-overlapping source groups from 78 source citations in the 4148-record cohort. Five outer folds were fixed before modeling, and each outer training partition was divided into three group-aware inner folds for tuning. No source_group, citation, body key or sample-group key appeared in both training and evaluation partitions of the corresponding fold. Grouping reduces the potential for source leakage [26], although it cannot recover geological identifiers absent from published data. Figure 1 contrasts analytical record counts with class-specific source support and concentration (Tables 1 and 2).

Table 1. Source structure of the source-group-held-out development cohort.

Host-Rock Class	Analyses	Publications	Source Groups	Localities	Body Keys	Sample Groups	Grains
Felsic/acidic	1591	34	34	64	64	126	Not recorded
Intermediate	1514	38	37	51	52	138	Not recorded
Mafic	658	9	9	17	10	70	Not recorded
Ultramafic	385	16	15	15	15	30	Not recorded

Counts within classes can sum above the overall unique totals because a publication or connected source may contribute to more than one class. Overall: 4148 analyses, 78 publication citations, 76 connected source groups, 129 locality strings, 123 observed body keys and 364 sample groups. Grain count is unavailable, not zero. Missing-value sentinels were excluded from locality and body counts. Original locality and body metadata were unavailable for 204 intermediate-class analyses, which remained grouped by publication. These counts describe distinct metadata keys, not independently verified geological bodies.

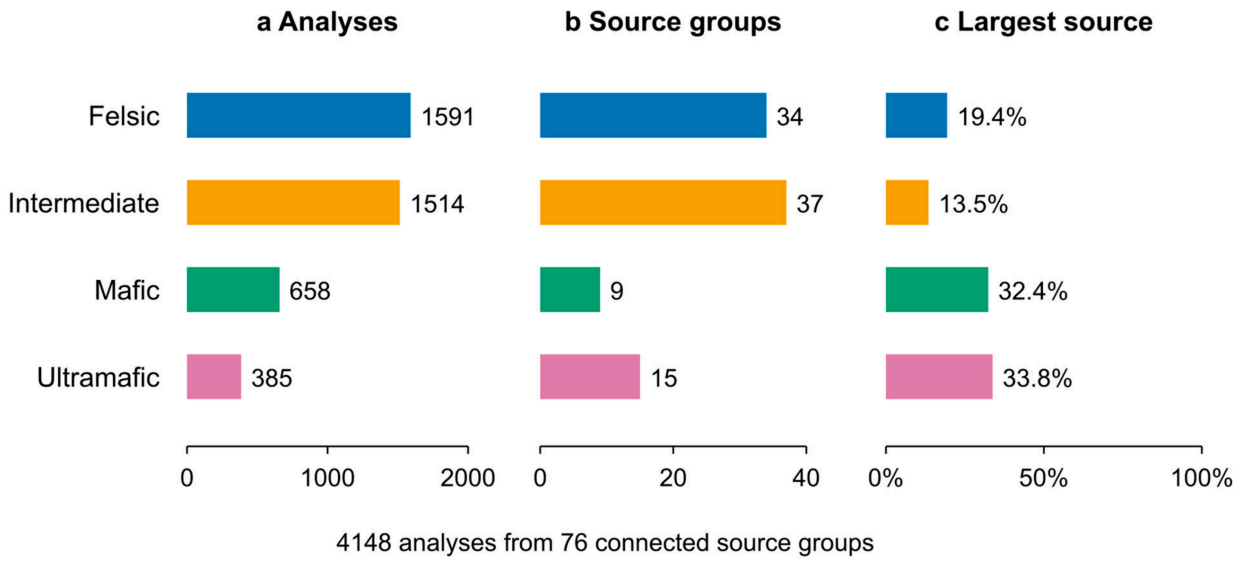


Figure 1. Source structure of the 4148-analysis evaluation cohort. (**a**) Number of apatite analyses by host-rock class. (**b**) Number of connected source groups contributing to each class. (**c**) Share of class-specific analyses contributed by the largest publication source. Colors identify the same classes throughout the figure. A connected source group can contribute to multiple classes, so class-specific group counts do not sum to the 76 unique groups. Record counts are not counts of independent geological sources.

Table 2. Frozen source-group outer-fold support.

Outer Fold	Analyses	Felsic	Intermediate	Mafic	Ultramafic	Test Source Groups
0	892	317	232	213	130	5
1	1163	319	566	180	98	14
2	764	319	236	156	53	18
3	676	319	247	58	52	19
4	653	317	233	51	52	20

Every development analysis is evaluated exactly once per seed, in an outer fold containing its entire source group. Fold-level record support is not the number of independent geological sources. The numbers of held-out mafic source groups in folds 0–4 were 1, 1, 2, 3 and 2; the corresponding ultramafic counts were 1, 1, 4, 4 and 5.

2.2. Feature Representations and Preprocessing

Raw comprises the 14 measured REE concentrations in ppm. Geo comprises nine deterministic features derived from them: chondrite-normalized La/Yb, La/Sm and Gd/Yb ratios; Eu and Ce anomalies; LREE and HREE sums; their ratio; and total REE. CI-chondrite constants follow McDonough and Sun [27]. Eu and Ce anomalies divide normalized Eu and Ce by the geometric means of their normalized neighboring REE: Sm and Gd for Eu, and La and Pr for Ce. In the archived feature code, LREE is the sum of La–Sm, whereas the code-defined feature labeled HREE is the sum of Eu–Lu, including Eu; total REE is the sum of these two features. This archived label is retained for reproducibility and should not be interpreted as a redefinition of the conventional geochemical HREE grouping. CLR subtracts the mean log concentration from each of the 14 log concentrations. Its 14 coordinates sum to zero and have rank 13; they represent a selected REE subcomposition rather than whole-rock composition. Before feature construction, all 14 REE concentrations were checked for completeness and nonpositive values. The frozen processed matrix contains no missing, zero or negative values, and no imputation, zero replacement or pseudo-count was applied. Publication-specific below-detection-limit flags and replacement metadata

are not encoded in the processed matrix, so the treatment of BDL values in the original source studies cannot be independently reconstructed from this dataset. ILR provides an orthonormal 13-coordinate log-ratio representation for sensitivity analysis. ILR coordinates use a sequential Helmert basis in the listed La-to-Lu element order [25]. Table 3 gives the explicit Geo definitions used in the frozen analysis.

Table 3. Definitions and numerical interpretation of the nine Geo inputs in the archived analysis.

Geo Input	Definition	Interpretation and Units
$(La/Yb)_N$	$\frac{La_N}{Yb_N}$	Relative normalized abundance contrast between La and Yb across the selected REE range. CI-normalized concentrations; dimensionless.
$(La/Sm)_N$	$\frac{La_N}{Sm_N}$	Relative normalized abundance contrast between La and Sm within the lighter-REE segment. CI-normalized concentrations; dimensionless.
$(Gd/Yb)_N$	$\frac{Gd_N}{Yb_N}$	Relative normalized abundance contrast between Gd and Yb within the heavier-REE segment. CI-normalized concentrations; dimensionless.
Eu/Eu*	$\frac{Eu_N}{\sqrt{Sm_N \times Gd_N}}$	Eu relative to the geometric mean of normalized Sm and Gd; 1 denotes no departure from that local geometric interpolation. CI-normalized concentrations; dimensionless.
Ce/Ce*	$\frac{Ce_N}{\sqrt{La_N \times Pr_N}}$	Ce relative to the geometric mean of normalized La and Pr; 1 denotes no departure from that local geometric interpolation. CI-normalized concentrations; dimensionless.
LREE sum	La + Ce + Pr + Nd + Sm	Absolute abundance sum of the five listed lighter REE. Raw concentrations; ppm.
HREE sum (archived label)	Eu + Gd + Tb + Dy + Ho + Er + Tm + Yb + Lu	Absolute abundance sum of the nine listed elements, including Eu, as implemented in the archive; the label is not a conventional HREE redefinition. Raw concentrations; ppm.
LREE/HREE (code-defined)	$\frac{\text{LREE sum}}{\text{HREE sum}}$	Relative abundance of La–Sm versus the code-defined Eu–Lu sum. Raw concentration sums; dimensionless.
Total REE	LREE sum + HREE sum	Total abundance of the 14 measured REEs; Eu is included exactly once. Raw concentrations; ppm.

N denotes CI-chondrite normalization [27]. The concentrations used as normalizers (ppm), in the order La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb and Lu, are 0.237, 0.613, 0.0928, 0.457, 0.148, 0.0563, 0.199, 0.0361, 0.246, 0.0546, 0.160, 0.0247, 0.161 and 0.0246. The archived HREE label includes Eu and is retained solely to identify the implemented predictor. These deterministic transformations do not add measured chemical information. The definitions reproduce the frozen implementation. Light blue shading marks the column headings only and does not encode any class or value.

Seven primary schemes were evaluated on identical analyses and folds: Raw (14 inputs), Geo (9), CLR (14), Raw+Geo (23), Raw+CLR (28), Geo+CLR (23) and Raw+Geo+CLR (37). ILR-only (13) and Raw+ILR (27) were supplementary schemes excluded from primary candidate selection. Because Geo, CLR and ILR are computed from Raw, these inputs differ in representation and potential inductive bias rather than in independently measured chemistry. Within each fit, StandardScaler was estimated only from the fit's training data for SVM, MLP, ResNet and FT-Transformer. RF, XGBoost, logistic regression and LDA used unstandardized numeric features in the reported run.

The nine Geo inputs summarize three aspects of the same measured REE vector: contrasts between selected normalized elements, local departures of Eu and Ce from

neighboring elements, and absolute abundance summaries (Table 3). Here, the subscript N denotes division by the corresponding CI-chondrite reference concentration [27], with both concentrations expressed in ppm. These are numerical descriptions of REE patterns, not independent measurements or uniquely diagnostic indicators of a magmatic process.

2.3. Models and Source-Aware Evaluation

2.3.1. Model Families and Comparison Roles

The primary model set covers distinct ways of learning relationships within the fixed-dimensional REE inputs [28–33]. SVM uses a radial basis function kernel to construct a nonlinear margin-based classifier [29]. RF aggregates randomized decision trees [28], whereas XGBoost (version 3.4.1) sequentially fits regularized trees using gradient information from the classification objective [32]. Their inclusion permits comparison of kernel learning, bagging and boosting under the same source-held-out partitions rather than assuming that one learning mechanism is best suited to apatite geochemistry.

The three neural models differ in how they combine the input variables. The implemented MLP stacks fully connected layers with batch normalization, ReLU and dropout before a four-class output layer [30]. The tabular ResNet adds residual fully connected blocks after an input projection, providing a comparison with the plain MLP [33]. The FT-Transformer maps each numerical variable to an affine feature token and combines those tokens with a class token through a Transformer encoder with GELU feed-forward activation [31,33]. Its layer-normalized class-token representation is passed to the classification head. These descriptions refer to the implementations in the experiment archive; they do not imply that the models reproduce every setting of the cited reference architectures.

Multinomial logistic regression and shrinkage LDA provide supplementary linear baselines. They test whether the nonlinear model families offer a clear predictive benefit under the implemented fitting procedures. Both baselines were evaluated on the same outer folds and seven representations but were excluded from the primary 42-candidate selection. The comparison is conditional on the recorded preprocessing and tuning budgets: in particular, the archived LR/LDA runs used unstandardized inputs, whereas the separate matched-source LR diagnostic used fold-specific standardization.

2.3.2. Hyperparameter Search and Nested Selection

The primary candidate set crossed SVM, RF, XGBoost, MLP, tabular ResNet and FT-Transformer with the seven feature schemes, yielding 42 candidates spanning kernel, tree-ensemble and neural-network methods. For each outer fold and candidate, an Optuna TPE sampler with a median pruner [34] optimized mean three-fold inner Macro-F1 using only outer training groups. SVM, RF and XGBoost received 50 trials per outer fold; MLP, ResNet and FT-Transformer received 100 trials to search their larger hyperparameter spaces. Trial budgets were fixed before outer-fold evaluation. Search spaces, best trials, fold scores and SQLite study histories are preserved. Logistic regression and LDA were evaluated as supplementary baselines with 50 trials per outer fold for each of the seven schemes; they were excluded from primary 42-way selection. ILR sensitivity was evaluated separately using the same six primary model families. Table 4 reports the search spaces actually executed. The TPE sampler used seed 20260718, and the median pruner used five startup trials. These ranges were checked against all 340 archived Optuna studies.

Table 4. Actual hyperparameter search spaces and trial budgets per outer fold and representation.

Model	Search Space	Trials
SVM	C: 10^{-2}–10^{3} (log); gamma: 10^{-5}–10 (log).	50
RF	Trees: 300–1200 (step 100); depth: {unlimited, 6, 10, 16, 24, 32}; split size: 2–20; leaf size: 1–10; max_features: {sqrt, log2, 0.5, 0.8, 1.0}.	50
XGBoost	Trees: 200–1200 (step 100); depth: 3–10; learning rate: 0.01–0.25 (log); child weight: 0.5–20 (log); row/column fractions: 0.6–1; alpha: 10^{-8}–10 (log); lambda: 10^{-3}–100 (log).	50
MLP	Hidden width: {64,128,256,384,512}; hidden layers: 1–4; batch: {64,128,256,512}; learning rate: 10^{-4}–5×10^{-3} (log); dropout: 0–0.4; weight decay: 10^{-7}–10^{-2} (log).	100
ResNet	Width: {64,128,256}; residual blocks: 1–4; hidden factor: {2,3,4}; batch: {64,128,256}; learning rate: 10^{-4}–5×10^{-3} (log); dropout: 0–0.4; weight decay: 10^{-7}–10^{-2} (log).	100
FT-Transformer	Token width: {16,32,64}; heads: {2,4,8}; encoder layers: 1–3; feed-forward factor: {2,3,4}; batch: {64,128,256}; learning rate: 10^{-4}–5×10^{-3} (log); dropout: 0–0.4; weight decay: 10^{-7}–0.3 (log).	100
Logistic regression	C: 10^{-3}–10^{2} (log); penalty fixed at L2.	50
LDA	Shrinkage: {auto,0.1,0.3,0.5,0.8}; solver fixed at lsqr.	50

Ranges and options are reproduced from the archived core.py. Continuous entries marked log are searched on a logarithmic scale; other continuous ranges are linear. Integer intervals use unit steps unless stated otherwise, and braces denote categorical options. Trials are budgeted separately for each model–representation pair in each outer fold. Unspecified software defaults are determined by the archived environment lock. The larger deep-model budgets do not constitute an equal-compute comparison. LR/LDA are supplementary and do not enter the primary 42-candidate selection. Light blue shading marks the column headings only and does not encode any class or value.

For each outer fold, the candidate with the highest inner mean Macro-F1 was recorded in a selection manifest before outer-fold evaluation. The 42 fixed-candidate outer predictions were then computed. The resulting selected_pipeline therefore estimates the nested model-and-feature selection procedure, not the maximum of candidate outer-test scores. The selected pairs for outer folds 0–4 were FT-Transformer+Raw+CLR, ResNet+CLR, ResNet+Raw+Geo+CLR, ResNet+Raw+Geo and SVM+Geo, respectively. Fixed-candidate outer predictions were retained for descriptive comparisons and were not used for a second selection stage. Figure 2 summarizes the separation between inner selection and outer evaluation.

For application to new data, the nested outer-fold results characterize the model-and-feature selection procedure and do not identify a single final fitted classifier. A deployment configuration would be selected by applying the same three-fold source-group-aware inner cross-validation to the complete 4148-record development dataset, comparing the same 42 primary model–feature combinations and tuning their hyperparameters with the stated budgets and mean Macro-F1 criterion. Any exact tie would be resolved using a fixed candidate order declared before this deployment selection. The configuration would be frozen using development cross-validation alone; outer-fold scores and new-data labels would not be used for selection. Its preprocessing would then be fitted on the complete development dataset and the selected model would be retrained on all development records. For a neural model, the epoch count would first be determined by the same training-internal early-stopping rule, followed by refitting from initialization for that fixed number of epochs on the complete development dataset, including the stopping subset. The resulting preprocessing and fitted classifier would be frozen and applied unchanged to

new data. This is a prospective deployment procedure: no single full-development model was trained or evaluated in the present study, and its performance is not reported here.

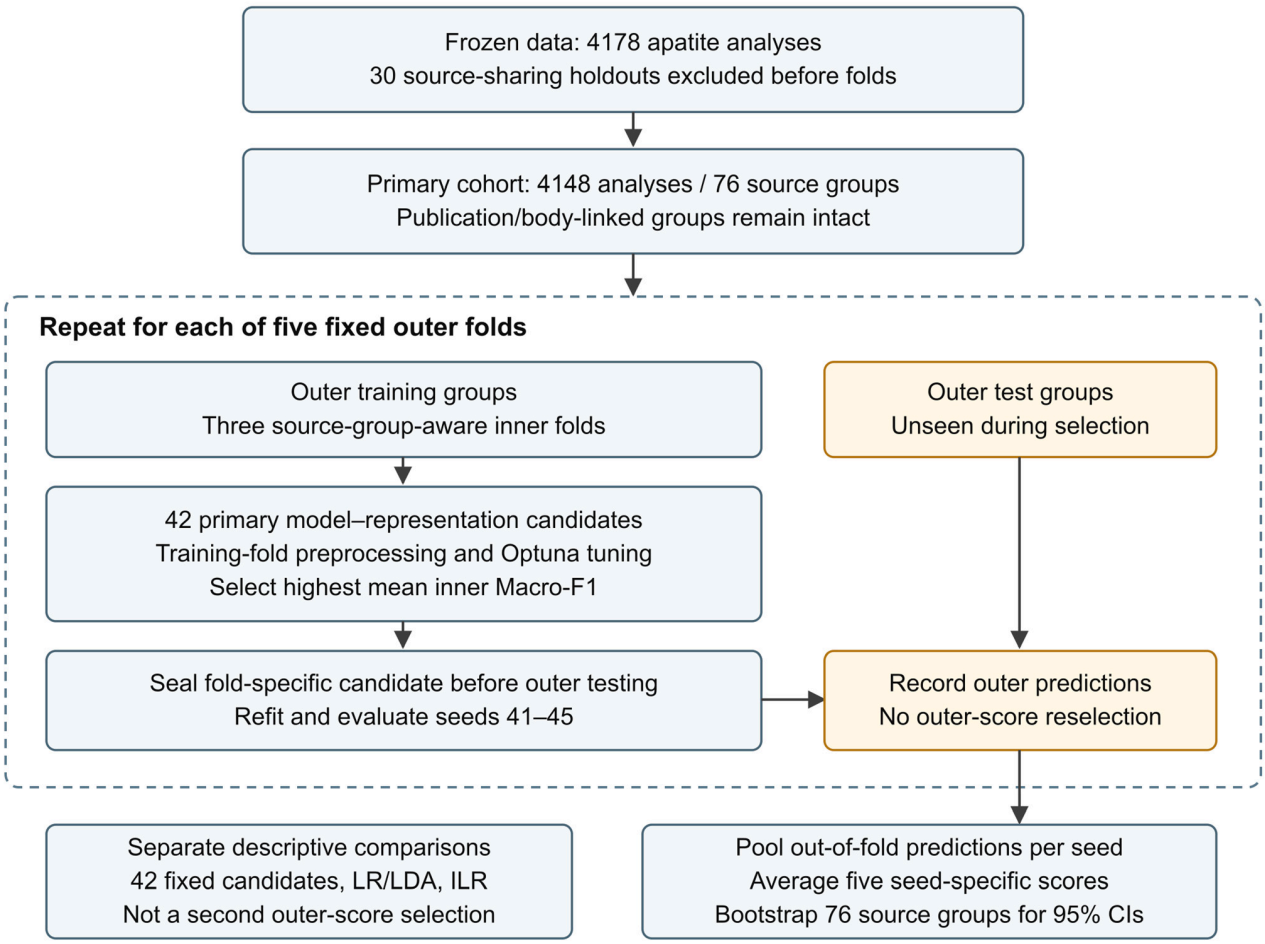


Figure 2. Source-group-aware nested evaluation. The 30 source-sharing holdouts were removed before assigning the five outer folds. Within each outer training partition, three inner source-group folds were used for tuning and selecting among 42 primary candidates. The fold-specific candidate was recorded before outer prediction. Five evaluation seeds were then used to estimate the performance of the selection procedure. Fixed-candidate, baseline and ILR comparisons were retained separately; no second selection was made from outer scores.

2.3.3. Training, Initialization and Repeated Evaluation

Five evaluation seeds (41–45) were used for every outer candidate and baseline. Python 3.12, NumPy 2.3.2 and PyTorch 2.8.0+cu128 random generators were seeded for each fit. Network layers used PyTorch default initialization, except that FT-Transformer tokenization weights and the class token were initialized from a zero-mean normal distribution with standard deviation 0.02, with zero token biases. Neural networks used ReLU-based MLP/ResNet blocks or attention-based FT-Transformer blocks, AdamW optimization and inverse-frequency-weighted cross-entropy [35]. A record-level stratified 15% subset of the current fit's training fold was used for internal early stopping, with a maximum of 100 epochs and a patience of 12. After selecting the epoch with the highest early-stopping Macro-F1, each model was refitted for that many epochs on its full current training fold. The early-stopping subset was drawn entirely from the current training fold and contained no inner-validation or outer-test records. Logistic regression used L2 regularization, lbfgs and balanced class weights; LDA used an lsqr solver with tuned shrinkage. RF and XG-

Boost used their recorded fold-specific settings. No hyperparameter tuning was based on outer-test scores.

The 15% record-level subset was used only to monitor optimization and determine the stopping epoch within the current training partition. Record-level stratification retains class coverage for this training monitor when some classes have few source groups; it is not intended to estimate transfer to unseen sources. Because the optimization and stopping subsets may share source groups, early-stopping Macro-F1 may be optimistic for new-source transfer and may influence the selected epoch or extent of overfitting. This distinction does not introduce leakage into the inner-validation or outer-test folds: their source groups remain separated from the current training partition, including the stopping subset, and their records or outcomes are not used to determine the epoch. Final performance is evaluated only on the source-group-held-out outer predictions. Refitting on the full current training partition restores the stopping records to model fitting without accessing held-out evaluation data.

2.3.4. Matched Diagnostic of Source Dependence

A matched-cohort diagnostic compared predictions for records from previously observed versus held-out sources. It used Raw inputs and a fixed-logistic-regression specification (C=1, balanced class weights), with the same preprocessing rule in both diagnostic arms: StandardScaler was fitted within each training fold. One five-fold stratified record-level split and the fixed five source-group folds were applied to the same 4148-record cohort. Only records whose source appeared in the record-level training fold were retained, and those same records were scored under both partition policies. This separate fixed-model diagnostic used fold-specific standardization, unlike the unscaled logistic-regression baseline in the main comparison. Its scores therefore should not be directly compared with that baseline as if the fitted preprocessing were identical.

2.4. Metrics, Uncertainty and Paired Inference

The primary endpoint was four-class Macro-F1 computed with all four class labels fixed, even when a resample lacked a class [36]. Accuracy, balanced accuracy, class-specific F1 and recall, class support, and confusion matrices were secondary descriptors. Pooled out-of-fold predictions were first scored separately for each of the five seeds and then averaged. For confidence intervals, 2000 percentile bootstrap resamples drew the 76 connected source groups with replacement; each draw retained all analyses and predictions belonging to its sampled groups. The same source draws were used for all candidate comparisons, so paired differences remained matched at the independent-source level. For source-level Friedman and Wilcoxon analyses, four-class Macro-F1 was calculated separately within each source group and averaged over the five seeds, retaining all four labels and assigning zero to undefined class scores. These equally weighted source-level scores differ from the pooled-record Macro-F1 used for point estimates and bootstrap intervals. A repeated-measures Friedman test described each model-within-scheme and scheme-within-model family. The archived v2 postprocessing script applied two-sided source-paired Wilcoxon signed-rank tests with Holm correction within the comparison families. The complete 42-cell matrix is descriptive; comparisons outside these families are exploratory. Bootstrap intervals are nominal 95% intervals without multiplicity adjustment; they should not be read as simultaneous confidence intervals for all comparisons. The archived predictions and postprocessing code permit independent recalculation. These intervals are conditional on the frozen outer-fold partition and fitted out-of-fold predictions; the bootstrap did not repeat tuning or training and does not capture the full variability of alternative source-group partitions.

2.5. Source-Balanced Descriptive Geochemistry

To connect predictive comparisons with the observed geochemistry, we added a source-balanced descriptive analysis of the same 4148-analysis cohort (Figure 3). No new observations were added, and no model was refitted for this analysis. Each of the 14 concentrations was divided by the corresponding CI-chondrite value [27]. Geo descriptors were calculated for each analysis using Table 3, before aggregation. For every source_group–class_label combination, we calculated the median of each normalized REE and descriptor. We then summarized these source–class medians within each host-rock class, assigning equal weight to each contributing source group. Thus, a source with many analyses did not receive greater weight in these descriptive summaries.

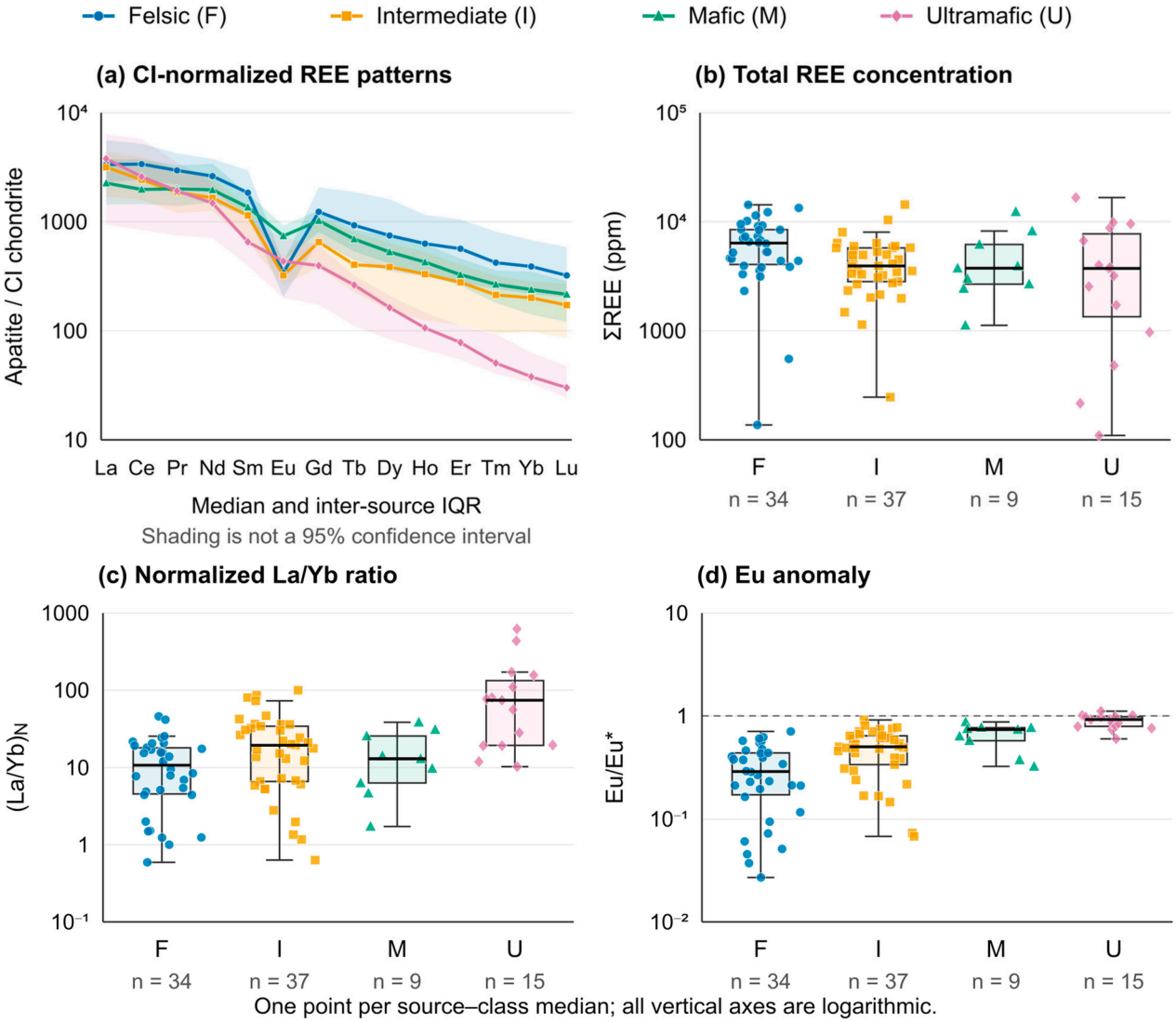


Figure 3. Source-balanced REE patterns and geochemical indicators for the four operational host-rock classes. Each descriptor was first summarized by its median within each source-group–class stratum. (**a**) Lines are the pointwise medians of CI-chondrite-normalized source–class REE medians; shaded bands span their interquartile ranges. (**b**–**d**) Total REE (ppm), normalized La/Yb and Eu/Eu*: points show individual source–class medians, boxes span the 25th–75th percentiles, and whiskers extend to the most extreme values within 1.5 interquartile ranges. All source–class medians, including values beyond whiskers, are shown. There are 34 felsic, 37 intermediate, 9 mafic and 15 ultramafic strata, drawn from 76 unique connected groups; some groups contribute to multiple classes. Quantiles were calculated on untransformed values; logarithmic axes are used only for display. Bands and boxes describe between-source spread, not confidence intervals or model importance.

The resulting 95 source–class strata comprised 34 felsic, 37 intermediate, nine mafic and 15 ultramafic strata, but only 76 unique connected source groups because some groups contribute to multiple classes. REE curves show pointwise medians and interquartile ranges across the source–class medians. For total REE, normalized La/Yb and Eu/Eu*, boxplots show the median and interquartile range, with whiskers at the most extreme observations within 1.5 interquartile ranges and all source–class medians overlaid. Quantiles were computed on the original values using linear interpolation; log scaling was used only for display. These summaries are descriptive and distinct from the pooled-record performance statistics and source-bootstrap confidence intervals. They were not used to choose features, tune models or redefine class labels. Figure source tables and the analysis scripts are provided in the Supplementary Figure Data and Code S2.

3. Results

3.1. Source Coverage and REE Characteristics

The 4148-record cohort included 658 mafic analyses from nine publications and 385 ultramafic analyses from sixteen publications. Source concentration remained substantial: the largest mafic source contributed 213 of 658 analyses (32.4%), and the largest ultramafic source contributed 130 of 385 (33.8%). Table 1 reports publication, locality, geological-body and sample-group counts by class; grain counts could not be recovered. All primary outer predictions were generated for source groups absent from the corresponding training fold. Figure 1 shows why analytical sample size and independent-source support should be considered separately.

Source-balanced REE summaries reveal both class contrasts and substantial overlap (Figure 3). The median of the source–class total-REE medians was 6380 ppm for felsic, 3937 ppm for intermediate, 3753 ppm for mafic and 3737 ppm for ultramafic apatite. The respective interquartile ranges were 4061–8478, 2827–5761, 2681–6196 and 1346–7749 ppm. The four ranges share an overlapping interval, so absolute REE abundance alone does not provide a clean separation of these classes in the source-balanced summaries.

Relative REE patterns provide a different view. The source-balanced median $(La/Yb)_N$ was 10.75, 19.33, 12.97 and 74.02 for felsic, intermediate, mafic and ultramafic apatite, respectively. The ultramafic class therefore had the greatest median normalized La-to-Yb contrast, but its interquartile range was broad (19.27–133.30) and overlapped those of intermediate and mafic sources. The corresponding Eu/Eu* medians were 0.290, 0.501, 0.737 and 0.923. Felsic sources had the strongest median negative Eu anomaly, whereas the ultramafic median was closest to unity. The plotted source distributions show that these central tendencies are not deterministic class boundaries. They describe the operational classes in this collection rather than a single magmatic differentiation sequence.

3.2. Source-Held-Out Performance and the Matched Diagnostic

The selected nested procedure achieved Macro-F1 0.556 (95% source-group bootstrap CI, 0.398–0.652), accuracy 0.559 (0.473–0.650) and balanced accuracy 0.564 (0.422–0.675). The intervals reflect variation among the 76 source groups, including limited minority-class source representation in some outer folds (Table 5).

The fixed-logistic-regression diagnostic retained 4144 records from 73 groups whose source was present in the record-level training fold; four records from singleton or otherwise unseen sources were excluded from both comparison arms. For the retained records, Macro-F1 was 0.653 (source-group bootstrap CI, 0.515–0.733) under record-level folds and 0.481 (0.357–0.588) under source-group-held-out folds. The paired source-group bootstrap difference was 0.172 (0.076–0.269). These are fixed-model diagnostic results, distinct from the tuned nested-procedure estimate.

Table 5. Nested source-group-held-out procedure: pooled out-of-fold performance.

Metric	Estimate	95% Source-Group CI
Macro-F1	0.556	0.398–0.652
Accuracy	0.559	0.473–0.650
Balanced accuracy	0.564	0.422–0.675
Felsic F1 (support 1591)	0.595	0.474–0.689
Intermediate F1 (support 1514)	0.504	0.376–0.622
Mafic F1 (support 658)	0.647	0.171–0.833
Ultramafic F1 (support 385)	0.479	0.149–0.720

Point estimates average five seed-specific pooled source-group-held-out scores. Confidence limits use 2000 bootstrap resamples of connected source groups, not individual analyses.

3.3. Model and Representation Comparisons

Table 6 and Figure 4 report source-group-held-out Macro-F1 for every fixed primary candidate. The full matrix is descriptive and was not used to select the final procedure by outer score. The response to a change in representation was not uniform across model families. ResNet increased from 0.477 with Raw to 0.577 with CLR and 0.602 with Geo+CLR, whereas RF decreased from 0.491 with Raw to 0.431 with CLR and 0.413 with Raw+Geo+CLR. MLP changed from 0.491 with Raw to 0.550 with Raw+CLR, but only to 0.502 with all three representations. These patterns describe model-dependent behavior; the larger values are not used to declare a globally optimal pipeline.

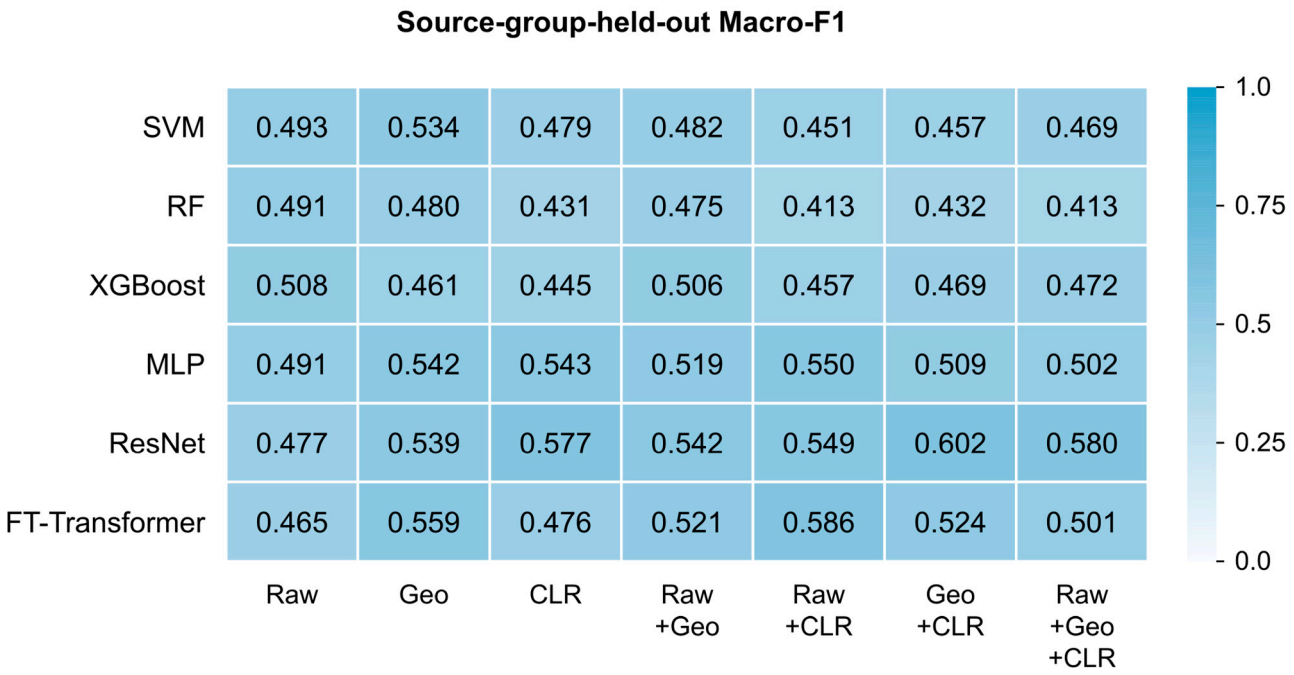


Figure 4. Descriptive source-group-held-out Macro-F1 for the six primary model families and seven REE representations. Each cell is the mean of five seed-specific scores calculated from pooled outer-fold predictions. Model and representation order is fixed, and the color scale spans the full 0–1 metric range. This matrix was not used to select a final pipeline on outer-fold performance; candidate selection was confined to the inner folds.

Across 13 prespecified source-level overall comparison families, two unadjusted Friedman p values were below 0.05 (models under CLR, 0.039; schemes within FT-Transformer, 0.028). Neither remained below 0.05 after Holm correction across the 13 families. None of the 231 within-family Wilcoxon pairwise comparisons remained below 0.05 after within-

family Holm adjustment. The confidence intervals overlap, and the paired comparisons do not support a reliable ranking of the candidates. Consequently, the visible cell-to-cell contrasts in Figure 4 are not interpreted as established architecture or representation advantages. Absence of a detected difference does not establish equivalence.

Table 6. Descriptive source-group-held-out Macro-F1 matrix for the 42 fixed candidates.

Model	Raw	Geo	CLR	Raw+Geo	Raw+CLR	Geo+CLR	Raw+Geo+CLR
SVM	0.493	0.534	0.479	0.482	0.451	0.457	0.469
RF	0.491	0.480	0.431	0.475	0.413	0.432	0.413
XGBoost	0.508	0.461	0.445	0.506	0.457	0.469	0.472
MLP	0.491	0.542	0.543	0.519	0.550	0.509	0.502
ResNet	0.477	0.539	0.577	0.542	0.549	0.602	0.580
FT-Transformer	0.465	0.559	0.476	0.521	0.586	0.524	0.501

Each value is the mean of five seed-specific pooled out-of-fold scores. This table does not select the final pipeline; uncertainty and paired comparisons are reported separately.

3.4. Simple Baselines and ILR Sensitivity

Figure 5 and Table 7 compare the primary models with simpler baselines. With Raw inputs, logistic regression achieved Macro-F1 0.573 (95% source-group CI, 0.436–0.672), with a numerically higher point estimate than the six primary models for that representation; this descriptive difference does not establish a reliable ranking. In an exploratory baseline comparison using Raw+Geo+CLR, ResNet achieved 0.580 (0.417–0.682) and logistic regression 0.566 (0.424–0.662); their paired source-bootstrap difference was 0.014 (−0.120 to 0.132). LDA had lower point estimates in these two displayed schemes. Results for all seven baseline schemes are provided in the prediction archive. ILR-only and Raw+ILR were excluded from primary selection. Across the six primary model families, ILR-only Macro-F1 ranged from 0.430 to 0.529, while Raw+ILR ranged from 0.455 to 0.588; the Raw+ILR estimate for ResNet was 0.588 (95% source-group CI, 0.427–0.682). ILR results are descriptive because no paired inferential comparisons for these sensitivity schemes were archived.

Table 7. Simple baselines and primary models under two common representations.

Model	Raw Macro-F1 [95% CI]	Raw+Geo+CLR Macro-F1 [95% CI]
SVM	0.493 [0.393, 0.588]	0.469 [0.325, 0.563]
RF	0.491 [0.354, 0.576]	0.413 [0.298, 0.497]
XGBoost	0.508 [0.378, 0.590]	0.472 [0.333, 0.574]
MLP	0.491 [0.377, 0.585]	0.502 [0.365, 0.598]
ResNet	0.477 [0.361, 0.574]	0.580 [0.417, 0.682]
FT-Transformer	0.465 [0.354, 0.552]	0.501 [0.361, 0.590]
Logistic regression	0.573 [0.436, 0.672]	0.566 [0.424, 0.662]
LDA	0.330 [0.266, 0.403]	0.480 [0.356, 0.567]

Confidence intervals resample the same 76 connected source groups. Baselines are supplemental, not candidates in the primary 42-way nested selection.

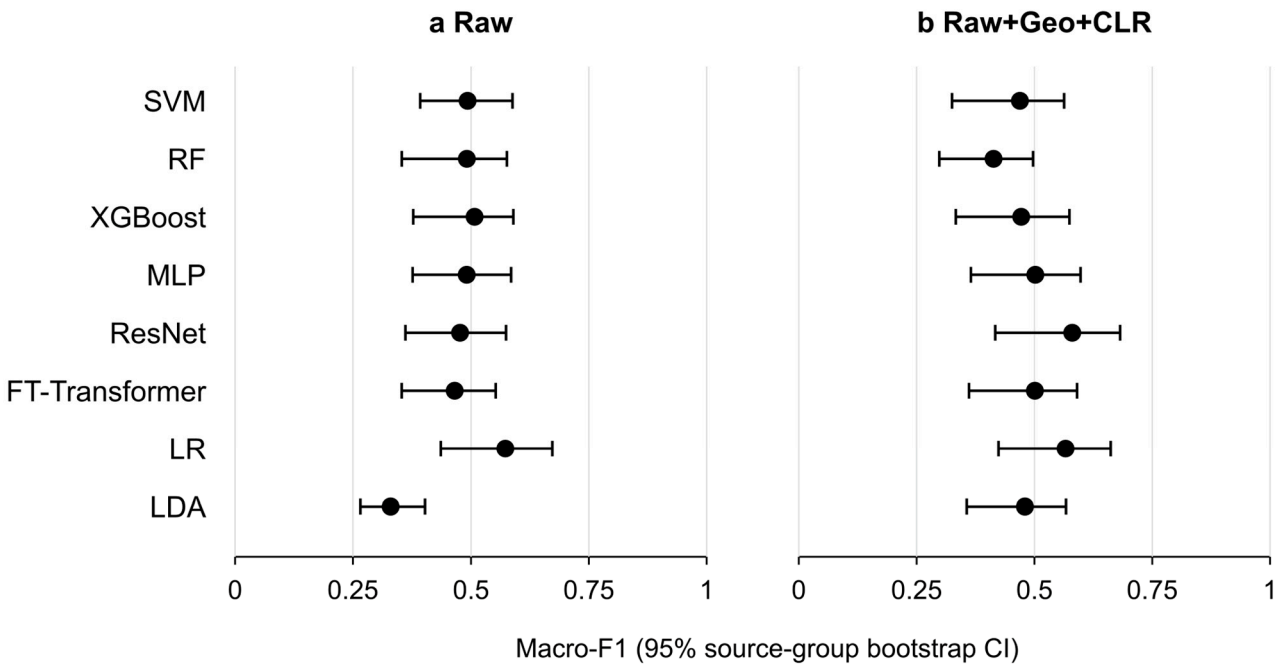


Figure 5. Source-group-held-out performance of the six primary models and the logistic-regression (LR) and linear-discriminant-analysis (LDA) baselines under (**a**) Raw and (**b**) Raw+Geo+CLR inputs. Points show the mean of five seed-specific pooled out-of-fold Macro-F1 scores. Horizontal bars are marginal 95% percentile intervals from 2000 bootstrap resamples of the 76 connected source groups, retaining all analyses and predictions from each sampled group. These are not simultaneous confidence intervals; overlap alone is not a paired significance test. LR and LDA used the unstandardized features of the archived baseline procedure.

3.5. Class-Specific Performance and Error Structure

Figure 6 resolves the overall score into the directions of misclassification. Averaged over seeds, 55.6% of felsic and 52.0% of intermediate analyses were assigned to their true classes. The dominant off-diagonal transitions were felsic to intermediate (40.2%) and intermediate to felsic (28.0%). Intermediate analyses were also assigned to mafic (11.8%) and ultramafic (8.2%) classes. These percentages were obtained by normalizing each seed's confusion matrix by the true-class support before averaging, not by treating repeated predictions as independent observations.

Mafic recall was 71.2%, with its largest error direction toward intermediate (13.6%). Ultramafic recall was 46.8%, with substantial assignments to mafic (24.7%) and intermediate (24.2%). The associated F1 estimates were 0.595 for felsic, 0.504 for intermediate, 0.647 for mafic and 0.479 for ultramafic. Source-bootstrap intervals were particularly wide for mafic (0.171–0.833) and ultramafic (0.149–0.720) F1, consistent with uncertainty from their limited independent-source support. The highest class point estimate should therefore not be interpreted as reliably superior transfer performance. The matrix summarizes the nested selection procedure, whose model–representation pair may differ across outer folds; it is not a comparison between a single Raw model and a single fused model.

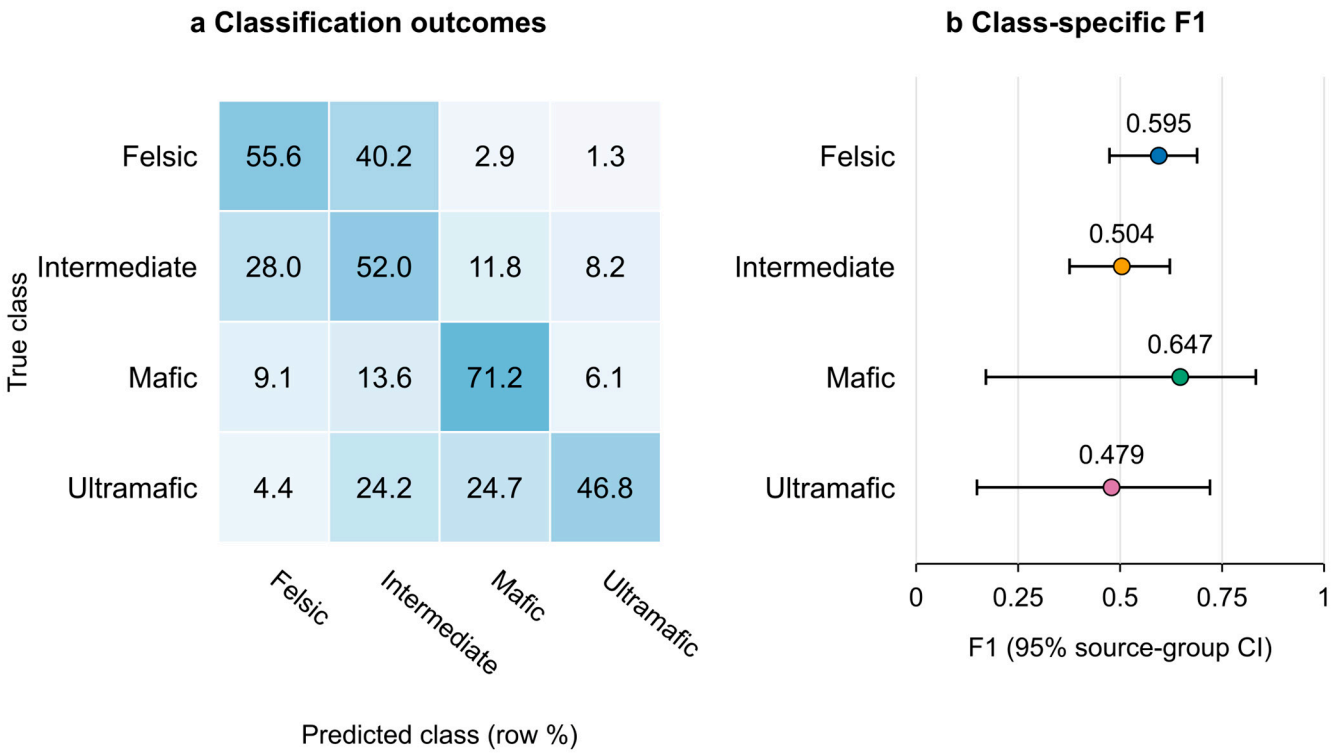


Figure 6. Error structure and class-specific performance of the nested selected procedure. (**a**) Row-normalized confusion matrix: predictions were selected using the candidate selected by inner cross-validation for each outer fold, normalized separately within each true class and seed, and then averaged over the five seeds. Entries are percentages and each unrounded row sums to 100%. (**b**) Mean seed-specific class F1 and marginal 95% source-group bootstrap intervals. The unique-analysis support is 1591 felsic, 1514 intermediate, 658 mafic and 385 ultramafic; repeated seeds do not increase the independent sample size. Confidence intervals resample the 76 connected source groups rather than individual records.

4. Discussion

4.1. Source Dependence and Generalization

The validation protocol determines which generalization question is answered. A record-level random split can place related apatite analyses from the same publication or geological body in both training and test partitions. Source-group-held-out nested validation instead assesses performance on sources absent during model fitting and selection. Its Macro-F1 estimate and wide source-group interval indicate substantial uncertainty in transfer to new geological settings. In the matched fixed-logistic-regression diagnostic, Macro-F1 was 0.653 under record-level folds and 0.481 under source-group-held-out folds for the same 4144 records from 73 source groups; the paired difference was 0.172 (95% source-group bootstrap CI, 0.076–0.269). This contrast is specific to the fixed diagnostic and does not establish a causal effect of split policy or generalize across model families. Independent minority-class sources remain relatively few, and a source group may contain multiple analyses or classes; neither a larger record count nor repeated evaluation seeds removes this dependence. Historical record-level results from a different cohort are not directly comparable.

Literature-compiled apatite REE datasets may also vary with analytical instruments, laboratories, detection limits, sampling strategies and geological context. Unit harmonization and duplicate removal address only part of this variation and cannot completely remove potential source effects. Such heterogeneity could influence performance on new collections, although this analysis does not isolate or attribute bias to any individual factor.

Source-group-aware validation assesses transfer across the constructed source groups, but cannot identify the contribution of each factor or eliminate unrecorded heterogeneity.

4.2. Model Complexity and the Role of Simple Baselines

The confidence intervals overlap, and the paired comparisons do not support a reliable ranking of the candidates. The 42-candidate comparison and simple baselines therefore do not establish a consistent advantage for deep architectures over a linear model. Logistic regression was competitive on Raw inputs, while the small point-estimate advantage of ResNet under Raw+Geo+CLR had a paired interval compatible with either benefit or harm. The absence of Holm-significant representation contrasts does not imply equivalence, but it limits claims of a reproducible combined-representation benefit in this dataset. Model and representation comparisons are most informative when interpreted through the prespecified nested procedure and source-level uncertainty rather than the largest descriptive matrix entry. Thus, the value of evaluating deep tabular architectures and derived geochemical representations in this task lies not in a demonstrated universal performance gain, but in establishing their behavior under a controlled source-group-held-out comparison.

This distinction matters when deploying a classifier on a new apatite collection. Increasing architecture capacity may change how interactions are fitted, but it does not increase the number of independent source groups. The difference between a complex model and a linear baseline is therefore a question to test under source separation, rather than an expected benefit of depth. The unstandardized LR/LDA baselines and unequal Optuna budgets also delimit the comparison: these results characterize the implemented procedures, not every possible implementation or equally optimized version of each model family.

For new-data use, the relevant configuration is the one selected within the complete development dataset by the procedure described in Section 2.3.2, followed by refitting on all development records. Choosing the most frequent outer-fold configuration or the largest value in the descriptive outer-score matrix would not implement that procedure. The reported nested estimate assesses selection followed by source-held-out evaluation; it is not an external validation of a single full-development classifier.

4.3. REE Patterns and the Effects of Representation

Figure 3 separates absolute abundance from relative REE pattern. The higher source-balanced median total REE of felsic apatite coexists with substantial overlap among all four classes. The stronger median negative Eu anomaly in felsic sources and higher normalized La/Yb in the operational ultramafic class provide more specific descriptions of how the source distributions differ. These observations motivate examining abundance, ratios and anomalies together, consistent with the use of apatite geochemistry in earlier discrimination studies [1,10]. However, the classes comprise different publications and geological settings. Neither a class median nor a ratio distribution establishes a unique crystallization process, oxidation state or mineralization history, and we do not infer such mechanisms from these plots.

Raw retains concentration magnitude, whereas CLR removes a common log-abundance component and expresses relative REE proportions. Geo ratios and anomalies encode selected aspects of those proportions, while the abundance sums retain magnitude. Combining them can alter the geometry presented to a model but cannot supply new measured chemistry. The contrasting Raw-to-CLR changes for ResNet and RF (Figure 4) therefore support a representation-dependent response, not a universally beneficial transformation. The source-balanced overlap in Figure 3 also cautions against expecting a single abundance or anomaly threshold to separate all sources. These descriptive plots and the

performance matrix answer different questions: the first describes the input population, whereas the second evaluates prediction after fitting.

CLR represents only the selected 14-REE subcomposition and has rank 13 because its coordinates sum to zero [24,25]. The orthonormal ILR sensitivity removes coordinate redundancy but retains dependence on the chosen element set and source coverage. The present plots do not provide model attribution. Correlated elements and deterministic derived variables can distribute apparent importance differently across representations, so an individual-element ranking would require an explicitly defined, stability-tested attribution analysis. No SHAP, permutation-importance or causal attribution result is claimed here. Source-balanced distributions instead supply an empirical geochemical context for the model comparisons without reinstating the former claim that fusion consistently improves performance.

4.4. Misclassification and Geological Applicability

The largest error exchange is between felsic and intermediate records (Figure 6), which are also represented by overlapping source-level abundance and normalized-ratio distributions (Figure 3). This coexistence is compatible with limited separability in the measured REE space, but the aggregate plots do not establish why any individual analysis was misclassified. The operational assignment of granodiorite and tonalite to the intermediate class further defines the boundary being learned. Its interpretation must follow the frozen lithological lookup rather than an assumed whole-rock compositional threshold.

Errors in the ultramafic class extend toward both mafic and intermediate labels, despite its larger median normalized La/Yb. Thus, a class-level geochemical contrast does not guarantee reliable record-level classification on held-out sources. The operational ultramafic category spans heterogeneous lithologies and contexts, and its limited source support yields a wide F1 interval. Likewise, the mafic class has the largest F1 point estimate but is represented by only nine connected groups, including folds with a single held-out mafic group. These results support reporting uncertainty for individual classes and avoiding claims that minority-class discrimination is nearly perfect.

For provenance applications, a prediction should be interpreted within the four operational labels and the chemical coverage of the training collection. Unseen lithologies, analytical settings or REE distributions may fall outside this scope. The current results do not establish a calibrated out-of-domain detector or an externally validated decision threshold. Additional independent geological sources would be needed to test such use, particularly for the underrepresented classes.

4.5. Limitations and Priorities for Further Validation

Several limitations remain. Grain identifiers could not be recovered consistently, and connected publication/body groups are proxies for independent geological sources. Locality and body metadata vary in precision. In the frozen provenance crosswalk, 41 archived source-file/worksheet/row tuples were shared by 82 records. A read-only post-archive comparison with the original GEOROC 2024-12 CSV found one unique full 14-REE-vector match for every affected analysis, with matching sample names; verified CSV positions are supplied in the supplement. The historical locator fields remain duplicated, and this check does not establish the accuracy of every other source metadata field. Source-specific below-detection-limit flags and handling metadata are also unavailable. The 30 source-sharing records set aside before fold construction were not evaluated as an independent external test in the formal prediction archive. After their exclusion, the primary mafic cohort contained gabbroic and doleritic records but no basalt records; transfer to basalt-hosted apatite was therefore not evaluated. Historical record-level results use a different cohort and are

contextual rather than a matched estimate of the effect of changing the partition. Moreover, the operational ultramafic category spans heterogeneous geological contexts, including carbonatite- and kimberlite-related rocks, mantle xenolith or metasomatic-mantle settings, and mafic–ultramafic intrusions (Table A1). It is represented by 15 connected source groups (Table 1) and has class-specific F1 of 0.479 (95% CI, 0.149–0.720; Table 5). This heterogeneity may affect class separability, but the present analysis does not isolate subtype effects or establish a cause for the observed performance. The class-level estimates therefore do not establish carbonatite-specific discrimination or general transfer to carbonatites. Consequently, inference is limited to this apatite 14-REE dataset, the four igneous host-rock classes, the stated source-group construction and the implemented model procedures.

Interpretability is a further limitation, particularly for the deep-learning models: their nonlinear interactions are not directly expressed as geochemical decision rules. Neither the descriptive REE plots nor the presence of logistic-regression baselines establishes which inputs drive individual predictions. No standardized-CLR coefficient analysis, grouped permutation importance, SHAP analysis or stability assessment of model attribution was performed. Correlated REEs, rank-deficient CLR coordinates and deterministic Geo transformations also complicate attribution to individual elements; a coefficient or importance score would not by itself establish a petrogenetic cause. Future work should examine standardized log-ratio coefficients in an explicitly specified linear model and grouped permutation importance for fitted models, with preprocessing confined to training data, correlated or derived inputs treated jointly, and attribution stability assessed across held-out source groups. Such work is needed before model predictions can be linked to specific geochemical controls.

5. Conclusions

Source-group-aware nested evaluation of 4148 apatite analyses from 76 connected groups yielded Macro-F1 0.556 (95% source-group bootstrap CI, 0.398–0.652) for four-class igneous host-rock classification. The same-record fixed-logistic-regression diagnostic gave higher Macro-F1 under record-level than source-held-out splitting (0.653 versus 0.481), illustrating source dependence in that diagnostic without isolating a universal causal effect of partition policy. Across the six primary model families and seven representations, multiplicity-adjusted source-level comparisons did not establish a consistent advantage for deeper architectures or combined representations; the Raw logistic-regression baseline remained competitive.

Source-balanced REE summaries add a geochemical view of these results. Total-REE distributions overlap across the four classes, while normalized La/Yb and Eu anomalies show differences in their source-level central tendencies and spread. The selected procedure nevertheless exhibits substantial felsic–intermediate confusion and uncertain transfer for mafic and ultramafic sources. These observations identify class overlap and source coverage as priorities for interpretation and future data collection. The study provides a reproducible comparison of model families and REE representations under source-aware validation, rather than evidence that greater model complexity or deterministic feature expansion necessarily improves classification of unseen geological sources.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app162010421/s1, Supplementary_Methods_and_Code_S1 (ZIP): Data provenance, host-rock label mapping, GEOROC record crosswalk, feature-reconstruction code and validation results, and statistical post-processing files; Supplementary_Figure_Data_and_Code_S2 (ZIP): PNG and SVG versions of Figures 1–6 in the main text, machine-readable source tables, figure-genertion scripts, and validation reports.

Author Contributions: Conceptualization, Y.Z.; methodology, Y.Z.; investigation, H.H.; formal analysis, H.H.; resources, F.H.; writing—original draft preparation, H.H.; writing—review and editing, Y.Z., C.L. and W.F.; supervision, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded jointly by the Deep Earth Probe and Mineral Resources Exploration-National Science and Technology Major Project (Grant No. 2025ZD1009303) and the Open Fund of Jiangxi Engineering Technology Research Center for Nuclear Geoscience Data Science and System Engineering (Grant No. JETRCNGDSS202207).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The formal experiment archive is not included in the publicly available supplementary files because of its size. It was submitted to the journal as non-published material for editorial and peer-review purposes and is available from the corresponding author upon reasonable request. Together, the formal experiment archive and Supplementary_Methods_and_Code_S1 contain the processed record-level dataset; analysis IDs and assigned citation, source-group and fixed-fold fields; source/class inventories; mapping and deduplication ledgers; the v2 runner, core implementation and configuration; per-record predictions; Optuna study databases; and statistical post-processing files. The analysis_id field is the unique record key. Forty-one archived source-file/worksheet/row locator tuples are duplicated across 82 analyses; these historical fields are preserved unchanged. Supplementary Methods and Code S1 provides a separate verified GEOROC 2024-12 CSV record-number crosswalk for all 82 analyses, based on unique complete 14-REE-vector matches and corroborating sample names, together with the raw CSV and verification code. It also provides the parameterized utility code/rebuild_features_from_work_dataset.py and its validation report. The utility reconstructs all Geo, CLR and ILR columns from the raw La–Lu fields in the formal experiment archive. Citation/reference and source-group assignments may recur across records. The source publications are identified in Appendix A/Table A1 and References [37–114]. Supplementary_Figure_Data_and_Code_S2 supplies PNG and SVG versions of Figures 1–6 in the main text, their machine-readable source tables, scripts for the source-balanced geochemistry and archived-prediction summaries, and validation reports. These materials were prepared without altering the frozen model inputs or rerunning model training.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A. Reproducibility and Analysis Files

The companion archive contains the frozen processed dataset (SHA-256 4c707eff0447a72fc606a1faaa0b26ac07b347e69cb6d83c56b7b3ff2dd40803), source-provenance and group manifests, outer-fold selection manifests, 340 Optuna SQLite databases, v2 predictions, the executable runner, core model implementation, configuration and server environment lock. The statistical script reads the unmodified experiment ZIP (SHA-256 d3a353aa434a89f0600aefc6dba49115cc00d3079cc57d620b222ffd762e594c) and produces the reported source-group confidence intervals and paired comparisons. Data-cleaning, rock-label mapping, exclusion and feature-construction scripts are included in the Supplementary Files. Evaluation seeds were 41–45; the split/HPO seed was 20260718. The server reported four NVIDIA RTX 5090 GPUs, driver 580.76.05, PyTorch 2.8.0+cu128 (CUDA 12.8 build) and Optuna 5.0.0; exact package versions are retained in SERVER_ENVIRONMENT_LOCK.txt. Historical data-preparation scripts document the collection workflow but still require intermediate inputs and author-specific path settings; they do not provide a self-contained reconstruction of literature collection, exclusion and cohort assembly. Supplementary_Methods_and_Code_S1 now includes the parameterized

rebuild_features_from_work_dataset.py entry point, which reconstructs and verifies all 36 Geo/CLR/ILR columns directly from the delivered raw La–Lu fields in the archived work_dataset. This closes the feature-construction check from delivered model inputs, but not the earlier raw-source collection pipeline. The addendum also distinguishes the driver-displayed CUDA capability from the archived PyTorch CUDA 12.8 build without altering the immutable experiment ZIP.

The source-level counts and corresponding primary data publications are listed in Table A1. Exact source-group assignments and fold-specific model parameters are supplied in machine-readable form. No global parameter set was selected across the five outer folds. The accompanying code, configuration, recorded studies and environment lock specify the implementation used to obtain the reported results. The S2 figure scripts document filtering, source–class aggregation and seed-wise confusion-matrix normalization. Their checks reproduce the selected-procedure performance estimates and intervals before visualizing them.

Table A1. Source identifiers, record counts, and primary data publications for the 4148-analysis cohort. Frozen provenance IDs are retained; author-name spellings and other bibliographic details are corrected against publication metadata. Original archival labels remain in the data files. F, felsic; I, intermediate; M, mafic; U, ultramafic. The 30 James Ross/Daye holdout records are excluded from this inventory.

Frozen Source Identifier	N	Host-Rock Class, n	Reference
[21356] LI HE (2014)	5	F 5	[37]
[21904] JIANG XIAO-YAN (2018)	71	F 71	[38]
[22246] XIE FUWEI (2018)	62	F 4; I 58	[39]
[22334] JIANG XIAO-YAN(2018) [GeoReM [11164]]	48	F 30; I 18	[40]
[22784] YU ZHI-QIANG (2019)	9	F 6; I 3	[41]
[23038] BRUAND E. (2019)	37	F 9; I 28	[42]
[23128] LIU ZHI-CHAO (2019)	60	F 60	[43]
[23144] ZAFAR T. (2019)	41	F 41	[44]
[23497] QU PAN (2019)	49	F 49	[45]
[23716] SUN SAI-JUN (2020)	20	F 20	[46]
[23994] AZADBAKHT Z. (2020)	78	F 65; I 13	[47]
[24394] YANG FAN (2020)	308	F 308	[48]
[24625] SUN JIN-FENG (2021)	92	F 92	[49]
[24991] XING KAI (2021)	20	F 12; I 8	[50]
[25249] ELIWA H. A. (2021)	29	F 29	[51]
[25478] ZHANG XIAO-MING (2021)	43	F 32; I 11	[52]
[25559] SEPIDBAR F. (2021)	32	F 9; I 23	[53]
[25693] JING YAN (2022)	40	F 40	[54]
[25726] ZHOU RUN-JIE (2022)	143	F 16; I 127	[55]
[25840] YUAN YU (2022)	6	F 2; I 4	[56]
[26205] XU LEI-LUO (2023)	96	F 96	[57]
[26251] MU MAO-SONG (2022)	25	F 25	[58]
[26341] ZHANG HUICHAO (2022)	15	F 15	[59]
[26423] ZHAO XU (2022)	78	F 78	[60]
[26426] FAN CHAOXI (2023)	14	F 9; I 5	[61]
[26565] LONG XIN-YU (2023)	181	F 116; I 65	[62]

Table A1. *Cont.*

Frozen Source Identifier	N	Host-Rock Class, n	Reference
[26664] ZHANG XIN (2023)	90	F 45; I 45	[63]
[26873] LI XIN (2023)	100	F 90; I 10	[64]
[26927] QUAN YI-KANG (2023)	53	F 21; I 32	[65]
[22108] CHEN ZUXING (2018)	18	F 18	[66]
[23733] NATHWANI C. L. (2020)	145	F 108; I 37	[67]
[25988] WANG HAIRUO (2022)	207	F 47; I 160	[68]
[25692] LI XI-YAO (2022)	14	F 14	[69]
[26239] SHARPE M. S. (2022)	9	F 9	[70]
[21289] LAURENT O.(2017) [GeoReM [10516]]	88	I 88	[71]
[21351] ZHANG CHAO(2013) [GeoReM [7259]]	5	I 3; M 2	[72]
[21391] RABBIA O. M. (2017)	4	I 4	[73]
[22731] XIE FUWEI (2018)	64	I 64	[74]
[22821] SUN SAI-JUN(2019) [GeoReM [12263]]	59	I 59	[75]
[24262] XIA YING (2020)	3	I 2; M 1	[76]
[24410] XIAO YUANYUAN (2020)	35	I 35	[77]
[24970] ZHANG LI-PENG (2021)	40	I 40	[78]
[26328] CAO JINGYA (2023)	14	I 14	[79]
[26425] ZHOU RUN-JIE (2022)	61	I 61	[80]
[26836] ZHU QINGBO (2023)	33	I 33	[81]
[25433] MARFIN A. E. (2021)	12	I 12	[82]
[25437] XU WEN-GANG (2015)	16	I 16	[83]
[15778] TANG MING (2012)	22	I 22	[84]
[26414] WU YA-DONG (2022)	204	I 204	[85]
[24549] WANG MENGXI (2020)	155	M 155	[86]
[17673] WU FU-YUAN (2013)	2	U 2	[87]
[17934] BÜHN B. (2001)	3	U 3	[88]
[18025] YE HAI-MIN(2013) [GeoReM [7366]]	40	U 40	[89]
[20286] WEIDENDORFER D. (2016)	1	U 1	[90]
[24067] VELÁZQUEZ SANTANA L. C. (2020)	130	U 130	[91]
[24751] YU KAIZHANG (2021)	5	U 5	[92]
[25106] SAVARD J. J. (2021)	4	U 4	[93]
[19041] CHENG ZHIGUO (2014)	6	U 6	[94]
[8274] BEDINI R. M. (1999)	1	U 1	[95]
[16609] ZHOU QIN (2010)	3	U 3	[96]
[18709] GONZALES-JIMENEZ J. M. (2014)	12	U 12	[97]
[22326] BALI E. (2018)	15	U 15	[98]
[22894] VILLASECA C. (2018)	32	U 32	[99]
[24017] LU JIANGGU (2020)	4	U 4	[100]
[27108] HOWARTH G. H. (2023)	29	U 29	[101]
[28431] YAN XIANG-YU (2024)	20	M 20	[102]
[28996] CAO YONGHUA (2024)	98	U 98	[103]
KIEFFER M. A. (2023)	213	M 213	[104]
Gordonbrook dolerite (2026)	31	M 31	[105]
Xing & Wang (2017) American Mineralogist 102, DOI 10.2138/am-2017-6028	41	M 41	[106]

Table A1. *Cont.*

Frozen Source Identifier	N	Host-Rock Class, n	Reference
Pedersen et al. (2021) Contributions to Mineralogy and Petrology 176:102, DOI 10.1007/s00410-021-01861-x	180	M 180	[107]
Rakhimov (2024) Minerals 14:1271, DOI 10.3390/min14121271	15	M 15	[108]
[28685] HAO LE-RAN (2024)	48	I 48	[109]
[28545] GUMSLEY ANNA (2024)	43	I 43	[110]
[28852] LI YU (2024)	29	I 29	[111]
[27186] YANG FEI (2024)	29	I 29	[112]
[27386] TANG PAN (2024)	12	I 12	[113]
[28723] BRUAND E. (2014)	49	I 49	[114]

References

1. Belousova, E.A.; Griffin, W.L.; O'Reilly, S.Y.; Fisher, N.I. Apatite as an indicator mineral for mineral exploration: Trace-element compositions and their relationship to host rock type. *J. Geochem. Explor.* **2002**, *76*, 45–69. [CrossRef]
2. Tan, H.M.R.; Huang, X.W.; Meng, Y.M.; Xie, H.; Qi, L. Multivariate statistical analysis of trace elements in apatite: Discrimination of apatite with different origins. *Ore Geol. Rev.* **2023**, *153*, 105269. [CrossRef]
3. Qiu, K.F.; Zhou, T.; Chew, D.; Hou, Z.-L.; Müller, A.; Yu, H.-C.; Lee, R.G.; Chen, H.; Deng, J. Apatite trace element composition as an indicator of ore deposit types: A machine learning approach. *Am. Mineral.* **2024**, *109*, 303–314. [CrossRef]
4. Li, X.M.; Zhang, Y.X.; Li, Z.K.; Zhao, X.-F.; Zuo, R.-G.; Xiao, F.; Zheng, Y. Discrimination of Pb–Zn deposit types using sphalerite geochemistry: New insights from machine learning algorithm. *Geosci. Front.* **2023**, *14*, 101580. [CrossRef]
5. Zou, S.; Chen, X.; Brzozowski, M.J.; Leng, C.-B.; Xu, D. Application of machine learning to characterizing magma fertility in porphyry Cu deposits. *J. Geophys. Res. Solid. Earth* **2022**, *127*, e2022JB024584. [CrossRef]
6. Hou, L.-L.; Wu, S.; Yi, J.-Z.; Ci, Q.; Chen, L.; Liu, X.; Wei, S.; A Wang Danzeng; Zheng, Y.; Liu, P. Discriminating deposit types using chlorite trace elements based on machine learning. *Earth Sci.* **2024**, *49*, 4303–4317. Available online: https://cjournal.hep.com.cn/1000-2383/CN/1159928821840535919 (accessed on 25 December 2024). [CrossRef]
7. Hong, S.; Zuo, R.; Huang, X.; Xiong, Y. Distinguishing IOCG and IOA deposits via random forest algorithm based on magnetite composition. *J. Geochem. Explor.* **2021**, *230*, 106859. [CrossRef]
8. Zhou, T.; Cai, Y.W.; An, M.G.; Zhou, F.; Zhi, C.-L.; Sun, X.-C.; Tamer, M. Visual interpretation of machine learning: Genetical classification of apatite from various ore sources. *Minerals* **2023**, *13*, 491. [CrossRef]
9. Cao, G.; Chen, H.; Zhang, Y.; Sun, W.; Zhao, J.; Zhao, H.; Wang, H. Primary controlling factors of apatite trace element composition and implications for exploration in orogenic gold deposits. *Geochem. Geophys. Geosyst* **2024**, *25*, e2024GC011574. [CrossRef]
10. Han, F.G.; Leng, C.B.; Chen, J.J.; Zhan, Y.-Y. Rock type discrimination by using trace elements of apatite based on the machine learning. *Bull. Mineral. Petrol. Geochem.* **2024**, *43*, 607–620, (In Chinese with English abstract). [CrossRef]
11. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2012; Volume 25, pp. 1097–1105.
12. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the NAACL-HLT 2019*; Association for Computational Linguistics: Minneapolis, MN, USA, 2019; pp. 4171–4186.
13. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; IEEE: Las Vegas, NV, USA, 2016; pp. 770–778. [CrossRef]
14. Fu, D.; Su, C.; Wang, W.; Yuan, R. Deep learning based lithology classification of drill core images. *PLoS ONE* **2022**, *17*, e0270826. [CrossRef] [PubMed]
15. Zhang, P.Y.; Sun, J.M.; Jiang, Y.J.; Gao, J.S. Deep learning method for lithology identification from borehole images. In *79th EAGE Conference and Exhibition 2017*; EAGE: Paris, France, 2017. [CrossRef]
16. Li, S.; Chen, J.; Liu, C.; Wang, Y. Mineral prospectivity prediction via convolutional neural networks based on geological big data. *J. Earth Sci.* **2021**, *32*, 327–347. [CrossRef]
17. Lyu, Y.; He, C.; Zhao, J.M.; Zhang, Z.J. Storage and analysis of tabular data of mineralogy, petrology and geochemistry: From on-premises to cloud. *Bull. Mineral. Petrol. Geochem.* **2025**, *44*, 542–556, (In Chinese with English abstract). [CrossRef]
18. Borisov, V.; Leemann, T.; Sessler, K.; Haug, J.; Pawelczyk, M.; Kasneci, G. Deep neural networks and tabular data: A survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *35*, 7499–7519. [CrossRef] [PubMed]

19. Grinsztajn, L.; Oyallon, E.; Varoquaux, G. Why do tree-based models still outperform deep learning on typical tabular data? In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA; New York, NY, USA, 2022; Volume 35, pp. 507–520.
20. Shwartz-Ziv, R.; Armon, A. Tabular data: Deep learning is not all you need. *Inf. Fusion.* **2022**, *81*, 84–90. [CrossRef]
21. Sun, C.; Wang, X.; Cheng, Y. A Manifold Learning Process Monitoring Framework for Nonlinear Industrial Systems. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 3566518. [CrossRef]
22. Sun, C.-Y.; Yin, Y.-Z.; Kang, H.-B.; Ma, H.-J. A Quality-Related Fault Detection Method Based on the Dynamic Data-Driven Algorithm for Industrial Systems. *IEEE Trans. Autom. Sci. Eng.* **2022**, *19*, 3942–3952. [CrossRef]
23. Sun, S.-S.; McDonough, W.F. Chemical and isotopic systematics of oceanic basalts: Implications for mantle composition and processes. In *Magmatism in the Ocean Basins*; Saunders, A.D., Norry, M.J., Eds.; Geological Society: London, UK; Special Publications, 1989; Special Publications Volume 42; pp. 313–345. [CrossRef]
24. Aitchison, J. *The Statistical Analysis of Compositional Data*; Chapman & Hall: London, UK, 1986.
25. Egozcue, J.J.; Pawlowsky-Glahn, V.; Mateu-Figueras, G.; Barceló-Vidal, C. Isometric logratio transformations for compositional data analysis. *Math. Geol.* **2003**, *35*, 279–300. [CrossRef]
26. Kaufman, S.; Rosset, S.; Perlich, C.; Stitelman, O. Leakage in data mining: Formulation, detection, and avoidance. *ACM Trans. Knowl. Discov. Data* **2012**, *6*, 15. [CrossRef]
27. McDonough, W.F.; Sun, S.-S. The composition of the Earth. *Chem. Geol.* **1995**, *120*, 223–253. [CrossRef]
28. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
29. Vapnik, V.N. *Statistical Learning Theory*; Wiley: New York, NY, USA, 1998.
30. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
31. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA; New York, NY, USA, 2017; Volume 30, pp. 5998–6008.
32. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; ACM: New York, NY, USA, 2016; pp. 785–794. [CrossRef]
33. Gorishniy, Y.; Rubachev, I.; Khrulkov, V.; Babenko, A. Revisiting deep learning models for tabular data. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA; New York, NY, USA, 2021; Volume 34, pp. 18932–18943.
34. Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*; ACM: New York, NY, USA, 2019; pp. 2623–2631. [CrossRef]
35. He, H.; Garcia, E.A. Learning from imbalanced data. *IEEE Trans. Knowl. Data Eng.* **2009**, *21*, 1263–1284. [CrossRef]
36. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process Manag.* **2009**, *45*, 427–437. [CrossRef]
37. Li, H.; Ling, M.X.; Ding, X.; Zhang, H.; Li, C.-Y.; Liu, D.-Y.; Sun, W.-D. The geochemical characteristics of Haiyang A-type granite complex in Shandong, eastern China. *Lithos* **2014**, *200–201*, 142–156. [CrossRef]
38. Jiang, X.Y.; Li, H.; Ding, X.; Wu, K.; Guo, J.; Liu, J.-Q.; Sun, W.-D. Formation of A-type granites in the Lower Yangtze River Belt: A perspective from apatite geochemistry. *Lithos* **2018**, *304–307*, 125–134. [CrossRef]
39. Xie, F.; Tang, J.; Lang, X.; Ma, D. The different sources and petrogenesis of Jurassic intrusive rocks in the southern Lhasa subterrane, Tibet: Evidence from the trace element compositions of zircon, apatite, and titanite. *Lithos* **2018**, *314–315*, 447–462. [CrossRef]
40. Jiang, X.Y.; Luo, J.C.; Guo, J.; Wu, K.; Zhang, Z.-K.; Sun, W.-D.; Xia, X.-P. Geochemistry of I- and A-type granites of the Qingyang–Jiuhuashan complex, eastern China: Insights into early cretaceous multistage magmatism. *Lithos* **2018**, *316–317*, 278–294. [CrossRef]
41. Yu, Z.Q.; Chen, W.F.; Chen, P.R.; Wang, K.-X.; Fang, Q.-C.; Tang, X.-S.; Ling, H.-F. Chemical composition and Sr isotopes of apatite in the Xiangshan A-type volcanic-intrusive complex, Southeast China: New insight into petrogenesis. *J. Asian Earth Sci.* **2019**, *172*, 66–82. [CrossRef]
42. Bruand, E.; Storey, C.; Fowler, M.; Heilimo, E. Oxygen isotopes in titanite and apatite, and their potential for crustal evolution research. *Geochim. Cosmochim. Acta* **2019**, *255*, 144–162. [CrossRef]
43. Liu, Z.C.; Wu, F.Y.; Liu, X.C.; Wang, J.-G.; Yin, R.; Qiu, Z.-L.; Ji, W.-Q.; Yang, L. Mineralogical evidence for fractionation processes in the Himalayan leucogranites of the Ramba Dome, southern Tibet. *Lithos* **2019**, *340–341*, 71–86. [CrossRef]
44. Zafar, T.; Leng, C.B.; Zhang, X.C.; Rehman, H.U. Geochemical attributes of magmatic apatite in the Kukaazi granite from western Kunlun orogenic belt, NW China: Implications for granite petrogenesis and Pb-Zn (-Cu-W) mineralization. *J. Geochem. Explor.* **2019**, *204*, 256–269. [CrossRef]
45. Qu, P.; Li, N.B.; Niu, H.C.; Yang, W.-B.; Shan, Q.; Zhang, Z.-Y. Zircon and apatite as tools to monitor the evolution of fractionated I-type granites from the central Great Xing'an Range, NE China. *Lithos* **2019**, *348–349*, 105207. [CrossRef]

46. Sun, S.J.; Zhang, R.Q.; Cong, Y.N.; Zhang, L.-P.; Sun, W.-D.; Li, C.-Y.; Ding, X. Analogous diagenetic conditions of dark enclave and its host granite derived by magma mixing: Evidence for a post-mixing magmatic process. *Lithos* **2020**, *356–357*, 105373. [CrossRef]
47. Azadbakht, Z.; Lentz, D.R.; McFarlane, C.R.M.; Whalen, J.B. Using magmatic biotite chemistry to differentiate barren and mineralized Silurian–Devonian granitoids of New Brunswick, Canada. *Contrib. Mineral. Petrol.* **2020**, *175*, 69. [CrossRef]
48. Yang, F.; Santosh, M.; Glorie, S.; Xue, F.; Zhang, S.; Zhang, X. Apatite geochronology and chemistry of Luanchuan granitoids in the East Qinling Orogen, China: Implications for petrogenesis, metallogenesis and exploration. *Lithos* **2020**, *378–379*, 105797. [CrossRef]
49. Sun, J.F.; Yang, J.H.; Zhang, J.H.; Yang, Y.-H.; Zhu, Y.-S. Apatite geochemical and Sr Nd isotopic insights into granitoid petrogenesis. *Chem. Geol.* **2021**, *566*, 120104. [CrossRef]
50. Xing, K.; Shu, Q.; Lentz, D.R. Constraints on the Formation of the Giant Daheishan Porphyry Mo Deposit (NE China) from Whole-Rock and Accessory Mineral Geochemistry. *J. Petrol.* **2021**, *62*, egab018. [CrossRef]
51. Eliwa, H.A.; Deevsalar, R.; Mahdy, N.M.; Kumar, S.; El-Gameel, K.; Zafar, T.; Khalaf, I.M.; Murata, M.; Ozawa, H.; Andresen, A.; et al. Field, textural, geochemical, and isotopic constraints on the origin and evolution of the magmatic microgranular enclaves from the Gharib Granitoid Complex, North Eastern Desert, Egypt. *Precambrian Res.* **2021**, *365*, 106380. [CrossRef]
52. Zhang, X.M.; Sun, C.Y.; Xu, W.L.; Yang, D.-B.; Tang, J. Geochemistry of apatites from Mesozoic granitoids in the northeastern North China Craton and their petrogenetic implications. *Lithos* **2021**, *402–403*, 106198. [CrossRef]
53. Sepidbar, F.; Ghorbani, G.; Zoheir, B.; Palin, R.M.; Homam, S.M.; Zafar, T.; Ma, J.; He, L. Coeval calc-alkaline and alkaline Cadomian magmatism in the Bafq, central Iran: Insights into their petrogenesis. *Lithos* **2021**, *406–407*, 106535. [CrossRef]
54. Jing, Y.; Yang, H.; Ge, W.C.; Dong, Y.; Ji, Z.; Bi, J.-H.; Zhou, H.-Y.; Xing, D.-H. When did the final closure occur of the eastern Paleo-Asian Ocean: Constraints from the latest Early–Middle Triassic adakitic granites in the southeastern Central Asian Orogenic Belt. *Gondwana Res.* **2022**, *103*, 146–171. [CrossRef]
55. Zhou, R.J.; Wen, G.; Li, J.W.; Jiang, S.-Y.; Hu, H.; Deng, X.-D.; Zhao, X.-F.; Yan, D.-R.; Wei, K.-T.; Cai, H.-A.; et al. Apatite chemistry as a petrogenetic–metallogenic indicator for skarn ore-related granitoids: An example from the Daye Fe–Cu–(Au–Mo–W) district, Eastern China. *Contrib. Mineral. Petrol.* **2022**, *177*, 23. [CrossRef]
56. Yuan, Y.; Guo, J.L.; Zong, K.; Feng, L.; Wang, Z.; Moynier, F.; Zhang, W.; Hu, Z.; Xu, H. Stable zirconium isotopic fractionation during alkaline magma differentiation: Implications for the differentiation of continental crust. *Geochim. Cosmochim. Acta* **2022**, *326*, 41–55. [CrossRef]
57. Xu, L.L.; Bi, X.W.; Zhang, X.C.; Huang, M.-L.; Liu, G. Mantle contribution to the generation of the giant Jinduicheng porphyry Mo deposit, Central China: New insights from combined in-situ element and isotope compositions of zircon and apatite. *Chem. Geol.* **2023**, *616*, 121238. [CrossRef]
58. Mu, M.S.; Yang, D.B.; Yang, H.T.; Wang, A.-Q.; Hao, L.-R.; Xu, W.-L.; Wang, F. Petrogenesis of late Paleoproterozoic post-collisional magmatism in southern north China Craton: Insights from geochemistry and Nd–Hf isotopic compositions of A-type granites. *Precambrian Res.* **2022**, *383*, 106887. [CrossRef]
59. Zhang, H.; Zhang, H.; Hou, Z.; Chai, P.; Zhou, K. Geochronology and Geochemistry of the Granite Xenolith within the Lamprophyre at the Zhenyuan Gold Deposit (Yunnan Province, SW China). *Acta Geol. Sin.—Engl. Ed.* **2022**, *96*, 477–489. [CrossRef]
60. Zhao, X.; Li, N.B.; Niu, H.C.; Wang, J.; Yan, S.; Yang, Y.-Y.; Fu, R.-X.; Huizenga, J.M. HREE enrichment during magmatic evolution recorded by apatite: Implication for the ion-adsorption HREE mineralization in South China. *Lithos* **2022**, *432–433*, 106896. [CrossRef]
61. Fan, C.; Xu, C.; Shi, A.; Smith, M.P.; Kynicky, J.; Wei, C. Origin of heavy rare earth elements in highly fractionated peraluminous granites. *Geochim. Cosmochim. Acta* **2023**, *343*, 371–383. [CrossRef]
62. Long, X.Y.; Tang, J.; Xu, W.L.; Sun, C.-Y.; Luan, J.-P.; Xiong, S.; Zhang, X.-M. Trace element and Nd isotope analyses of apatite in granitoids and metamorphosed granitoids from the eastern Central Asian Orogenic Belt: Implications for petrogenesis and post-magmatic alteration. *Geosci. Front.* **2023**, *14*, 101517. [CrossRef]
63. Zhang, X.; Xie, J.; Huang, S.; Ge, L.; Li, Q.; Liu, J.; Yan, J. Early Cretaceous Ningguo-Guangde weakly fractionated I-type granitoids, Jiangnan tungsten belt, China: Implications for petrogenesis and W Mo mineralization. *J. Geochem. Explor.* **2023**, *249*, 107211. [CrossRef]
64. Li, X.; Zheng, Y.C.; Shen, Y.; Wu, C.-D.; Xu, P.-Y.; Yang, Y.-S.; Wang, L.; Hou, Z.-Q. Petrogenesis of Sn-related granitoids and implications for the formation of the world-class Gejiu Sn district, South China: Insights from whole-rock and accessory mineral geochemistry. *Lithos* **2023**, *448–449*, 107166. [CrossRef]
65. Quan, Y.; Yang, D.; Yan, X.; Wang, A.; Hao, L.; Yang, H.; Wang, F.; Xu, W. Petrogenesis of Mesozoic granitoids in the northeastern North China Craton: Constraints from apatite trace elements and in-situ Nd isotopic data. *Lithos* **2023**, *450–451*, 107190. [CrossRef]

66. Chen, Z.; Zeng, Z.; Wang, X.; Yin, X.; Chen, S.; Guo, K.; Lai, Z.; Zhang, Y.; Ma, Y.; Qi, H.; et al. U-Th/He dating and chemical compositions of apatite in the dacite from the southwestern Okinawa Trough: Implications for petrogenesis. *J. Asian Earth Sci.* **2018**, *161*, 1–13. [CrossRef]
67. Nathwani, C.L.; Loader, M.A.; Wilkinson, J.J.; Buret, Y.; Sievwright, R.H.; Hollings, P. Multi-stage arc magma evolution recorded by apatite in volcanic rocks. *Geology* **2020**, *48*, 323–327. [CrossRef]
68. Wang, H.; Cai, K.; Sun, M.; Xia, X.-P.; Lai, C.-K.; Li, P.; Wan, B.; Zhang, Z. Apatite as a magma redox indicator and its application in metallogenic research. *Lithos* **2022**, *422–423*, 106749. [CrossRef]
69. Li, X.Y.; Li, S.; Suo, Y.; Huang, F.; Wang, P.; Luan, S.; Zhou, J. High-silica rhyolites in the terminal stage of massive Cretaceous volcanism, SE China: Modified crustal sources and low-pressure magma chamber. *Gondwana Res.* **2022**, *102*, 133–150. [CrossRef]
70. Sharpe, M.S.; Barker, S.J.; Rooyakkers, S.M.; Wilson, C.J.N.; Chambefort, I.; Rowe, M.C.; Schipper, C.I.; Charlier, B.L.A. A sulfur and halogen budget for the large magmatic system beneath Taupō volcano. *Contrib. Mineral. Petrol.* **2022**, *177*, 95. [CrossRef]
71. Laurent, O.; Zeh, A.; Gerdes, A.; Villaros, A.; Gros, K.; Słaby, E. How do granitoid magmas mix with each other? Insights from textures, trace element and Sr–Nd isotopic composition of apatite and titanite from the Matok pluton (South Africa). *Contrib. Mineral. Petrol.* **2017**, *172*, 80. [CrossRef]
72. Zhang, C.; Ma, C.; Holtz, F.; Koepke, J.; Wolff, P.E.; Berndt, J. Mineralogical and geochemical constraints on contribution of magma mixing and fractional crystallization to high-Mg adakite-like diorites in eastern Dabie orogen, East China. *Lithos* **2013**, *172–173*, 118–138. [CrossRef]
73. Rabbia, O.M.; Correa, K.J.; Hernández, L.B.; Ulrich, T. "Normal" to adakite-like arc magmatism associated with the El Abra porphyry copper deposit, Central Andes, Northern Chile. *Int. J. Earth Sci.* **2017**, *106*, 2687–2711. [CrossRef]
74. Xie, F.; Tang, J.; Chen, Y.; Lang, X. Apatite and zircon geochemistry of Jurassic porphyries in the Xiongcun district, southern Gangdese porphyry copper belt: Implications for petrogenesis and mineralization. *Ore Geol. Rev.* **2018**, *96*, 98–114. [CrossRef]
75. Sun, S.J.; Yang, X.Y.; Wang, G.J.; Sun, W.-D.; Zhang, H.; Li, C.-Y.; Ding, X. In situ elemental and Sr-O isotopic studies on apatite from the Xu-Huai intrusion at the southern margin of the North China Craton: Implications for petrogenesis and metallogeny. *Chem. Geol.* **2019**, *510*, 200–214. [CrossRef]
76. Xia, Y.; Wang, Q.; Zhu, D.C.; Ernst, R.E.; Zhang, S.; Liu, D.; Zhao, Z.-D. Intermediate rocks in the Comei large igneous provinces produced by amphibole crystallization of tholeiitic basaltic magma. *Lithos* **2020**, *374–375*, 105731. [CrossRef]
77. Xiao, Y.; Chen, S.; Niu, Y.; Wang, X.; Xue, Q.; Wang, G.; Gao, Y.; Gong, H.; Kong, J.; Shao, F.; et al. Mineral Compositions of Syn-collisional Granitoids and their Implications for the Formation of Juvenile Continental Crust and Adakitic Magmatism. *J. Petrol.* **2020**, *61*, egaa038. [CrossRef]
78. Zhang, L.; Hu, Y.; Deng, J.; Wang, J.; Wang, K.; Sui, Q.; Chen, Y.; Wu, K.; Sun, W. Genesis and mineralization potential of the Late Cretaceous Chemen granodioritic intrusion in the southern Gangdese magmatic belt, Tibet. *J. Asian Earth Sci.* **2021**, *217*, 104829. [CrossRef]
79. Cao, J.; Yang, X.; Yang, S.; Zhong, C.; Wang, Y. Records of apatite for multiple injections of magmas in adakitic plutons: A case study of Mesozoic plutons in the Shatanjiao region of the Tongling ore cluster, south China. *J. Asian Earth Sci.* **2023**, *242*, 105507. [CrossRef]
80. Zhou, R.J.; Wen, G.; Li, J.W.; Cao, K.; Wei, K.-T.; Cai, H.-A.; Shang, S.-C.; Sun, Y.; Fan, L.; Li, B.-C.; et al. Petrogenesis and metal fertility of Yinzu pluton in the Daye district, eastern China: Insights from whole-rock and mineral (zircon, apatite and amphibole) geochemistry. *Lithos* **2022**, *432–433*, 106898. [CrossRef]
81. Zhu, Q.; Zhao, X.; Hong, W.; Jin, G.; Huang, W.; Wang, C.; Liu, K.; Gao, T.; Wen, H. Geochronology, Hf isotope and trace element of zircon and apatite for neoproterozoic granodiorites in the eastern Jiangnan Orogen: Implications for the neoproterozoic tectonic evolution. *Lithos* **2023**, *446–447*, 107134. [CrossRef]
82. Marfin, A.E.; Radomskaya, T.A.; Ivanov, A.V.; Kamenetsky, V.S.; Kamenetsky, M.B.; Yakich, T.Y.; Gertner, I.F.; Kamo, S.L.; Ernst, R.E.; Bryanskiy, N.V.; et al. U–Pb Dating of Apatite, Titanite and Zircon of the Kingash Mafic–Ultramafic Massif, Kan Terrane, Siberia: From Rodinia Break-up to the Reunion with the Siberian Craton. *J. Petrol.* **2021**, *62*, egab049. [CrossRef]
83. Xu, W.G.; Fan, H.R.; Hu, F.F.; Santosh, M.; Yang, K.-F.; Lan, T.-G. In situ chemical and Sr–Nd–O isotopic compositions of apatite from the Tongshi intrusive complex in the southern part of the North China Craton: Implications for petrogenesis and metallogeny. *J. Asian Earth Sci.* **2015**, *105*, 208–222. [CrossRef]
84. Tang, M.; Wang, X.L.; Xu, X.S.; Zhu, C.; Cheng, T.; Yu, Y. Neoproterozoic subducted materials in the generation of Mesozoic Luzong volcanic rocks: Evidence from apatite geochemistry and Hf–Nd isotopic decoupling. *Gondwana Res.* **2012**, *21*, 266–280. [CrossRef]
85. Wu, Y.D.; Yang, J.H.; Sun, J.F.; Wang, H.; Zhou, B.-Q.; Xu, L.; Wu, B. Revisiting the Late Jurassic adakitic rocks in the Yanshan fold and thrust belts, North China Craton: Partial melts from thickened continental crust? *Lithos* **2022**, *430–431*, 106885. [CrossRef]
86. Wang, M.; Wang, C.Y. Crystal Size Distributions and Trace Element Compositions of the Fluorapatite from the Bijigou Fe–Ti Oxide-Bearing Layered Intrusion, Central China: Insights for the Expulsion Processes of Interstitial Liquid from Crystal Mush. *J. Petrol.* **2020**, *61*, egaa069. [CrossRef]

87. Wu, F.Y.; Arzamastsev, A.A.; Mitchell, R.H.; Li, Q.-L.; Sun, J.; Yang, Y.-H.; Wang, R.-C. Emplacement age and Sr–Nd isotopic compositions of the Afrikanda alkaline ultramafic complex, Kola Peninsula, Russia. *Chem. Geol.* **2013**, *353*, 210–229. [CrossRef]
88. Bühn, B.; Wall, F.; Le Bas, M.J. Rare-earth element systematics of carbonatitic fluorapatites, and their significance for carbonatite magma evolution. *Contrib. Mineral. Petrol.* **2001**, *141*, 572–591. [CrossRef]
89. Ye, H.M.; Li, X.H.; Lan, Z.W. Geochemical and Sr–Nd–Hf–O–C isotopic constraints on the origin of the Neoproterozoic Qieganbulake ultramafic–carbonatite complex from the Tarim Block, Northwest China. *Lithos* **2013**, *182–183*, 150–164. [CrossRef]
90. Weidendorfer, D.; Schmidt, M.W.; Mattsson, H.B. Fractional crystallization of Si-undersaturated alkaline magmas leading to unmixing of carbonatites on Brava Island (Cape Verde) and a general model of carbonatite genesis in alkaline magma suites. *Contrib. Mineral. Petrol.* **2016**, *171*, 43. [CrossRef]
91. Velázquez Santana, L.C.; McLeod, C.L.; Blakemore, D.; Shaulis, B.; Hill, T. Bolivian hornblendite cumulates: Insights into the depths of Central Andean arc magmatic systems. *Lithos* **2020**, *370–371*, 105618. [CrossRef]
92. Yu, K.; Liu, Y.; Foley, S.F.; Zhu, Y.; Hu, Z.; Zong, K.; Chen, H. Transformation from oxidized to reduced alkaline magmas in the northern North China Craton. *Lithos* **2021**, *390–391*, 106104. [CrossRef]
93. Savard, J.J.; Mitchell, R.H. Petrology of ijolite series rocks from the Prairie Lake (Canada) and Fen (Norway) alkaline rock-carbonatite complexes. *Lithos* **2021**, *396–397*, 106188. [CrossRef]
94. Cheng, Z.; Zhang, Z.; Santosh, M.; Hou, T.; Zhang, D. Carbonate- and silicate-rich globules in the kimberlitic rocks of northwestern Tarim large igneous province, NW China: Evidence for carbonated mantle source. *J. Asian Earth Sci.* **2014**, *95*, 114–135. [CrossRef]
95. Bedini, R.M.; Bodinier, J.L. Distribution of incompatible trace elements between the constituents of spinel peridotite xenoliths: ICP-MS data from the East African Rift. *Geochim. Cosmochim. Acta* **1999**, *63*, 3883–3900. [CrossRef]
96. Zhou, Q.; Wu, F.Y.; Chu, Z.Y.; Ge, W.C. Isotopic compositions of mantle xenoliths and age of the lithospheric mantle in Yitong, Jilin Province. *Acta Petrol. Sin.* **2010**, *26*, 1241–1264, (In Chinese with English abstract).
97. González-Jiménez, J.M.; Villaseca, C.; Griffin, W.L.; O'rEilly, S.Y.; Belousova, E.; Ancochea, E.; Pearson, N.J. Significance of ancient sulfide PGE and Re–Os signatures in the mantle beneath Calatrava, central Spain. *Contrib. Mineral. Petrol.* **2014**, *168*, 1047. [CrossRef]
98. Bali, E.; Hidas, K.; Guðfinnsson, G.H.; Kovács, Z.; Török, K.; Román-Alpiste, M.J. Zircon and apatite-bearing pyroxene hornblendite mantle xenolith from Hungary, Carpathian–Pannonian region. *Lithos* **2018**, *316–317*, 19–32. [CrossRef]
99. Villaseca, C.; Belousova, E.A.; Barfod, D.N.; González-Jiménez, J.M. Dating metasomatic events in the lithospheric mantle beneath the Calatrava volcanic field (central Spain). *Lithosphere* **2019**, *11*, 192–208. [CrossRef]
100. Lu, J.; Tilhac, R.; Griffin, W.L.; Zheng, J.; Xiong, Q.; Oliveira, B.; O'REilly, S.Y. Lithospheric memory of subduction in mantle pyroxenite xenoliths from rift-related basalts. *Earth Planet. Sci. Lett.* **2020**, *544*, 116365. [CrossRef]
101. Howarth, G.H.; Costin, G.; Peel, C.J.; Qashani, Z. Petrology of an apatite-calcite segregationary kimberlite from the Andriesfontein Pipe, Kaapvaal Craton. *Lithos* **2023**, *458–459*, 107341. [CrossRef]
102. Yan, X.Y.; Yang, D.B.; Xu, W.L.; Quan, Y.-K.; Wang, A.-Q.; Hao, L.-R. Controls on apatite compositional variations in magmatic rocks: A case study of Mesozoic mafic-felsic intrusions from the eastern North China Craton. *J. Asian Earth Sci.* **2024**, *265*, 106111. [CrossRef]
103. Cao, Y.H.; Zhou, L.Q.; Brzozowski, M.; Wang, C.Y. Apatite and pyroxene as records of magmatic-hydrothermal processes and platinum-group mineral (PGM) formation in the Wengeqi mafic-ultramafic intrusion, Inner Mongolia, China: The role of oxidized, H_2O-rich magmas. *Chem. Geol.* **2024**, *670*, 122359. [CrossRef]
104. Kieffer, M.A.; Dare, S.A.S.; Namur, O. The use of trace elements in apatite to trace differentiation of a ferrobasaltic melt in the Sept-Iles Intrusive Suite, Quebec, Canada: Implications for provenance discrimination. *Geochim. Cosmochim. Acta* **2023**, *342*, 169–197. [CrossRef]
105. Spier, C.A.; Rosenbaum, G.; Milan, L.A.; Gomes, I.V.; Mulder, J. Devonian magmatism in the southern New England Orogen: Evidence from dolerite intrusions within the Gordonbrook Serpentinite (New South Wales, Australia). *Aust. J. Earth Sci.* **2026**. Epub ahead of print. [CrossRef]
106. Xing, C.M.; Wang, C.Y. Cathodoluminescence images and trace element compositions of fluorapatite from the Hongge layered intrusion in SW China: A record of prolonged crystallization and overprinted fluid metasomatism. *Am. Mineral.* **2017**, *102*, 1390–1401. [CrossRef]
107. Pedersen, J.M.; Ulrich, T.; Whitehouse, M.J.; Kent, A.J.R.; Tegner, C. The volatile and trace element composition of apatite in the Skaergaard intrusion, East Greenland. *Contrib. Mineral. Petrol.* **2021**, *176*, 102. [CrossRef]
108. Rakhimov, I.R. Apatite in pegmatoid gabbro (Khudolaz Differentiated Complex, Southern Urals): Electron-microprobe and LA-ICP-MS geochemical study in context of petrogenesis. *Minerals* **2024**, *14*, 1271. [CrossRef]
109. Hao, L.R.; Yang, D.B.; Wang, A.Q.; Quan, Y.; Yan, X.; Wang, S. Diverse apatite geochemical compositions in Early Paleozoic granitoids of the North Qinling Orogen, China: Insights into their petrogenesis and magma sources. *Lithos* **2024**, *480–481*, 107663. [CrossRef]

110. Gumsley, A.; Słaby, E.; Dey, S.; Chew, D.; Gmochowska, W.; Wudarska, A.; Gumsley, A.; Szopa, K.; Krzykawski, T.; Marciniak-Maliszewska, B.; et al. The evolution of Archean crust undisturbed for three billion years: A case study from the Bahalda Pluton, Singhbhum Craton, eastern India. *Precambrian Res.* **2024**, *406*, 107385. [CrossRef]
111. Li, Y.; Sun, C.Y.; Xu, W.L.; Zhang, X.M. In situ geochemistry of apatite: Petrogenetic and tectonic interpretations of Jurassic felsic magmatism in the Yanbian area, NE China. *Lithos* **2024**, *486–487*, 107764. [CrossRef]
112. Yang, F.; Wu, G.; Chen, G.Z.; Li, S.; Li, Y.; Zhang, T.; Chen, Y. Petrogenesis and implications for tin mineralization of the Beidashan granitic pluton, southern Great Xing'an Range, NE China: Constraints from whole-rock and accessory mineral geochemistry. *J. Asian Earth Sci.* **2024**, *259*, 105883. [CrossRef]
113. Tang, P.; Tang, J.X.; Wang, L.Q.; Lin, B.; Li, F.; Qi, J.; Wang, M.; Xiong, Y.; Xie, J.; Tao, G. Apatite and zircon geochemistry deciphers difference in the nature of ore-forming magma in the Bangpu porphyry Mo–Cu deposit, Tibet. *J. Asian Earth Sci.* **2024**, *264*, 106049. [CrossRef]
114. Bruand, E.; Storey, C.; Fowler, M. Accessory mineral chemistry of high Ba–Sr granites from northern Scotland: Constraints on petrogenesis and records of whole-rock signature. *J. Petrol.* **2014**, *55*, 1619–1651. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.