

Probabilistic three-dimensional mineralogical characterisation of drill core using μ XRF and XCT: application to the Nautanen North IOCG deposit

Milena Nasretdinova

`milena.nasretdinova@adelaide.edu.au`

University of Adelaide <https://orcid.org/0009-0009-5447-028X>

Francisco Cuadra-Amaro

University of Tasmania

Sampriti Basak

Boliden Mineral AB

Tobias Hermansson

Boliden Mineral AB

Peter Dowd

University of Adelaide

Chaoshui Xu

University of Adelaide

Research Article

Keywords: μ XRF, XCT, Mineralogical Characterisation, Nautanen North, IOCG, Geometallurgy

Posted Date: September 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11189416/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Probabilistic three-dimensional mineralogical characterisation of drill core using μ XRF and XCT: application to the Nautanen North IOCG deposit

Milena Nasretdinova^{1,2*}, Francisco Cuadra-Amaro³, Sampriti Basak⁴, Tobias Hermansson⁴, Peter Dowd^{1,2}, Chaoshui Xu^{1,2}

1. Australian Research Council Industrial Transformation Training Centre for Integrated Operations for Complex Resources, Adelaide University, Adelaide 5005, Australia
2. College of Engineering and Information Technology, Adelaide University, Adelaide 5005, Australia
3. Centre for Ore Deposit and Earth Sciences, University of Tasmania, Hobart 7000, Australia
4. Exploration R&D Department, Boliden Mineral AB, Boliden, Sweden

**Correspondence: milena.nasretdinova@adelaide.edu.au*

Abstract

Accurate mineralogical characterisation of drill cores is important for quantifying variability and supporting geometallurgical interpretations. Conventional assays provide bulk elemental concentrations but do not preserve mineral distributions, whereas individual non-destructive methods provide only partial mineralogical information. In this study, a probabilistic integrated framework was developed for three-dimensional mineralogical mapping, using four drill-core samples from Nautanen North, northern Sweden, as a case study. Micro-X-ray fluorescence (μ XRF) maps classified with the Advanced Mineral Identification and Characterization System (AMICS) were co-registered with corresponding X-ray computed tomography (XCT) surfaces to derive mineralogical labels, which were used as two-dimensional supervision for a deep learning segmentation model, with XCT responses as the input. Detailed mineralogy was consolidated into four XCT-supported classes: Matrix silicates, Quartz, Garnet and Fe-bearing sulphides and oxides. Generalisation was evaluated using nested leave-one-sample-out validation, and predicted class probabilities were calibrated to better reflect prediction reliability. The selected support-aware calibrated ensemble achieved a mean accuracy of 0.899 and a mean macro Dice coefficient of 0.583. Matrix silicates transferred most consistently, whereas the Fe-bearing sulphides and oxides class remained difficult to discriminate because of its heterogeneous and overlapping XCT responses. The calibrated ensemble was subsequently applied to predict the mineralogical classes within a half-core XCT volume. Predictions remained spatially coherent across the evaluated XCT volume, with high confidence, low normalised predictive entropy and high spatial agreement. The framework demonstrates the potential to extend targeted surface mineralogical observations into three-dimensional predictions of mineralogical classes, accompanied by class probabilities and uncertainty fields, that can support geometallurgical domain definition and the selection of suitable samples for additional mineralogical analysis or metallurgical testing.

Keywords: μ XRF; XCT; Mineralogical Characterisation; Nautanen North; IOCG; Geometallurgy

1. Introduction

Geometallurgy provides an integrated framework for linking geological, mineralogical and textural variability within an orebody to downstream processing behaviour, economic value and operational risk (Dominy et al., 2018). Its implementation commonly relies on spatial models that use measurable ore properties to predict the processing characteristics of individual mining blocks (Williams and Richardson, 2004; Lamberg, 2011; Lund et al., 2013). By integrating geological domains with mineralogical, textural and metallurgical information, geometallurgy can reduce uncertainty, improve resource utilisation and support robust mine planning and operation (van den Boogaart and Tolosana-Delgado, 2018; Lishchuk et al., 2020). However, conventional testwork is costly and time-consuming, while sample compositing and homogenisation can obscure variability relevant to mineral processing. Geometallurgical characterisation therefore requires objective, reproducible and minimally destructive measurements that can be acquired across large sample populations and spatially integrated along drill holes and within orebody models.

Mineralogical variability is particularly important because an element present at a given bulk grade may be distributed among different minerals, alteration assemblages and textural settings (Ehrig et al., 2014). Conventional geochemical assays quantify bulk elemental composition but do not directly identify mineral hosts or preserve their spatial relationships. Element-to-mineral conversion and quantitative X-ray diffraction can improve bulk mineralogical estimates but generally do not retain the spatial arrangement of individual phases (Parian et al., 2015). Mineral abundance, grain sizes, associations, connectivity and textural overprinting can influence ore breakage, liberation, separation and fluid-mineral interactions during processing (Pérez-Barnuevo et al., 2018; Voigt et al., 2019; Merrill-Cifuentes et al., 2022; Butcher et al., 2023; Pereira et al., 2023). Spatially resolved mineralogical measurements are therefore needed to complement conventional assays and preserve heterogeneity that may otherwise be obscured by sample homogenisation (Evans et al., 2015; Warlo et al., 2021). Quantitative texture analysis can further support geometallurgical characterisation by describing the spatial arrangement and relationships of mineralogical features (Nguyen, 2013).

Geological drill-core logging provides routine descriptions of lithology, mineralogy, alteration and texture, but the resulting information is generally qualitative, labour-intensive to acquire and influenced by the experience and subjective interpretation of the responsible geologist (Simán et al., 2021; Houshmand et al., 2022). Automated mineralogical techniques based on scanning electron microscopy coupled with energy-dispersive X-ray spectroscopy (SEM-EDS), including Quantitative Evaluation of Minerals by Scanning Electron Microscopy (QEMSCAN), Mineral Liberation Analysis (MLA), and the TESCAN Integrated Mineral Analyzer (TIMA), can quantify mineral compositions, associations, grain-size distributions and liberation characteristics (Gottlieb et al., 2000; Gu, 2003; Hrstka et al., 2018; Schulz et al., 2020). However, these methods generally require polished surfaces, and analytical cost and acquisition time increase with spatial resolution, hence restricting the analysed area and the number of samples. Two-dimensional section analysis is also subject to stereological effects and may poorly

represent preferentially oriented, clustered or rare minerals, particularly in rocks with complex mineralogy and overprinting relationships (Lätti and Adair, 2001).

Micro-X-ray fluorescence (μ XRF) scanning provides an alternative approach for characterising relatively large drill-core surfaces with minimal sample preparation (Harraden et al., 2025). Modern μ XRF systems use focused X-ray beams to collect spatially resolved elemental spectra from heterogeneous and irregularly shaped samples, commonly without the coating and polishing required for electron-based analyses (Flude et al., 2017). Depending on the instrument configuration and analytical conditions, μ XRF can map major and minor elements at spatial resolutions ranging from tens to hundreds of micrometres (μm) and can analyse larger surfaces than conventional SEM-based automated mineralogy (Menzies et al., 2019; Menzies et al., 2022; Viana et al., 2025). The resulting elemental and X-ray intensity data can be converted into categorical mineral maps using systems such as the Advanced Mineral Identification and Characterization System (AMICS). These maps preserve the spatial distribution of mineralogical features and may reveal variations in grain size, mineral orientation, alteration bands, veins and mineral associations. Previous studies have reported agreement between μ XRF-derived chemical data and conventional assays and between AMICS mineral classifications and phases identified by quantitative X-ray diffraction (XRD) (Månbro and Parian, 2023). Supervised and unsupervised machine learning methods have also been investigated for automating the interpretation of μ XRF elemental and mineral maps (Harraden et al., 2025).

However, μ XRF–AMICS mineralogy should not be regarded as error-free ground truth. Classification is based primarily on elemental spectra rather than crystallographic structure, making chemically similar minerals and polymorphs difficult to distinguish. X-ray penetration below the analysed surface may also produce mixed spectra where adjacent minerals contribute to a single measurement. Mineral libraries may consequently require geological validation, manual refinement and consolidation of unknown or mixed measurements into broader classes. Furthermore, in coarse-grained, patchy or anisotropic rocks, mineral proportions measured on one surface may differ from those of the surrounding volume. Thus, μ XRF–AMICS provides spatially detailed mineralogical supervision, but retains uncertainties associated with spectral ambiguity, interaction volume, classification-library construction and the representativeness of the analysed surface relative to the surrounding volume.

X-ray computed tomography (XCT) provides complementary information by imaging the three-dimensional internal structure of a specimen non-destructively (Kyle and Ketcham, 2015). Radiographic projections acquired from multiple angles are reconstructed into a volumetric grid of voxels, each assigned a value related to the X-ray linear attenuation of the material (Cnudde and Boone, 2013). Attenuation depends principally on density, elemental composition and incident X-ray energy, allowing minerals or mineral assemblages to be differentiated where their responses are sufficiently distinct (Ketcham and Carlson, 2001; Wang and Miller, 2020). XCT can be applied across a broad range of specimen sizes and resolutions, enabling complete drill-core sections, hand specimens and smaller subsamples to be examined while preserving

their spatial relationships (Baker et al., 2012; Warlo et al., 2021). It has consequently been investigated for three-dimensional grain-size measurement, texture quantification, structural analysis and geometallurgical classification (Evans et al., 2015; Becker et al., 2016; Guntoro et al., 2020; Voigt et al., 2020; Nasretdinova et al., 2026).

Despite its volumetric advantages, XCT does not directly identify minerals (Evans et al., 2015; Warlo et al., 2021). Beam hardening, streaks, cone-beam effects, detector-response variation, image blurring and partial-volume effects can broaden the attenuation distribution of individual phases (Kyle and Ketcham, 2015; Hanna and Ketcham, 2017; Warlo et al., 2021). Differences in chemical composition within the same mineral phase, including variations within individual grains, may cause further variability in XCT attenuation. Attenuation also varies with X-ray energy, meaning that settings that optimise contrast between one pair of minerals may reduce contrast between others (Godel, 2013; Reyes et al., 2017). Mineral-specific attenuation calculations can assist in evaluating theoretical separability before scanning, but cannot account fully for reconstruction artefacts, compositional variation or partial-volume mixing in measured data (Bam et al., 2020). Fixed greyscale or attenuation thresholds may be effective where mineral phases have distinct attenuation values and grains are large relative to voxel size, but become unreliable in ores with fine mineral grains, heterogeneous mineral compositions or strongly intergrown mineral phases (Evans et al., 2015; Wang and Miller, 2020). Edge-detection and watershed methods can improve grain separation where boundaries are distinct but cannot resolve minerals with intrinsically overlapping attenuation (Godel, 2013; Guntoro et al., 2019). Machine-learning methods can incorporate contextual information and improve segmentation of complex datasets, although they require representative training data, computational resources and rigorous validation (Wang and Miller, 2020; Kuva et al., 2025).

Surface mineralogical imaging and XCT provide complementary information for multimodal ore characterisation. Automated mineralogy or μ XRF-AMICS can provide mineralogical constraints on an exposed surface, while XCT provides the surrounding three-dimensional structural and attenuation context. XCT has also been combined with X-ray diffraction tomography, X-ray fluorescence tomography and related methods, although many applications remain restricted to small specimens (Wildenschild and Sheppard, 2013; Laforce et al., 2017). At drill-core scale, integrated XCT-XRF systems and correlative XCT-automated-mineralogy workflows offer a more practical route for preserving three-dimensional geological context while adding compositional constraints (Bergqvist et al., 2019; Butcher, 2020). However, the techniques measure different physical properties and operate at different resolutions and interaction volumes; consequently, even after careful co-registration, corresponding pixels may not represent exactly equivalent physical sampling volumes or quantitative mineral proportions (Warlo et al., 2021).

The geological value of multimodal imaging was demonstrated by Sayab et al. (2025), who combined three-dimensional XCT with two-dimensional μ XRF elemental and mineral imaging of oriented drill cores from the Rajapalot Au-Co system in Finland. Building on previous multidimensional imaging studies (Sayab et al., 2015, 2016; Butcher,

2020; Menzies and Boehm, 2021), XCT was used to visualise sulphide distributions relative to folds, foliation and alteration bands, while μ XRF provided mineralogical and chemical constraints at the exposed surface. Integration with structural data enabled interpretation of mineralising-fluid pathways and cobalt enrichment, demonstrating the value of spatially linking two-dimensional mineralogical observations with three-dimensional XCT volumetric information (Sayab et al., 2025). Warlo et al. (2021) similarly simplified detailed automated-mineralogy maps into broader XCT-resolvable classes, including quartz-albite, silicate-carbonate and sulphide-oxide groups. Large molybdenite and scheelite grains could be traced volumetrically, whereas finer phases were less reliably resolved and sulphides and oxides were susceptible to over-segmentation. Together, these studies highlight both the value of integrating two-dimensional mineralogical information with three-dimensional XCT and the need to simplify detailed mineralogy into classes that can be resolved reliably from XCT.

Image analysis and machine learning provide a means of progressing from visual correlation to automated classification. Viana et al. (2025) combined μ XRF-derived minerals and X-ray intensity maps with image patches, grey-level co-occurrence descriptors, dimensionality reduction and unsupervised clustering to classify drill-core textures. Combining features from both map types improved the separation of major textural groups, demonstrating that compositional and intensity-based information are complementary. However, simple clustering may prioritise pixel-intensity contrasts and struggle with mixed or transitional textures, supporting the use of deep feature extraction and soft or probabilistic classification for heterogeneous geological imagery (Viana et al., 2025).

Deep convolutional neural networks are well suited to semantic image segmentation because they learn hierarchical features from local intensities and wider spatial context (Roy et al., 2018; Wang et al., 2021). The U-Net architecture combines an encoding pathway for feature extraction with a decoding pathway that restores spatial details (Ronneberger et al., 2015) and has subsequently been applied to porous media, digital rocks and geological image segmentation (Wang et al., 2021; Phan et al., 2021). This ability to combine intensity, spatial form, texture and neighbourhood relationships makes deep segmentation more suitable than fixed thresholding where mineral XCT responses overlap. Kuva et al. (2025) demonstrated the closest published analogue by training a modified U-Net on a co-registered two-dimensional μ XRF–AMICS mineral map and XCT slice, then applying the model to successive XCT slices to generate three-dimensional mineral predictions. The architecture incorporated multiscale features and attention mechanisms and used cross-validation to reduce overfitting under limited ground-truth availability (Ronneberger et al., 2015; Roy et al., 2018; Kuva et al., 2025). The model performed best for abundant, spatially coherent classes, while rare or finely distributed minerals remained difficult to resolve because of limited training support, irregular geometry, overlapping XCT responses and registration sensitivity. These results demonstrate the feasibility of transferring two-dimensional mineralogical supervision into three-dimensional XCT predictions while emphasising the importance of registration quality, class balance and uncertainty-aware interpretation (Kuva et al., 2025).

The Nautanen North deposit in northern Sweden provides a challenging test case for three-dimensional XCT-based mineralogical characterisation because of its strong hydrothermal alteration, deformation and complex mineral associations (Drejing-Carroll et al., 2023). Repeated alteration and mineralisation produced heterogeneous assemblages of silicates, garnet, magnetite and sulphides, with minerals occurring in multiple textural associations and locally as intergrowths or inclusions (Waara, 2015; Melfou et al., 2023). Consequently, individual minerals cannot be assumed to occupy unique, non-overlapping XCT attenuation ranges, and detailed mineralogy may need to be represented by broader XCT-supportable classes where separation is unreliable (Evans et al., 2015; Warlo et al., 2021).

Previous studies have therefore addressed different components of the problem but leave some important methodological gaps. Correlative automated mineralogy and XCT have demonstrated the value of three-dimensional context but also the difficulty of quantitative segmentation and cross-scale comparison (Butcher, 2020; Warlo et al., 2021). Combined XCT and μ XRF imaging has supported structural and ore-genetic interpretation without automatically propagating mineral labels throughout the volume (Sayab et al., 2025). Automated analysis of μ XRF mineral and intensity maps has improved objective texture classification but remains two-dimensional and highlights the limitations of hard clustering for transitional domains (Viana et al., 2025). Continuous XCT workflows have demonstrated scalable, depth-resolved and geometallurgically relevant texture characterisation but do not directly identify minerals (Nasretdinova et al., 2026). Although deep-learning transfer from two-dimensional μ XRF-AMICS mineralogy to three-dimensional XCT has been demonstrated (Kuva et al., 2025), the requirements for robust applications to heterogeneous drill core have not yet been resolved with respect to registration quality, class imbalance, cross-sample generalisation, probability calibration and uncertainty, and the spatial consistency and interpretability of volumetric predictions.

To address these gaps, this study develops and evaluates a framework for probabilistic three-dimensional mineralogical characterisation of drill core using μ XRF-AMICS and XCT data, with the Nautanen North deposit as a case study. Two-dimensional AMICS mineral maps are co-registered with corresponding XCT surface images and used as mineralogical supervision for a U-Net segmentation model in which XCT provides the input. The study aims to: (i) establish a reproducible co-registration procedure; (ii) define XCT-supportable mineralogical classes that balance mineralogical details, geometallurgical relevance, training support and XCT separability; (iii) assess class-specific performance and cross-sample generalisation; (iv) calibrate predicted class probabilities and quantify predictive uncertainties; (v) evaluate support-aware ensemble strategies and identify a transferable modelling configuration; and (vi) apply the validated framework to volumetric XCT data and assess the spatial consistency of the resulting three-dimensional predictions of mineralogical classes. Rather than assuming that XCT uniquely identifies every mineral voxel, the study evaluates which classes can be mapped consistently, where grouped mineral phases are more defensible, and how predictive uncertainty should constrain geological and geometallurgical interpretations.

2. Geological setting

The Nautanen North deposit is located in the Gällivare-Malmberget area of northern Norrbotten, Sweden, within the northern part of the Fennoscandian Shield (Fig. 1). The district forms a major Paleoproterozoic metallogenic province containing iron oxide-apatite, porphyry-style Cu–Au and iron oxide copper-gold (IOCG) mineral systems, including Kiruna, Malmberget, Aitik and Nautanen North (Wanhainen, 2005; Bergman and Weihed, 2020; Drejing-Carroll et al., 2023). The regional bedrock comprises Archaean basement overlain by Paleoproterozoic volcanic, volcanoclastic and sedimentary successions that were subsequently intruded, metamorphosed and repeatedly deformed during the Svecokarelian orogenic evolution of northern Sweden (Bergman and Weihed, 2020; Drejing-Carroll et al., 2023). This geological history produced strongly modified and compositionally heterogeneous rocks in which primary lithological variability is superimposed by multiple generations of hydrothermal alteration and mineralisation.

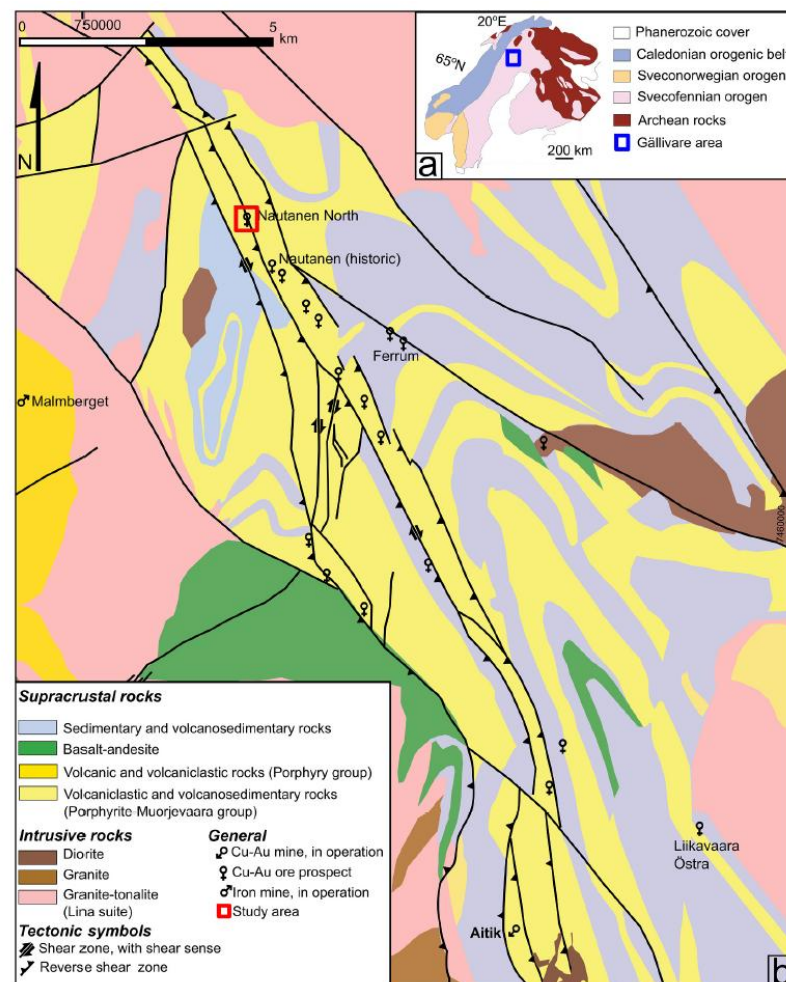


Figure 1. Geological setting of the Nautanen North deposit, northern Sweden: (a) Simplified geological map of the Fennoscandian Shield; (b) Geological map of the Gällivare area. Reproduced from Melfou et al. (2023)

Nautanen North is located within the Nautanen deformation zone, a regional NNW–SSE-striking composite shear system that hosts several Cu–Au occurrences in the Gällivare area. Rocks within the deformation zone are characterised by penetrative foliation,

folding, shearing and pervasive hydrothermal alteration, which commonly obscure primary volcanic and volcanoclastic textures. The deposit is localised between bounding shear zones, with higher-grade Cu mineralisation concentrated in mineralised shear bands, veins, stockworks and sulphide-rich breccias. In the northern part of the deposit, mineralised breccias form a vertically stacked, relay-like system between steeply dipping structures, illustrating the strong structural control on hydrothermal fluid flow and sulphide precipitation (Bauer et al., 2022; Drejing-Carroll et al., 2023). These relationships generate pronounced spatial variations in mineral abundance, grain geometry and mineral associations over relatively short distances, all of which are relevant to mineralogical and geometallurgical characterisation at drill-core scale.

The principal host rocks comprise amphibole-biotite gneiss, biotite gneiss, banded biotite gneiss, and sericite-garnet gneiss and schist. Amphibole-biotite and biotite gneisses contain variable proportions of biotite, quartz and plagioclase, whereas banded biotite gneiss is characterised by alternating quartz-feldspar-rich and coarse biotite-rich bands, locally containing garnet and magnetite. Sericite-garnet gneiss and schist occur towards the central and western parts of the system and form the dominant host to much of the higher-grade mineralisation. These rocks contain sericite, garnet, quartz, plagioclase, K-feldspar, biotite and chlorite, with garnet occurring as abundant porphyroblasts ranging from sub-millimetre scale to more than 10 cm in diameter. The host-rock sequence is interpreted predominantly as hydrothermally modified mafic to intermediate igneous protoliths rather than as a simple primary stratigraphic succession (Drejing-Carroll et al., 2023). Consequently, minerals such as plagioclase, K-feldspar, biotite, quartz and garnet record both primary rock composition and subsequent alteration, providing mineralogical contrasts that are relevant to the spatial characterisation of the ore system.

Hydrothermal alteration developed through several partly overlapping stages, progressing from early sodic alteration through Na-Ca-Fe, high-temperature (HT) Ca-Fe, HT Ca-K-Fe and HT K-Fe assemblages to lower-temperature (LT) K-Fe alteration and later veining. Repeated overprinting and transitional assemblages are common. Early sodic alteration substantially modified primary feldspars, whereas the subsequent Ca-Fe and K-Fe alteration stages introduced or redistributed minerals including biotite, K-feldspar, garnet, apatite and Fe-rich oxide phases. HT Ca-Fe alteration is characterised by amphibole + magnetite $\pm$ garnet, with accessory apatite and locally chalcopyrite and pyrite, whereas HT Ca-K-Fe and HT K-Fe assemblages contain variable combinations of biotite, K-feldspar, magnetite and garnet. The proximal LT K-Fe assemblage is dominated by sericite + K-feldspar + garnet, with variable biotite and Fe sulphides and accessory calcite and apatite (Drejing-Carroll et al., 2023). Magnetite formed during more than one alteration stage: an early magnetite-apatite association is interpreted as a possible distal expression of the regional iron oxide-apatite system, whereas later Fe-oxide mineralisation overlapped with the development of Cu- and Fe-sulphide mineralisation (Drejing-Carroll et al., 2023).

Copper is hosted predominantly in chalcopyrite and is associated with pyrite, pyrrhotite and magnetite, with minor molybdenite and bornite. Chalcopyrite occurs as disseminated grains, foliation-parallel aggregates, vein and shear fillings, stockworks

and within sulphide-rich breccia matrices. Mineralisation is spatially associated particularly with HT and LT K-Fe alteration, and textural relationships indicate progressive replacement of earlier magnetite and Fe sulphides by chalcopyrite in more strongly mineralised domains (Drejing-Carroll et al., 2023). Apatite occurs in several alteration assemblages and may be associated with magnetite, amphibole, K-feldspar, chalcopyrite and sericite, with local replacement by epidote, allanite or chlorite (Melfou et al., 2023). Calcite is less abundant but occurs as an accessory phase within LT K-Fe alteration and locally in structurally controlled veins. Quartz occurs both as a major component of host-rock assemblages and within later veins, further contributing to the mineralogical and textural heterogeneity of the system.

Garnet is particularly relevant to the present study because it occurs across several alteration and structural settings and exhibits substantial variability in grain size and internal texture. Garnet at Nautanen North is dominated by spessartine-almandine compositions and may occur as syntectonic porphyroblasts within Ca-Fe alteration as well as abundantly within sericite-garnet gneiss and LT K-Fe-altered rocks (Waara, 2015; Drejing-Carroll et al., 2023). Individual porphyroblasts range from approximately millimetre scale to centimetre scale and may contain inclusions of quartz, magnetite, tourmaline and sulphides. Garnet-rich compositional bands can also occur together with pyrite and chalcopyrite. Its variable abundance, coarse and locally inclusion-rich texture, and association with altered and mineralised domains therefore make garnet a useful mineralogical marker for evaluating whether XCT-based classification can capture both mineral response and spatial context rather than relying solely on isolated attenuation values.

From a geometallurgical perspective, the Nautanen North system is therefore characterised by spatially variable mixtures of gangue, alteration and ore-related minerals rather than discrete mineral domains with simple one-to-one relationships between mineral identity and rock type. The mineral phases represented in this study span host-rock constituents, alteration products and mineralisation-related phases. Their proportions, grain dimensions, intergrowths and spatial associations change with host rock, alteration facies, deformation and proximity to mineralisation. This heterogeneity provides a geologically and geometallurgically relevant test of whether two-dimensional mineralogical supervision can be transferred to XCT for probabilistic three-dimensional characterisation. Importantly, the geological relationships do not imply that each mineral occupies a unique XCT attenuation range; rather, compositional variability, intergrowth, replacement and mineral inclusions are expected to generate overlapping responses. The geological setting therefore supports an uncertainty-aware and context-dependent modelling approach rather than deterministic assignment of individual XCT intensity intervals to mineral identities.

3. Materials and methods

3.1 μ XRF data acquisition, AMICS mineral mapping and mineral-label preparation

The exposed cut surfaces of the drill-core samples were analysed by μ XRF using a Bruker M4 Tornado Plus system. The instrument uses a focused rhodium X-ray source and

silicon drift detectors to acquire spatially resolved X-ray spectra across heterogeneous sample surfaces. Measurements were acquired at a nominal spatial resolution of 100 μm using an anode current of 600 μA . Comparable rock and drill-core studies have demonstrated the applicability of μXRF to high-resolution elemental and mineralogical characterisation of relatively large rock surfaces (Menzies et al., 2022; Månbro and Parian, 2023; Viana et al., 2025; Kuva et al., 2025).

The measured μXRF spectrum at each 100 μm grid cell reflects the elemental composition within the corresponding X-ray interaction volume and therefore provides spatial information on compositional variability across the exposed surface. Mineral classification was subsequently performed using AMICS, in which measured μXRF responses were compared with a mineral reference database to generate spatially resolved mineral classifications. As part of the classification workflow, histogram-based assessment was implemented for each mineral class to examine the distribution of the associated μXRF responses and their consistency with the expected geochemical characteristics.

The resulting mineral maps provided categorical, spatially resolved mineral labels while preserving mineral distributions and associated textural relationships across the exposed drill-core surfaces. Because μXRF measurements represent a finite interaction volume, responses acquired close to mineral boundaries may contain contributions from more than one phase. Mineral classification is also based principally on elemental composition rather than crystallographic structure, which can limit discrimination between chemically similar phases. White mica, including muscovite and sericite, was not resolved as a separate AMICS phase. Because μXRF does not detect hydrogen and chemically similar K-bearing silicates can be difficult to distinguish using elemental spectra alone, some white mica may be represented within the phases grouped as Matrix silicates.

The identified mineral phases comprised plagioclase feldspar, biotite, K-feldspar, garnet, quartz, chalcopyrite, pyrite, apatite, calcite and Fe oxides, together with a small proportion of unknown classifications. The specific mineralogical identity of the Fe-oxide category was not independently verified by optical microscopy or SEM and was therefore retained as the broader category of Fe oxides. AMICS-derived mineral abundances were reported as weight percentage (Wt %) and surface area percentage (Area %). Surface area percentage was used to quantify the relative representation of each mineral phase in the original maps and to provide a preliminary indication of the amount of labelled data available for subsequent modelling (Fig. 2). Because these values describe the original μXRF –AMICS maps, the exact class proportions available for model training were determined after class harmonisation, resolution matching, masking and purity filtering.

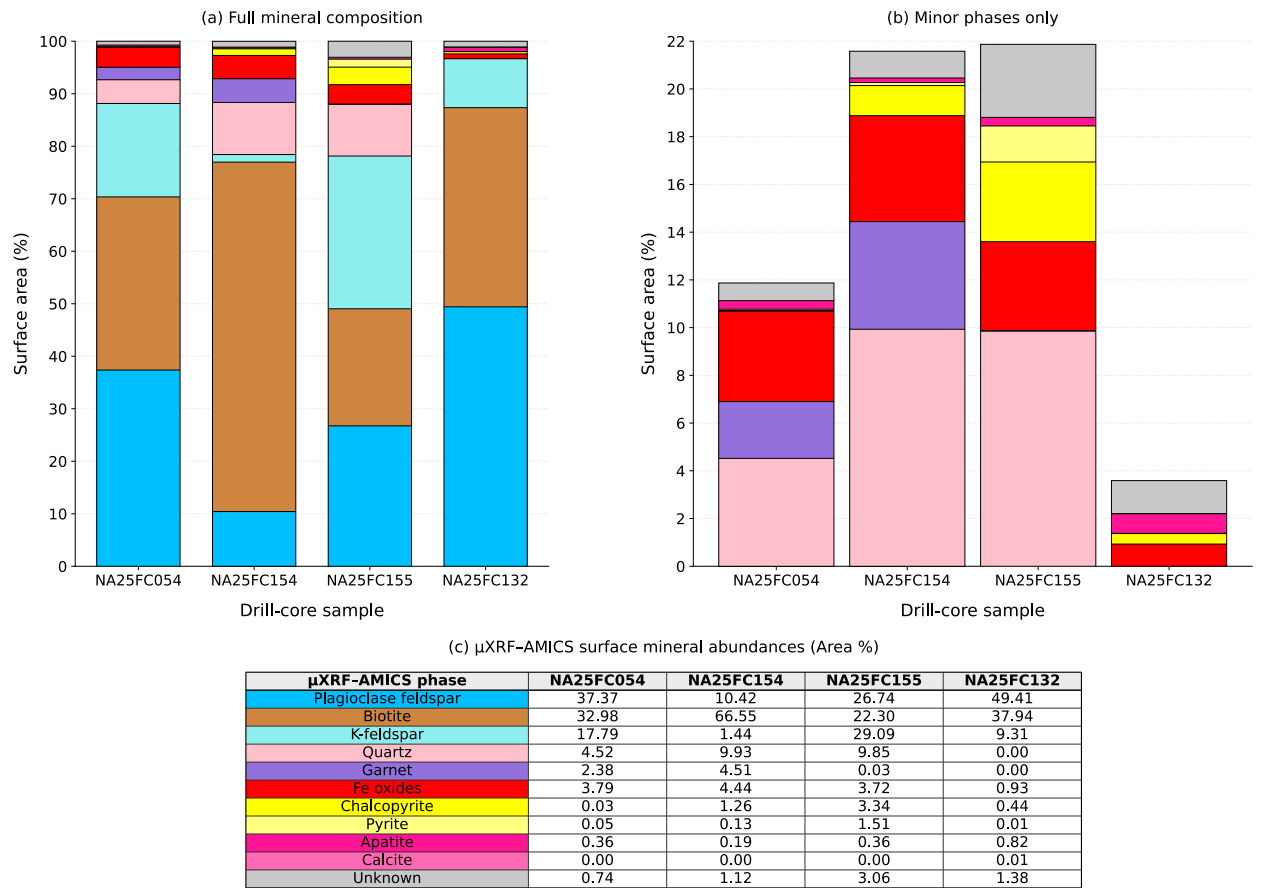


Figure 2. μ XRF-AMICS-derived surface mineral abundances for the four drill-core samples: (a) Complete mineral compositions; (b) Expanded view of minor mineral phases; (c) Tabulated surface area percentages by mineral phase and sample

The four mapped surfaces exhibited substantial differences in mineralogical composition and consequently in the amount of supervision available for individual mineral phases. Plagioclase, biotite and K-feldspar were the dominant mineral phases across the mapped surfaces, although their relative proportions varied markedly among samples. Quartz accounted for 4.52%, 9.93% and 9.85% of the mapped surfaces of NA25FC054, NA25FC154 and NA25FC155, respectively, but was not identified on the mapped surface of NA25FC132. Garnet accounted for 2.38%, 4.51% and 0.03% in the same three samples, respectively, and was also not identified in NA25FC132. Fe oxides, chalcopyrite and pyrite were present on all four mapped surfaces but occurred in lower and variable proportions, accounting for 0.93-4.44%, 0.03-3.34% and 0.01-1.51%, respectively.

Apatite and calcite provided insufficient labelled support for reliable class-specific modelling. Apatite occupied only 0.19-0.82% of the mapped surfaces, whereas calcite was effectively absent, occupying 0-0.01%. These phases were therefore excluded from model supervision. Their exclusion reflected the limited number of labelled grid cells in this dataset and did not imply a lack of geological significance within the Nautanen North mineral system.

Before resolution matching, the retained AMICS mineral labels were harmonised into four XCT-supportable classes (Table 1) based on three considerations: available labelled-grid support, expected XCT separability, and preservation of geologically and

geometallurgically meaningful mineralogical contrasts. The final classes comprised Matrix silicates (plagioclase feldspar, biotite and K-feldspar), Quartz, Garnet, and Fe-bearing sulphides and oxides (Fe oxides, chalcopyrite and pyrite). This harmonisation reduced fragmentation of the mineralogical supervision and avoided unsupported phase-specific separation where XCT responses were expected to overlap. The four classes nevertheless remained substantially unequal in abundance.

Table 1. Harmonisation of μ XRF–AMICS mineral phases into XCT-supportable model classes and the rationale for class treatment

Model class	AMICS mineral phases	Geological, geometallurgical and modelling rationale
Matrix silicates	Plagioclase + Biotite + K-feldspar	Dominant host-rock and alteration-related silicates. Combined to avoid unsupported mineral-specific separation where XCT responses overlap and to represent the dominant silicate matrix relevant to rock composition and texture. This class may also include white mica that could not be resolved separately using the available μ XRF–AMICS classification.
Quartz	Quartz	Important host-rock, gangue and vein-forming mineral. Retained separately to test whether XCT attenuation and spatial context can distinguish quartz from the surrounding silicate matrix.
Garnet	Garnet	Geologically significant alteration mineral with variable abundance, locally coarse grain size and inclusion-rich textures. Retained separately as a geologically and geometallurgically relevant XCT target.
Fe-bearing sulphides and oxides	Fe oxides + Chalcopyrite + Pyrite	Alteration- and mineralisation-related Fe-bearing sulphide and oxide phases. Combined because their representation is uneven among samples and their XCT attenuation responses may overlap, making reliable phase-specific discrimination difficult at the available resolution. Interpreted as a grouped response rather than unique identification of individual oxide or sulphide minerals.
Excluded	Apatite + Calcite	Geologically meaningful accessory phases associated with alteration and veining but insufficiently represented in the available μ XRF–AMICS surfaces for reliable class-specific model development (0.19-0.82 Area % for apatite, and 0-0.01 Area % for calcite).
Ignored	Pores + Shadows + Unknown	Do not represent sufficiently reliable mineralogical labels for supervised learning. These categories were therefore masked and did not contribute to model training, retained-class support calculations or performance evaluation.

The original mineral maps had a nominal spatial resolution of 100 μ m. To establish categorical mineral-label support at the nominal XCT resolution of 200 μ m, the maps were converted using non-interpolative majority voting. For each resulting 200 μ m grid cell, the harmonised class represented by the greatest number of corresponding 100 μ m AMICS grid cells was retained, and the label purity was calculated as:

$$purity_i = \frac{\max_m(n_{im})}{4} \quad (1)$$

where n_{im} is the number of original 100 μ m AMICS grid cells assigned to harmonised mineral class m within block i at the resolution of 200 μ m, which comprises four 100 μ m AMICS grid cells arranged as a 2×2 block. Blocks with a purity below 0.75 were excluded from model supervision. AMICS grid cells assigned to the AMICS categories of Pores, Shadows or Unknown, together with cells outside the sample region, were also excluded. This procedure reduced categorical ambiguity at mineral boundaries while avoiding interpolation between discrete mineral labels. The class proportions available for model

training were determined from the retained 200 μm block labels and could therefore differ slightly from the original surface abundances because of majority voting and purity filtering.

3.2 XCT data acquisition

X-ray computed tomography data were acquired from the drill core using the Orexplore GeoCore X10 system (Orexplore, 2021). The scans were acquired at 120 kVp, and the reconstructed data used in this study had a nominal isotropic voxel size of 200 μm . The reconstructed volumes provide a three-dimensional representation of X-ray attenuation of the scanned drill core. Each XCT voxel records an attenuation-related response controlled principally by the density and effective atomic number of the material at the applied X-ray energy (Kyle and Ketcham, 2015). Mineral phases with sufficient contrast in physical and chemical properties may therefore exhibit different reconstructed XCT responses. However, attenuation distributions can overlap substantially among minerals, particularly where phases have similar attenuation characteristics or occur as fine intergrowths and inclusions (Evans et al., 2015; Warlo et al., 2021). The visibility of mineralogical features in the reconstructed volume consequently depends on both the attenuation contrast between constituent phases and their dimensions relative to the 200 μm voxel size. The XCT attenuation-related response was therefore not interpreted as a direct or unique indicator of mineral identity.

At the 200 μm voxel scale, voxels intersecting mineral boundaries may contain contributions from more than one mineral phase, producing intermediate attenuation values; this is known as the partial-volume effect. Reconstruction-related artefacts may separately distort the reconstructed attenuation field, particularly near strongly attenuating phases (Joseph and Spital, 1982; Kyle and Ketcham, 2015; Willemink and Noël, 2019). These mixed or distorted attenuation values cannot always be assigned uniquely to a single constituent mineral. Accordingly, the continuous attenuation field was retained throughout the subsequent analysis rather than segmented into predefined mineral classes or interpreted using fixed, non-overlapping attenuation-response thresholds.

This treatment of the XCT data was central to the modelling strategy. The XCT data provide three-dimensional attenuation and spatial-context information, whereas the mineralogical labels represent only the corresponding two-dimensional μXRF –AMICS-mapped surfaces. The subsequent framework therefore learns probabilistic relationships between the co-registered datasets. This allows attenuation response and spatial context to contribute jointly to mineral-class prediction without assuming unique attenuation-response ranges for individual mineral phases.

Paired μXRF –AMICS and XCT datasets were available for four drill-core samples from three drill holes (Table 2). Following the resolution-matching procedure described above, the 100 μm μXRF –AMICS maps were represented on a 200 μm categorical label grid corresponding to the XCT voxel spacing.

Table 2. Summary of paired μXRF –AMICS and XCT datasets used in this study

Sample	Drill hole	Depth interval (m)	μ XRF-AMICS grid spacing before resolution matching	AMICS label grid after resolution matching	XCT voxel size
NA25FC054	DH-A	663.77-664.68	100 μ m	200 μ m	200 μ m
NA25FC132	DH-B	1000.60-1001.52	100 μ m	200 μ m	200 μ m
NA25FC154	DH-C	1046.90-1047.86	100 μ m	200 μ m	200 μ m
NA25FC155	DH-C	1065.40-1066.33	100 μ m	200 μ m	200 μ m

Within these samples, XCT datasets covered the complete volumes of the core segments, whereas the corresponding μ XRF-AMICS maps represented regions of the exposed cut-core surfaces approximately 181-183 mm long (Fig. 3). The μ XRF-AMICS datasets therefore provided spatially restricted two-dimensional mineralogical supervision within substantially larger three-dimensional XCT datasets. In addition, as μ XRF-AMICS and XCT datasets differed in extent and initial orientations, geometric co-registration was required to establish the spatial correspondence between the two-dimensional mineral labels and the corresponding XCT surface before model development.

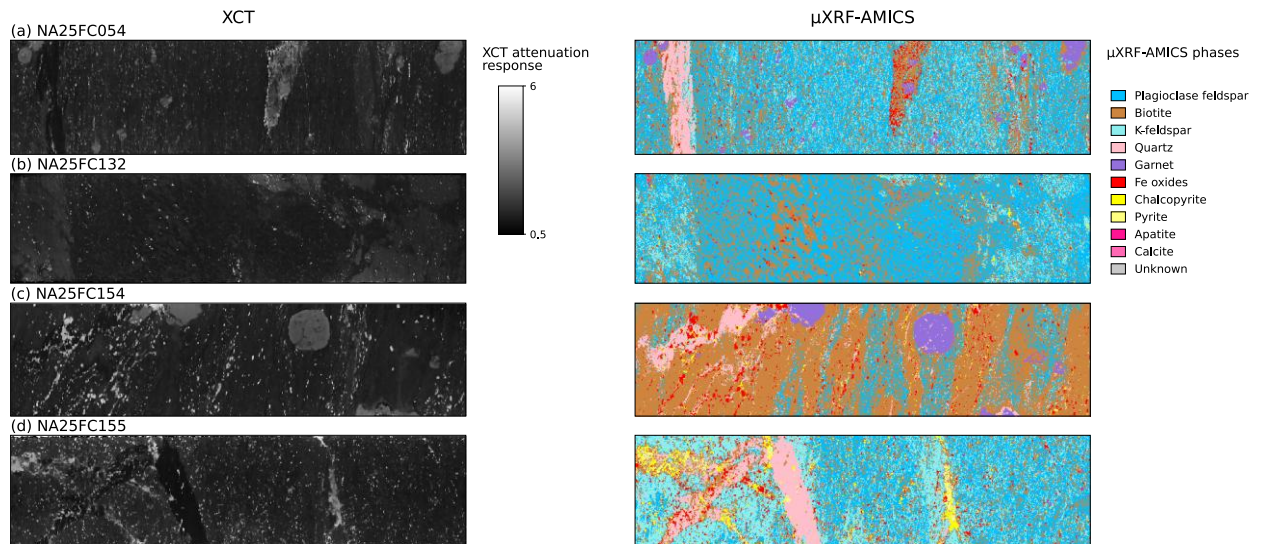


Figure 3. Paired XCT data and corresponding μ XRF-AMICS mineral maps for the four drill-core samples used in this study

3.3 Probabilistic XCT-based mineralogical mapping and validation workflow

Following mineral-label harmonisation and resolution matching, a probabilistic segmentation workflow was developed to learn the relationship between two-dimensional μ XRF-AMICS mineralogical supervision and the corresponding XCT response. The workflow comprised strict surface co-registration, preparation of supervised XCT-label patches, semantic segmentation using AMICS-derived mineral labels as supervision and XCT data as the input, nested leave-one-sample-out validation, probability calibration, support-aware ensembling and subsequent application to verified inward XCT planes. The methodology maintained a strict distinction between mineralogical supervision and model input: μ XRF-AMICS supplied class labels during model development, whereas XCT provided the only input used for held-out evaluation and volumetric inference.

3.3.1 Geometric co-registration and supervised patch dataset preparation

The co-registration procedure was designed to establish strict spatial correspondence between the μ XRF–AMICS mineral labels and the XCT representation of the exposed longitudinal cut surfaces of the drill cores. In contrast to interval-scale or weakly paired multimodal integration, each retained 200 μ m grid cell used for supervision was required to contain both a grouped AMICS mineral label and the corresponding XCT response. Registration incorporated the known sample depths, the bottom-up depth convention of the XCT reconstruction, the orientation of **each** μ XRF–AMICS map, the geometry of the exposed cut surface and spatially corresponding mineralogical and structural features.

To identify the exposed longitudinal cut surface within the XCT volume, transverse XCT sections were segmented using Otsu thresholding (Otsu, 1979), separating the drill-core body from the surrounding image background. Morphological processing and connected-component analysis were then applied, and the largest connected region corresponding to the drill core was retained. The boundary representing the exposed longitudinal cut surface was identified from the segmented core and used to define its geometric position within the XCT volume. Together, these steps reduced the risk that isolated segmented objects outside the drill core or internal low-attenuation regions would interfere with identification of the exposed cut-surface boundary.

Although the exposed cut surface was nominally planar, it was not necessarily parallel to the XCT voxel grid and showed slight inclination or local geometric variation in some samples. Direct extraction at the identified core boundary could therefore include boundary voxels containing attenuation contributions from both the drill core and the surrounding image background. The XCT sampling surface was positioned slightly inward from the identified cut surface, towards the core interior, to reduce boundary-related partial-volume contributions while retaining close spatial correspondence with the physical surface represented by the μ XRF–AMICS mineral map (Fig. 4).

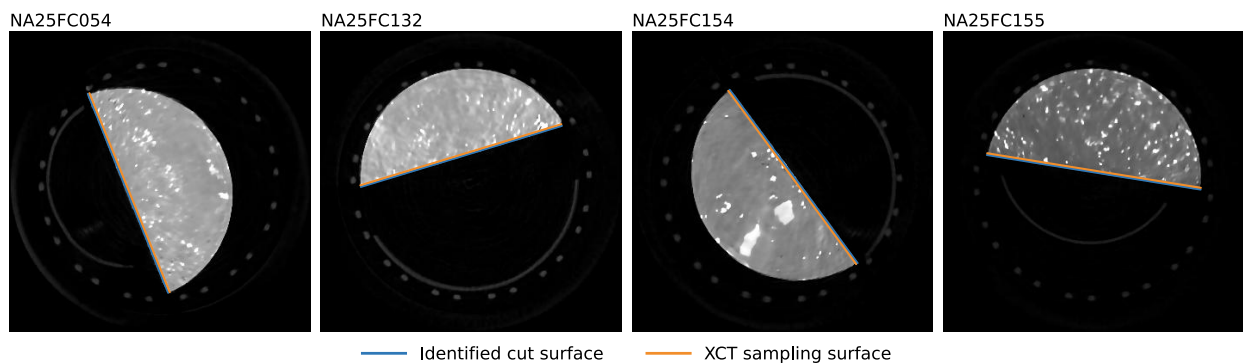


Figure 4. Transverse XCT sections showing the identified cut surface and the corresponding inward XCT sampling surface for the four analysed samples

Following surface extraction, the XCT representation was geometrically matched to the corresponding 200 μ m AMICS label grid. Registration was performed using orientation, angular and translational adjustments established from the spatial correspondence between the two datasets. The XCT data used as model input were not resized, warped or spatially interpolated during registration. Instead, the categorical AMICS mineral

labels and associated masks were transferred to the XCT coordinate system by nearest-neighbour coordinate mapping, thereby preserving their categorical character and preventing the generation of artificial intermediate mineral classes.

Registration quality was assessed from the correspondence between XCT and AMICS features using registered overlays, major mineralogical boundaries, spatial edge correspondence and depth consistency. Candidate registrations were additionally evaluated using attenuation- and edge-based correspondence metrics. Final acceptance was based on geological and spatial verification in addition to numerical registration scores. Registered surfaces were not automatically resized to impose dimensional agreement; any dimensional mismatch between the accepted XCT surface and the corresponding AMICS geometry triggered reassessment of the registration rather than resampling.

Therefore, a grid cell was included in the final supervision mask (M_{sup}) only when it had a valid geometric mapping between AMICS and XCT (M_{map}), satisfied the AMICS purity criterion described in Section 3.1 (M_{purity}), lay outside a broad mineral-boundary exclusion region ($M_{non-boundary}$), and had a valid XCT response (M_{XCT}):

$$M_{sup} = M_{map} \cap M_{purity} \cap M_{non-boundary} \cap M_{XCT} \quad (2)$$

Thus, only grid cells satisfying all four conditions were eligible for model development and evaluation.

The boundary exclusion mask referred to transitions between broad regions assigned to different harmonised AMICS mineral classes within the drill core, rather than the outer edge between rock and surrounding air. Grid cells outside the sample were excluded separately from grid cells without reliable mineral labels. To identify broad class transitions without excluding every local label change, the 200 μm AMICS labels were summarised using a 5×5 -cell block-majority representation (1.0×1.0 mm per block). Grid cells along the resulting transitions, expanded by one 200 μm grid cell, were excluded from supervision and evaluation. At these interfaces, a small residual offset between the AMICS and XCT images could pair an XCT response with the label of the neighbouring mineral class. Labelled grid cells away from these transitions remained eligible.

Each registered surface was represented on a rectangular 200 μm image grid containing the irregular outline of the drill core. Square patches of 128×128 grid cells were extracted from this grid; the core outline did not need to be square. Patches were placed at a stride of 32 cells (6.4 mm) in both row and column directions, so adjacent patches overlapped by 96 cells (19.2 mm), or 75% of their width. At grid cells covered by multiple patches, predictions were averaged during full-surface reconstruction to reduce discontinuities at patch edges. Each patch covered a 25.6×25.6 mm area of the image grid and could contain grid cells outside the rock or without an eligible mineral label. Only grid cells satisfying M_{sup} contributed to the training loss and held-out performance metrics, and a patch was retained only if it contained at least 256 such cells. The final dataset contained 130 patches for NA25FC054, 104 for NA25FC132, 130 for NA25FC154 and 104 for NA25FC155. The

corresponding numbers of unique 200 μm grid cells satisfying M_{sup} were 86,848, 103,823, 132,489 and 71,207, respectively.

For each patch, the XCT values, grouped mineral labels, valid-label mask, sample identifier and original row and column coordinates were retained. Class-support calculations counted each eligible 200 μm surface grid cell once, regardless of how many patches contained it. The reconstructed class-probability vector was required to sum to one at each covered grid cell.

The resulting paired XCT patches and μXRF –AMICS mineral labels formed the geometrically registered, quality-filtered and spatially referenced dataset used for model development and evaluation.

3.3.2 XCT response characteristics of the grouped mineralogical classes

Before model training, the relationship between the grouped μXRF –AMICS mineralogical classes and the corresponding XCT responses was characterised using the unique co-registered, labelled cells on the 200 μm surface grid. The analysis used the aligned XCT responses before fold-specific scaling so that the resulting distributions represented the original XCT-response values rather than model-normalised inputs. Grid cells excluded by the final mineral-label or XCT-validity masks were not included.

XCT-response distributions were evaluated separately for Matrix silicates, Quartz, Garnet and Fe-bearing sulphides and oxides within each sample. A minimum of 500 labelled grid cells was required for density estimation to avoid interpreting unstable distributions based on sparse support. Sample-class combinations below this threshold were retained in the support summary but excluded from density estimation. This criterion affected Garnet in NA25FC155, for which only three retained labelled grid cells were available. When a mineralogical class was absent from a sample, no sample-specific distribution was estimated. The 500-cell threshold was applied only to the descriptive density analysis.

For the pooled mineral-class XCT-response comparison, eligible sample-specific distributions were weighted equally before combination so that samples containing larger labelled mineral areas did not dominate the visual representation. In parallel, pooled descriptive statistics were calculated from the complete set of unique labelled grid cells for each mineralogical class to summarise the overall central tendency and range of XCT responses. Sample-specific support and the 5th, 25th, 50th, 75th and 95th percentiles were also calculated to assess between-sample variability.

Pairwise separability of the grouped mineralogical classes was quantified using the empirical histogram-intersection overlap coefficient. For each pair of mineralogical classes, their XCT-response histograms, normalised to unit area, were compared over a common response range. The minimum probability associated with the two distributions was then summed across corresponding histogram bins. The resulting coefficient ranges from 0, indicating no distributional overlap, to 1, indicating identical distributions. This analysis was descriptive and was used to characterise overlap among the XCT-response distributions associated with the AMICS-defined mineralogical groups; the resulting

distributions were not used to define attenuation thresholds or as additional model inputs.

3.3.3 Hypercolumn-scSE U-Net architecture and model training

The segmentation model used a hypercolumn-scSE U-Net architecture trained using co-registered μ XRF–AMICS mineral labels as supervision and XCT as the model input (Fig. 5). The U-Net comprised an encoding pathway that progressively extracted higher-level spatial features and a decoding pathway that restored full-resolution predictions through skip connections (Ronneberger et al., 2015). Hypercolumns were constructed by upsampling decoder feature maps and concatenating features from multiple spatial scales, enabling the final classifier to integrate fine-scale mineral boundaries with broader textural and neighbourhood information (Hariharan et al., 2015). Concurrent spatial and channel squeeze-and-excitation (scSE) blocks were incorporated to adaptively emphasise informative feature channels and spatial regions while suppressing less informative responses (Roy et al., 2018).

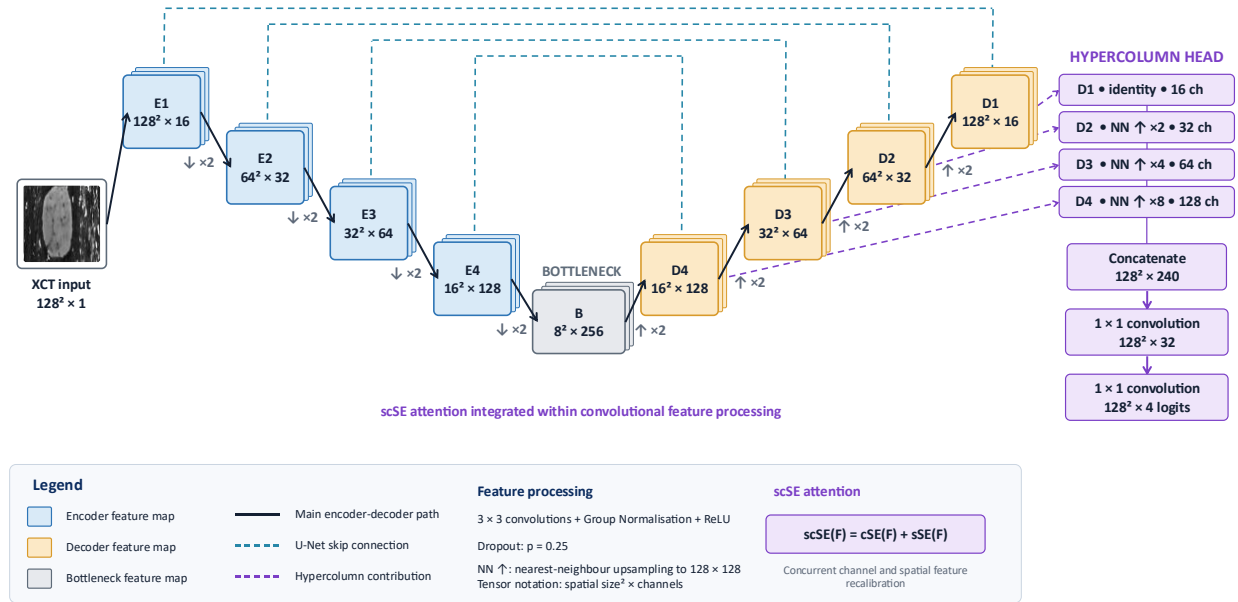


Figure 5. Hypercolumn-scSE U-Net architecture used in this study

The final network accepted a single XCT input channel and used 16 base feature channels, Group Normalisation (Wu and He, 2018) and a dropout rate of 0.25. The encoder channel progression was 16, 32, 64, 128 and 256 channels at the bottleneck, with corresponding decoder stages of 128, 64, 32 and 16 channels. Decoder feature maps from the four spatial scales were combined at the full 128 × 128 resolution, with the lower-resolution decoder feature maps upsampled using nearest-neighbour interpolation before concatenation into a 240-channel hypercolumn, followed by a 1 × 1 convolution reducing the representation to 32 channels and a final 1 × 1 convolution producing four output logits. The four outputs corresponded to Matrix silicates, Quartz, Garnet and Fe-bearing sulphides and oxides. Classes with insufficient training support for an individual inner model were suppressed so that unsupported class responses did not contribute to its predictions.

For each inner configuration, XCT values were scaled using the 1st and 99th percentiles estimated from valid XCT grid cells of the corresponding training samples only. The transformed values were clipped to the interval $[0, 1]$, and the fitted scaling parameters were applied unchanged to the associated validation and held-out samples. This ensured that preprocessing parameters were estimated without access to the held-out data.

The training objective combined masked class-weighted cross-entropy loss (L_{WCE}) and masked Dice loss (L_{Dice}) with equal weighting:

$$L = 0.5L_{WCE} + 0.5L_{Dice} \quad (3)$$

Only grid cells satisfying the supervised-label mask defined in Section 3.3.1 contributed to the loss calculation. Class weights were estimated independently from the unique labelled 200 μm grid cells in the training samples of each inner configuration. For supported classes, weights were derived from the square root of the ratio between the median supported-class grid-cell count and the corresponding class count and were constrained to the range of 0.25-8.0.

Weighted patch sampling increased the probability of selecting patches containing minority classes, with a class considered present within a patch when at least 16 labelled 200 μm grid cells belonged to that class. This sampling strategy increased exposure to minority mineral classes without altering the underlying number of unique labelled surface grid cells.

Model parameters were optimised using AdamW with an initial learning rate of 10^{-4} and a weight decay of 0.01 (Loshchilov and Hutter, 2019). The learning rate was reduced by a factor of 0.5 following four epochs without validation improvement, to a minimum of 10^{-6} . Training was allowed to continue for a maximum of 80 epochs, with early stopping after 12 epochs without improvement. A batch size of 8 was used, and the gradient norm was clipped at 5.0.

Data augmentation comprised horizontal and vertical reflections, rotations by multiples of 90° , and conservative perturbation of the scaled XCT response. Multiplicative gain was sampled between 0.98 and 1.02 and an additive offset between -0.01 and 0.01 was applied. Training and inference were performed using 32-bit floating-point arithmetic. Automatic mixed precision was disabled following numerical instability encountered during preliminary model development.

3.3.4 Nested leave-one-sample-out validation, checkpoint selection and performance assessment

Generalisation was evaluated using four outer leave-one-sample-out folds. Within each outer fold, one complete sample was excluded as the test sample, while the remaining three samples were rotated through three inner validation configurations, each comprising two training samples and one validation sample. Consequently, each outer held-out prediction was generated from three independently trained networks. The outer held-out sample did not contribute to model fitting, XCT scaling, checkpoint selection, probability calibration or ensemble weighting. Its XCT data were introduced only during final held-out prediction, while its mineral labels were used only for subsequent

performance evaluation. Across the complete nested experiment, this procedure produced 12 trained networks from four outer folds and three inner configurations per fold.

Sample-level separation was used instead of random grid-cell- or patch-level splitting because neighbouring patches derived from the same registered surface are spatially correlated. Random partitioning of such patches between training and testing would therefore introduce spatial dependence between the datasets and could produce an overly optimistic estimate of model generalisation.

Checkpoint selection used a minority-sensitive validation criterion in which 25% of the score was assigned to the Matrix-silicate Dice coefficient and 75% to the mean Dice coefficient of the supported non-matrix classes represented in the validation sample:

$$S_{val} = 0.25D_{Matrix} + 0.75D_{Non-matrix} \quad (4)$$

where D_{Matrix} is the Dice coefficient for Matrix silicates and $D_{Non-matrix}$ is the mean Dice coefficient of the supported non-matrix classes represented in the validation sample. A class was considered supported by an inner model when at least 100 unique labelled training grid cells were available for that class. This weighting emphasised performance for the less abundant non-matrix classes while retaining a contribution from the dominant Matrix-silicate class.

Held-out surface performance was evaluated using overall accuracy and class-specific precision, recall, Dice coefficient and intersection over union (IoU). For class k :

$$Precision_k = \frac{TP_k}{TP_k + FP_k} \quad (5)$$

$$Recall_k = \frac{TP_k}{TP_k + FN_k} \quad (6)$$

$$Dice_k = \frac{2TP_k}{2TP_k + FP_k + FN_k} \quad (7)$$

$$IoU_k = \frac{TP_k}{TP_k + FP_k + FN_k} \quad (8)$$

where TP_k , FP_k and FN_k denote the numbers of true-positive, false-positive and false-negative grid cells for class k , respectively. Macro precision, recall, Dice and IoU were calculated across classes represented by at least 100 labelled 200 μm grid cells in each held-out μXRF –AMICS label map. The Dice coefficient refers to an individual class, whereas macro Dice denotes the unweighted arithmetic mean of the Dice coefficients across supported classes. Classes that were absent or represented by fewer than 100 labelled 200 μm grid cells were excluded from both the macro metrics and class-specific quantitative interpretation. Unless stated otherwise, summary metrics across outer folds were reported as unweighted arithmetic means, giving each held-out sample equal influence.

Probability quality was additionally evaluated using the multiclass Brier score (Brier, 1950), expected calibration error (ECE), normalised predictive entropy and maximum class probability. The multiclass Brier score quantified the squared difference between the predicted probability vector and the corresponding one-hot mineral label:

$$Brier = \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K (p_{ik} - y_{ik})^2 \quad (9)$$

where N is the number of valid labelled 200 μm grid cells, K is the number of model classes, p_{ik} is the predicted probability of class k at grid cell i , and y_{ik} is the corresponding one-hot target. The score was calculated over valid labelled 200 μm grid cells only.

Expected calibration error (ECE) was calculated using 15 equal-width confidence bins over the interval $[0, 1]$. For each evaluated labelled 200 μm grid cell, prediction confidence was defined as the maximum predicted class probability. Within each populated bin, mean prediction confidence was compared with the fraction of correctly classified labelled grid cells. ECE was calculated as:

$$ECE = \sum_{b=1}^B \frac{n_b}{N} |acc(b) - conf(b)| \quad (10)$$

where B is the number of confidence bins, n_b is the number of evaluated labelled grid cells in bin b , N is the total number of evaluated labelled grid cells, $acc(b)$ is the fraction correctly classified within the bin, and $conf(b)$ is its mean predicted confidence.

Differences between grouped AMICS surface proportions and the corresponding predicted surface proportions were quantified for both hard-class predictions and mean predicted probabilities. For the probability-based measure, the mean absolute error in probability-derived class proportions was calculated as the mean absolute difference, across represented classes, between the reference surface fraction and the mean predicted class probability; values were subsequently averaged across outer folds.

3.3.5 Probability calibration and support-aware ensembling

For each outer held-out sample, the three independently trained inner models were combined using class-specific training-support information. Five ensemble configurations were evaluated: the original equally averaged uncalibrated ensemble; support-aware equal and validation-weighted uncalibrated ensembles; and corresponding support-aware equal and validation-weighted ensembles after probability calibration. The final workflow retained the support-aware, equally averaged, calibrated ensemble.

An inner model was permitted to contribute to class k only when at least 100 unique spatially labelled training grid cells of that class were available in its training samples. For inner model r , this condition was represented by a binary support indicator s_{rk} , where $s_{rk} = 1$ when the support criterion was satisfied and $s_{rk} = 0$ otherwise.

For class k at grid cell i , the support-aware class-specific ensemble average was calculated as:

$$q_{ik} = \frac{\sum_{r=1}^R s_{rk} w_{rk} p_{ik}^{(r)}}{\sum_{r=1}^R s_{rk} w_{rk}} \quad (11)$$

where $R = 3$ is the number of inner models, $p_{ik}^{(r)}$ is the probability predicted by inner model r , s_{rk} is its class-support indicator and w_{rk} is its class-specific ensemble weight. For the selected equal-weighted ensemble, $w_{rk} = 1$ for every eligible model-class combination. For the validation-weighted variants, w_{rk} was derived from the class-

specific Dice coefficient achieved by inner model r on its inner-validation sample. When at least 100 labelled validation grid cells were available for class k , w_{rk} was set to the corresponding validation Dice coefficient, subject to a minimum weight of 0.05. When fewer than 100 validation grid cells were available, the mean Dice coefficient across supported classes represented in that inner-validation sample was used as a fallback. Classes unsupported in the training samples of the model retained a weight of zero. The denominator in Eq. (11) normalised the resulting weights across eligible inner models separately for each class. In every outer fold, at least one inner model satisfied the support criterion for each retained class; therefore, the denominator in Eq. (11) was non-zero.

Because different classes could be supported by different subsets of inner models, the resulting class-specific averages were renormalised across the supported classes:

$$p_{ik}^{(ens)} = \frac{q_{ik}}{\sum_{j=1}^K q_{ij}} \quad (12)$$

so that the probabilities of eligible classes summed to one at every predicted grid cell. This renormalisation ensured a valid class-probability vector when different subsets of inner models supported different classes.

Probability calibration was performed independently for each inner model using single-parameter logit rescaling (Guo et al., 2017), fitted exclusively to its inner-validation predictions. This procedure divides the model logits by a single fitted scaling parameter before the softmax transformation. Let S_r denote the set of classes supported by inner model r . For class $k \in S_r$, the calibrated probability was calculated as:

$$p_{ik}^{(r, T_r)} = \frac{\exp(\frac{z_{ik}^{(r)}}{T_r})}{\sum_{j \in S_r} \exp(\frac{z_{ij}^{(r)}}{T_r})} \quad (13)$$

where $z_{ik}^{(r)}$ is the logit predicted for grid cell i and class k by inner model r , and $T_r > 0$ is the scalar logit-rescaling parameter fitted for that model. Classes not included in S_r were assigned zero probability. Consequently, the calibrated probabilities summed to one across the classes supported by each model.

For each inner model, T_r was selected by minimising the multiclass negative log-likelihood using inner-validation labelled 200 μm grid cells whose reference classes were supported by that model. Inner-validation labelled 200 μm grid cells belonging to classes unsupported by the corresponding model were excluded from logit-rescaling fitting because their predicted probability was constrained to zero. Logit-rescaling optimisation was restricted to the interval 0.1–3.0 and used a maximum of 250,000 validation grid cells. Labels from the outer held-out sample were not used to estimate the logit-rescaling parameter.

Following calibration, the predictions from the three inner models were combined using the support-aware procedure defined in Eqs. (11) and (12). For the calibrated configurations, $p_{ik}^{(r)}$ in Eq. (11) was replaced by $p_{ik}^{(r, T_r)}$. The support-aware equally averaged calibrated ensemble was retained for subsequent evaluation on the outer held-out samples and for three-dimensional XCT inference.

3.3.6 Exact-surface reproduction and three-dimensional XCT inference

Before extending XCT-based predictions beneath the labelled surface, the exact XCT surface representation used during the original held-out evaluation was regenerated from the underlying XCT volume. This step ensured that three-dimensional inference began from the same geometrical and numerical surface on which the model had previously been evaluated rather than from an approximately corresponding XCT plane.

The regenerated XCT input was processed using the corresponding held-out ensemble, fold-specific XCT scaling parameters, class-support masks and fitted logit-rescaling parameters without model retraining. Reproduced class-probability fields, hard-class predictions, normalised predictive entropy and maximum-probability fields were compared with the previously stored held-out predictions. Volumetric inference proceeded only after numerical reproduction of the original surface predictions had been established.

Following successful surface reproduction, the trained ensemble was applied to parallel XCT sampling planes extending inward from the verified model surface of NA25FC154. Sampling positions were evaluated at 0.2 mm spacing, corresponding to the nominal XCT resolution. Plane-specific connected-rock masking retained 118 planes with slab-centre distances of 0.3 to 23.7 mm inward from the verified model surface. No additional μ XRF-AMICS information or mineralogical labels were supplied during volumetric inference; XCT remained the sole prediction input.

A plane-specific valid-rock mask was used to exclude surrounding air, scanner components and other non-core regions from volumetric interpretation. Candidate rock grid cells were identified using the stored XCT rock-response threshold of 0.1, and only the largest contiguous rock region was retained within each plane. Contiguity was defined using eight-neighbour connectivity, whereby cells sharing either an edge or a corner were considered connected. A threshold-sensitivity assessment using XCT responses from 0.05 to 0.15 produced comparable estimates of the rear core boundary. Beyond 20 mm inward from the verified model surface, an operational minimum rock-coverage criterion of 8% of the rectangular prediction domain was applied. This criterion prevented planes containing only small residual rock regions near the curved rear boundary from being retained. The first plane falling below this criterion occurred at 23.9 mm, making 23.7 mm the last retained plane.

For regional reporting, planes with slab-centre distances of 0.3-20.1 mm formed the stable interior zone, whereas retained planes from 20.3 to 23.7 mm formed the boundary-proximal zone because they lay within the >20 mm boundary-monitoring region and showed rapidly decreasing connected-rock support. This division was a geometric, support-based reporting convention rather than a measured threshold of mineralogical accuracy.

For each evaluated plane, overlapping 128×128 patch predictions were reconstructed to complete the prediction surface by arithmetic averaging at grid cells covered by more than one patch. For each covered cell, probabilities from all contributing patches were summed and divided by the number of contributing patches.

The resulting probability volumes had dimensions $N \times 4 \times H \times W$, where N is the number of evaluated planes, H and W are the spatial dimensions of the reconstructed surface, and the four channels correspond to Matrix silicates, Quartz, Garnet and Fe-bearing sulphides and oxides. Hereafter, p_{ik} denotes the final calibrated ensemble probability assigned to class k at cell i . Each cell was assigned to the class with the highest predicted probability:

$$\hat{y}_i = \arg \max p_{ik}, \quad k \in \{1, \dots, K\}, \quad K = 4 \quad (14)$$

where $K = 4$ is the number of model classes.

The analysis mask for each plane was defined by the intersection of the plane-specific connected-rock mask, finite XCT model inputs and valid patch-reconstruction coverage. For spatial-consistency comparisons between pairs of retained inward planes, metrics were calculated only at grid coordinates identified as valid rock in both planes. Restricting each comparison to this common rock support prevented changes in the core cross-sectional area near the curved rear boundary from being incorrectly interpreted as changes in mineral-class predictions.

3.3.7 Uncertainty, spatial consistency and transition metrics

Predictive uncertainty was characterised using normalised Shannon entropy and maximum class probability. For $K = 4$ classes, normalised predictive entropy at grid cell i was calculated as:

$$H_i = -\frac{1}{\ln K} \sum_{k=1}^K p_{ik} \ln(p_{ik} + \varepsilon) \quad (15)$$

where p_{ik} is the predicted probability of class k at grid cell i , and ε is a small numerical constant included to avoid evaluation of the logarithm at zero. Normalisation by $\ln K$ constrained entropy to the interval $[0, 1]$, with lower values indicating a more concentrated probability distribution and higher values indicating greater predictive uncertainty.

Maximum class probability was calculated as:

$$C_i = \max p_{ik}, \quad k \in \{1, \dots, K\} \quad (16)$$

and provided a complementary measure of prediction confidence.

Spatial consistency was evaluated only at coordinates identified as valid rock in both compared planes. Let Ω_{ab} denote the intersection of the connected valid-rock masks for planes a and b , and let $M_{ab} = |\Omega_{ab}|$ denote the number of common valid-rock grid cells. Pairwise hard-class agreement was calculated as:

$$A(a, b) = \frac{1}{M_{ab}} \sum_{i \in \Omega_{ab}} 1[\hat{y}_i^a = \hat{y}_i^b] \quad (17)$$

where $\hat{y}_i^a$ and $\hat{y}_i^b$ are the hard-class predictions at coordinate i in planes a and b , respectively, and the indicator function equals one when the predictions are identical.

Changes in the complete probability vector between planes were quantified using mean absolute probability change:

$$\Delta p(a, b) = \frac{1}{M_{ab}K} \sum_{i \in \Omega_{ab}} \sum_{k=1}^K |p_{ik}^a - p_{ik}^b| \quad (18)$$

Class-specific stability was evaluated using self-retention:

$$R_k(a, b) = \frac{\sum_{i \in \Omega_{ab}} 1[\hat{y}_i^a = k \wedge \hat{y}_i^b = k]}{\sum_{i \in \Omega_{ab}} 1[\hat{y}_i^a = k]} \quad (19)$$

The numerator represents grid cells assigned to class k in both planes, whereas the denominator represents grid cells assigned to class k in the shallower reference plane a . For adjacent-plane analysis, b was the next deeper plane. Self-retention was calculated only when at least one common valid-rock grid cell was assigned to class k in the shallower plane, ensuring that the denominator in Eq. (19) was non-zero. Regional class-specific self-retention was weighted by the number of class- k grid cells in the shallower plane. Longer-range agreement was calculated using the first retained plane as a and a selected deeper plane as b .

Regional means of predictive entropy and maximum class probability were weighted by the number of valid-rock grid cells in each plane. Regional adjacent-plane agreement was weighted by the number of common valid-rock grid cells in each plane pair. Valid-rock support was additionally reported as the number of connected valid-rock grid cells and as a fraction of the complete rectangular prediction domain. These measures quantified spatial continuity and predictive behaviour without treating the unlabelled core interior as independent mineralogical ground truth.

Together, the workflow progressed from strictly co-registered two-dimensional μ XRF-AMICS supervision through model training using XCT as the model input, nested sample-level validation, probability calibration, support-aware ensembling, exact-surface reproduction, three-dimensional XCT inference, uncertainty and spatial-consistency assessment (Fig. 6).

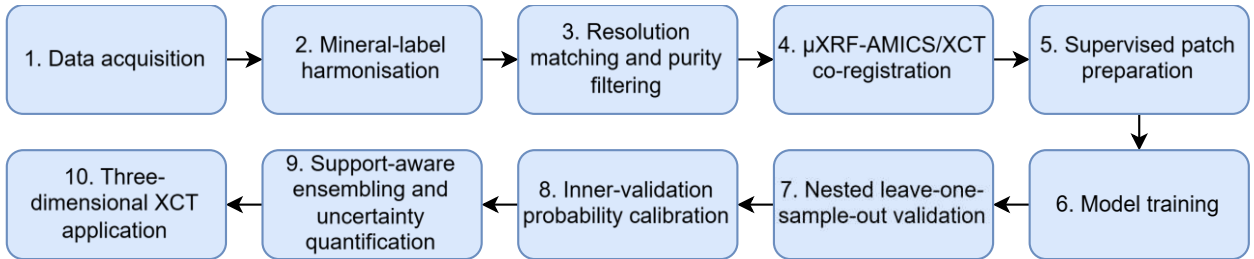


Figure 6. Complete analytical workflow for probabilistic three-dimensional mineralogical characterisation

4. Results

4.1 XCT response characteristics of the grouped mineralogical classes

The four grouped mineralogical classes exhibited distinct but partially overlapping XCT attenuation responses (Fig. 7). Across the pooled labelled dataset, Garnet showed the highest characteristic attenuation, with a median XCT attenuation value of 2.233, compared with 1.501 for Fe-bearing sulphides and oxides, 1.241 for Matrix silicates and 1.034 for Quartz. Garnet occupied a comparatively concentrated higher-attenuation

range, whereas Fe-bearing sulphides and oxides displayed a substantially broader response extending across attenuation values associated with the other mineralogical groups.

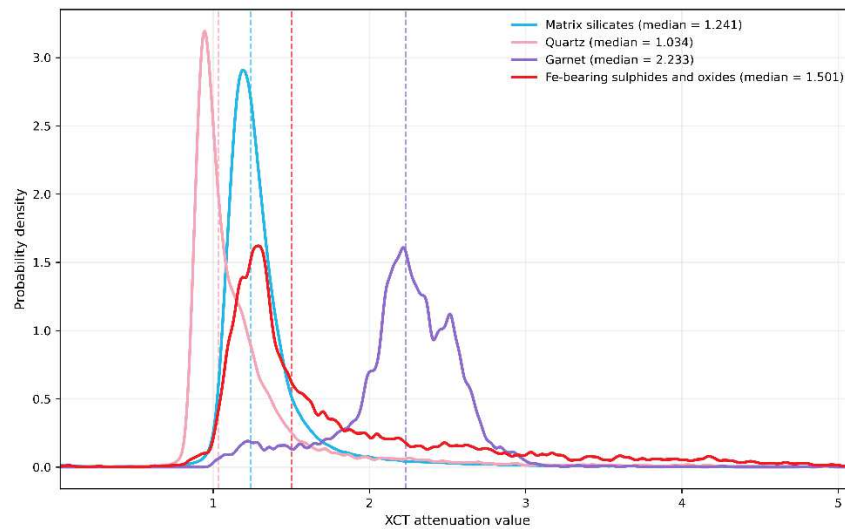


Figure 7. Class-conditioned XCT attenuation distributions for Matrix silicates, Quartz, Garnet and Fe-bearing sulphides and oxides derived from co-registered μ XRF–AMICS-labelled 200 μ m grid cells

The attenuation distributions were therefore not uniquely separable by mineralogical class. The greatest pairwise distribution overlap occurred between Matrix silicates and Fe-bearing sulphides and oxides, with an overlap coefficient of 0.667 (Fig. 8). Matrix silicates and Quartz showed an overlap coefficient of 0.483, while Quartz and Fe-bearing sulphides and oxides showed a similar overlap of 0.461. Garnet was more clearly differentiated from Matrix silicates and Quartz, with overlap coefficients of 0.165 and 0.161, respectively, although moderate overlap of 0.311 remained between Garnet and Fe-bearing sulphides and oxides.

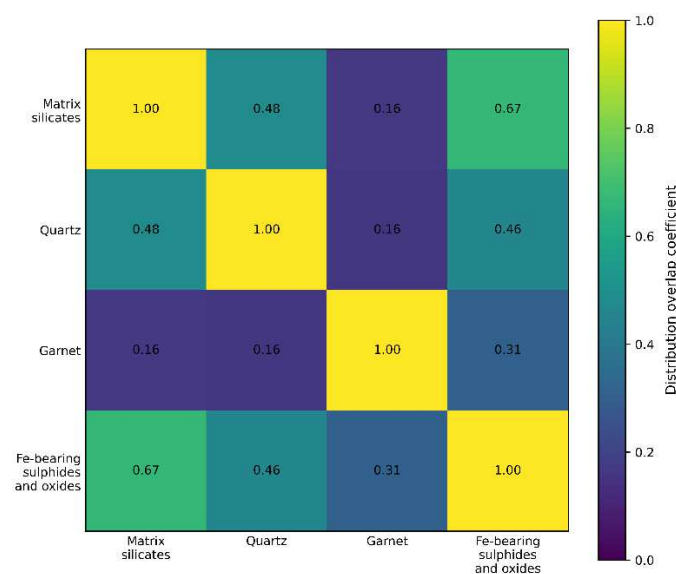


Figure 8. Pairwise empirical distribution-overlap coefficients for the four grouped mineralogical classes

The XCT responses conditioned on four mineralogical classes also varied between samples (Fig. 9). The median XCT attenuation for Matrix silicates ranged from 1.205 in NA25FC132 to 1.278 in NA25FC155. The Fe-bearing sulphides and oxides class exhibited greater between-sample variability, with median values ranging from 1.278 in NA25FC132 to 1.718 in NA25FC054. Quartz had median values of 1.011, 1.185 and 0.941 in NA25FC054, NA25FC154 and NA25FC155, respectively. Garnet showed consistently higher attenuation where adequately represented, with median values of 2.283 in NA25FC054 and 2.216 in NA25FC154.

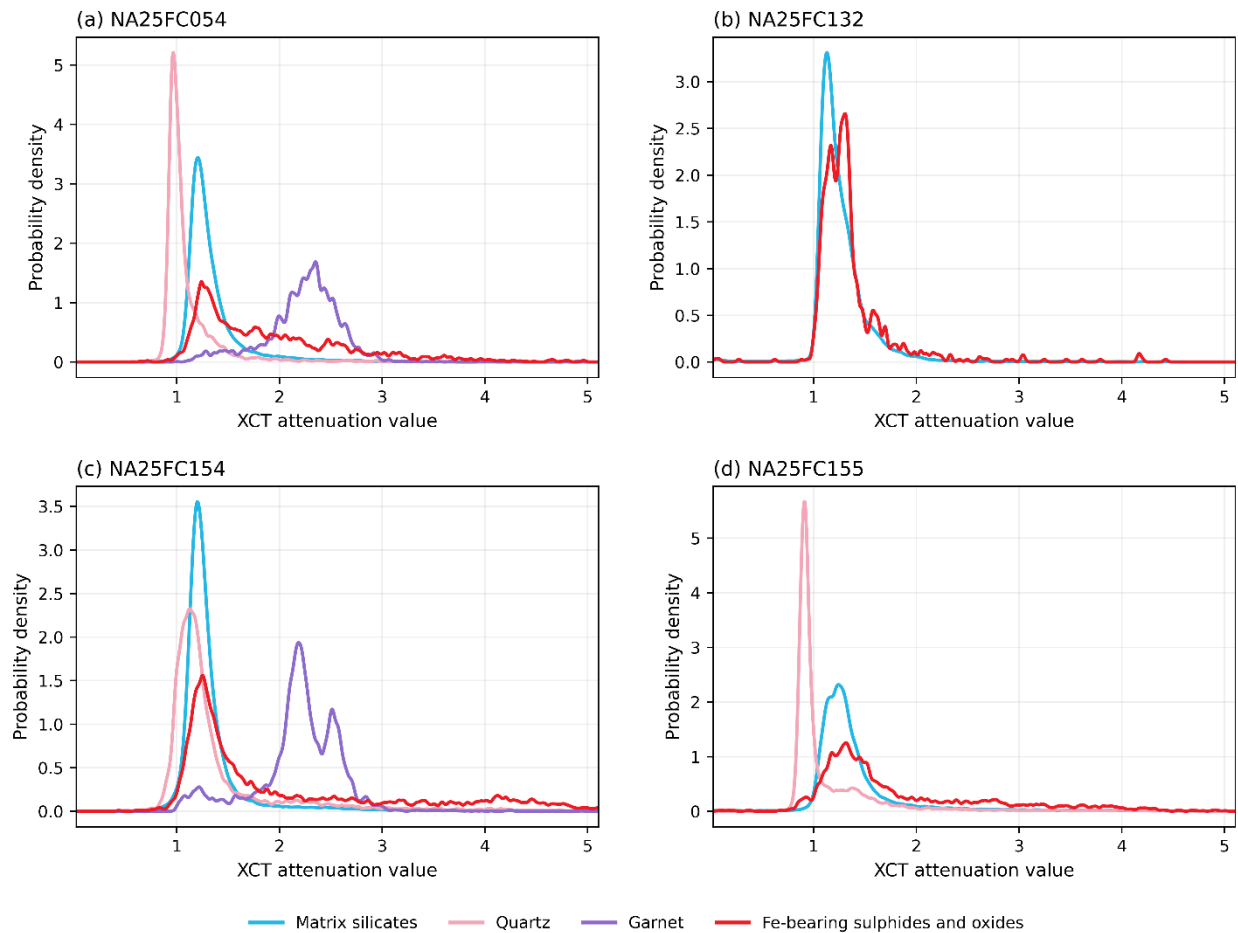


Figure 9. Sample-specific XCT attenuation distributions conditioned on four mineralogical classes for NA25FC054, NA25FC132, NA25FC154 and NA25FC155

The number of labelled grid cells varied substantially among the drill-core samples. Each labelled grid cell represented one unique valid mineral label on the co-registered 200 μm grid. NA25FC132 contained no labelled grid cells for Quartz or Garnet, while Garnet in NA25FC155 was represented by only three labelled grid cells. In contrast, NA25FC054 and NA25FC154 contained substantial numbers of labelled grid cells for all four model classes (Table 3). Class-sample combinations with absent or minimal labelled representation were excluded from class-specific performance assessment rather than treated as prediction failures.

Table 3. Number of unique valid labelled grid cells for the grouped mineralogical classes in the co-registered 200 μm μXRF –AMICS and XCT dataset

Grouped class	NA25FC054	NA25FC132	NA25FC154	NA25FC155	Total labelled grid cells
Matrix silicates	75,028	103,044	112,485	54,023	344,580
Quartz	6,366	0	8,631	12,694	27,691
Garnet	3,202	0	6,399	3	9,604
Fe-bearing sulphides and oxides	2,252	779	4,974	4,487	12,492

The XCT-response distributions showed systematic contrasts among the grouped mineralogical classes but did not permit unique deterministic separation of the four classes. In particular, the pronounced overlap between Fe-bearing sulphides and oxides and Matrix silicates demonstrated why fixed, mutually exclusive XCT-response thresholds were insufficient. These distributional characteristics therefore supported probabilistic classification, in which reconstructed XCT responses were interpreted together with spatial context.

4.2 Outer leave-one-sample-out prediction performance

Cross-sample prediction performance was evaluated across the four outer leave-one-sample-out folds. In each fold, one complete drill-core sample was withheld from model fitting and subsequently predicted from its XCT responses (Fig. 10). Spatial comparisons between the grouped μ XRF-AMICS label maps and the corresponding held-out predictions showed that the spatial patterns of the dominant mineral classes were reproduced across all four outer folds. Matrix-silicate regions displayed the closest agreement between the label maps and predictions, whereas disagreements were concentrated in regions occupied by minority classes and across complex mineralogical transitions. Larger Quartz- and Garnet-rich structures present in the label maps were also reproduced in several held-out predictions, although the degree of agreement varied among drill-core samples and mineral classes. The normalised predictive entropy fields were spatially heterogeneous across the analysed surfaces. Lower entropy occurred predominantly where one mineral class had a clearly dominant predicted probability, whereas higher entropy occurred where probabilities were distributed more evenly among competing classes. This spatial variation demonstrated that model confidence was not uniform across the hard-class predictions.

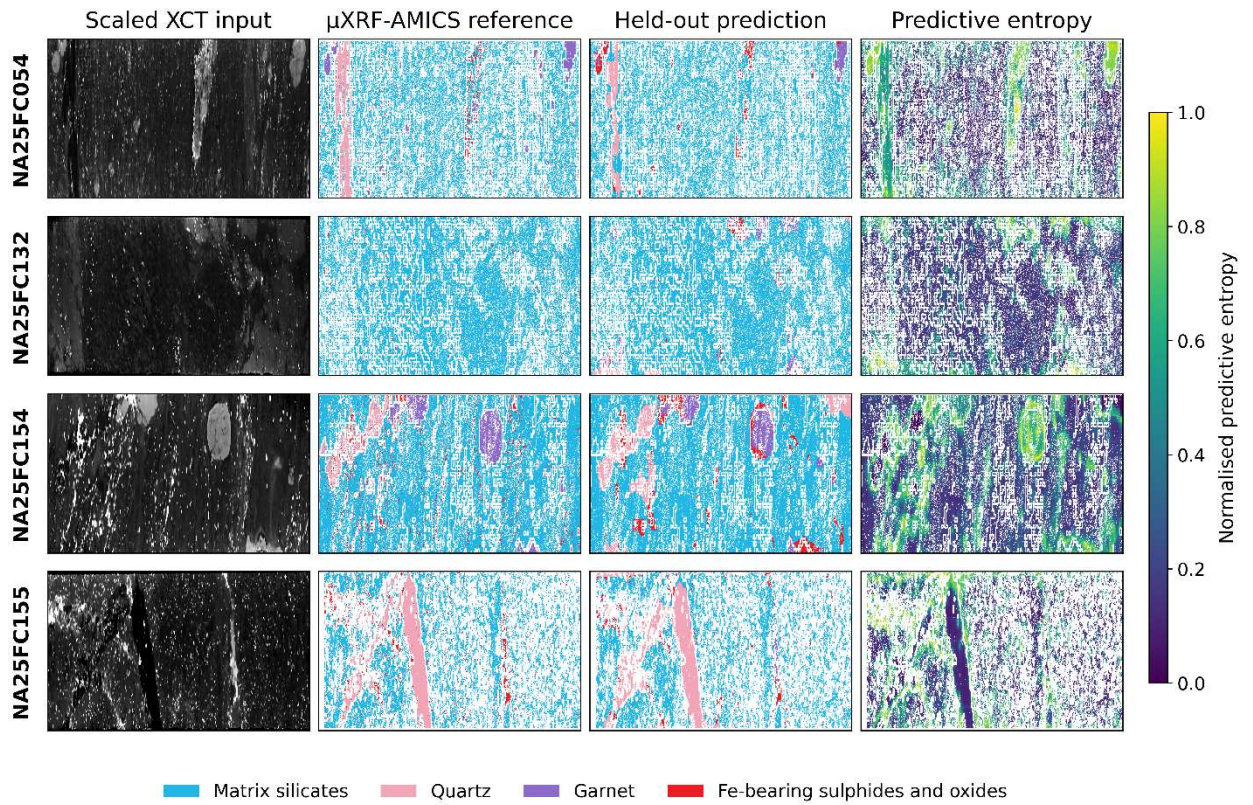


Figure 10. Held-out prediction results, showing XCT input, grouped μ XRF–AMICS reference, held-out hard-class prediction and normalised predictive entropy

Across the four outer-fold held-out evaluations, accuracy ranged from 0.863 to 0.960, with a mean of 0.899, while macro Dice averaged 0.583 (Table 4). In total, the four outer evaluations comprised 394,367 labelled 200 μ m grid cells.

Table 4. Held-out performance of the selected calibrated ensemble for each drill-core sample

Sample	Evaluation grid cells	Accuracy	Macro Dice	Brier score	ECE
NA25FC054	86,848	0.910	0.589	0.138	0.031
NA25FC132	103,823	0.960	0.506	0.075	0.070
NA25FC154	132,489	0.864	0.583	0.220	0.019
NA25FC155	71,207	0.863	0.654	0.213	0.030
Total grid cells / mean metrics	394,367	0.899	0.583	0.161	0.038

NA25FC054 achieved an overall accuracy of 0.910 and a macro Dice of 0.589, calculated across all four eligible classes. NA25FC154 achieved an accuracy of 0.864 and a macro Dice of 0.583, also calculated across all four eligible classes. NA25FC155 showed a similar overall accuracy of 0.863 and achieved a macro Dice of 0.654 across three eligible classes. NA25FC132 achieved the highest overall accuracy of 0.960, although its macro Dice of 0.506 was calculated across only two eligible classes. Because the set of eligible classes differed among the held-out samples, the fold-level macro Dice values should be

interpreted together with the class-specific results discussed in Section 4.3 below rather than as directly equivalent four-class summaries.

The calibrated class-probability maps provide information beyond that shown by the discrete hard-class predictions (Fig. 11). For each grid cell, the hard-class map shows only the class with the highest calibrated predicted probability. The probability maps instead retain the predicted probability assigned to every class, including alternative classes that were not selected. They therefore indicate both the model's confidence in the selected class and the degree of competition among classes, helping to identify ambiguous predictions.

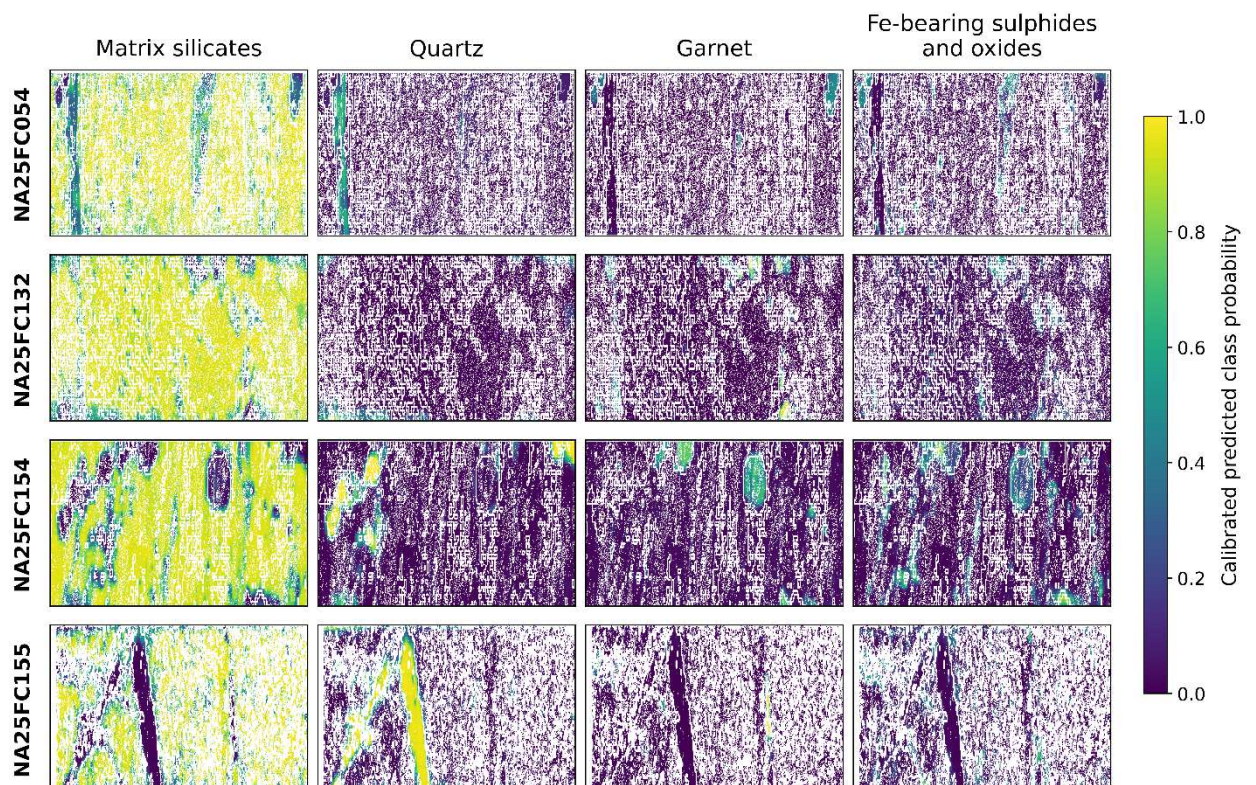


Figure 11. Held-out calibrated class-probability maps for Matrix silicates, Quartz, Garnet and Fe-bearing sulphides and oxides in each independently held-out sample

4.3 Class-specific generalisation

Held-out prediction performance differed substantially among the grouped mineralogical classes, as shown in Fig. 12.

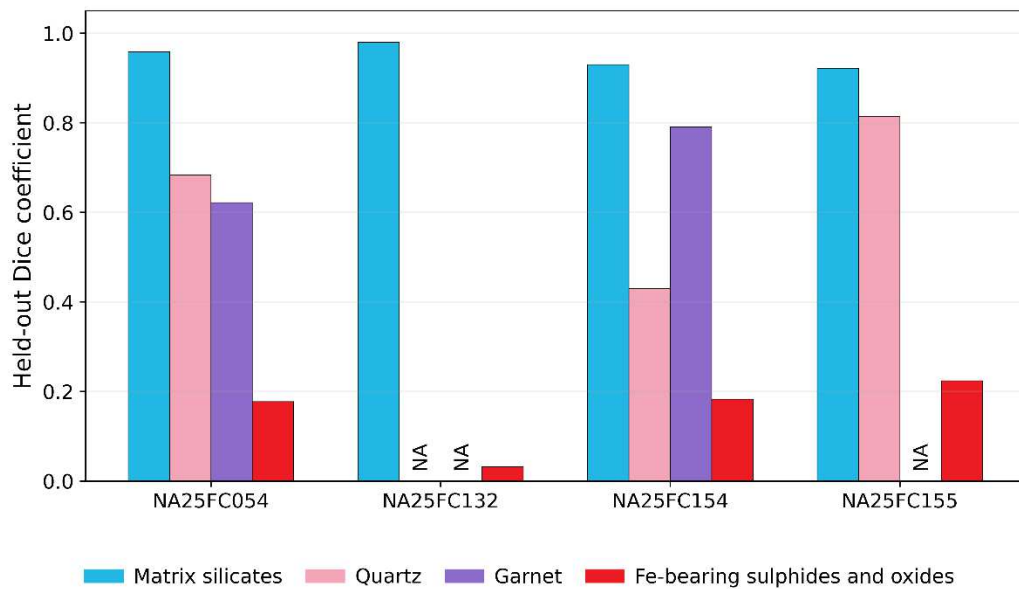


Figure 12. Class-specific outer-fold generalisation. NA indicates classes with insufficient held-out labelled-grid support for quantitative interpretation

Matrix silicates showed the strongest and most consistent cross-sample generalisation, with Dice coefficients of 0.957, 0.980, 0.930 and 0.922 for NA25FC054, NA25FC132, NA25FC154 and NA25FC155, respectively. The consistently high Dice coefficients indicate that the model identified the dominant Matrix silicates class across all four independently held-out drill-core samples. Quartz also showed transferable discrimination where sufficient labelled-grid support was available, although performance varied considerably between samples. Quartz Dice coefficients were 0.653 for NA25FC054, 0.430 for NA25FC154 and 0.815 for NA25FC155. Quartz could not be quantitatively evaluated in NA25FC132 because no labelled Quartz grid cells were present in its held-out reference map.

Garnet achieved Dice coefficients of 0.565 in NA25FC054 and 0.791 in NA25FC154. These results are consistent with the comparatively distinctive higher-attenuation Garnet response identified in Section 4.1. Garnet was absent from the NA25FC132 reference map and represented by only three labelled 200 μm grid cells in NA25FC155. These cases were therefore considered insufficient for quantitative interpretation rather than being treated as zero prediction performance. The Fe-bearing sulphides and oxides class showed consistently weaker generalisation than the other adequately supported classes. Dice coefficients were 0.182 for NA25FC054, 0.032 for NA25FC132, 0.183 for NA25FC154 and 0.224 for NA25FC155. Importantly, these low values occurred despite held-out labelled-grid support ranging from 779 to 4,974 labelled 200 μm grid cells. Detailed class-specific metrics are summarised in Table 5. This behaviour is consistent with the broad, strongly overlapping attenuation response of this class described in Section 4.1.

Table 5. Class-specific generalisation. Held-out labelled-grid support, Dice coefficient, precision, recall and intersection-over-union for each grouped mineralogical class and sample. NA indicates classes with insufficient held-out labelled-grid support for quantitative interpretation

Sample	Class	Reference 200 μm grid cells	Dice	Precision	Recall	IoU
NA25FC054	Matrix silicates	75,028	0.957	0.927	0.988	0.917
	Quartz	6,366	0.653	0.927	0.505	0.485
	Garnet	3,202	0.565	0.942	0.404	0.394
	Fe-bearing sulphides and oxides	2,252	0.182	0.190	0.175	0.100
NA25FC132	Matrix silicates	103,044	0.980	0.993	0.967	0.961
	Quartz	0	NA	NA	NA	NA
	Garnet	0	NA	NA	NA	NA
	Fe-bearing sulphides and oxides	779	0.032	0.032	0.032	0.016
NA25FC154	Matrix silicates	112,485	0.930	0.923	0.937	0.870
	Quartz	8,631	0.430	0.461	0.403	0.274
	Garnet	6,399	0.791	0.876	0.721	0.654
	Fe-bearing sulphides and oxides	4,974	0.183	0.174	0.193	0.101
NA25FC155	Matrix silicates	54,023	0.922	0.922	0.922	0.855
	Quartz	12,694	0.815	0.776	0.857	0.688
	Garnet	3	NA	NA	NA	NA
	Fe-bearing sulphides and oxides	4,487	0.224	0.314	0.174	0.126

Attenuation overlap alone does not fully determine prediction performance because the network also incorporates local and multi-scale spatial information. However, the Fe-bearing sulphides and oxides class combines Fe oxides, chalcopyrite and pyrite, whose relative abundances, grain sizes and textural associations vary among the drill-core samples. This produces sample-dependent XCT attenuation and spatial characteristics that also overlap those of Matrix silicates. Consequently, features learned for the Fe-bearing sulphides and oxides class in one sample were less transferable than those of the more consistently expressed Matrix silicates class.

The class-specific results also clarify interpretation of the sample-level accuracies reported in Section 4.2. High overall accuracy was driven primarily by robust prediction of the dominant Matrix-silicate response and did not imply equivalent performance for other mineralogical classes. Quartz and Garnet retained meaningful cross-sample discrimination where adequately represented, whereas Fe-bearing sulphides and oxides remained the principal source of class ambiguity.

4.4 Probability calibration and uncertainty

Post-hoc probability calibration using single-parameter logit rescaling substantially improved agreement between predicted confidence and the observed fraction of correct predictions, while producing only minor changes in hard-class discrimination (Fig. 13).

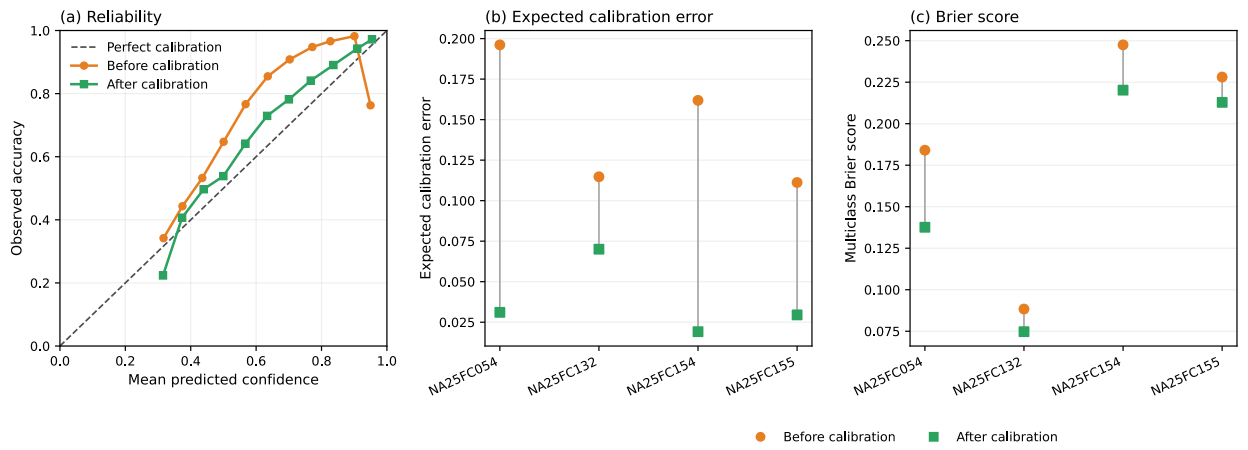


Figure 13. Probability calibration and reliability: (a) Reliability diagram comparing mean predicted confidence with the fraction of correctly classified labelled grid cells within each confidence bin, before and after single-parameter logit rescaling; (b) Expected calibration error for each held-out drill-core sample; (c) Multiclass Brier score before and after calibration

Mean ECE decreased from 0.146 before calibration to 0.038 after calibration, corresponding to a reduction of approximately 74% and indicating closer agreement between predicted confidence and the observed fraction of correct predictions. Mean multiclass Brier score decreased from 0.187 to 0.161, whereas mean macro Dice changed only slightly from 0.586 to 0.583. Thus, probability calibration principally improved the probability estimates rather than hard-class discrimination.

ECE decreased in every outer fold: from 0.196 to 0.031 in NA25FC054, from 0.115 to 0.070 in NA25FC132, from 0.162 to 0.019 in NA25FC154 and from 0.111 to 0.030 in NA25FC155. Multiclass Brier score also decreased in every fold, demonstrating that the improvement was observed across all four held-out drill-core samples rather than being driven by a single sample (Table 6).

Table 6. Probability calibration comparison. Expected calibration error, multiclass Brier score, macro Dice, mean normalised predictive entropy and mean absolute error in probability-derived class proportions before and after probability calibration

Metric	Before calibration	After calibration
Expected calibration error	0.146	0.038
Multiclass Brier score	0.187	0.161
Macro Dice	0.586	0.583
Mean normalised predictive entropy	0.533	0.313
Mean absolute error in probability-derived class proportions	0.083	0.038

The reliability diagram illustrates how this improvement arose (Fig. 13a). Before calibration, most populated confidence bins lay above the 1:1 line, meaning that the fraction of correct predictions within those bins exceeded their mean predicted confidence and indicating predominantly under-confident probability estimates. After calibration, the bin estimates lay closer to the 1:1 line, indicating closer agreement between predicted confidence and empirical correctness. The final downward point

represents the fraction of correct predictions within the highest pre-calibration confidence bin; the connecting line is included only as a visual guide and does not imply a continuous or monotonic relationship.

Calibration was accompanied by a reduction in mean normalised predictive entropy from 0.533 to 0.313 for the support-aware equally averaged ensemble. This reduction reflected redistribution of probability mass towards the most likely classes following single-parameter logit rescaling. Because macro Dice decreased only slightly, the entropy reduction indicated sharper probability distributions rather than increased discriminatory capacity. The calibration improvement was demonstrated separately by the reductions in expected calibration error, multiclass Brier score and inner-validation negative log-likelihood.

The mean absolute error in probability-derived class proportions decreased from 0.083 before calibration to 0.038 after calibration. Thus, single-parameter logit rescaling improved both the agreement between predicted confidence and observed correctness and the agreement between probability-derived and μ XRF-AMICS-labelled class proportions, while producing only a small change in hard-class segmentation performance.

These results support retention of calibrated class probabilities and prediction uncertainties as primary model outputs. In contrast to hard classifications, the probability fields preserve the degree of competition among alternative grouped mineralogical responses and therefore provide a more informative representation where XCT responses overlap.

4.5 Cross-sample transferability and ensemble selection

Comparison of the five ensemble configurations showed that support-aware aggregation primarily improved class-balanced discrimination, whereas probability calibration produced the dominant improvement in the quality of probability estimates (Table 7). The original equally averaged uncalibrated ensemble achieved a mean macro Dice of 0.548. Introducing support-aware equal averaging increased mean macro Dice to 0.586 while maintaining overall accuracy at approximately 0.90.

Table 7. Performance comparison of the evaluated ensemble configurations across the outer leave-one-sample-out folds

Ensemble configuration	Accuracy	Macro Dice	Brier score	ECE	Selected
Original equal, uncalibrated	0.898	0.548	0.185	0.138	No
Support-aware equal, uncalibrated	0.902	0.586	0.187	0.146	No
Support-aware validation-weighted, uncalibrated	0.902	0.586	0.188	0.150	No
Support-aware equal, calibrated	0.899	0.583	0.161	0.038	Yes
Support-aware validation-weighted, calibrated	0.899	0.585	0.161	0.038	No

The two calibrated support-aware configurations occupied a distinct region of the discrimination-calibration space (Fig. 14). The support-aware equally averaged calibrated configuration achieved a mean overall accuracy of 0.899, macro Dice of 0.583, Brier score of 0.161 and ECE of 0.038. The support-aware validation-weighted calibrated configuration produced very similar performance, with an accuracy of 0.899, macro Dice of 0.585, Brier score of 0.161 and ECE of 0.038.

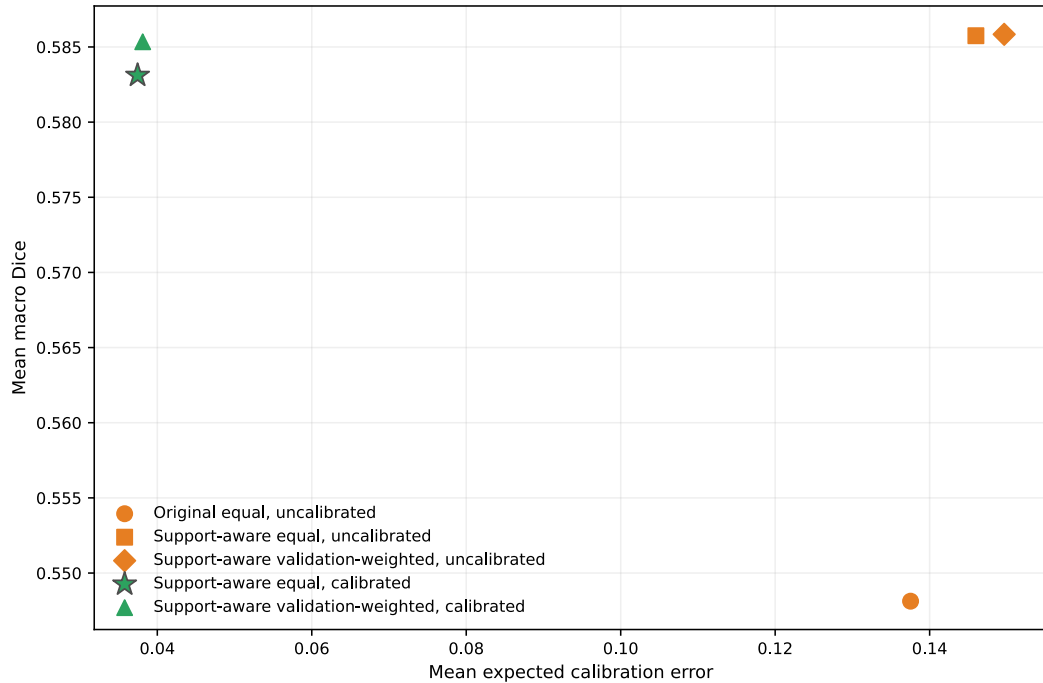


Figure 14. Discrimination-calibration trade-off across ensemble variants. Mean macro Dice versus mean expected calibration error for the five evaluated ensemble configurations. The support-aware equally averaged calibrated ensemble is identified as the selected configuration

Although the validation-weighted, calibrated configuration achieved a marginally higher mean macro Dice than the equally averaged configuration (0.585 compared with 0.583), the two configurations had identical accuracy and Brier scores at the reported three-decimal precision. Both calibrated configurations had an ECE of 0.038 at the reported three-decimal precision. The support-aware, equally averaged, calibrated ensemble was therefore retained as the primary configuration because it achieved essentially equivalent accuracy, macro Dice, Brier score and ECE to the validation-weighted configuration, while avoiding additional dependence on validation-derived ensemble weights.

The nested leave-one-sample-out design maintained separation between model development, probability calibration and outer-fold evaluation (Table 8). For each outer fold, the designated test sample was excluded while the remaining samples were further divided into inner training and validation subsets used during model development and estimation of the single-parameter logit rescaling parameter. Consequently, the reported performance represents transfer to physically independent drill-core samples rather than random partitioning of neighbouring grid cells within the same specimen.

Table 8. Nested leave-one-sample-out design and inner-validation logit-rescaling results. The outer held-out sample was excluded from all model-development and calibration decisions

Outer held-out sample	Inner validation sample	Inner training samples	Fitted scaling parameter, T_r	Validation NLL before	Validation NLL after	Calibration grid cells
NA25FC054	NA25FC132	NA25FC154, NA25FC155	0.173	0.582	0.133	103,823
NA25FC054	NA25FC154	NA25FC132, NA25FC155	1.017	0.381	0.381	126,090
NA25FC054	NA25FC155	NA25FC132, NA25FC154	0.624	0.501	0.426	71,207
NA25FC132	NA25FC054	NA25FC154, NA25FC155	0.673	0.314	0.267	86,848
NA25FC132	NA25FC154	NA25FC054, NA25FC155	0.935	0.508	0.506	132,489
NA25FC132	NA25FC155	NA25FC054, NA25FC154	0.829	0.424	0.412	71,207
NA25FC154	NA25FC054	NA25FC132, NA25FC155	0.438	0.376	0.231	83,646
NA25FC154	NA25FC132	NA25FC054, NA25FC155	0.572	0.578	0.470	103,823
NA25FC154	NA25FC155	NA25FC054, NA25FC132	0.851	0.397	0.387	71,207
NA25FC155	NA25FC054	NA25FC132, NA25FC154	0.519	0.397	0.262	86,848
NA25FC155	NA25FC132	NA25FC054, NA25FC154	0.777	0.366	0.344	103,823
NA25FC155	NA25FC154	NA25FC054, NA25FC132	0.935	0.466	0.465	132,489

In summary, the four outer-fold experiments demonstrated that μ XRF–AMICS-supervised XCT-based mineralogical relationships could transfer between independent samples, although the strength of transfer was strongly class dependent. Matrix silicates showed consistently high generalisation, Quartz and Garnet retained useful discrimination where sufficient labelled support was available, and Fe-bearing sulphides and oxides remained substantially more ambiguous. The support-aware equally averaged calibrated configuration was therefore retained as the final modelling strategy for subsequent volumetric application.

4.6 Three-dimensional application and spatial consistency of predicted mineralogical classes

Following the surface-based mineralogical validation, the selected calibrated modelling configuration was applied to the XCT volume of NA25FC154 to examine the three-dimensional spatial behaviour of the predicted grouped mineralogical classes. NA25FC154 was selected because all four model classes were represented with substantial μ XRF–AMICS support at the validated surface. Volumetric inference used the NA25FC154 outer-fold ensemble, for which NA25FC154 had been excluded from model

training, XCT scaling, checkpoint selection, probability calibration and ensemble weighting. The mask-aware volumetric application retained 118 parallel XCT planes extending from 0.3 to 23.7 mm inward from the validated surface at 0.2 mm spacing. No additional μ XRF-AMICS information was introduced during this stage.

Within the stable interior, the predicted hard-class maps retained the principal spatial organisation observed near the validated surface (Fig. 15a). Matrix silicates remained the dominant class, while larger predicted regions assigned to Quartz, Garnet and the Fe-bearing sulphides and oxides class persisted across multiple successive planes. Changes in categorical predictions generally developed progressively with inward distance rather than as widespread discontinuous rearrangements between neighbouring planes. The corresponding entropy maps showed coherent spatial structure, with elevated uncertainty concentrated around class boundaries and internally heterogeneous regions (Fig. 15b). At greater inward distances, the coloured areas became progressively smaller because parallel sampling planes intersected less rock as they approached the curved rear boundary of the half-core. At 20.1 and 23.7 mm, the connected rock mask occupied approximately 59% and 13% of the rectangular prediction domain, respectively. White regions therefore represent grid cells outside the rock mask rather than missing predictions. The 20.1 mm plane was treated as the last stable-interior plane, beyond which the results were treated as boundary-proximal.

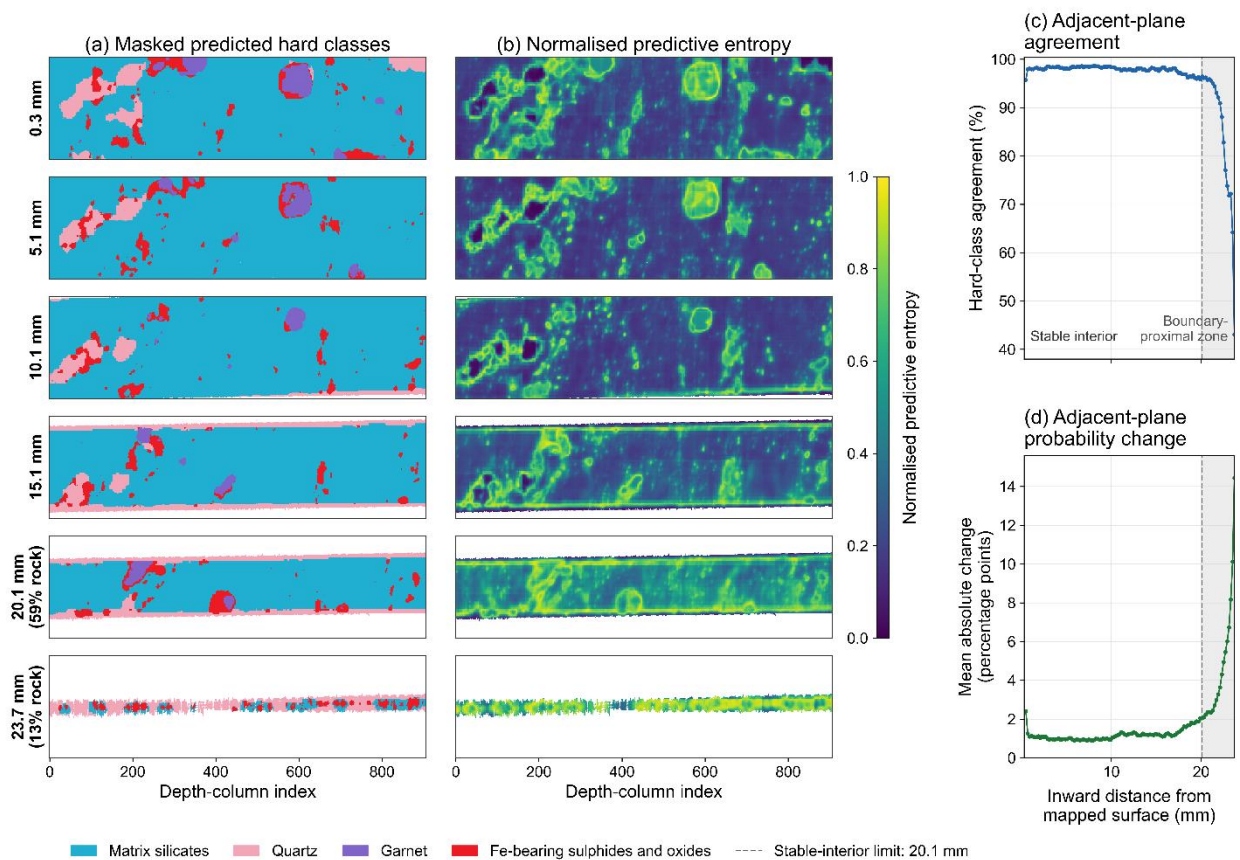


Figure 15. Three-dimensional application of the calibrated model and spatial consistency of predicted mineralogical classes in NA25FC154: (a, b) Predicted grouped mineralogical classes and normalised predictive entropy at representative inward distances. White areas lie outside the rock mask, and percentages indicate the retained

valid-rock fraction; (c, d) Adjacent-plane hard-class agreement and mean absolute probability-vector change, reported in percentage points. The dashed line at 20.1 mm separates the stable interior planes from the shaded boundary-proximal zone

The complete mask-retained sequence contained 20,003,727 valid-rock voxels across 118 planes (Table 9). Within the stable interior interval from 0.3 to 20.1 mm, 18,537,066 valid-rock voxels were retained, corresponding to a mean fraction of valid-rock voxels of 0.909 relative to the rectangular prediction domain. The corresponding mean value for the boundary-proximal zone was 0.400.

Within the stable interior zone, the voxel-weighted mean maximum class probability was 0.847 and the mean normalised predictive entropy was 0.360. Weighted adjacent-plane hard-class agreement was 0.980, and agreement for every adjacent-plane pair was at least 0.957. Hard-class agreement between the first retained plane (0.3 mm) and the plane at 20.1 mm, evaluated over 119,606 common valid-rock grid cells, was 0.662. This lower long-range agreement was compatible with the high adjacent-plane agreement because incremental changes accumulated across successive planes.

Prediction behaviour changed substantially in the boundary-proximal zone. Mean maximum class probability decreased to 0.635, normalised predictive entropy increased to 0.673 and weighted adjacent-plane agreement decreased to 0.880. The final transition at 23.7 mm had an adjacent-plane agreement of approximately 0.430. These changes coincided with rapidly diminishing valid-rock support and were therefore interpreted primarily as boundary-proximal sensitivity rather than evidence of a genuine mineralogical transition.

Table 9. Mask-aware volumetric support, uncertainty and spatial consistency in NA25FC154

Region	Inward distance (mm)	Planes	Valid-rock voxels	Valid fraction	Mean normalised predictive entropy	Mean maximum probability	Weighted adjacent-plane agreement
All mask-retained sequence	0.3-23.7	118	20,003,727	0.832	0.383	0.831	0.973
Stable interior zone	0.3-20.1	100	18,537,066	0.909	0.360	0.847	0.980
Boundary-proximal zone	20.3-23.7	18	1,466,661	0.400	0.673	0.635	0.880

The probability fields showed corresponding spatial behaviour: mean absolute probability change remained low through most of the stable-interior planes but increased sharply within the boundary-proximal zone, particularly near the final retained planes (Fig. 15d).

Because Matrix silicates dominated the predicted volume, overall adjacent-plane agreement could overstate the stability of minority mineralogical classes. Class-specific self-retention was therefore evaluated separately within the stable interior zone (Table 10). Matrix silicates showed the greatest adjacent-plane self-retention at 0.988, followed by Quartz at 0.964, Garnet at 0.931 and Fe-bearing sulphides and oxides at 0.867. The Fe-

bearing sulphides and oxides class had the lowest self-retention, consistent with its weaker cross-sample discrimination and broader overlapping XCT responses. Class proportions in the boundary-proximal zone were not interpreted directly because declining rock support, partial-volume effects and increasing uncertainty affected the predicted response distributions.

Table 10. Class-specific predicted response and spatial retention within the stable interior zone of NA25FC154

Class	Mean probability	Hard-class fraction	Adjacent-plane self-retention
Matrix silicates	0.774	0.860	0.988
Quartz	0.091	0.073	0.964
Garnet	0.053	0.023	0.931
Fe-bearing sulphides and oxides	0.082	0.044	0.867

These results show that the model, validated using independently held-out μ XRF–AMICS surfaces, could be applied to volumetric XCT data to provide spatially coherent predictions of grouped mineralogical classes throughout the stable interior of the evaluated core volume, together with class-probability and uncertainty fields. The spatial-consistency metrics supported continuity of the predictions between successive inward planes. Direct physical validation of the interior predictions would require independent mineralogical measurements at depth and was beyond the scope of this study.

5. Discussion

The results presented above demonstrate that XCT measurements contain reproducible information related to mineralogy, while also showing that mineralogical interpretation cannot be reduced to simple attenuation thresholds. The four μ XRF–AMICS-derived grouped mineralogical classes occupied different XCT response distributions, but with substantial overlaps between several classes. Garnet exhibited the most distinctive higher-attenuation response, whereas Matrix silicates, Quartz and Fe-bearing sulphides and oxides showed considerable overlaps. The strongest overlap occurred between Matrix silicates and Fe-bearing sulphides and oxides. This behaviour is expected because XCT attenuation reflects the combined X-ray interaction of the analysed materials rather than mineral identity directly. Different minerals can therefore exhibit comparable responses, while variations in mineral composition, grain geometry, partial-volume effects and neighbouring phases can broaden the response associated with a nominal mineralogical class (Evans et al., 2015; Wang and Miller, 2020; Warlo et al., 2021).

The sample-specific distributions further demonstrate that the relationship between attenuation and mineralogy is not completely invariant among samples. This is particularly evident for the Fe-bearing sulphides and oxides, whose response was broader and more variable between samples than that of the dominant Matrix-silicate class. A classifier based exclusively on fixed attenuation intervals would consequently be vulnerable to both class overlap and sample-specific shifts in XCT responses. The

convolutional framework developed in this work instead learns from XCT responses together with local and multi-scale spatial context. The resulting predictions should therefore be interpreted as learned XCT responses associated with the μ XRF–AMICS-defined mineralogical groups, rather than as direct identification of minerals from attenuation alone.

The mineral-class grouping was also central to the modelling framework. The Matrix silicates class comprised plagioclase, biotite and K-feldspar, whereas the Fe-bearing sulphides and oxides class comprised Fe oxides, chalcopyrite and pyrite. The Matrix silicates class may also include unresolved white mica. This potential within-class ambiguity does not affect the four-class modelling framework but limits mineral-species-level interpretation within the Matrix silicates class. These grouped classes were defined on the basis of geometallurgical relevance, available labelled support and expected XCT separability, rather than an assumption that their constituent minerals were physically identical. Quartz and Garnet were retained as separate classes where their occurrence and XCT responses provided sufficient support for more specific prediction. The resulting four-class representation balanced mineralogical specificity against the statistical support required for transferable modelling (Warlo et al., 2021; Kuva et al., 2025).

A major strength of the study was the validation at the level of independent samples rather than randomly partitioned surface grid cells. Spatially neighbouring XCT grid cells are strongly correlated, and conventional random training-test splits would allow closely related portions of the same mineralogical structures to occur in both datasets. The leave-one-sample-out design provided a substantially more demanding test because each evaluation specimen was physically excluded from the corresponding model development process. The nested structure additionally separated outer testing from inner model selection and probability calibration. The resulting performance therefore provided support for the transfer of learned XCT–mineralogical relationships between samples rather than the reproduction of spatial patterns already represented during training.

The difference between overall accuracy and class-specific performance was particularly important. Overall accuracy remained high across the four outer folds, but this metric was strongly influenced by the dominant Matrix silicates class. Matrix silicates achieved consistently high Dice values across all held-out samples, indicating that their XCT responses and spatial characteristics were comparatively transferable. Quartz also showed useful transferability, although its performance varied between samples, indicating sensitivity to differences in occurrence, morphology and surrounding mineralogical context.

Garnet showed moderate to good transferability where sufficient labelled grid cells were available. This result was consistent with its comparatively distinctive higher-attenuation XCT response. However, the absence of Garnet from NA25FC132 and its representation by only three labelled grid cells in NA25FC155 illustrated an important limitation of the mineralogical dataset. Accordingly, class-specific performance was not quantified when the held-out label map contained insufficient representation of a class, and such cases were reported as insufficiently supported rather than as failed predictions.

The Fe-bearing sulphides and oxides class was substantially more difficult to generalise. Its low Dice values persisted despite meaningful labelled support, indicating that the weakness cannot be attributed solely to data scarcity. This weaker performance likely reflected a combination of broad and overlapping XCT responses, internal mineralogical heterogeneity, limited diversity of independent μ XRF–AMICS supervision, partial-volume effects and residual registration uncertainty (Kuva et al., 2025). Fe oxides, chalcopyrite and pyrite are compositionally and texturally distinct minerals, and their XCT expressions can vary according to composition, grain size, morphology and association with surrounding phases. Combining these minerals increases labelled support and produces a geometallurgically meaningful response of the Fe-bearing sulphides and oxides, but it also increases intra-class variability. The resulting prediction should therefore be treated as a response of the Fe-bearing sulphides and oxides rather than a quantitative determination of an individual sulphide or oxide mineral.

The probabilistic formulation was particularly important under these conditions. When two or more classes had overlapping XCT characteristics, a categorical maximum-probability decision obscured the degree of ambiguity between them. The calibrated probability maps retained this information and therefore provided a more complete representation of model behaviour. The single-parameter logit rescaling improved agreement between predicted confidence and the observed fraction of correctly classified labelled grid cells, while leaving hard-class discrimination almost unchanged (Guo et al., 2017). This distinction was important because calibration did not make previously inseparable minerals distinguishable; instead, it improved the statistical interpretation of confidence in the existing predictions.

The reduction in the mean absolute error between probability-derived and μ XRF–AMICS-labelled class proportions after calibration was also relevant to future drill-core interpretation because probability fields could be spatially aggregated without assigning every uncertain grid cell to a single class. Hard-class proportions remained useful descriptive outputs, whereas mean class probabilities retained information from ambiguous regions. The associated entropy field further identified areas in which multiple mineralogical interpretations remained plausible. Consequently, calibrated probabilities and uncertainty estimates were more informative than hard classifications alone for uncertainty-aware characterisation.

The ensemble comparison reinforces this interpretation. Support-aware aggregation improved class-balanced discrimination by preventing models with insufficient training support for a particular mineralogical class from contributing equally to that class. Calibration then produced the main improvement in the quality of probability estimates. The additional use of validation-derived ensemble weighting provided little practical advantage over equal averaging. Given the limited number of independent specimens available in each inner validation configuration, retaining the simpler equally averaged calibrated ensemble avoids introducing unnecessary dependence on small validation subsets while preserving essentially equivalent performance.

Application of the selected calibrated ensemble throughout the NA25FC154 XCT volume showed that the predicted grouped mineralogical classes remained spatially coherent

across much of the available half-core thickness. Within the stable interior interval from 0.3 to 20.1 mm, weighted adjacent-plane agreement was 0.980, mean maximum class probability was 0.847 and mean normalised predictive entropy was 0.360. Because adjacent planes were separated by only 200 μm , high agreement between neighbouring planes was expected. It was therefore considered together with agreement between the first retained plane and the plane at 20.1 mm, class-specific self-retention and uncertainty. Matrix silicates showed the greatest self-retention, whereas the Fe-bearing sulphides and oxides class was less stable, consistent with the class-specific differences identified during independent surface validation.

A key limitation was that mineralogical supervision was available only at the two-dimensional μXRF –AMICS surfaces, whereas the model was applied throughout a three-dimensional XCT volume. Co-registration allowed mineralogical information from the analysed surfaces to supervise the corresponding XCT responses, and the outer leave-one-sample-out evaluation provided direct validation at independently held-out surfaces. Application to the interior planes subsequently inferred grouped mineralogical classes from the learned relationships. Independent verification of these interior predictions would require mineralogical measurements from within the core volume, which were beyond the scope of this study (Butcher, 2020; Warlo et al., 2021; Kuva et al., 2025).

From a mining and beneficiation perspective, the developed framework provides a route for extending information from targeted mineralogical measurements into larger non-destructively scanned drill-core volumes. In general, automated mineralogy provides high-value compositional information but is commonly available only for selected surfaces or samples, whereas XCT supplies continuous information on internal attenuation responses and spatial structures. Training a probabilistic model using carefully co-registered mineralogical supervision therefore provides a mechanism for translating sparse mineralogical observations into broader XCT-based mineralogical-response information. The resulting probability and uncertainty fields could support geometallurgical domain definition, selection of intervals for additional mineralogical or metallurgical testing and identification of internal variability potentially relevant to comminution, liberation and downstream separation behaviour. The practical value lies not in replacing detailed mineralogical analysis, but in directing further investigation and integrating mineralogical-response variability with other geological and geometallurgical measurements. Future integration with comminution, liberation and recovery data is required before direct process-performance predictions can be made. Retaining calibrated probabilities and uncertainty remains essential because these outputs distinguish relatively confident predictions from regions where overlapping XCT responses permit multiple plausible mineralogical interpretations.

Several limitations remained. Mineralogical supervision was available from only four paired μXRF –AMICS and XCT surfaces. Within each outer fold, each network was trained on two samples and validated on the third development sample. The four samples came from three drill holes, with NA25FC154 and NA25FC155 both originating from DH-C. Although each outer test sample was excluded from model development,

the available data assessed transfer between specimens rather than fully independent transfer between drill holes. The same four outer folds were also used to compare ensemble configurations; consequently, the retained configuration requires confirmation using an additional independent dataset. Within the scope of this study, mask-aware volumetric inference was demonstrated for one representative sample, NA25FC154, selected because all four model classes had substantial labelled representation at its validated surface. Although the total number of labelled grid cells was large, neighbouring cells within each surface were spatially correlated and represented a limited range of geological and mineralogical associations. A large number of grid cells therefore did not provide the same diversity as a larger number of independent samples. This limitation was particularly relevant to Quartz, Garnet and the Fe-bearing sulphides and oxides class, which were absent or unevenly represented among the available surfaces. The resulting training data captured only a restricted range of grain sizes, textures, mineral associations and XCT responses for these minority classes, which likely contributed to their more variable cross-sample performance. Additional paired μ XRF–AMICS and XCT samples from different drill holes, alteration domains and mineralisation styles would broaden the training distribution, improve minority-class representation and provide a stronger assessment of model transferability.

The accuracy of spatial co-registration was another important consideration because μ XRF–AMICS provided supervision on a 200 μ m categorical grid. Misalignment near mineral boundaries can introduce incorrect training labels even when both datasets are individually accurate. The strict registration and quality-control procedures used in this study reduced this source of uncertainty, but registration error cannot be eliminated completely when independently acquired datasets are combined.

The Fe-bearing sulphides and oxides class represents the clearest target for future methodological improvement. Additional mineralogical supervision could establish whether its constituent phases (Fe oxides, chalcopyrite and pyrite) could be modelled separately or whether their XCT responses remained intrinsically ambiguous under the available acquisition conditions. Additional sensing modalities could also be explored if they provided complementary information capable of resolving attenuation overlaps. Any increase in mineralogical specificity should, however, be supported by sufficient independent labelled data rather than by creating statistically under-represented classes.

Overall, this study demonstrated that sparse two-dimensional mineralogical supervision could establish transferable probabilistic relationships between grouped mineralogical classes and XCT responses and identified the conditions under which those relationships transferred reliably. Transfer was strongest for classes with consistent XCT characteristics and adequate training representation, whereas overlapping and internally heterogeneous classes retained greater uncertainty. Explicit representation of this uncertainty was therefore essential for scientifically defensible XCT-based mineralogical characterisation.

6. Conclusions

In this study, a probabilistic framework was developed that integrates co-registered μ XRF–AMICS mineralogical information with XCT data to predict grouped

mineralogical classes in drill cores. The mineralogy in the case study was represented by Matrix silicates, Quartz, Garnet and Fe-bearing sulphides and oxides, while XCT attenuation responses served as model inputs.

The results demonstrate that XCT responses contain information systematically associated with the grouped mineralogical classes, but with substantial overlaps between several mineral groups. Garnet displayed a comparatively distinctive response, whereas Matrix silicates, Quartz and Fe-bearing sulphides and oxides showed greater overlaps. This limits the suitability of deterministic classification thresholds and supports probabilistic modelling that incorporates both XCT responses and spatial context.

Nested leave-one-sample-out validation demonstrated that the learned relationships can transfer between independent drill-core samples, although transferability is strongly class dependent. Matrix silicates were predicted most consistently, while Quartz and Garnet showed useful generalisation where sufficient labelled-grid support was available. Fe-bearing sulphides and oxides remained more difficult to discriminate because of their broad and internally heterogeneous XCT responses and the limited diversity of available mineralogical supervision.

Probability calibration improved the reliability and interpretability of the predicted class probabilities without materially altering the underlying categorical discrimination. Retaining calibrated probability and uncertainty information is therefore important where mineralogical classes cannot be uniquely distinguished from XCT data.

Application of the surface-validated framework to 118 mask-retained XCT planes extending from 0.3 to 23.7 mm inward from the verified XCT model surface in NA25FC154 produced spatially coherent predictions of grouped mineralogical classes with probability and uncertainty fields. Within the stable interior interval from 0.3 to 20.1 mm, weighted adjacent-plane agreement reached 0.980, while mean maximum class probability and mean normalised predictive entropy were 0.847 and 0.360, respectively. Prediction confidence and spatial agreement decreased near the curved rear boundary as valid-rock support diminished. These results demonstrated the feasibility of mask-aware volumetric deployment and provided useful insight into spatially consistent mineralogical-class patterns within the core interior. Independent mineralogical measurements at depth will be needed to physically validate the interior predictions.

Overall, the developed framework provides a reproducible basis for using surface mineralogical observations to generate three-dimensional predictions of grouped mineralogical classes from non-destructive XCT measurements while explicitly retaining prediction uncertainties. Future work should prioritise additional performance assessment of the proposed framework using a more comprehensive set of paired μ XRF–AMICS and XCT samples spanning a wider range of drill holes, alteration domains, mineralisation styles and minority mineralogical classes to improve training diversity and provide a stronger assessment of transferability.

Author Contributions

Conceptualization: M.N., P.D., and C.X.; Methodology: M.N.; Data acquisition: F.C.-A., S.B., T.H.; Writing – original draft preparation: M.N.; Writing – review and editing: M.N., F.C.-A., S.B., T.H., P.D., and C.X.; Supervision: P.D. and C.X. All authors have read and agreed to the published version of the manuscript.

Funding

This research was supported by the Australian Research Council Industrial Transformation Training Centre for Integrated Operations for Complex Resources (ARC ITTC IOCR; project number IC190100017) and funded by universities, industry and the Australian Government.

Data Availability Statement

The data used in this study cannot be shared publicly due to commercial sensitivity and confidentiality agreements. The scripts developed by the authors are available from the corresponding author upon reasonable request.

Acknowledgments

The authors gratefully acknowledge Boliden Mineral AB for providing access to the data used in this study.

Conflicts of Interest

The authors declare no conflict of interest.

References

- Baker, D.R., Mancini, L., Polacci, M., Higgins, M.D., Gualda, G.A.R., Hill, R.J., Rivers, M.L., 2012. An introduction to the application of X-ray microtomography to the three-dimensional study of igneous rocks. *Lithos* 148, 262–276. <https://doi.org/10.1016/j.lithos.2012.06.008>
- Bam, L.C., Miller, J.A., Becker, M., 2020. A mineral X-ray linear attenuation coefficient tool (MXLAC) to assess mineralogical differentiation for X-ray computed tomography scanning. *Minerals* 10, 441. <https://doi.org/10.3390/min10050441>
- Bauer, T.E., Lynch, E.P., Sarlus, Z., Drejing-Carroll, D., Martinsson, O., Metzger, N., Wanhainen, C., 2022. Structural controls on iron oxide copper-gold mineralization and related alteration in a Paleoproterozoic supracrustal belt: Insights from the Nautanen deformation zone and surroundings, northern Sweden. *Economic Geology* 117, 327–359. <https://doi.org/10.5382/econgeo.4862>
- Becker, M., Jardine, M.A., Miller, J.A., Harris, M., 2016. X-ray computed tomography — A geometallurgical tool for 3D textural analysis of drill core? In: *Proceedings of the Third AusIMM International Geometallurgy Conference (GeoMet 2016)*, Perth, Australia, 15–17 June 2016. Australasian Institute of Mining and Metallurgy, Melbourne, pp. 231–240.

- Bergman, S., Weihed, P., 2020. Archean (>2.6 Ga) and Paleoproterozoic (2.5–1.8 Ga), pre- and syn-orogenic magmatism, sedimentation and mineralization in the Norrbotten and Överkalix lithotectonic units, Svecokarelian orogen. Geological Society, London, Memoirs 50, 27–81. <https://doi.org/10.1144/M50-2016-29>
- Bergqvist, M., Landström, E., Hansson, A., Luth, S., 2019. Access to geological structures, density, minerals and textures through novel combination of 3D tomography, XRF and sample weight. ASEG Extended Abstracts 2019, 1–3. <https://doi.org/10.1080/22020586.2019.12073146>
- Brier, G.W., 1950. Verification of forecasts expressed in terms of probability. Monthly Weather Review 78, 1–3. [https://doi.org/10.1175/1520-0493\(1950\)078%3C0001:VOFEIT%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(1950)078%3C0001:VOFEIT%3E2.0.CO;2)
- Butcher, A.R., 2020. Upscaling of 2D mineralogical information to 3D volumes for geoscience applications using a multi-scale, multi-modal and multi-dimensional approach. IOP Conference Series: Materials Science and Engineering 891, 012006. <https://doi.org/10.1088/1757-899X/891/1/012006>
- Butcher, A.R., Dehaine, Q., Menzies, A.H., Michaux, S.P., 2023. Characterisation of ore properties for geometallurgy. Elements 19, 352–358. <https://doi.org/10.2138/gselements.19.6.352>
- Cnudde, V., Boone, M.N., 2013. High-resolution X-ray computed tomography in geosciences: A review of the current technology and applications. Earth-Science Reviews 123, 1–17. <https://doi.org/10.1016/j.earscirev.2013.04.003>
- Dominy, S.C., O'Connor, L., Parbhakar-Fox, A., Glass, H.J., Purevgerel, S., 2018. Geometallurgy — A route to more resilient mine operations. Minerals 8, 560. <https://doi.org/10.3390/min8120560>
- Drejing-Carroll, D., Hitzman, M.W., Coller, D., 2023. Geology of the Nautanen North Cu–Au–Ag–(Mo) deposit, Norrbotten, Sweden. Economic Geology 118, 1765–1794. <https://doi.org/10.5382/econgeo.5019>
- Ehrig, K., Liebezeit, V., Smith, M., Macmillan, E., Lower, C., 2014. Geologists and the value chain—How material characterisation by modern mineralogy can optimise design and operation of processing facilities. In: Proceedings of the 9th International Mining Geology Conference, Adelaide, Australia, 18–20 August 2014, pp. 5–13.
- Evans, C.L., Wightman, E.M., Yuan, X., 2015. Quantifying mineral grain size distributions for process modelling using X-ray micro-tomography. Minerals Engineering 82, 78–83. <https://doi.org/10.1016/j.mineng.2015.03.026>

- Flude, S., Haschke, M., Storey, M., 2017. Application of benchtop micro-XRF to geological materials. *Mineralogical Magazine* 81, 923–948.
<https://doi.org/10.1180/minmag.2016.080.150>
- Godel, B., 2013. High-resolution X-ray computed tomography and its application to ore deposits: From data acquisition to quantitative three-dimensional measurements with case studies from Ni–Cu–PGE deposits. *Economic Geology* 108, 2005–2020.
<https://doi.org/10.2113/econgeo.108.8.2005>
- Gottlieb, P., Wilkie, G., Sutherland, D., Ho-Tun, E., Suthers, S., Perera, K., Jenkins, B., Spencer, S., Butcher, A., Rayner, J., 2000. Using quantitative electron microscopy for process mineralogy applications. *JOM* 52, 24–25. <https://doi.org/10.1007/s11837-000-0126-9>
- Gu, Y., 2003. Automated scanning electron microscope based mineral liberation analysis: An introduction to JKMRC/FEI Mineral Liberation Analyser. *Journal of Minerals and Materials Characterization and Engineering* 2, 33–41.
<https://doi.org/10.4236/jmmce.2003.21003>
- Guntoro, P.I., Tiu, G., Ghorbani, Y., Lund, C., Rosenkranz, J., 2019. Application of machine learning techniques in mineral phase segmentation for X-ray microcomputed tomography (μ CT) data. *Minerals Engineering* 142, 105882.
<https://doi.org/10.1016/j.mineng.2019.105882>
- Guntoro, P.I., Ghorbani, Y., Butcher, A.R., Kuva, J., Rosenkranz, J., 2020. Textural quantification and classification of drill cores for geometallurgy: Moving toward 3D with X-ray microcomputed tomography (μ CT). *Natural Resources Research* 29, 3547–3565. <https://doi.org/10.1007/s11053-020-09685-5>
- Guo, C., Pleiss, G., Sun, Y., Weinberger, K.Q., 2017. On calibration of modern neural networks. In: *Proceedings of the 34th International Conference on Machine Learning*. *Proceedings of Machine Learning Research* 70, 1321–1330.
<https://proceedings.mlr.press/v70/guo17a.html>
- Hanna, R.D., Ketcham, R.A., 2017. X-ray computed tomography of planetary materials: A primer and review of recent studies. *Geochemistry* 77, 547–572.
<https://doi.org/10.1016/j.chemer.2017.01.006>
- Hariharan, B., Arbeláez, P., Girshick, R., Malik, J., 2015. Hypercolumns for object segmentation and fine-grained localization. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 447–456.
<https://doi.org/10.1109/CVPR.2015.7298642>

- Harraden, C.L., Steiner, A.P., Barker, S.L.L., Bouzari, F., 2025. Mineralogy from μ XRF elemental maps: Comparison of supervised and unsupervised methods. *Geochemistry: Exploration, Environment, Analysis* 25, geochem2024-050. <https://doi.org/10.1144/geochem2024-050>
- Houshmand, N., Goodfellow, S., Esmaili, K., Ordóñez-Calderón, J.C., 2022. Rock type classification based on petrophysical, geochemical and core imaging data using machine and deep learning techniques. *Applied Computing and Geosciences* 16, 100104. <https://doi.org/10.1016/j.acags.2022.100104>
- Hrstka, T., Gottlieb, P., Skála, R., Breiter, K., Motl, D., 2018. Automated mineralogy and petrology — Applications of TESCAN Integrated Mineral Analyzer (TIMA). *Journal of Geosciences* 63, 47–63. <https://doi.org/10.3190/jgeosci.250>
- Joseph, P.M., Spital, R.D., 1982. The effects of scatter in X-ray computed tomography. *Medical Physics* 9, 464–472. <https://doi.org/10.1118/1.595111>
- Ketcham, R.A., Carlson, W.D., 2001. Acquisition, optimization and interpretation of X-ray computed tomographic imagery: Applications to the geosciences. *Computers & Geosciences* 27, 381–400. [https://doi.org/10.1016/S0098-3004\(00\)00116-3](https://doi.org/10.1016/S0098-3004(00)00116-3)
- Kuva, J., Jooshaki, M., Jolis, E.M., Sammaljärvi, J., Siitari-Kauppi, M., Jankovský, F., Zuna, M., Bischoff, A., Sardini, P., 2025. Characterizing heterogeneous rocks in 3D with a multimodal deep learning approach — Implications for transport simulations. *Tomography of Materials and Structures* 7, 100055. <https://doi.org/10.1016/j.tmater.2025.100055>
- Kyle, J.R., Ketcham, R.A., 2015. Application of high-resolution X-ray computed tomography to mineral deposit origin, evaluation and processing. *Ore Geology Reviews* 65, 821–839. <https://doi.org/10.1016/j.oregeorev.2014.09.034>
- Laforce, B., Masschaele, B., Boone, M.N., Schaubroeck, D., Dierick, M., Vekemans, B., Walgraeve, C., Janssen, C., Cnudde, V., Van Hoorebeke, L., Vincze, L., 2017. Integrated three-dimensional microanalysis combining X-ray microtomography and X-ray fluorescence methodologies. *Analytical Chemistry* 89, 10617–10624. <https://doi.org/10.1021/acs.analchem.7b03205>
- Lamberg, P., 2011. Particles — The bridge between geology and metallurgy. In: *Proceedings of the Conference in Minerals Engineering, Luleå, Sweden, 8–9 February 2011*. Luleå University of Technology, Luleå.
- Lätti, D., Adair, B.J.I., 2001. An assessment of stereological adjustment procedures. *Minerals Engineering* 14, 1579–1587. [https://doi.org/10.1016/S0892-6875\(01\)00176-5](https://doi.org/10.1016/S0892-6875(01)00176-5)

- Lishchuk, V., Koch, P.-H., Ghorbani, Y., Butcher, A.R., 2020. Towards integrated geometallurgical approach: Critical review of current practices and future trends. *Minerals Engineering* 145, 106072. <https://doi.org/10.1016/j.mineng.2019.106072>
- Loshchilov, I., Hutter, F., 2019. Decoupled weight decay regularization. In: *Proceedings of the 7th International Conference on Learning Representations (ICLR)*. <https://openreview.net/forum?id=Bkg6RiCqY7>
- Lund, C., Lamberg, P., Lindberg, T., 2013. Practical way to quantify minerals from chemical assays at Malmberget iron ore operations—An important tool for the geometallurgical program. *Minerals Engineering* 49, 7–16. <https://doi.org/10.1016/j.mineng.2013.04.005>
- Månbro, C., Parian, M., 2023. Chemical and mineralogical characterisation of iron ore drill core using μ XRF. In: *Proceedings of the Digital Conference in Minerals Engineering, Luleå, Sweden, 7–8 February 2023*. Luleå University of Technology, Luleå.
- Melfou, M., Richard, A., Tarantola, A., Villeneuve, J., Carr, P., Peiffert, C., Mercadier, J., Dean, B., Drejing-Carroll, D., 2023. Tracking the origin of metasomatic and ore-forming fluids in IOCG deposits through apatite geochemistry: Nautanen North deposit, Norrbotten, Sweden. *Lithos* 438–439, 106995. <https://doi.org/10.1016/j.lithos.2022.106995>
- Menzies, A.H., Gloy, G., Scheller, S., Álvarez, E., Tagle, R., 2019. Automated mineralogy: Analysis of REE samples using micro-XRF. *Applied Earth Science* 128, 54–55. <https://doi.org/10.1080/25726838.2019.1605713>
- Menzies, A., Boehm, S., 2021. Dual source X-ray and electron SEM system: Elemental mapping of an epithermal gold-bearing sample from Karangahake, New Zealand. *Microscopy and Microanalysis* 27, 456–460. <https://doi.org/10.1017/S1431927621002130>
- Menzies, A.H., Tagle, R., Reinhardt, F., Hirschle, C., Schellkopf, L.J., Kelly, N., 2022. High-speed micro-XRF analysis of rock samples and drill cores. *Microscopy and Microanalysis* 28(S1), 650–651. <https://doi.org/10.1017/S1431927622003117>
- Merrill-Cifuentes, J., Cracknell, M.J., Escolme, A., 2022. Unsupervised textural classification of rocks in large imagery datasets. *Minerals Engineering* 180, 107496. <https://doi.org/10.1016/j.mineng.2022.107496>
- Nasretdinova, M., Dowd, P., Xu, C., 2026. X-ray computed tomography for geometallurgical ore texture analysis: Implications for comminution response. *Minerals Engineering* 249, 110814. <https://doi.org/10.1016/j.mineng.2026.110814>

- Nguyen, K., 2013. A new texture analysis technique for geometallurgy. In: Proceedings of the Second AusIMM International Geometallurgy Conference, Brisbane, Australia, pp. 187–190.
- Orexplore, 2021. GeoCore X10. Orexplore Technologies Ltd. Available at: <https://www.orexplore.com/what-we-do/geocore-x10/> (accessed 2 September 2026).
- Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9, 62–66. <https://doi.org/10.1109/TSMC.1979.4310076>
- Parian, M., Lamberg, P., Möckel, R., Rosenkranz, J., 2015. Analysis of mineral grades for geometallurgy: Combined element-to-mineral conversion and quantitative X-ray diffraction. Minerals Engineering 82, 25–35. <https://doi.org/10.1016/j.mineng.2015.04.023>
- Pereira, L., Schach, E., Tolosana-Delgado, R., Frenzel, M., 2023. All about particles: Modelling ore behaviour in mineral processing. Elements 19, 359–364. <https://doi.org/10.2138/gselements.19.6.359>
- Pérez-Barnuevo, L., Lévesque, S., Bazin, C., 2018. Drill core texture as geometallurgical indicator for the Mont-Wright iron ore deposit, Quebec, Canada. Minerals Engineering 122, 130–141. <https://doi.org/10.1016/j.mineng.2018.03.020>
- Phan, J., Ruspini, L.C., Lindseth, F., 2021. Automatic segmentation tool for 3D digital rocks by deep learning. Scientific Reports 11, 19123. <https://doi.org/10.1038/s41598-021-98697-z>
- Reyes, F., Lin, Q., Udoudo, O., Dodds, C., Lee, P.D., Neethling, S.J., 2017. Calibrated X-ray micro-tomography for mineral ore quantification. Minerals Engineering 110, 122–130. <https://doi.org/10.1016/j.mineng.2017.04.015>
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015. Lecture Notes in Computer Science 9351. Springer, Cham, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
- Roy, A.G., Navab, N., Wachinger, C., 2018. Concurrent spatial and channel ‘squeeze & excitation’ in fully convolutional networks. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (Eds.), Medical Image Computing and Computer Assisted Intervention—MICCAI 2018. Lecture Notes in Computer Science 11070. Springer, Cham, pp. 421–429. https://doi.org/10.1007/978-3-030-00928-1_48

- Sayab, M., Suuronen, J.-P., Hölttä, P., Aerden, D., Lahtinen, R., Kallonen, A.P., 2015. High-resolution X-ray computed microtomography: A holistic approach to metamorphic fabric analyses. *Geology* 43, 55–58. <https://doi.org/10.1130/G36250.1>
- Sayab, M., Suuronen, J.-P., Molnár, F., Villanova, J., Kallonen, A., O'Brien, H., Lahtinen, R., Lehtonen, M., 2016. Three-dimensional textural and quantitative analyses of orogenic gold at the nanoscale. *Geology* 44, 739–742. <https://doi.org/10.1130/G38074.1>
- Sayab, M., Menzies, A., Palin, R.M., Butcher, A.R., Cook, N., Kuva, J., Dehaine, Q., 2025. Structural controls on cobalt mineralisation during regional metamorphism: An example from the Rajapalot area, Finland. *Mineralium Deposita* 60, 213–231. <https://doi.org/10.1007/s00126-024-01315-x>
- Schulz, B., Sandmann, D., Gilbricht, S., 2020. SEM-based automated mineralogy and its application in geo- and material sciences. *Minerals* 10, 1004. <https://doi.org/10.3390/min10111004>
- Simán, F., Jansson, N.F., Kampmann, T., Liwicki, F., 2021. Rock classification with machine learning: A case study from the Zinkgruvan Zn–Pb–Ag deposit, Bergslagen, Sweden. In: 2021 Swedish Artificial Intelligence Society Workshop (SAIS), pp. 1–5. <https://doi.org/10.1109/SAIS53221.2021.9483959>
- van den Boogaart, K.G., Tolosana-Delgado, R., 2018. Predictive geometallurgy: An interdisciplinary key challenge for mathematical geosciences. In: Daya Sagar, B.S., Cheng, Q., Agterberg, F. (Eds.), *Handbook of Mathematical Geosciences*. Springer, Cham, pp. 673–686. https://doi.org/10.1007/978-3-319-78999-6_33
- Viana, A.Z., Månbro, C., Jooshaki, M., Parian, M., 2025. Automated ore texture classification using μ XRF imaging and unsupervised machine learning: Correlation with surface hardness. *Minerals Engineering* 234, 109744. <https://doi.org/10.1016/j.mineng.2025.109744>
- Voigt, M.J., Miller, J., Bbosa, L., Govender, R.A., Bradshaw, D., Mainza, A., Becker, M., 2019. Developing a 3D mineral texture quantification method of drill core for geometallurgy. *Journal of the Southern African Institute of Mining and Metallurgy* 119, 347–353. <https://doi.org/10.17159/2411-9717/590/2019>
- Voigt, M., Miller, J.A., Mainza, A.N., Bam, L.C., Becker, M., 2020. The robustness of the gray-level co-occurrence matrices and X-ray computed tomography method for the quantification of 3D mineral texture. *Minerals* 10, 334. <https://doi.org/10.3390/min10040334>

- Waara, S., 2015. Garnet occurrence and its relationship to mineralization at the Nautanen deposit, northern Sweden. MSc thesis, Luleå University of Technology, Luleå, Sweden.
- Wang, Y., Miller, J.D., 2020. Current developments and applications of micro-CT for the 3D analysis of multiphase mineral systems in geometallurgy. *Earth-Science Reviews* 211, 103406. <https://doi.org/10.1016/j.earscirev.2020.103406>
- Wang, Y.D., Shabaninejad, M., Armstrong, R.T., Mostaghimi, P., 2021. Deep neural networks for improving physical accuracy of 2D and 3D multi-mineral segmentation of rock micro-CT images. *Applied Soft Computing* 104, 107185. <https://doi.org/10.1016/j.asoc.2021.107185>
- Wanhainen, C., 2005. On the origin and evolution of the Palaeoproterozoic Aitik Cu–Au–Ag deposit, northern Sweden: A porphyry copper–gold ore modified by multistage metamorphic-deformational, magmatic-hydrothermal and IOCG-mineralising events. PhD thesis, Luleå University of Technology, Luleå, Sweden.
- Warlo, M., Bark, G., Wanhainen, C., Butcher, A.R., Forsberg, F., Lycksam, H., Kuva, J., 2021. Multi-scale X-ray computed tomography analysis to aid automated mineralogy in ore geology research. *Frontiers in Earth Science* 9, 789372. <https://doi.org/10.3389/feart.2021.789372>
- Wildenschild, D., Sheppard, A.P., 2013. X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems. *Advances in Water Resources* 51, 217–246. <https://doi.org/10.1016/j.advwatres.2012.07.018>
- Willeminck, M.J., Noël, P.B., 2019. The evolution of image reconstruction for CT—From filtered back projection to artificial intelligence. *European Radiology* 29, 2185–2195. <https://doi.org/10.1007/s00330-018-5810-7>
- Williams, S.R., Richardson, J.M., 2004. Geometallurgical mapping: A new approach that reduces technical risk. SGS Minerals Services Technical Paper 2004-01, SGS Minerals Services.
- Wu, Y., He, K., 2018. Group normalization. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (Eds.), *Computer Vision—ECCV 2018. Lecture Notes in Computer Science* 11217. Springer, Cham, pp. 3–19. https://doi.org/10.1007/978-3-030-01261-8_1