Dear Editor,

Thank you for your handling of our manuscript. We also sincerely thank the two reviewers and Prof. Darby for their very useful comments and suggestions. These comments and suggestions are helped to improve the quality of our manuscript, and therefore we have given them full consideration in the revised draft. For your convenient review, all changes and additions in the revised manuscript have been highlighted in blue text. Below is our point-by-point response to the comments raised by the reviewers and community (Dennis Darby).

Sincerely yours

Zhengquan Yao on behalf of all co-authors

Reply to the reviewers' comments (RC1) on "Circum-Arctic Sediment PROvenance Database (CASPROD): A database of mineralogy and geochemistry for the Circum-Arctic surface sediments" by Yao et al.

I have reviewed the manuscript that presents the Circum-Arctic Sediment PROvenance Database (CASPROD). This database combines four key provenance proxies: Sr–Nd isotopes, detrital zircon U–Pb ages, clay mineral assemblages, and detrital mineral proportions from 4308 sampling stations across the Arctic Ocean. By synthesizing data from 117 previously published studies and incorporating newly available datasets from previously underrepresented regions, the authors have developed an important community resource for Arctic sediment provenance research and paleoenvironmental reconstructions. The manuscript is well-written and scientifically rigorous. I believe it is suitable for publication in ESSD, after minor revisions.

Reply:

We sincerely thank the reviewer for the positive evaluation and feedback on our manuscript. We are very encouraged by the recognition of the CASPROD database as a valuable community resource for Arctic sediment provenance and paleoenvironmental research. We have carefully considered all the comments and suggestions provided below and have revised the manuscript accordingly to further improve its quality.

Comment 1:

DIVA is an appropriate method for spatial interpolation here, but I think the manuscript would be benefit from a brief discussion on the spatial uncertainty of the interpolated dataset. Regions such as parts of the Canada Basin and Canadian Arctic Archipelago, have sparse sampling, so their interpolated values might carry relatively high uncertainty. Pointing out these potential limitations would be important for anyone who intends to run quantitative analyses on this database.

Reply:

We agree with the reviewer that discussing spatial uncertainty is crucial for the appropriate quantitative use of the CASPROD database, particularly in regions with sparse sampling coverage. To address this, we have added a paragraph in Section 2.6 (Data interpolation) of the revised manuscript (*Line 254-262*). We point out that while the DIVA (Data-Interpolating Variational Analysis) method effectively accounts for topographic barriers (such as coastlines and ridges), the reliability of the interpolated grids remains inherently constrained by regional data density. We have highlighted the Canada Basin and parts of the Canadian Arctic Archipelago as specific areas where lower sampling density results in higher interpolation uncertainty, and we caution users to exercise care when performing localized quantitative analyses in these zones.

Revisions in TEXT:

"Although the DIVA method successfully integrates topographic features and complex coastlines to optimize spatial continuity, the accuracy of the interpolated values remains highly dependent on data density. In well-sampled regions, such as the Chukchi Sea, Laptev Sea, and parts of the Barents-Kara Seas, the grid cells represent robust local signals. Conversely, in regions with sparse sampling coverage, most notably parts of the deep Canada Basin and portions of the Canadian Arctic Archipelago, the interpolated values carry relatively higher uncertainty. Consequently, the interpolated results in these data-sparse or unsampled areas should be treated with caution and utilized primarily as a qualitative reference."

Comment 2:

The Arctic environment is changing rapidly, such as, intensified riverine discharge, sea-ice retreat, enhanced permafrost thaw... For this reason, surface sediments collected at different times might reflect partially different source configurations. I advise the authors to add a brief comment on the potential temporal heterogeneity within the compiled dataset, and how users might consider this when interpreting the data.

Reply:

We agree with the reviewer that the rapid, modern environmental changes in the Arctic (e.g., accelerated sea-ice retreat, enhanced permafrost thaw, and changing riverine fluxes) raise interesting questions regarding potential temporal heterogeneity in surface sediment samples. However, we suggest that this temporal heterogeneity does not significantly bias the baseline spatial patterns established in CASPROD, making its impact negligible for broader geological and paleoenvironmental applications. To clarify this for database users, we have added a paragraph addressing this potential temporal heterogeneity in Section 2.3 (*Dataset description*) of the revised manuscript (*Line 195-201*).

Revisions in TEXT:

"Although the depositional ages of these surface and core-top samples may vary slightly, they are all concentrated within the Holocene. Over this relatively short geological timescale, the primary circum-Arctic sediment source configurations and geological basements have remained remarkably stable. Consequently, any minor variation in depositional age among these surface and core-top sediments is practically negligible when utilizing CASPROD as a baseline to track provenance shifts over longer geological timescales, such as Quaternary glacial-interglacial cycles."

Reply to the reviewers' comments (RC2) on "Circum-Arctic Sediment PROvenance Database (CASPROD): A database of mineralogy and geochemistry for the Circum-Arctic surface sediments" by Yao et al.

This manuscript presents the first circum-Arctic surface sediment provenance database (CASPROD), which systematically integrates multi-proxy datasets including Sr-Nd isotopes, detrital zircon U-Pb ages, clay mineral assemblages, and detrital mineral compositions from 4308 sampling stations. The dataset is rich in volume and broad in spatial coverage. The study has a clear objective and holds significant scientific significance for understanding sediment source-to-sink processes in the Arctic, as well as for reconstructing past ice-sheet dynamics, palaeoceanographic circulation, and sea-ice variability. The database thus fills a critical gap caused by previously fragmented and non-standardized sediment provenance data in the Arctic region, and provides a reliable baseline for future palaeoceanographic and paleoclimate research in the Arctic region. Given the scientific importance and the quality of the compiled dataset, I strongly recommend publication after minor revisions as outlined below.

Reply:

We are grateful to the reviewer for positive assessment of our work. Your validation of the database's scientific significance, data volume, and spatial coverage is highly appreciated. We have carefully addressed each of the points raised in the detailed comments below, which we believe have significantly enhanced the quantality of the manuscript.

Comment 1:

The current introduction focuses primarily on Sr-Nd isotopes, detrital zircon U-Pb ages, clay minerals, and detrital mineral compositions. However, in Arctic sediment provenance studies, iron oxide minerals have also been used to trace sediment sources and transport pathways, particularly for distinguishing ice-rafted debris and current-transported materials from different source regions. I suggest that the authors briefly mention other potential provenance indicators, e.g., iron oxides in the introduction. This would help to contextualize the chosen proxy system and further highlight the comprehensiveness and rationale of the approach adopted in this study.

Reply:

We agree with the reviewer's valuable suggestion. Iron oxide minerals indeed represent an important provenance tracer in Arctic sediment provenance research, particularly for tracking sea-ice drift pathways and matching ice-rafted debris (IRD) to specific circum-Arctic source regions. In fact, this constructive point was also emphasized by Prof. Dennis Darby, whose pioneering work has established iron oxides as a foundational tool in Arctic oceanography. In the Introduction in the revised manuscript, we highlight the significance of Fe-oxide minerals as critical provenance indicators alongside radiogenic isotopes, geochronology, and bulk mineralogy (*Line 83-85; 96-99*).

Revisions in TEXT:

"A range of provenance tracers, such as Sr-Nd isotopes, clay mineral assemblages, detrital mineral compositions, detrital zircon U-Pb ages and Fe-oxide minerals, provide complementary constraints on sediment sources in the Arctic Ocean."

"In addition, Fe-oxide minerals have been widely utilized to constrain sediment provenance in the Arctic Ocean and to reconstruct past sea-ice conditions (e.g., Darby, 2014; Darby et al., 2015; Tripati and Darby, 2018; Cofield and Darby, 2025)."

Comment 2:

Section 3.2 provides a relatively detailed interpretation of the provenance implications of the dataset. For a data paper, this section could be shortened to focus more on the spatial patterns, major trends, and key findings revealed by the database. Overly extended discussion of individual proxy interpretations may be better placed in follow-up research papers. This adjustment would better highlight the contribution of the dataset itself and align with the conventions of data journals.

Reply:

We agree with the reviewer's suggestion. To address this, we have thoroughly revised Section 3.2 (*Circum-Arctic provenance implications*) by removing overly extended discussions and interpretations. This structural tightening successfully highlights the value of the database itself as a community benchmark while leaving highly localized, detailed mechanism interpretation for follow-up research papers.

Comment 3:

Figure 1: this figure includes red and blue arrows indicating ocean circulation patterns, but these are not mentioned in the figure caption. Please add a clear description of these arrows in the caption and include the relevant literature sources.

Reply:

We thank the reviewer for pointing out this omission. We have revised the figure caption for Figure 1 to include a description of the red, blue and green arrows representing the primary ocean circulation systems.

Revised Caption for Figure 1:

"Figure 1. Overview map of the Arctic Ocean illustrating regional subdivisions and circum-Arctic lithological frameworks. The study area is divided into the central Arctic Ocean and the Arctic marginal seas. The geological background is adapted from Harrison et al. (2011). Major magmatic and volcanic belts and cratons (Akinin et al., 2020) are indicated by arrows and annotated with colored boxes. BS denotes the Bering Strait. Major surface ocean circulation pathways are denoted by colored arrows, where red arrows indicate the Transpolar Drift, dark blue arrows indicate the Beaufort Gyre, and green arrows denote regional coastal currents (Timmermans and Marshall, 2020)."

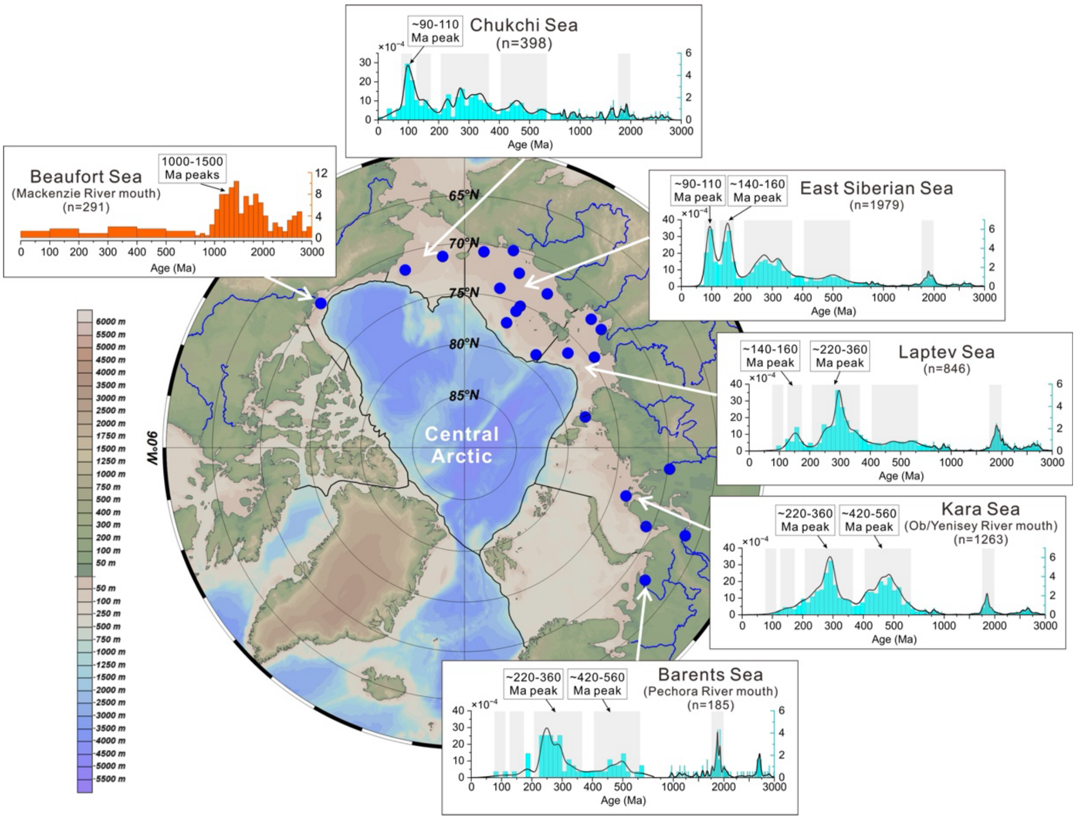
Comment 4:

Figure 3: In figure 3, the depth legend appears to be incorrectly labeled. Ocean bathymetry values should be negative numbers. Please correct this.

Reply:

We thank the reviewer for catching this oversight. We have corrected the depth legend in Figure 3 (See below), modifying the labels for ocean bathymetry to negative values (e.g., -50 m to -5500 m).

Revised Figure 3 in TEXT:



Chukchi Sea
(n=398)
~90-110 Ma peak
Beaufort Sea
(Mackenzie River mouth)
(n=291)
1000-1500 Ma peaks
East Siberian Sea
(n=1979)
~90-110 Ma peak
~140-160 Ma peak
Laptev Sea
(n=846)
~140-160 Ma peak
~220-360 Ma peak
Kara Sea
(Ob/Yenisey River mouth)
(n=1263)
~220-360 Ma peak
~420-560 Ma peak
Barents Sea
(Pechora River mouth)
(n=185)
~220-360 Ma peak
~420-560 Ma peak
Age (Ma)
Central Arctic
65°N
70°N
75°N
80°N
85°N
90°W

Figure 3

Reply to comments from community (Dennis Darby) on "Circum-Arctic Sediment PROvenance Database (CASPROD): A database of mineralogy and geochemistry for the Circum-Arctic surface sediments" by Yao et al.

Comment 1:

This paper provides a useful listing of provenance techniques applicable to the Arctic Ocean. It suffers for two major omissions: 1) The preprint ignores the Fe grain chemical fingerprinting method used in a dozen or more papers. 2) It avoids any scrutiny of the listed methods as to their accuracy. While the precision of several listed methods might be high, when dealing with multiple sources of ice-rafted sediment in the Arctic Ocean, the accuracy can be quite low. The Fe grain fingerprinting method has been shown to have an accuracy of less than 2% error of misidentifying the correct source or provenance.

Comment 2:

The database by Yao et al. misses an important component of provenance studies. This is the relative accuracy of source determinations. Users would greatly benefit from knowing how accurate each provenance method is. This does not mean how precise each method is!

Reply:

Thank you very much for your constrictive comments and suggestions. In response to your first point, we have added a brief review of the Fe-oxide method into the Introduction, highlighting its extensive application in Arctic research (*Page 3-4, Line 84-106*).

Revisions in the TEXT:

A range of provenance tracers, such as Sr-Nd isotopes, clay mineral assemblages, detrital (i.e., light and heavy) mineral compositions, detrital zircon U-Pb ages, major and minor elements, and Fe-oxide minerals, provide complementary constraints on sediment sources in the Arctic Ocean. Sr-Nd isotopes effectively distinguish sediments derived from juvenile orogenic belts and ancient Archean-Proterozoic cratons (Bazhenova et al., 2017; Maccali et al., 2018; Li et al., 2023), whereas detrital zircon U-Pb age spectra offer robust geochronological fingerprints that discriminate between North American and Eurasian basement affinities (Fedo et al., 2003; Feng et al., 2025). Clay mineral assemblages (e.g., Naidu and Mowatt, 1983; Stein et al., 1994, 2004; Wahsner et al., 1999; Thiry, 2000; Viscosi-Shirley et al., 2003a; Vogt and Knies, 2009; Saukel et al., 2010; Jang et al., 2023), detrital mineral assemblages (e.g., Vogt, 1996, 1997; Behrends, 1999; Stein, 2008; Wang et al., 2022), and major and minor elements (e.g., Schoster et al., 2000; Viscosi-Shirley et al., 2003a, 2003b) may reflect source-rock provenance, transport processes, and past varying climatic regimes. Fagel et al. (2014) have used mineral assemblages, trace element compositions and Nd and Pb isotope signatures to identify the detrital sediment provenance and to interpret changes in the relative contribution of the different source-areas in relationship to paleoenvironmental conditions. In addition, the chemical fingerprint of major and minor elements in specific Fe-oxide minerals, an approach developed by Darby and co-workers (Darby and Bischof, 1996; Darby, 2003; Darby et al., 2015), has been widely used to determine quite accurately sediment provenances in the Arctic Ocean and to reconstruct past sea-ice conditions (e.g., Darby, 2014; Tripati and Darby, 2018; Cofield and Darby, 2025).

To address the reviewer's concerns regarding the relative accuracy of source determinations, we have added an additional section (*3.3 Evaluating the discriminability of provenance proxies*) in the revised manuscript (*Page 20-23, Line 442-514*).

In this new section, we evaluate the relative accuracy of the provenance proxies. As we detail in the text, while single-grain techniques like Fe-oxide fingerprinting can be evaluated using traditional cross-validation accuracy, the integrated-signal/bulk proxies compiled in CASPROD (e.g., Sr-Nd isotopes, clay/detrital minerals, zircon U-Pb ages) yield distributional signatures of sediment mixtures. Therefore, we evaluate these proxies based on their discriminability—quantified as the signal-to-noise ratio (SNR) between inter-source differences and intra-source natural heterogeneity. We provided an honest evaluation of where these methods suffer from low accuracy/discriminability due to regional signal overlaps, thereby underscoring the need for multi-proxy cross-validation as intended by this database.

We believe these additions directly address your concerns and provide users with a transparent assessment to better guide their application of these provenance methods.

Revisions in the TEXT:

3.3 Evaluating the discriminability of provenance proxies

Provenance accuracy is method-dependent and is linked to the analytical unit of a given proxy. For single-grain techniques like Fe-oxide chemical fingerprinting, accuracy is typically evaluated through cross-validation against a reference to quantify incorrect source assignments (Darby et al., 2015). However, this traditional concept of accuracy is not directly applicable to the integrated-signal proxies compiled in CASPROD (e.g., Sr-Nd isotopes, clay and detrital minerals, and zircon U-Pb ages), which yield bulk or distributional signatures of sediment mixtures rather than discrete grain labels. Instead, we evaluate these proxies based on discriminability, defined here as the signal-to-noise ratio (SNR) between inter-source differences and intra-source variability (Figure 7). A high SNR (generally >1) indicates that the compositional contrast between two sources substantially outweighs their internal natural heterogeneity, allowing for reliable provenance discrimination, with larger values yielding greater discriminability.

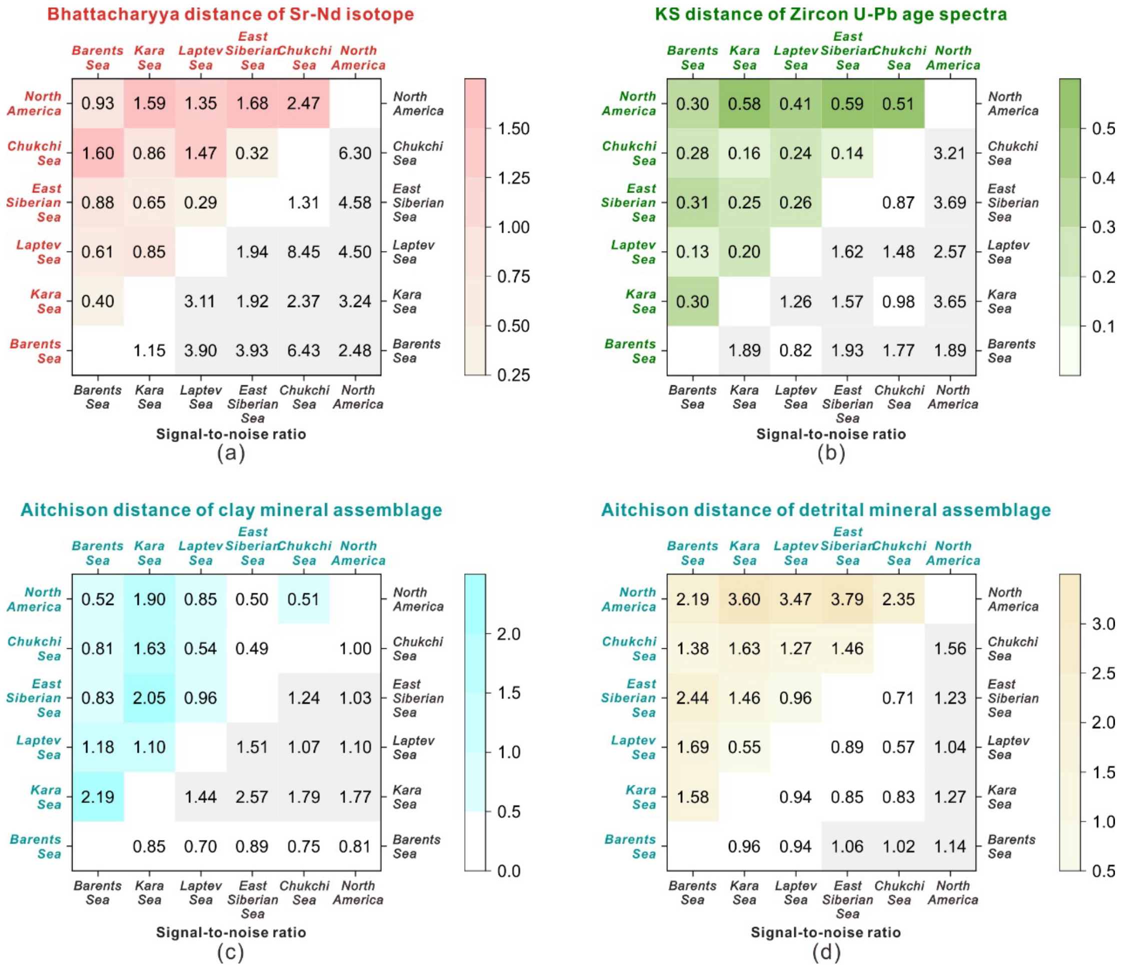


Figure 7. Distance matrix heatmaps illustrating the discriminability of four provenance proxies. In each panel, the upper and lower triangle display pairwise inter-source distances and the corresponding signal-to-noise ratios (SNR), respectively. SNR is defined as the ratio of inter-source distance to intra-source variability. Gray shaded cells denote high SNR values (>1), indicating that the corresponding two source regions can be effectively differentiated by the given proxy. KS: Kolmogorov–Smirnov.

To evaluate pairwise source-region distances and intra-source heterogeneity, we selected statistical metrics tailored to the specific data structure of each proxy. For Sr-Nd isotopes, we applied the Bhattacharyya distance (Bhattacharyya, 1946) to measure source separation in the bivariate εNd–$^{87}Sr/^{86}Sr$ space, as it accommodates substantial differences in data dispersion among source areas without assuming identical covariance structures. For detrital zircon U-Pb age spectra, we used the Kolmogorov–Smirnov (KS) distance. As a distribution-free metric quantifying the maximum vertical difference between two cumulative age-distribution functions, the KS distance is highly sensitive to overall spectral shape and serves as a standard for nonparametric age-spectrum comparisons (Vermeesch, 2012). For clay mineral and detrital mineral assemblages, we applied the Aitchison distance to measure compositional differences between source regions, because percentage of minerals are subject to the constant-sum constraint and conventional distance metrics would introduce spurious correlations (Garzanti et al., 2012). Across all proxies, intra-source heterogeneity is represented by the median distance derived from 1000 bootstrap resampling iterations within each source group.

The calculated SNR values reveals that low discriminability typically stems from substantial signal overlap between specific regions (Figure 7). For instance, detrital zircon U-Pb age spectra cannot effectively differentiate the Kara, Chukchi, and East Siberian seas, nor can they separate the Barents Sea and Laptev Sea (Fig. 7b). This overlap occurs because these marginal seas collectively drain extensive Phanerozoic magmatic provinces, such as the Siberian Traps and the Okhotsk–Chukotka Volcanic Belt, yielding highly convergent age distributions. Similarly, clay mineral assemblages show limited discriminability for the Barents Sea (Figure. 7c), reflecting pronounced intra-regional spatial variability driven by complex hydrodynamics and the mixing of diverse local endmembers. Detrital mineral assemblages also struggle to distinguish Eurasian interior sources, particularly the Kara, Laptev, and East Siberian seas (Figure. 7d). Sr-Nd isotopes, despite generally exhibiting high SNR values (>1), show limited discriminability between the East Siberian and Chukchi seas, or the Barents and Kara seas (Figure. 7a), underscoring the challenge of differentiating provenance with similar geological affinities.

These overlaps highlight the inherent limitations of relying on any single proxy to decipher a complex, multi-source sedimentary system like the Arctic Ocean. Accurate circum-Arctic provenance reconstruction requires a multi-proxy approach (e.g., Stein, 2008; Myers and Darby, 2022). Thus, in several studies multi-proxy approaches have been used to identify circum-Arctic sediment provenances more accurately, e.g. the combination of heavy minerals, clay minerals and major and minor elements (Schoster et al., 2000; Viscosi-Shirley et al., 2003a, 2003b) or the combination of heavy minerals, clay minerals, and chemical fingerprinting of major and minor elements in Fe-oxide minerals (Myers and Darby, 2022).

CASPROD also establishes a multi-proxy, cross-validation framework to overcome the ambiguity introduced by these overlapping signatures. By synthesizing geochemical, geochronological, and mineralogical datasets, this integrated framework provides multiple constraints on endmembers and substantially improve the accuracy of provenance identification. Furthermore, to achieve higher accuracy when interpretating highly complex mixed signals, particularly regarding ice-rafted debris, we encourage database users to couple CASPROD with other approaches, such as the Fe-oxide grain chemical fingerprinting method (Darby, 2003; Darby et al.; 2015; Myers and Darby, 2022).