Point-to-point Response to the Reviewers' Comments

Responses to the comments from the 1st Reviewer

This manuscript presents an exceptionally comprehensive soil geochemical dataset that addresses a critical gap in global biogeochemical databases by systematically characterizing 1,300+ samples across 30 mountain regions spanning five climatic zones in China. The authors' methodological rigor is evident in their stratified sampling design across three pedogenic horizons (A, B, and C), standardized analytical protocols for 24 macro- and microelements, and integration with ancillary environmental variables including climatic indices, vegetation parameters, and human activity factor. The dataset's particular strength lies in its unprecedented spatial coverage of montane ecosystems, combined with vertical resolution that captures pedogenic gradients crucial for understanding soil formation processes and biogeochemical cycling.

Overall, the authors' efforts in assembling this high-resolution, multi-horizon, and climatically contextualized soil dataset are timely and scientifically significant for researchers in soil science, biogeochemistry, ecology, and Earth system modeling. Moreover, the manuscript is generally well organized, and it is suitable for publication in the journal after some minor revisions. Please find my comments below.

Responses: We appreciate your agreement with the significance of our study and the helpful suggestions to improve the manuscript. We carefully revised the manuscript based on your comments and suggestions. Please see our point-to-point response to your comments below.

Specific comments:

I recommend the authors should store the valuable data in the Zendo website.

Responses: We sincerely thank you for the suggestion regarding data archiving. Our dataset has been deposited in the National Tibetan Plateau/Third Pole Environment Data Center, and a corresponding DOI has been generated and included in the manuscript (https://doi.org/10.11888/Terre.tpdc.302620 or https://cstr.cn/18406.11.Terre.tpdc.302620). This data repository has been used by many publications in Earth System Science Data (e.g., Li et al., 2024; Jin et al., 2023; Ma et al., 2024). Importantly, this platform provides DOI-linked data access, similar to Zenodo, and allows free and direct data download without registration. Accordingly, we do not repeat the submitting our dataset on this platform.

References:

Jin, Y., Wang, H., Xia, J., Ni, J., Li, K., Hou, Y., Hu, J., Wei, L., Wu, K., Xia, H., and Zhou, B.:

TiP-Leaf: a dataset of leaf traits across vegetation types on the Tibetan Plateau, Earth Syst. Sci. Data, 15, 25-39, https://doi.org/10.5194/essd-15-25-2023, 2023.

Li, X., Jin, H., Feng, Q., Wu, Q., Wang, H., He, R., Luo, D., Chang, X., Şerban, R.-D., and Zhan, T.: An integrated dataset of ground hydrothermal regimes and soil nutrients monitored in some previously burned areas in hemiboreal forests in Northeast China during 2016–2022, Earth Syst. Sci. Data, 16, 5009–5026, https://doi.org/10.5194/essd-16-5009-2024, 2024.

Ma, Y., Xie, Z., Chen, Y., Liu, S., Che, T., Xu, Z., Shang, L., He, X., Meng, X., Ma, W., Xu, B., Zhao, H., Wang, J., Wu, G., and Li, X.: Dataset of spatially extensive long-term quality-assured land–atmosphere interactions over the Tibetan Plateau, Earth Syst. Sci. Data, 16, 3017-3043, https://doi.org/10.5194/essd-16-3017-2024, 2024.

Line 123: Replace “was” with “were”. Please check other grammar issues in the manuscript.

Responses: We are very grateful for your detailed suggestions. We have corrected the sentence on line 123, replacing "was" with "were" as suggested. Following your advice, we have also performed a thorough proofread of the entire manuscript to identify and correct other grammatical issues and typos. We appreciate your help in improving the quality of our paper.

Line 132: Please specify the extraction method for pH measurement (e.g., water, KCl, or $CaCl_2$). This is essential for comparability with other pH datasets and can influence interpretation of cation exchange and element mobility.

Responses: Thank you for your valuable comment. We have revised the manuscript to specify the pH measurement method. The pH was determined using a water extraction method (with a 1:2.5 soil-to-water ratio). This clarification has been added to the Methods section (Line 132). The revised description is as follows:

Soil pH was measured using a pH meter (Mettler-Toledo FE28, Switzerland) after shaking the soil samples with deionized water at a 1:2.5 soil-to-water ratio.

Lines 154-158: The calculation of the Chemical Index of Alteration (CIA) should be more explicitly explained. Please clarify how CaO* was estimated, and whether the method has followed that of Nesbitt & Young (1982) directly or been corrected.

Responses: Thank you for pointing out a clearer explanation of the CIA calculation. In the revised manuscript, we provided a more detailed and explicit description of the method. The Chemical Index of Alteration (CIA) was calculated using the widely accepted formula proposed by Nesbitt & Young (1982):

$$CIA = \frac{Al_2O_3}{(Al_2O_3 + Na_2O + K_2O + CaO^*)} \times 100$$

where all oxides are expressed in molar proportions. As you rightly noted, CaO* should reflect only the amount of Ca derived from silicate minerals, excluding contributions from carbonates, phosphates, or exchangeable forms. To address this, we adopted a correction method following McLennan (1993), which has been applied in numerous geochemical studies to improve the reliability of CIA values. Specifically, CaO* was estimated as follows: when the measured CaO content is less than or equal to that of Na_2O, CaO* is assumed to be equal to the measured CaO; when the measured CaO content is greater than that of Na_2O, CaO* is assumed to be equal to Na_2O. We have updated the Methods section accordingly to reflect this correction.

References:

Nesbitt, H. W., and Young, G. M.: Early Proterozoic climates and plate motions inferred from major element chemistry of lutites, Nature, 299, 715-717, https://doi.org/10.1038/299715a0, 1982.

McLennan, S. M.: Weathering and Global Denudation, J. Geol., 101, 295-303, https://doi.org/10.1086/648222, 1993.

Lines 164-165: The strict coordination has been carried out, but it was not clearly defined. Does this refer to harmonization of sampling protocols across sites, or post-hoc statistical adjustments (e.g., normalization, transformation, unit standardization) to ensure cross-site comparability?

Response: Thank you for raising this important point. The “rigorous harmonization procedures” mentioned in the manuscript refers specifically to the harmonization of sampling protocols and laboratory measurement procedures across all sites. All soil samples were collected following a unified field sampling protocol and analyzed using consistent laboratory methods and instrumentation to ensure comparability of the physical and chemical data across different mountain regions. We would like to clarify that no post-hoc statistical adjustments (such as normalization, transformation, or unit conversion) were applied to the raw data. The consistency in methodology at both the field and laboratory stages eliminate the need for such adjustments and ensures that the observed variations reflect actual environmental differences rather than methodological artifacts. We have revised the manuscript to clarify this point and avoid misunderstanding:

The dataset integrates information from extensive field surveys, laboratory analyses, high-resolution satellite-derived vegetation indices, and ancillary environmental data compiled from national and global databases. To ensure data consistency and comparability across sites, all soil samples were collected following standardized sampling protocols and analyzed using uniform laboratory procedures and instrumentation.

Line 103: The manuscript would benefit from a concise description of the statistical or visualization methods used to generate Figures 2-6. This addition will help readers better interpret the trends and distributions presented.

Response: Thank you for your helpful suggestion. In the revised manuscript, we have added a new subsection titled "2.5 Statistical analysis" in the Materials and Methods section to provide a clear and concise description of the statistical and visualization methods used throughout the study. This addition aims to improve transparency and help readers better understand the analytical approaches and interpretation of the trends and patterns presented in the results. The added content is as follows:

2.5 Statistical analysis

All statistical analyses were conducted using R (version 4.3.1). To test differences in element concentrations among soil horizons, we employed linear mixed-effect models using the "lmer" function from the "lme4" package, where soil horizon was treated as a fixed factor and sampling site as a random factor. Regression analyses were conducted to examine the spatial distribution characteristics of each element. To explore the compositional differences in elemental assemblages across soil horizons and to assess the influence of environmental variables on soil element variation, redundancy analysis (RDA) was conducted using the "rda" function in the "vegan" package. Correlation analyses were conducted separately for each soil horizon to identify horizon-specific relationships between elemental concentrations and environmental drivers. Furthermore, simple linear regression was employed to quantify the individual explanatory power (R^2) of each environmental variable for each element. The cumulative explanatory power of all environmental factors was also calculated to evaluate their combined influence on element variation.

Line 260: The authors provided horizon-level sampling and vertical stratification but did not elucidate the implications for soil development modeling. Given the presence of C-horizon data and CIA indices, this dataset could serve as a valuable benchmark for soil formation modeling (e.g., using SoilGen or CLORPT frameworks). A short paragraph in Section 4 may highlight this point.

Response: Thank you for this insightful comment. Recognizing the potential value of horizon-specific data and weathering indices (e.g., CIA), we have added a short paragraph in the subsection of "Potential applications of the dataset" (Section 4) to highlight how this dataset could be used to support soil formation modeling efforts. The added content is as follows:

In addition, the inclusion of horizon-specific data (O, A, and C horizons), weathering indices, and lithological information provides valuable input for soil formation and rock weathering models. Process-based models like SoilGen or conceptual frameworks such as CLORPT (climate, organisms, relief, parent material, and time) can benefit from the dataset's vertical resolution and environmental coverage to simulate pedogenesis, profile evolution, and mineral nutrient release across climate gradients. Accordingly, the dataset can serve as a regional benchmark for calibrating and validating long-term soil development models, particularly in mountainous regions where such data are scarce yet critically needed.

Line 316: Add a sentence summarizing the dataset structure (e.g., file formats, variable descriptions, metadata schema) to assist users in quickly understanding how to work with the data.

Response: Thank you for this thoughtful suggestion. In response, we have revised the corresponding section of the manuscript to include a brief summary of the dataset structure.

5 Data availability

The database is freely accessible via the National Tibetan Plateau/Third Pole Environment Data Center at https://doi.org/10.11888/Terre.tpdc.302620 or https://cstr.cn/18406.11.Terre.tpdc.302620 (Wu et al., 2025b). The dataset provides comprehensive information for each sample, including mountain affiliation, geographical coordinates, climatic characteristics, vegetation type, soil type, parent rock type, normalized difference vegetation index, atmospheric nitrogen deposition rates, soil physicochemical properties, chemical weathering indices, and concentrations of 24 soil elements. The data are stored in Excel spreadsheet format, accompanied by a separate data documentation file that describes variable names, units, and definitions.

Line 249: The value "Fe (>200%)" as explanatory power in redundancy analysis seems inconsistent (R^2 cannot exceed 100%). Please double-check this statement or clarify if it refers to cumulative variance.

Response: Thank you for the valuable feedback. The y-axis of Figure 6 represents the cumulative explanatory power of all environmental variables for each individual element. Therefore, the total explanatory value may exceed 100%. To avoid confusion, we have clarified this point in the Statistical Analysis subsection of the Methods section and revised the figure legend accordingly.

Fig. 6 Explanation of elemental variation by environmental factors based on regression modelling. Columns with different colors represent different environmental variables. Total height of each bar indicates the cumulative explanatory power. MAP, mean annual precipitation; MAT, mean annual temperature; AI, aridity index; DIN, dissolved inorganic nitrogen; NDVI, normalized difference vegetation index; CIA, chemical index of alteration

Tables 1 and 2: Several abbreviations used in these tables (e.g., MAT, MAP) are not defined within the table notes. As tables should be interpretable independently of the main text, please add a legend or footnotes explaining all abbreviations.

Response: Thank you for your feedback. To improve the clarity of Tables 1 and 2, we improved the legends and defined all abbreviations used in the tables.

Figures 2 and 3: Both figures lack x-axis labels, which impairs interpretability. Ensure all figures include complete and clear axis annotations, including units.

Response: Thank you for raising this important question. We have revised Figures 2 and 3 to include complete and clear x-axis labels, along with appropriate units where applicable. These additions improve the readability and interpretability of the figures. The revised figure is shown below.

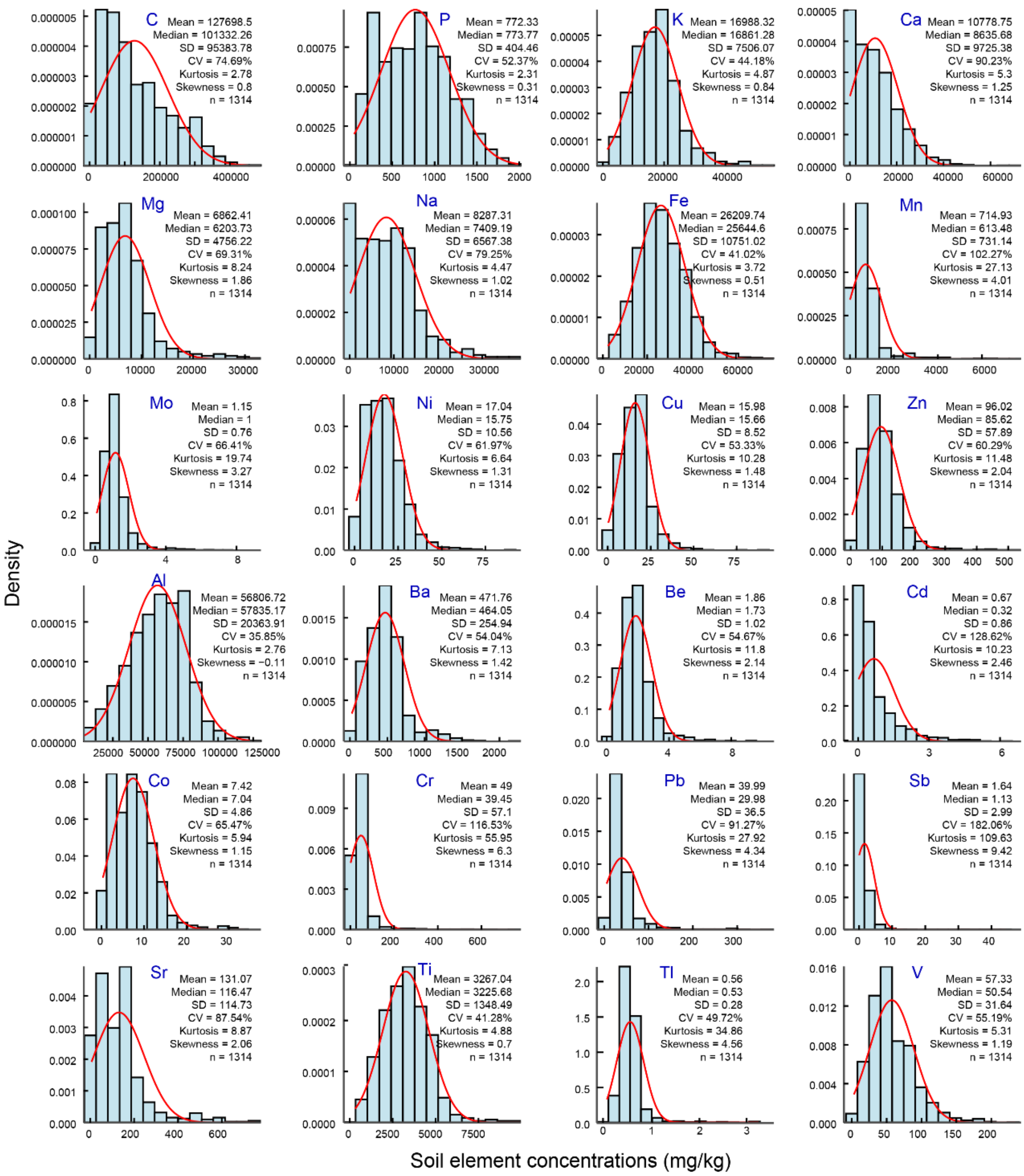

Fig. 2 Frequency distribution of soil elements across the China's mountains. Red curve on each histogram represents the fitted normal distribution. The statistical parameters of the corresponding element are annotated in the upper left of each sub-figure.

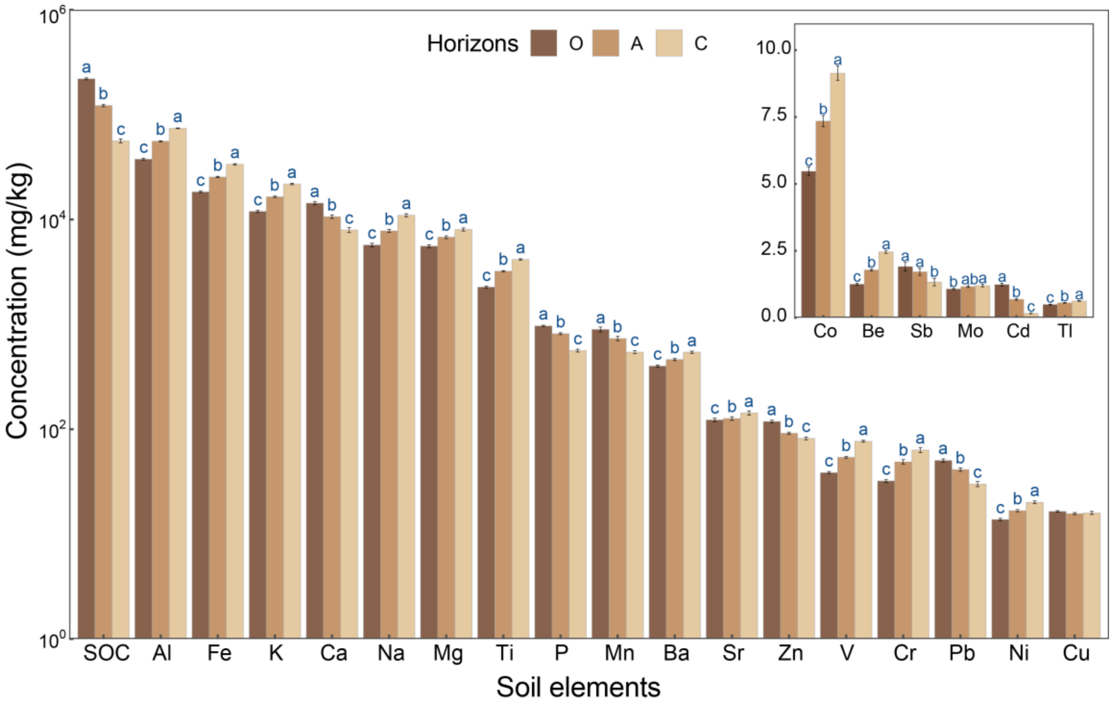

Fig. 3 Mean concentrations of 24 elements across different soil horizons. Lowercase letters indicate significant differences in each element among soil horizons ($p < 0.05$), and error bars represent the standard error.

Responses to the comments from the 2nd Reviewer

Wu and coauthors present a robust and geographically extensive dataset of more than 20 soil elements derived from 1,314 samples across 30 mountain regions in China. This dataset covers diverse bioclimatic zones and three soil development horizons, which offers a valuable vertical and horizontal resolution for understanding the large-scale biogeochemical patterns. This comprehensive, spatially-explicit dataset from mountainous regions is timely and necessary considering the sampling difficulty and terrain complexity. The methodology alongside the rigorous quality control, open-access availability, and comprehensive metadata can significantly increase the reusability and scientific value of this dataset. In general, the manuscript is well-organized and written in fluent academic English, and scientifically rigorous. There are some concerns on the current manuscript that may help further clarify and improve the dataset's accessibility and documentation. Please find my specific comments and suggestions below.

Responses: We appreciate very much for your positive comments on our manuscript, and we also thank you for the valuable and professional suggestions to improve the manuscript. According to your comments and suggestions, we have made careful corrections and improvements of our manuscript. The detailed revisions are listed below.

Please specify the sampling time in this work, which will help well use the dataset.

Responses: Thank you for your valuable comment. The soil sampling was conducted between July 2012 and March 2013. This information has now been explicitly added to the main text to enhance clarity.

The sampling strategies need to be described more specific, considering such a large spatial scale and soil stratification. Were the samples composited from multiple subsamples or taken as single cores? How many replicates were collected per horizon at each site? Were replicate samples analyzed separately or composited before analysis? This information will help to assess the spatial resolution and statistical robustness of the dataset.

Responses: Thank you very much for the thoughtful and constructive comments regarding the sampling strategies. At each site or altitude of a mountain, three soil profiles were excavated by hand, and then each soil horizon was carefully divided. Each soil sample collected was mixed by subsamples taken from a horizon. During laboratory analysis, these replicate samples were analyzed separately rather than composited, ensuring the reliability of the data and enabling robust statistical estimation, including standard error calculation. We now add and clarify this information in the section of Soil sampling. The revised content is as follows:

Sampling campaigns were conducted at 166 sites spanning 30 mountains between July 2012 and March 2013. In each mountain, sites were selected based on the altitude and dominant vegetation types. At each site, the geographic coordinate was recorded using a GPS device (eTrex Venture, USA). Three replicate plots (10 m × 10 m) were randomly established per site, spaced approximately 50 m apart to account for spatial heterogeneity. In each plot, soil profiles were manually excavated down to the parent material horizon. Soil horizons were delineated in the field based on morphological characteristics following the Chinese Soil Taxonomy (Chinese Soil Taxonomy Research Group, 2001; Yang et al., 2023). Horizon boundaries were determined through visual and tactile assessments (e.g., color, texture, consistency, moisture, and root distribution). Horizons were typically classified as O (organic), A (surface mineral), and C (parent material) horizons. For each profile, the name, code, depth range, and diagnostic features were recorded. Soil samples were collected sequentially from bottom to up within each profile to avoid cross-pollution, with composite samples formed by homogenizing subsamples from each horizon.

The authors have emphasized lithogenic and biogenic controls on soil elemental patterns in this study, and relevant lithology data have been used in their prior publications (e.g., Wu et al., 2025; Yang et al., 2022). However, such information is not included in the dataset. I strongly encourage the authors to incorporate this information as an additional column in the main dataset or in the supplementary materials. This will substantially enhance the dataset's applicability in Earth system modeling.

Responses: We appreciate your insightful feedback regarding the dataset. We fully agree that incorporating lithological information will significantly enhance the applicability of the dataset for Earth system modeling. In response to the comment, we have added parent material (lithology) data for each sampling site to the dataset. Specifically, parent material data for the SOTER geological reservoir at a scale of 1:1000000 in Chinese provinces (1990) were obtained from the National Earth System Science Data Center (http://www.geodata.cn). The revised dataset is available at the following link: https://doi.org/10.11888/Terre.tpdc.302620 or https://cstr.cn/18406.11.Terre.tpdc.302620

In Figures 5 and 6, the abbreviation "AI" (aridity index) is not defined. Please ensure that all variables and indices (e.g., AI, CIA, NDVI) are spelled out at first mention, including in figure captions and abstract, to support clarity for multidisciplinary readers.

Responses: Thank you for pointing out this important issue. We have carefully checked the entire manuscript, including the abstract, figure captions, and main text, to ensure that all abbreviations are fully spelled out at their first mention. The figure captions for Figures 5 and 6 have also been updated accordingly to improve clarity for readers.

Carefully check all the figures to ensure that axis labels, units, and legends are present, standardized, and clearly legible. Some figures appear to lack axis units or use inconsistent font sizes. Improving figure formatting will significantly enhance the readability and usability of the manuscript.

Responses: Thank you for your feedback on Figures. We have carefully reviewed and standardized all figures to ensure that axis labels, units, and legends are complete, consistent, and clearly legible. Font sizes and formatting have been adjusted uniformly across all figures to improve readability and overall presentation quality.

Although the DOI is cited, the manuscript would benefit from explicitly stating the name of the data hosting platform (i.e., "National Tibetan Plateau Data Center") and providing a summary of available file formats (e.g., .CSV, .XLSX) and data structure.

Responses: Thank you very much for your valuable feedback. In response, we have revised the Data Availability section to explicitly specify the name of the data hosting platform and to provide a summary of the available file formats and data structure. The revised text is as follows:

The database is freely accessible via the National Tibetan Plateau/Third Pole Environment Data Center at https://doi.org/10.11888/Terre.tpdc.302620 or https://cstr.cn/18406.11.Terre.tpdc.302620 (Wu et al., 2025b). The dataset provides comprehensive information for each sample, including mountain affiliation, geographical coordinates, climatic characteristics, vegetation type, soil type, parent rock type, normalized difference vegetation index, atmospheric nitrogen deposition rates, soil physicochemical properties, chemical weathering indices, and concentrations of 24 soil elements. The data are stored in Excel spreadsheet format, accompanied by a separate data documentation file that describes variable names, units, and definitions.

In the dataset files, columns such as “Vegetation” and “Horizons” use abbreviated codes. Please ensure these codes are clearly documented in the metadata or in a separate codebook/readme file.

Responses: Thank you for your helpful suggestion. We have updated the dataset documentation to include detailed explanations of all abbreviated codes used in the dataset files. Full definitions are now provided in the revised metadata file and the accompanying dataset description to enhance clarity and ensure ease of use for data users. The revised dataset file is available at the following link: https://doi.org/10.11888/Terre.tpdc.302620 or https://cstr.cn/18406.11.Terre.tpdc.302620

In addition to its clear value for biogeochemical modeling and soil quality assessment, the dataset also offers considerable potential for applications in soil development and weathering modeling. The inclusion of vertically stratified soil horizons, chemical weathering indices, and a range of environmental covariates, combined with the recommended addition of lithological data, provide a strong basis to simulate pedogenesis and mineral nutrient weathering and release across climate gradients. These potential applications are needed by highlighting them in the discussion to better reflect the broader relevance of the dataset.

Responses: We sincerely appreciate your insightful comments regarding the broader applicability of our dataset. We fully agree that beyond its demonstrated value for biogeochemical modeling and soil quality assessment, the dataset also holds significant potential for applications in soil development and weathering modeling. Following your suggestion, we have added a dedicated discussion of these potential applications in the revised manuscript, highlighting how the inclusion of vertically stratified horizons, chemical weathering indices, and lithological data can support process-based models of pedogenesis and nutrient release. This addition aims to clarify the dataset's broader relevance and enhance its value to researchers working on long-term soil development, especially in mountainous regions where such data are scarce. We thank the reviewer again for helping us improve the manuscript in this important aspect. The added content is as follows:

In addition, the inclusion of horizon-specific data (O, A, and C horizons), weathering indices, and lithological information provides valuable input for soil formation and rock weathering models. Process-based models like SoilGen or conceptual frameworks such as CLORPT (climate, organisms, relief, parent material, and time) can benefit from the dataset's vertical resolution and environmental coverage to simulate pedogenesis, profile evolution, and mineral nutrient release across climate gradients. Accordingly, the dataset can serve as a regional benchmark for calibrating and validating long-term soil development models, particularly in mountainous regions where such data are scarce yet critically needed.

Responses to the comments from the 3rd Reviewer

This manuscript describes the multi-elemental composition of soil collected in locations in different mountain regions across China. It also describes the associated environmental parameters in these locations using remote sensing data. The authors then used these environmental parameters, along with geospatial and climatic data, to explain the variations in the soil multi-elemental composition.

The database potentially contains important data which can be used for various applications as outlined by the authors in Sec. 4 of the manuscript. However, there are several important pieces of information that are missing in the manuscript, as well as serious concerns regarding data quality and completeness of the database, which I have described below.

Responses: We sincerely appreciate your recognition of the potential value of our dataset and also thank you for the constructive and professional comments and suggestions regarding data quality, completeness, and missing information. We have carefully considered the comments raised and made specific revisions to both the manuscript and the dataset documentation. Please see the point-by-point response and corresponding modifications below.

Database

It is better to store the database in a more open-access platform like Zenodo that does not prompt users of data to log-in.

Responses: Thank you very much for the helpful suggestion. Although the National Tibetan Plateau/Third Pole Environment Data Center suggests users to log in, data can be freely and directly downloaded via the DOI link without registration. The platform ensures open access and stable sharing through permanent DOI-based links, similar to Zenodo. Moreover, this data repository has been widely used in Earth System Science Data publications (e.g., Li et al., 2024; Jin et al., 2023; Ma et al., 2024), demonstrating its compatibility with the journal's data policy. Therefore, we retain our dataset on this platform and trust that this can also meet the journal's open-access requirements.

References:

Jin, Y., Wang, H., Xia, J., Ni, J., Li, K., Hou, Y., Hu, J., Wei, L., Wu, K., Xia, H., and Zhou, B.: TiP-Leaf: a dataset of leaf traits across vegetation types on the Tibetan Plateau, Earth Syst. Sci. Data 15, 25-39, https://doi.org/10.5194/essd-15-25-2023, 2023.

Li, X., Jin, H., Feng, Q., Wu, Q., Wang, H., He, R., Luo, D., Chang, X., Şerban, R.-D., and Zhan, T.: An integrated dataset of ground hydrothermal regimes and soil nutrients monitored in some previously burned areas in hemiboreal forests in Northeast China during 2016-2022,

Earth Syst. Sci. Data 16, 5009-5026, https://doi.org/10.5194/essd-16-5009-2024, 2024.

Ma, Y., Xie, Z., Chen, Y., Liu, S., Che, T., Xu, Z., Shang, L., He, X., Meng, X., Ma, W., Xu, B., Zhao, H., Wang, J., Wu, G., and Li, X.: Dataset of spatially extensive long-term quality-assured land–atmosphere interactions over the Tibetan Plateau, Earth Syst. Sci. Data 16, 3017-3043, https://doi.org/10.5194/essd-16-3017-2024, 2024.

The database lacks metadata that explains what the columns mean. The units are also not given. The authors stated that a "Description of the dataset.docx" document accompanies the dataset, but it is not included when I downloaded the database several times. All these render the data in the database practically useless.

Responses: We sincerely thank you for pointing out this important issue. This file was previously provided as a supplementary attachment under the Additional Information section on the data repository webpage, which may have caused it to be overlooked during download. In the revised version, all abbreviations and measurement units used in the dataset have been carefully documented in the file of "Description of the dataset.docx". To address this, we have uploaded the data description file together with the main dataset in a single compressed package, ensuring that all necessary metadata are conveniently available to users. A detailed explanation of the abbreviations and units can be found in the updated version, as illustrated in the table below.

Abbreviations	Full name	Unit
O	Organic horizon	-
A	Surface mineral horizon	-
C	Parent material horizon	-
BF	Broadleaf forest	-
CBF	Coniferous-broadleaf mixed forest	-
CF	Coniferous forest	-
MAP	Mean annual precipitation	mm
MAT	Mean annual temperature	°C
AI	Aridity index	-
N deposition	Atmospheric nitrogen deposition	kg N ha^{-1} yr^{-1}
NDVI	Normalized difference vegetation index	-
Depth	Thickness of soil	cm

BD	Soil bulk density	g cm^{-3}
CIA	Chemical alteration index	-
SOC	Soil organic carbon	mg kg^{-1}
Al	Aluminum	mg kg^{-1}
Ba	Barium	mg kg^{-1}
Be	Beryllium	mg kg^{-1}
Ca	Calcium	mg kg^{-1}
Fe	Iron	mg kg^{-1}
K	Potassium	mg kg^{-1}
Mg	Magnesium	mg kg^{-1}
Mn	Manganese	mg kg^{-1}
Na	Sodium	mg kg^{-1}
Sr	Strontium	mg kg^{-1}
Ti	Titanium	mg kg^{-1}
V	Vanadium	mg kg^{-1}
Zn	Zinc	mg kg^{-1}
P	Phosphorus	mg kg^{-1}
Cr	Chromium	mg kg^{-1}
Co	Cobalt	mg kg^{-1}
Ni	Nickel	mg kg^{-1}
Cu	Copper	mg kg^{-1}
Mo	Molybdenum	mg kg^{-1}
Cd	Cadmium	mg kg^{-1}
Sb	Antimony	mg kg^{-1}
Tl	Thallium	mg kg^{-1}
Pb	Lead	mg kg^{-1}

There are unrealistic values in the database. For example, the bulk densities are too low even in the surface mineral soil. I would not expect the values to be below 0.5 g/cm^3, assuming this is the correct unit based on the data description paper. Another example is the moisture content which reaches up to over 200%. How were these values calculated? It would help if the authors explicitly state the measurement protocols and equations used.

Responses: Thank you for the careful review and pointing out the values of bulk density and moisture content. Our samples were collected along soil development profiles, including the organic horizon (O), surface mineral horizon (A), and parent material horizon. The relatively low bulk density values (e.g., <0.5 g/cm^3) are primarily concentrated in the surface layers, especially in the O horizon of the forest soils, which are rich in organic matter and characterized by loose, low-density materials. This phenomenon has been well documented in studies of forest soils. For example, Ostrowska et al. (2010) reported bulk density values ranging from 0.10 to 0.16 g/cm^3 in the O horizon of forest soils in Poland. Zhou et al. (2016) measured bulk density ranging between 0.15 and 0.24 g/cm^3 in the O horizon of Gongga Mountain. Furthermore, due to the thinness and high organic matter contents in the O and A horizons, it was not feasible to collect samples using traditional stainless-steel cutting rings. Instead, for these surface horizons, we used a volumetric excavation method: small pits with known volumes were excavated to collect the density samples. The pit volume was verified by refilling with water of known volume (e.g., Maynard & Curran, 2006). For the mineral horizons, bulk density was measured using the standard cylindrical core method with stainless-steel rings. This dual approach ensures accurate volume determination across diverse soil layers with different physical properties.

For the soil moisture content, we used the standard gravimetric method:

$$Moisture\ (\%) = \frac{Fresh\ weight - Dry\ weight}{Dry\ weight} \times 100\%$$

Based on this formula, values exceeding 100% are common in organic-rich surface soils with high water retention capacity, particularly in the O horizon. Such results are consistent with observations from many forest ecosystems and reflect the natural variability of organic surface soils. We have included a clearer explanation of the measurement protocols and calculation methods in the revised manuscript and the updated data description file to improve clarity for users (Lines 133-137, Lines 139-141).

References:

Maynard, D. G., and Curran, M. P.: Bulk density measurement in forest soils, Soil sampling and methods of analysis, 863-869, https://doi.org/10.1201/978142000527, 2007.

Ostrowska, A., Porebska, G., and Kanafa, M.: Carbon Accumulation and Distribution in

Profiles of Forest Soils, Pol. J. Environ. Stud. 19(6), 1307-1315, 2010.

Zhou, J., Wu, Y., Bing, H., Yang, Z., Wang, J., Sun, H., Sun, S. and Luo, J.: Variations in soil phosphorus biogeochemistry across six vegetation types along an altitudinal gradient in SW China, Catena 142, 102-111, https://doi.org/10.1016/j.catena.2016.03.004, 2016.

What does depth mean in the database, and if I am right to assume that this represents the incremental depth, why are some organic layers lie below the surface mineral layer? Or does depth mean thickness here?

Responses: Thank you for your valuable comment. The values in this column represent soil layer thickness, not absolute or incremental depth. This was explained in the previous version of the data description file, but we acknowledge that the use of the term “depth” may have caused confusion. To improve clarity and avoid misinterpretation, we have revised the dataset to rename this column as “Thickness”, and we have clearly restated this definition in the revised data description file. We appreciate your helpful suggestion, which has enabled us to make this important clarification.

Why are some values missing for depth and bulk densities? This has to be explained by the authors.

Responses: Thank you for your careful review. The missing values for soil depth and bulk density primarily result from field and logistical limitations encountered during sampling and measurement. And we would like to clarify that the column "soil depth" in the dataset refers to horizon thickness. We have corrected its labeling in the revised dataset for improved clarity. Many of the sampling sites were located in high-elevation or steep-sloped areas with shallow, stony, or poorly developed soils, where it was difficult to clearly define the thickness of fragmented or discontinuous horizons. In such cases, rather than assigning potentially inaccurate or interpolated values, we retained these fields as missing to accurately reflect field conditions and ensure transparency in the dataset. Regarding bulk density, due to the extremely challenging and labor-intensive nature of soil sampling in remote mountain environments, we were only able to collect one bulk density sample per soil horizon at each site, without technical replicates. In our original dataset structure, the bulk density values were stored in the same table as other soil property data that did include replicates. This may have caused confusion or missing fields during data integration. To address this issue, we have now separated the bulk density measurements into a dedicated data file named "Soil bulk density", in which each row represents a unique sample per soil layer without replicated values. We have updated the metadata file and the dataset structure accordingly to make these details more explicit and user-friendly. The updated dataset is available at the following DOI: https://doi.org/10.11888/Terre.tpdc.302620 or https://cstr.cn/18406.11.Terre.tpdc.302620.

It will also help if the authors include the soil type for each soil layer when available. This helps make sense of the elemental composition, CIA, and many other variables contained in the database.

Response: We appreciate your insightful suggestion regarding the inclusion of soil type information. In response, we have incorporated soil type data into our dataset to enhance its interpretability and usability. Specifically, we obtained soil classification data from the 1:1000000 Soil Map of the People's Republic of China, which was compiled and published by the National Soil Census Office in 1995. These data were accessed through the Resource and Environmental Science Data Center (http://www.resdc.cn). The updated dataset is available at the following DOI: https://doi.org/10.11888/Terre.tpdc.302620 or https://cstr.cn/18406.11.Terre.tpdc.302620.

The dataset does not contain the uncertainties of reported values.

Response: We sincerely thank you for this important point. In this dataset, we did not report measurement uncertainties primarily because all elemental and physicochemical analyses were conducted in certified laboratories following standardized national protocols, with strict quality assurance and quality control (QA/QC) procedures, including the use of blanks, standard reference materials, and analytical replicates. These procedures ensured high analytical precision and reproducibility. Moreover, we have provided detailed descriptions of the QA/QC protocols and the analytical precision of key measurements in the "Physicochemical analysis" section of the methods. For example, we reported the relative standard deviations (RSDs) and recovery rates of standard reference materials used in ICP-OES and ICP-MS measurements. We believe this additional information will help users better assess the reliability and quality of the dataset.

An existing dataset of soil properties has recently been published in ESSD (https://doi.org/10.5194/essd-17-517-2025). How does your dataset compare to this?

Responses: Thank you very much for pointing out the recently published dataset by Shi et al. (2025). We have carefully reviewed their study and would like to clarify the distinctions between our dataset and theirs in terms of data type, sampling strategy, element coverage, and research focus.

The dataset of Shi et al. is a model-derived product based on digital soil mapping techniques and machine learning, utilizing over 11,000 legacy soil profiles primarily from the 1970s-1980s. Their product provides gridded data (90 m resolution) for 23 soil physical and chemical properties across six fixed depth intervals (0-5 m, 5-15, 15-30, 30-60, 60-100, 100-200), focusing mainly on agricultural and land surface modeling applications. The dataset is based on standardized depth layers and primarily includes properties such as pH, bulk density, porosity, organic carbon, CEC, and major nutrients (N, P, K), but does not include concentrations of most soil elements (e.g., trace metals and micronutrients).

In contrast, our dataset is based on newly collected soil samples from 166 sites across 30 mountain regions in China, with a total of over 1,300 samples. Importantly, sampling was conducted by pedogenic horizons (O, A, C), rather than fixed-depth increments, to better capture soil development processes in mountainous environments. Our dataset focuses on the measured concentrations of 24 elements, including macronutrients (e.g., K, Ca, Mg), micronutrients (e.g., Fe, Mn, Cu, Mo), and trace metals (e.g., Cd, Pb, Sb), which are currently underrepresented in national-scale datasets. Furthermore, while the dataset of Shi et al. aimed to support generalized land surface modeling, our dataset is designed to fill critical data gaps in mountain ecosystem studies, with a particular focus on elemental stratification, biogeochemical modeling, and responses to environmental gradients (e.g., climate, altitude, parent material).

Therefore, we believe our dataset offers original, fine-resolution, multi-element observational data in mountain soils. It fills a critical data gap for understanding element cycling and environmental responses and provides a valuable foundation for biogeochemical modeling, ecosystem assessments, and global change research in mountainous regions.

Manuscript

L1 – A more appropriate title would be "Multi-Element dataset of soil profiles across (diverse) climatic zones in China's mountains" since the focus is on the soil.

Responses: Thank you very much for your valuable suggestion regarding the manuscript title. We agree that the revised title you proposed more accurately reflects the focus of our dataset on soil profiles. Accordingly, we have adopted your recommendation and revised the manuscript title to: "Multi-element dataset of soil profiles across climatic zones in China's mountains."

L17-19 – This has to be stated the other way around as this is a data description paper. Talk about how the dataset could contribute to the better understanding of China's mountain ecosystems. Also, include a statement about how this is related to soil.

Responses: Thank you for the insightful suggestion. We agree that the sentence structure should better reflect the nature of a data descriptor paper. In response, we have revised the opening sentence of the abstract to emphasize the relevance and contribution of the dataset to mountain ecosystem studies, with a clear focus on soils. The revised version is as follows:

Datasets of soil multi-element concentrations are essential for advancing our understanding of ecological functioning and responses to global change in mountain regions. However, the paucity of such datasets represents a fundamental impediment to accurately assess and predict biogeochemical processes in these sensitive ecosystems.

L103 – The materials and methods lack the details needed especially in the analytical part. It also lacks a description of the statistical process that does not prepare readers about what to expect in the following parts of the paper.

Responses: We sincerely appreciate for your constructive comments. In the revised manuscript, we have improved the section of Materials and methods by providing details on sampling procedures and physicochemical analyses (Lines 116-121, Lines 126-130, Lines 133-137, Lines 139-143). Moreover, we add a subsection of Statistical analysis to clearly describe the methods used for data interpretation (Lines 175-186). These revisions aim to better understand the workflow of sample collection, laboratory analysis, and subsequent data processing. The revised Soil physical and chemical analyses and Statistical analysis sections are as follows:

Soil physical and chemical analyses

Soil moisture content was determined by oven-drying the soils at 105°C to constant mass, which was calculated using the following formula:

$$Moisture\ (\%) = \frac{Fresh\ weight - Dry\ weight}{Dry\ weight} \times 100\%$$

Soil pH was measured using a pH meter (Mettler-Toledo FE28, Switzerland) after shaking the soil samples with deionized water at a 1:2.5 soil-to-water ratio. Soil organic carbon (SOC) concentration was determined by a CE400 elemental analyzer (Elementar vario ISOTOPE cube, Germany), after removing carbonates with 5% HCl. Soil samples for element analysis were digested with concentrated HNO_3, HF, and $HClO_4$ (Bing et al., 2022). The concentrations of major elements (Al, Ba, Ca, Fe, K, Mg, Mn, Na, P, Sr, Ti, V, and Zn) in the digests were determined using an inductively coupled plasma atomic emission spectrometry (ICP-AES, Optima 2000, USA), and the concentrations of trace elements (Cd, Co, Cr, Cu, Mo, Ni, Pb, Sb, and Tl) were determined using an inductively coupled plasma mass spectrometry (ICP-MS, Agilent 7700x, USA), with SPEXTM serving as the standard solution. Quality control was ensured by analyzing replicates, blanks, and reference material (BW07405, China). The recovery of the reference material was routinely within the range of 95-105%, and the precision and accuracy of the analyses were < 5% (relative standard deviation).

Statistical analysis

All statistical analyses were conducted using R (version 4.3.1). To test differences in element concentrations among soil horizons, we employed linear mixed-effect models using the “lmer” function from the “lme4” package, where soil horizon was treated as a fixed factor and sampling site as a random factor. Regression analyses were conducted to examine the spatial distribution characteristics of each element. To explore the compositional differences in elemental assemblages across soil horizons and to assess the influence of environmental variables on soil element variation, redundancy analysis (RDA) was conducted using the “rda” function in the “vegan” package. Correlation analyses were conducted separately for each soil horizon to identify horizon-specific relationships between elemental concentrations and environmental drivers. Furthermore, simple linear regression was employed to quantify the individual explanatory power (R^2) of each environmental variable for each element. The cumulative explanatory power of all environmental factors was also calculated to evaluate their combined influence on element variation.

L106 – This figure includes a disputed territory which is not relevant to the data because the authors did not sample in these areas. The authors are therefore advised to either remove the disputed territory from the map or include a statement regarding this in the figure caption. Make sure that it is aligned with ESSD's policy on neutrality regarding jurisdictional claims. I leave it to the journal editors to decide on this.

Response: We sincerely thank you for pointing this out. In accordance with the editor’s suggestion and the journal’s policy on neutrality regarding jurisdictional claims, we have revised the figure to focus only on the regions where sampling was actually conducted. The revised Figure is as follows:

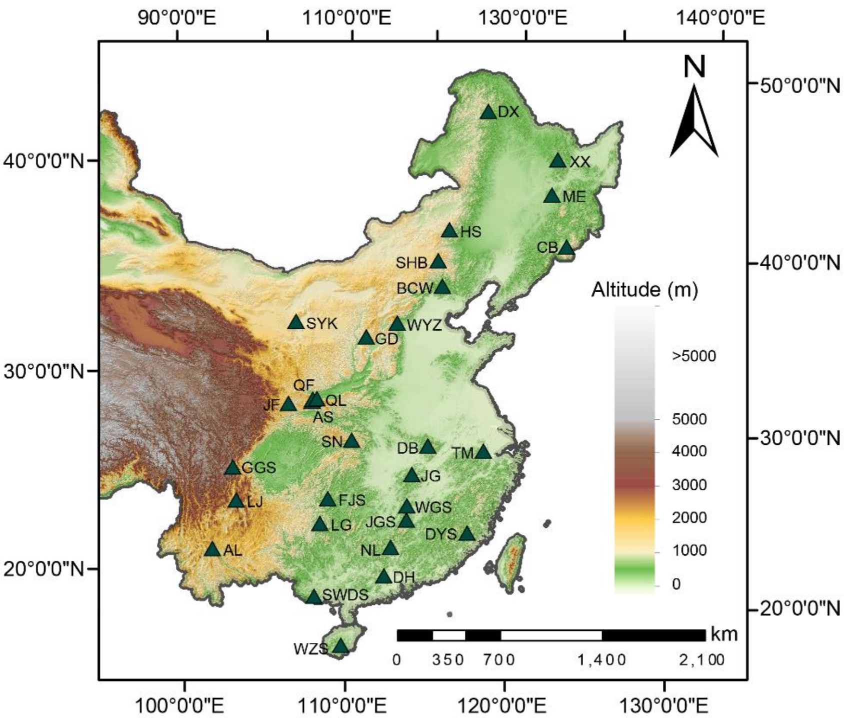

Fig. 1 Geographic distribution of the 30 China's mountains. AL, Mt. Ailao; AS, Mt. Ao; BCW, Mt. Baicaowa; CB, Mt. Changbai; DB, Mt. Dabie; DH, Mt. Dinghu; DX, Mt. Daxinganling; DYS, Mt. Daiyun; FJS, Mt. Fanjing; GD, Mt. Guandi; GGS, Mt. Gongga; HS, Mt. Han; JF, Mt. Jifeng; JG, Mt. Jiugong; JGS, Mt. Jinggang; LJ, Mt. Luoji; LG, Mt. Leigong; LJ, Mt. Luoji; ME, Mt. Maoer; NL, Mt. Nanling; QF, Mt. Qingfengxia; QL, Mt. Qinling; SHB, Mt. Saihanba; SN, Mt. Shennongjia; SWDS, Mt. Shiwandashan; SYK, Mt. Suyukou; TM, Mt. Tianmu; WGS, Mt. Wugong; WYZ, Mt. Wuyuezhai; WZS, Mt. Wuzhi; XX, Mt. Xiaoxinganling.

L113 – Provide the dates when the sampling was done.

Response: Thank you for your feedback. The soil sampling was conducted from July 2012 to March 2013. This information has been clearly incorporated into the main text to improve clarity.

L135 – How much soil was used in the analyses and how many replicates?

Response: We appreciate for your great comments. For each soil horizon at every sampling site, we collected three replicate samples. Each replicate was a composite sample, thoroughly mixed from subsamples taken in each horizon. In total, 1,314 soil samples were collected and sent to the laboratory for analysis. During laboratory procedures, all replicates were analyzed separately rather than being pooled, ensuring the reliability of the data. We have revised the manuscript to include a more detailed description of the sampling procedure to enhance clarity (Lines 118-121, Lines 126-130).

L156 – It is not clear how the oxide values were obtained from the elemental analysis. We're certain ratios used? If so, it has to be clearly stated in the manuscript.

Response: Thank you very much for the professional comment. In our study, the CIA was calculated based on the molar proportion of oxides estimated from the total elemental concentrations using the widely accepted formula proposed by Nesbitt & Young (1982):

$$CIA = \frac{Al_2O_3}{(Al_2O_3 + Na_2O + K_2O + CaO^*)} \times 100$$

Specifically, oxide contents were derived by converting elemental concentrations to their corresponding oxide forms using standard molecular weights, a common practice in many studies (e.g., Ochoa-Hueso et al., 2021). Regarding CaO*, we acknowledge that calcium can be present in multiple forms, including silicates, carbonates, phosphates, and exchangeable forms. To ensure that the CIA reflects only the contribution from silicate weathering, we applied the correction method proposed by McLennan (1993), which has been widely adopted in geochemical studies. This method estimates CaO* as follows: If the measured CaO concentration is less than or equal to Na_2O, CaO* is assumed to be equal to the measured CaO. If CaO > Na_2O, CaO* is set equal to Na_2O. This approach helps to minimize the influence of non-silicate calcium on the CIA calculation and thus improves the reliability of the weathering index. We have revised the Methods section accordingly to clarify this procedure (Lines 168-173).

It is also worth noting that CIA is not a central focus of this dataset. Our primary aim is to provide a comprehensive multi-element concentration dataset across vertically stratified soil horizons in mountain regions, which can support a wide range of research, including but not limited to weathering assessments. Nevertheless, we included the CIA to offer a useful reference for users interested in assessing the degree of chemical weathering.

References:

Nesbitt, H. W., and Young, G. M.: Early Proterozoic climates and plate motions inferred from major element chemistry of lutites, Nature 299, 715-717, https://doi.org/10.1038/299715a0, 1982.

McLennan, S. M.: Weathering and Global Denudation, J. Geol. 101, 295-303, https://doi.org/10.1086/648222, 1993.

Ochoa-Hueso, R., Plaza, C., Moreno-Jimenez, E., Delgado-Baquerizo, M.: Soil element coupling is driven by ecological context and atomic mass, Ecol. Lett. 24, 319-326., https://doi.org/10.1111/ele.13648, 2021.

L173-177 – This doesn't mean anything and is already given information for almost all soil types.

Response: We appreciate for your comment and understand that elemental abundance patterns are generally consistent across most soil types. However, our intention in presenting the concentration ranges and mass ratios is to provide a comprehensive overview of the relative abundance and variability of the 24 elements specifically within mountain soils across diverse climatic and geological settings. While such information may seem well-established for common soil types, it remains underreported for high-resolution, multi-element datasets derived from standardized profiles in mountainous ecosystems.

L246 – How can you have explanation by environmental factors exceeding 100%?

Response: Thank you for raising this important point. In Fig. 6, the y-axis represents the cumulative explanatory power of all environmental variables for each individual element. Therefore, the total explanatory value exceeding 100% is mathematically possible. To avoid confusion, we have clarified this aspect in the "Statistical analysis" section of the revised manuscript and have also updated the figure legend accordingly.