Lithology
shale
mudstone
mud
ooze
marl
clay
slate
argillite
meta-argillite
pelite
metapelite
claystone

Table S1: List of valid lithologies. Any samples with a lithology that did not match one in this list—or lacked lithology information altogether—were removed from the dataset during preliminary filtering.

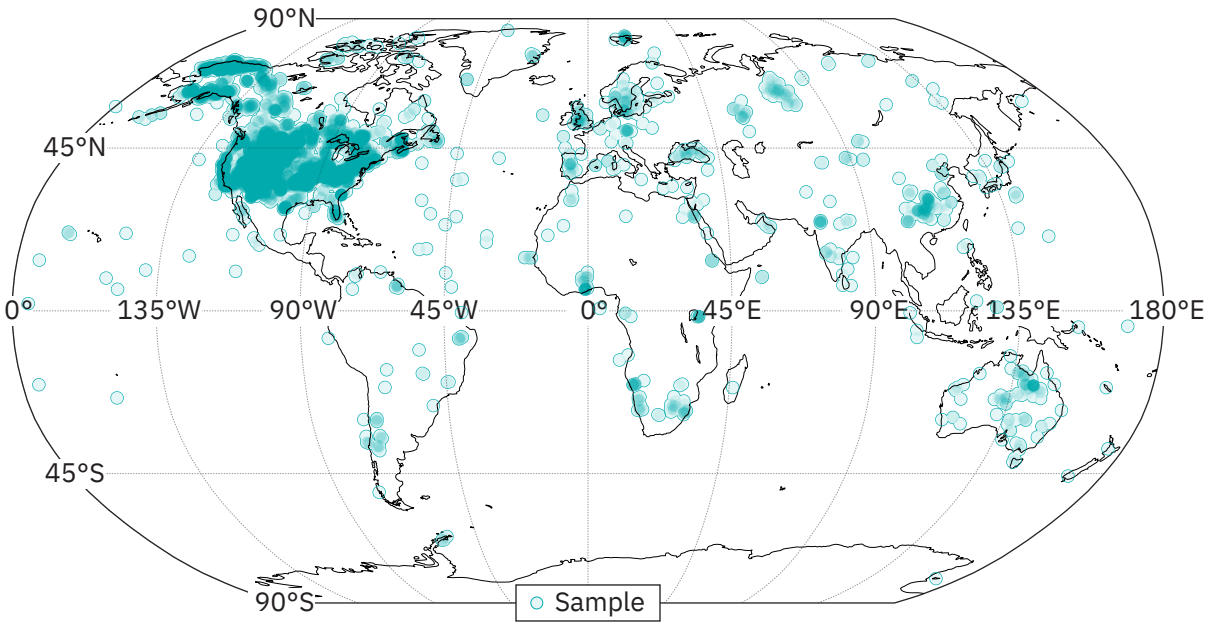

Figure S1: **Map showing locations of samples from the SGP database.** Note the uneven spatial distribution—more data are available from North America than the middle of Africa, for example.

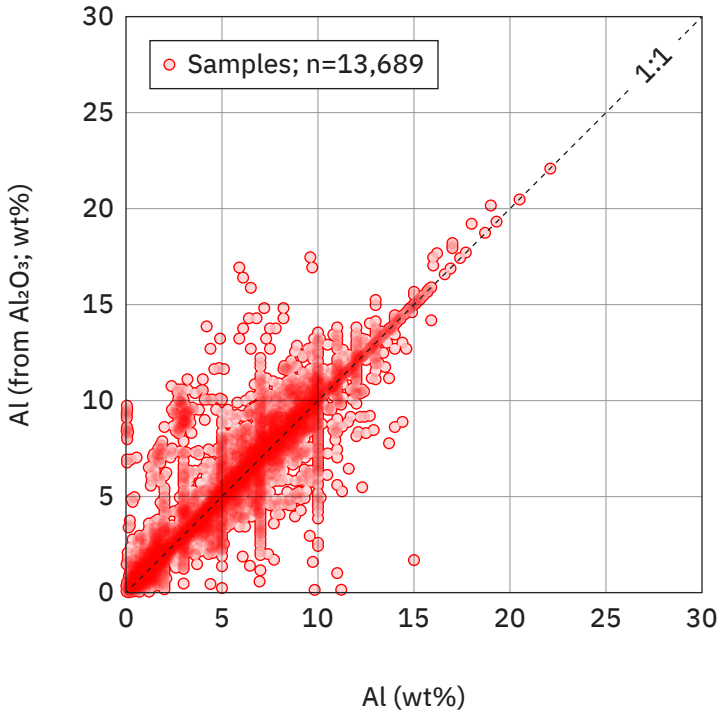

Figure S2: **Al contents, in wt%, derived from 2 different analytes per sample.** Values that fall off the 1:1 line suggest that both analytes do not provide the same answer. Importantly, most discordant values are from USGS-NGDB data with unknown methods. Additionally, many values that fall on the 1:1 line are from USGS-CMIBS and, since the data source reported both converted oxide and elemental values, are not true measurement replicates. Appropriate filtering methodologies can address both issues and provide robust data for analysis (for example, when discordant values and USGS-CMIBS samples are removed, the remaining data have an R^2 of 0.98).

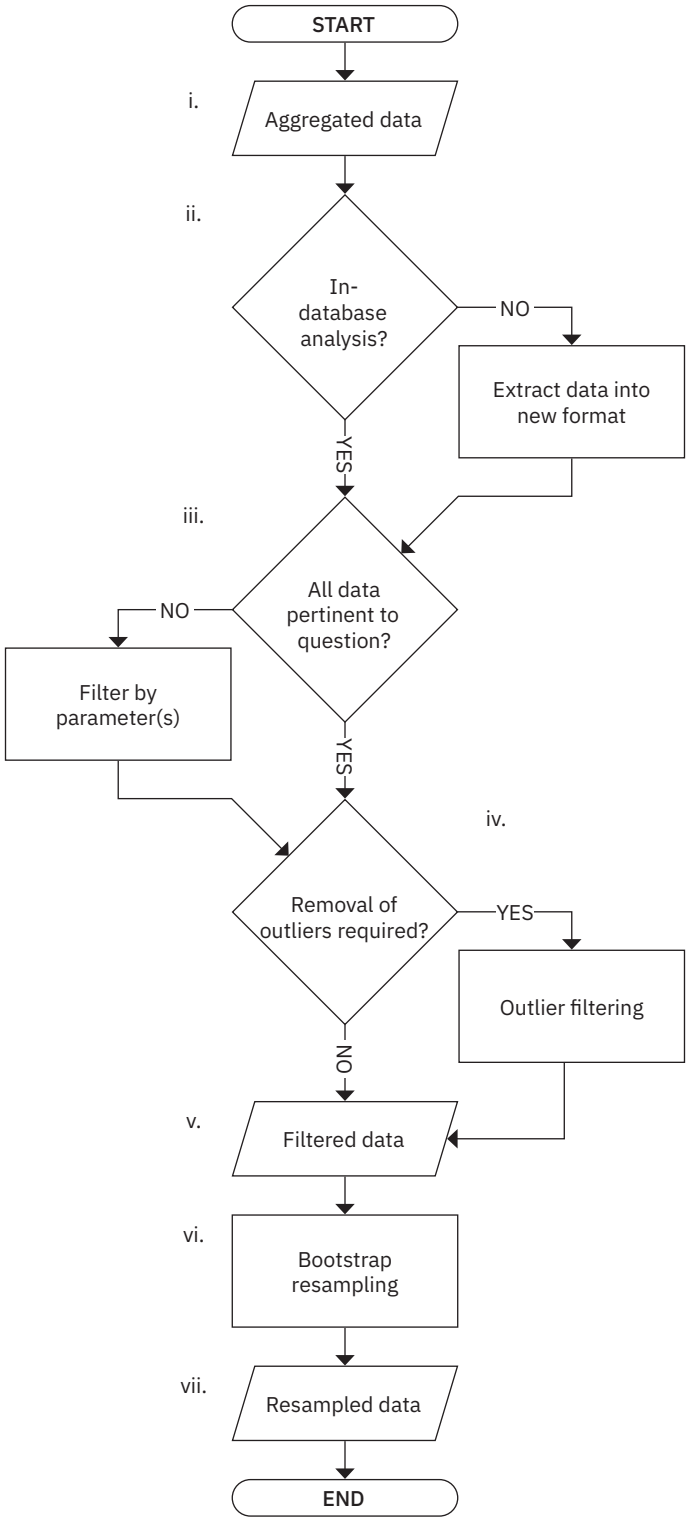

Figure S3: **Flowchart depicting procedure for extracting temporal trends from compiled data.** Crucially, this workflow is generalizable. The implementation of each step is up to individual researchers.

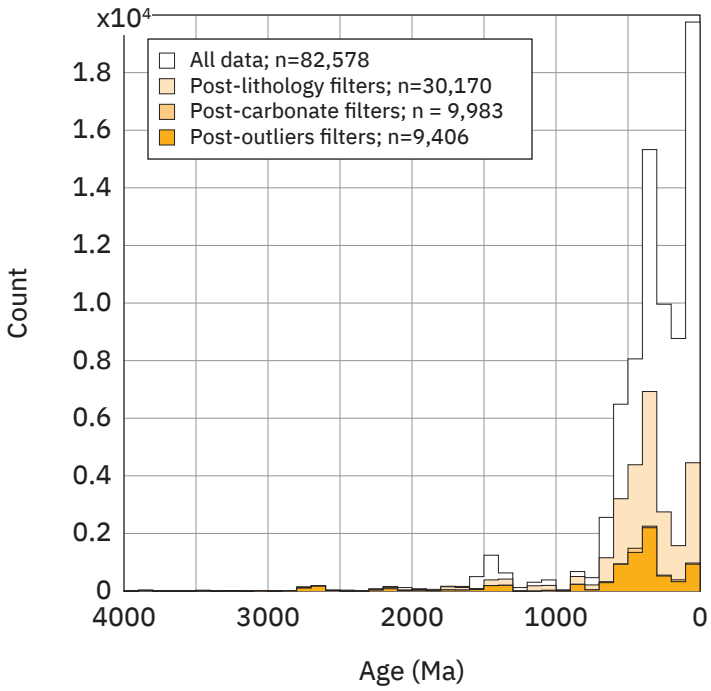

Figure S4: **Histograms depicting the effects of progressive filtering on the SGP database.** Through various filters described in the text, most of the SGP data were excluded from consideration for resampling (and, in turn, the resulting trends). Careful filtering of geochemical datasets is crucial for producing meaningful local and/or global trends from the data. Importantly, while we determined that these filtering steps were key for our study, choices about how and what to filter certainly are question-dependent.

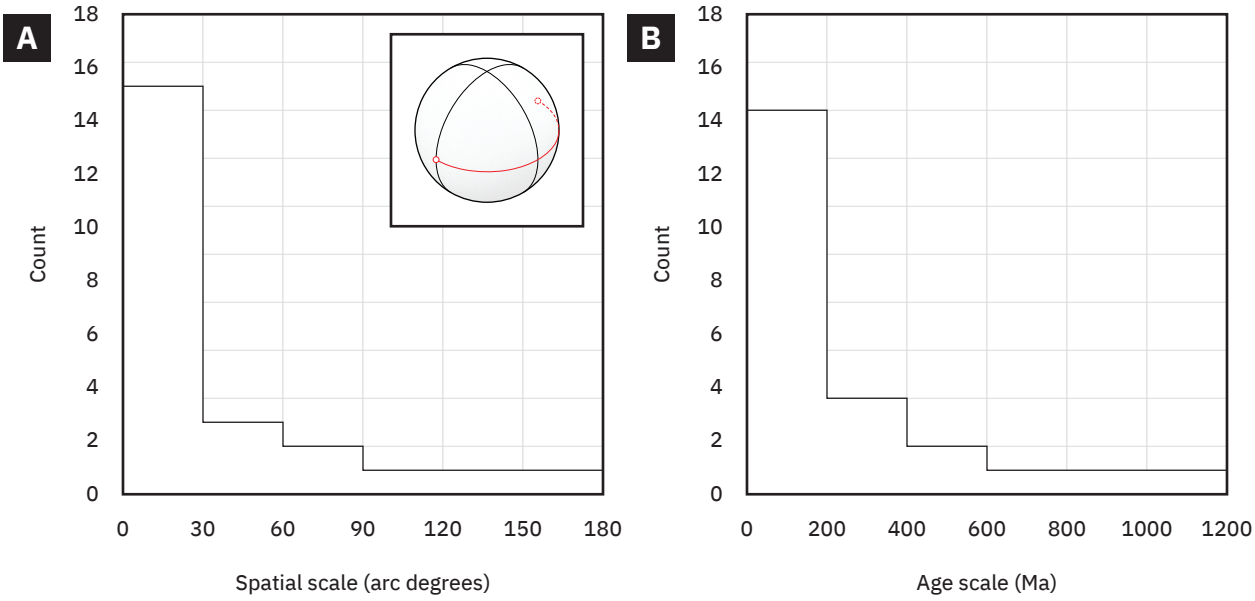

Figure S5: **Distributions of spatial and age scales used in the sensitivity tests in this paper. A.** Histogram depicting spatial scales (i.e., $scale_{spatial}$) in arc degrees. Inset, an illustration showing an arc of length 180 degrees on a sphere that is representative of Earth. **B.** Histogram of age scales (i.e., $scale_{temporal}$) in millions of years (Ma).

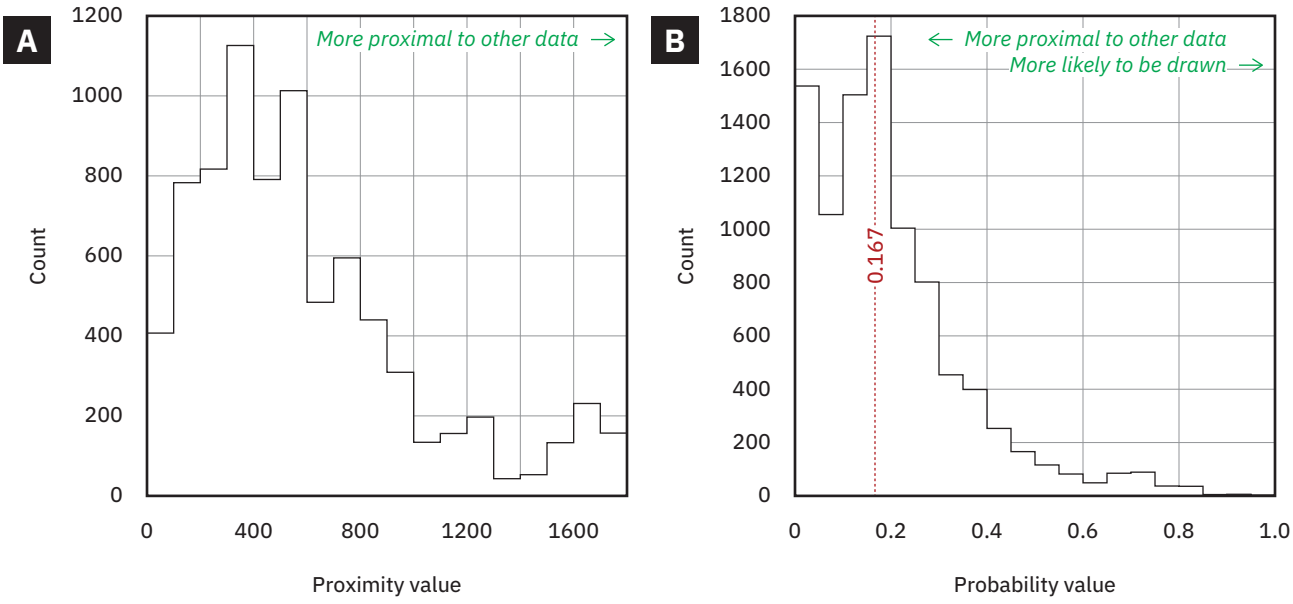

Figure S6: **Distribution of proximity values and resulting probability values for a $scale_{age}$ of 0.5 and $scale_{temporal}$ of 10. A.** Raw proximity values for all samples in the final filtered dataset. **B.** The final probability values, generated by using proximity values. The red dashed line shows the median probability value for all samples.

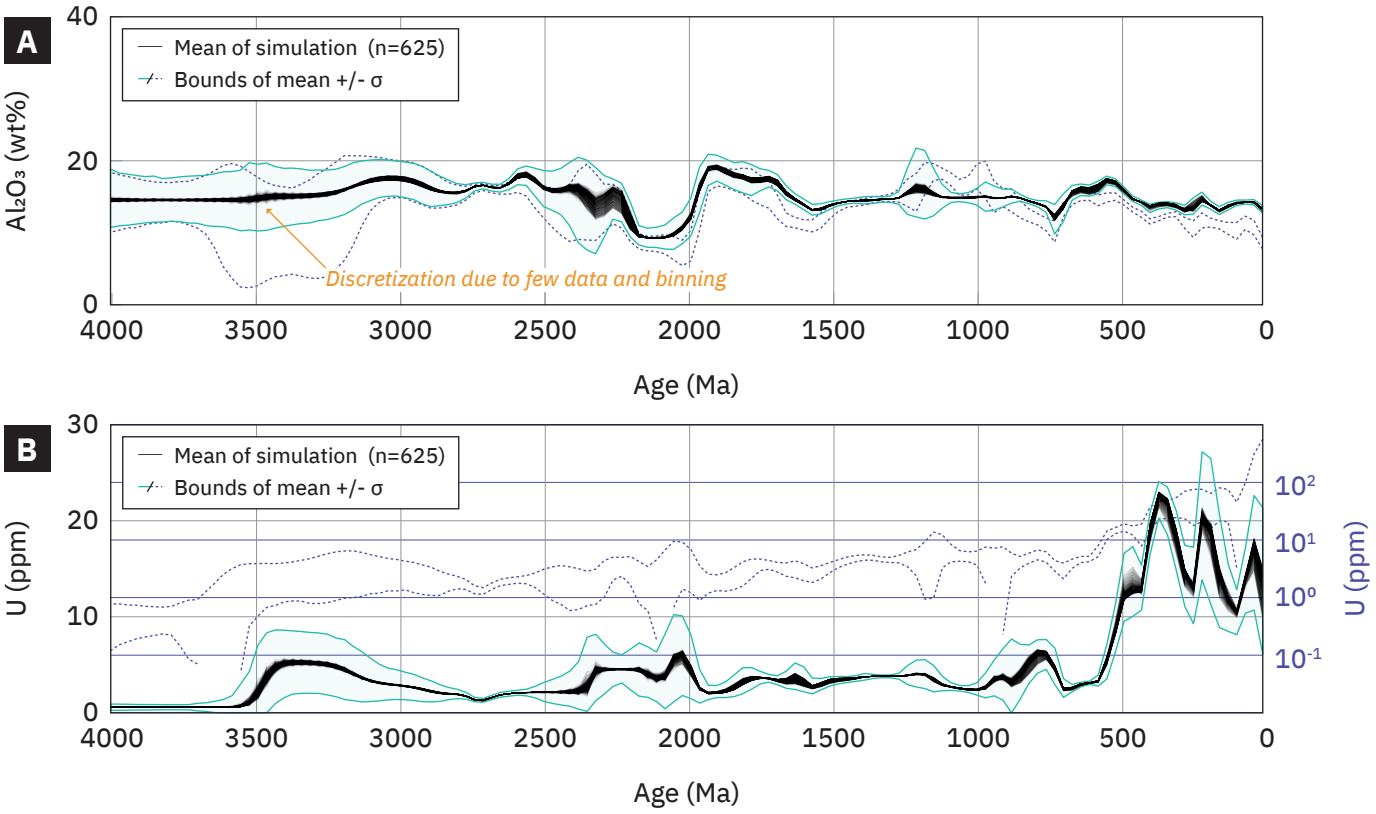

Figure S7: **Results of parameter sensitivity tests.** 625 iterations were run using various combinations of $scale_{age}$ and $scale_{temporal}$ (see Fig. S5). **A.** Means (black lines) and upper and lower error bounds (green lines, with shading between the bounds) for Al_2O_3 . **B.** Means and upper and lower error bounds for U. For both **A** and **B**, dashed blue lines are the upper and lower error bounds resulting from a pre-carbonate-and-outliers-filter dataset. Note the log scale for the right y-axis in **B**. Discontinuities in U bounds are the result of the lower bound equaling 0.