https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

Open Access Earth System Science Data Discussions

The global database of deep-time marine nitrogen isotope data

Yong Du[1], Huyue Song[1,*], Thomas J. Algeo[1,2,3,4], Hui Zhang[1], Jianwei Peng[1], Yuyang Wu[1,5], Jiankang Lai[1], Xiang Shu[1], Hanchen Song[1], Lai Wei[6], Jincheng Zhang[7], Eva E. Stüeken[8], Stephen E. Grasby[9], Jacopo Dal Corso[1], Xiaokang Liu[1], Daoliang Chu[1], Li Tian[1], Qingzhong Liang[7], Xinchuan Li[7], Hong Yao[7], Haijun Song[1]

[1]State Key Laboratory of Geomicrobiology and Environmental Changes, School of Earth Sciences, China University of Geosciences, Wuhan 430074, China

[2]State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan 430074, China

[3]Department of Geosciences, University of Cincinnati, Cincinnati, OH 45221-0013, USA

[4]State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Chengdu University of Technology, Chengdu 610059, China

[5]College of Marine Science and Technology, China University of Geosciences, Wuhan 430074, China

[6]Shool of Future Technology, China University of Geosciences, Wuhan 430074, China

[7]School of Computer Science, China University of Geosciences, Wuhan 430074, China

[8]School of Earth & Environmental Sciences, University of St Andrews, St Andrews KY16 9AL, UK

[9]Geological Survey of Canada, Natural Resources Canada, Calgary, Alberta T2L 2A7, Canada

Correspondence to: Huyue Song (hysong@cug.edu.cn)

Abstract. Stable nitrogen isotope records preserved in marine sediments provide critical insights into Earth's climate history and biospheric evolution. Although numerous studies have documented nitrogen isotope ($\delta^{15}N$) records across varied geological ages (Archean to Recent) and paleogeographic settings, the scientific community remains constrained by the absence of a standardized database to systematically investigate their spatiotemporal evolution. Here, we present the database of Deep-time Sediment Nitrogen Isotopes in Marine Systems (DSMS-NI), a comprehensive global compilation of $\delta^{15}N$ data and associated geochemical parameters, spanning a vast collection of sediment samples dating from the Recent to the Archean. This database encompasses 71 040 $\delta^{15}N$ records derived from 424 publications, systematically organized with 29 metadata fields categories (e.g., chronostratigraphic ages, coordinates, lithology, metamorphic grade, sedimentary facies, references) encompassing 1 927 829 metadata. This repository further incorporates 130 proxy data fields, including 285 715

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

geochemical data spanning total organic carbon (TOC), total nitrogen (TN), and organic carbon isotopes ($\delta^{13}C_{org}$), major and trace elements and iron species. These integrated parameters enable evaluation of sample fidelity and factors influencing $\delta^{15}N$ signatures. The DSMS-NI database will facilitate research across key geological intervals such as the Permian-Triassic boundary and the Cretaceous ocean anoxic events. Researchers can leverage temporal and paleogeographic information, alongside geochemical data, to conduct spatiotemporal analyses, thereby uncovering changes in deep-time marine nitrogen cycles and paleoenvironmental conditions. The database is open-access via the Geobiology portal (https://geobiologydata.cug.edu.cn/, last access: 30 April 2025), allowing users to access data and submit new entries to ensure continuous updates and expansion. This resource represents a vital foundation for studies in paleoclimate, paleoenvironment, and geochemistry, offering essential data for understanding long-term Earth-system processes. The data files described in this paper are available at https://doi.org/10.5281/zenodo.15117375 (Du et al., 2025a)

1 Introduction

Nitrogen, as an essential nutrient and redox-sensitive element, plays a crucial role in biological evolution and environmental climate changes (Ader et al., 2016; Pellerin et al., 2024). Typically, nitrogen isotope compositions are reported as a relative deviation of sample's isotopic ratio relative to that of atmospheric N_2, expressed in per mille (‰) as $\delta^{15}N = (R_{sample}/R_{AIR\text{-}N2} - 1) \times 1000$ ‰ where $R = {}^{15}N/{}^{14}N$. The $\delta^{15}N$ record has become one of the primary tools for tracing the evolution of the nitrogen cycle and reconstructing redox conditions through deep time (Algeo et al., 2014; Sahoo et al., 2023; Du et al., 2024; Moretti et al., 2024). Advances in analytical techniques have facilitated rapid growth in the application of $\delta^{15}N$ for paleoenvironmental studies in recent decades (Fig. 1; Zhong et al., 2023). Given nitrogen's short marine residence time of approximately 3000 years, which leads to regionally variable and rapidly shifting patterns (Gruber and Galloway, 2008), high-resolution $\delta^{15}N$ datasets with detailed temporal and spatial coverage are critical for elucidating nitrogen cycle dynamics through Earth history.

Existing compilations of deep-time marine $\delta^{15}N$ records exhibit significant limitations in term of temporal coverage and metadata compliance. Previous efforts have focused specifically on Precambrian to investigate the origins of microbial nitrogen metabolism and redox evolution during the Great Oxidation Event (Thomazo et al., 2011; Stüeken et al., 2016, 2024; Kipp et al., 2018; Uveges et al., 2025). Other studies have targeted Phanerozoic systems (Algeo et al., 2014) or specific intervals

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

such as the Paleozoic (Koehler et al., 2019), Cenozoic (Tesdal et al., 2013), Cambrian (Wang et al., 2018; Liu et al., 2020), and Triassic (Sun et al., 2024) to analyze key biological and environmental events. The largest compilation of data from pre-Cenozoic records contains fewer than 8000 $\delta^{15}N$ entries, with much of the data being repetitive across different datasets (Stüeken et al., 2024). In contrast, Tesdal et al. (2013) compiled up to 33 352 entries, but all of these records are from the past 6 million years. Moreover, these repositories often fail to adhere systematically to the FAIR (Findable, Accessible, Interoperable, Reusable) data principles (Wilkinson et al., 2016) and offer limited metadata categories. Typically, they provide only broad geologic ages, lithology, and metamorphic grades, while lacking essential metadata such as paleogeographic coordinates, depositional environments, and high-resolution chronostratigraphy (Table 1). Current metadata-rich databases that follow FAIR principles remain limited to fewer than 3000 $\delta^{15}N$ entries (e.g., Farrell et al., 2021; Lai et al., 2025), highlighting the urgent need for a rigorously standardized, spatiotemporally comprehensive $\delta^{15}N$ database.

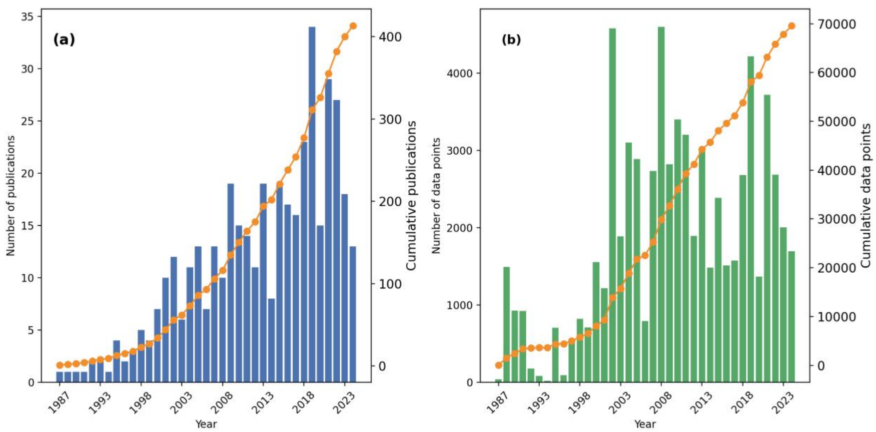

Figure 1. Temporal trends in (a) nitrogen isotope publications and (b) $\delta^{15}N$ data entries in the DSMS-NI database. Vertical bars denote annual publication/dataset counts, while dots connected by lines represent cumulative totals over the years.

The DSMS-NI database, a repository of deep-time sediment nitrogen isotopes in marine systems spanning Earth history, aims to address this need. The DSMS-NI database is a part of the broader GBDB (Geobiology Database) project, which aims to build a comprehensive database of biotic and biogeochemical evolution throughout time and to explore the mechanisms driving these evolutionary processes. By integrating detailed metadata, DSMS-NI provides a valuable resource for studying

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

nitrogen cycle evolution and paleoenvironmental conditions across a range of temporal and spatial scales. This compilation provides an extensive survey of $\delta^{15}N$ records for all marine sediment types, with a particular emphasis on data predating the Cenozoic Era. Derived from 424 peer-reviewed publications and publicly available datasets, it currently encompasses 71 040 discrete $\delta^{15}N$ measurements for various components (e.g., bulk rock, shell-bound, kerogen). In addition, it includes roughly 285 715 associated data points for carbon, sulfur isotopes, and major and trace element concentrations reported alongside the $\delta^{15}N$ values. Each entry is linked to a comprehensive set of standardized metadata, ensuring consistency and facilitating robust data analyses. Our goal is to make DSMS-NI a dynamic, evolving database that improves over time, with data visualizations updated concurrently on the Geobiologydata website (https://geobiologydata.cug.edu.cn/, last access: 30 April 2025).

Table 1 Overview of deep-time $\delta^{15}N$ complication.

Data Source	Number of $\delta^{15}N$ record	Metadata	Spatial range	Temporal range
Tesdal et al. (2013)	33 352	Fine age; Modern coordinate; Site	Global	Neogene to Present
Algeo et al. (2014), restricted access	6006	Broad age; Formation	Global	Ediacaran to Present
Stüeken et al. (2016)	6449	Broad age; Formation; Lithology; Metamorphic grade	Global	Since the Paleoarchean
Stüeken et al. (2024)	10 584	Broad age; Formation; Lithology; Metamorphic grade	Global	Since the Eoarchean
Kipp et al. (2018)	6468	Broad age; Formation; Lithology; Metamorphic grade	Global	Since the Paleoarchean
Koehler et al. (2019)	2454	Crude age; Formation; Lithology; Metamorphic grade	Global	Paleozoic
Farrell et al. (2021), SGP database	840	Broad age; Modern coordinate	Global	Paleozoic and Eidacaran
Lai et al. (2025), DM-SED database	2561	Fine age; Modern coordinate; Paleo-coordinate ; Site; Formation; Depositional environments;	Global	Since the Neoproterozoic

Lithology; Metamorphic grade

Version 0.0.1 of the DSMS-NI database is available in CSV format on Zenodo (https://doi.org/10.5281/zenodo.15117375), and dynamic updates will be maintained on the GeoBiology website. The following sections provide a comprehensive overview of the database compilation methods, data structure, and details of the dataset, including data sources, selection criteria, and definitions of metadata fields. Additionally, we analyze the temporal and spatial trends of $\delta^{15}N$ within the dataset, discuss potential applications and limitations, and outline the foundation for the database's continuous development and scientific utility.

2 Compilation methods

2.1 Data compilation

An extensive search was conducted based on published articles, reports, theses, and datasets to gather all available literature on deep-time nitrogen isotopes. Initially, a keyword-based search combining geological period and nitrogen isotope was performed on Google Scholar, yielding over 3000 relevant literature sources after removing duplicates. A significant portion of the articles, however, only discussed previously published $\delta^{15}N$ data, rather than presenting newly measured data, which were manually excluded from the data compilation. Additionally, geochemical databases such as PANGAEA (https://www.pangaea.de/, last access: April 1 2025), EarthChem (https://www.earthchem.org/, last access: April 1 2025), SGP (https://sgp-search.io/, last access: April 1 2025), and NOAA (https://www.ncei.noaa.gov/, last access: April 1 2025) were queried to ensure comprehensive coverage of dataset sources (Diepenbroek et al., 2002; Gard et al., 2019; Farrell et al., 2021). Where overlaps existed between datasets and publications, journal articles were prioritized as the primary data sources. Further filtering excluded studies on non-marine sediments, entries lacking essential metadata (e.g., geological age, latitude and longitude), and a limited number of Cenozoic records with inaccessible data. Ultimately, the curated dataset includes 424 valid sources published between 1983 and 2024, representing a comprehensive compilation of nitrogen isotope records for deep-time marine sediments.

Data from each publication were stored in various formats, including tables within the main text, supplementary files, or shared databases. Data extracted from tables and supplementary files were initially processed by computer algorithms, followed by manual verification and supplementation. For databases, data files were downloaded manually. In cases where publications did not provide direct

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

data, data points were extracted from figures using GetData Graph Digitizer (ver. 2.24), and these entries were labeled as "plot" in the Notes section. Each publication was then organized into an individual data file with clear labeling of sources and unique site identifiers. These files were subsequently merged into a master dataset based on standardized column headers. In the final master dataset, additional metadata were curated, including geological age, latitude and longitude, lithology, depositional facies, and metamorphic grade. High-resolution ages and paleo-coordinates were calculated and converted where applicable.

2.2 Data selection and quality control

Given that biogeochemical and paleoenvironmental studies based on nitrogen isotopes require the assessment of the depositional environment and post-depositional alteration, geochemical data apart from $\delta^{15}N$ are crucial (Tribovillard et al., 2006; Robinson et al., 2012). Therefore, we collected other contemporaneously published geochemical data of the same samples as $\delta^{15}N$ from the literature relevant to the formations in our database. All available data from each research site were included as comprehensively as possible, rather than excluding entries solely due to the absence of $\delta^{15}N$ values. This approach allows for the potential interpolation of the time-series data. However, geochemistry fields with fewer than 100 data points in the final compilation were excluded due to their limited analytical utility, such as Mo and Fe isotopes.

To ensure the reliability and applicability of the data, each entry underwent a rigorous screening and evaluation process. Initially, we assessed the data source and its spatiotemporal context. All studies included in the database were required to report verified $\delta^{15}N$ values with clear data provenance and well-defined spatiotemporal information. Data entries lacking traceable sources were excluded. Similarly, entries without precise geographic or temporal information were not considered. Data from highly heterogeneous geological settings with insufficient sampling resolution were also filtered out to minimize potential biases stemming from spatial variability in lithology or metamorphic grade. For studies reporting multiple measurements across a range of water depths, depositional facies, lithologies, or fossil shells, each measurement was recorded as an independent entry. The $\delta^{15}N$ values for bulk rock, decarbonated rock, and fossil shells were classified as primary entries, while values for other components, such as kerogen, clay-bound nitrogen, and porphyrins, were categorized solely as secondary entries. Only primary entries were analyzed in the data visualizations presented later in this study.

Only $\delta^{15}N$ obtained through standardized, widely accepted techniques were included in the

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

database. These primarily consist of elemental analyzer-isotope ratio mass spectrometry methods applied to bulk rock, decarbonated fractions, or kerogen (Song et al., 2023), as well as denitrifier-based mass spectrometry methods for microfossils (Farmer et al., 2021). Studies employing non-standard or unvalidated methods, such as stepwise combustion (Ishida et al., 2017), were excluded. Data from highly metamorphosed settings (e.g., hydrothermal alteration), terrestrial lakes and rivers, modern organisms and their metabolic products, and liquid phases were flagged and omitted from the database (e.g., Bebout et al., 1999; Chase et al., 2019; Xia et al., 2022). For data from the same site but at different depths or lithologies, or for measurements of different components at the same water depth, or repeated measurements of the same sample under varying conditions, each entry was recorded separately to accurately capture variability. Statistical methods were applied to detect potential outliers in nitrogen isotope values. Isotopic values falling outside the expected range for a given geological context (e.g., extreme $\delta^{15}N$ values in specific sedimentary rocks) were flagged for further review. If these anomalies could not be reasonably explained by contextual data, they were excluded from the final compilation.

Metadata on paleo-coordinates, depositional setting, lithology, and metamorphic grade are included wherever available. Entries are not excluded due to missing metadata, as these can potentially be supplemented in future research. When such metadata are not directly reported in the literature, we attempt to estimate them using supplementary data or external sources, such as paleogeographic reconstructions. For entries where metadata cannot be determined, blank values are assigned.

3 Data summary

Since nitrogen isotope studies in sediments began in the late 1980s, the number of published studies has shown an accelerating growth trend, doubling approximately every decade. This trend is mirrored by a steady increase in data volume, with an average annual addition of around 2720 data points over the past two decades (Fig. 1). However, the rate of data growth slightly lags behind that of publications, largely because early Ocean Drilling Program (ODP) and Integrated Ocean Drilling Program (IODP) projects contributed substantial datasets within individual publications (e.g., Liu et al., 2008). Ocean drilling remains a vital component of the database, covering geological intervals since the Cretaceous. Some early drilling data were not initially publicly accessible and have been supplemented through existing literature compilations, particularly the substantial dataset from Tesdal et al. (2013), along with enriched metadata.

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

The DSMS-NI database comprises a total of 29 metadata fields and 130 proxy data fields, organized into five primary categories (Table 2): (1) Sampling location, (2) Age information, (3) Geochemical data, (4) Lithological characteristics, and (5) references. For clarity and consistency throughout this data descriptor, the term "entries" refers to individual proxy values and their associated metadata (i.e., rows), while "fields" denote the metadata attributes recorded for each entry (i.e., columns).

Table 2 Field names and descriptions.

Field name	Description
Sample ID and location fields	
SampleID	Unique sample identification code, as originally published
SiteName	Name of the drill core site or section
SampleName	Author denoted title for the sample (often non-unique, e.g., numbered)
Location1	Detailed location of the data collection site
Location2	Country or ocean of the data collection site
Latitude	Modern latitude of collection site rounded to two decimals; negative values indicate the Southern Hemisphere (decimal degrees)
Longitude	Modern longitude of the collection site rounded to two decimals; negative values indicate the Western Hemisphere (decimal degrees)
Paleolatitude	Palaeolatitude of collection site rounded to two decimals; negative values indicate the Southern Hemisphere (decimal degrees)
Paleolongitude	Palaeolongitude of the collection site rounded to two decimals; negative values indicate the Western Hemisphere (decimal degrees)
Age fields	
Era	The geologic era, in reference to GTS v202309
Period	The geologic period, in reference to GTS v202309
Epoch	The geologic epoch, in reference to GTS v202309
Stage	The geologic stage, in reference to GTS v202309
Age	Age, in reference to GTS v202309
Formation	Geologic formation name
Unit	Specific geologic event layers
RelativeDepth	Stratigraphic height or depth (m)

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

Petrological characteristic fields	
Lithology	Lithological name of the sample, as originally published
LithType	Lithology type of sample (e.g., carbonate, siliciclastic)
MetamorphicGrade	The degree to which the rock has undergone transformation due to heat and pressure conditions
Setting	Depositional environment (e.g., epeiric, bathyal)
WaterDepth	Estimated depositional water depth of the data collection site
Data fields	
Isotopes	The isotope composition expressed in per mille (‰) as δ (e.g., $\delta^{15}N$, $\delta^{13}C$)
Elements	The concentration of elements within rocks (e.g., TN, P, Fe, Cu, Ce)
RockEval	Proxies of hydrocarbon potential measured by pyrolysis method (e.g., S1, OI, Tmax)
FeSpecies	Concentrations and ratios of different iron species in rocks (e.g., Fepy, Fehr/Fet)
Reference fields	
FirstAuthor	The last name of the first author of the original publication
Year	The year of the original publication
Title	The title of the original publication
Reference	The formated reference of the original publication
DOI	The DOI of the original publication
DataSource	The repository hosting the data except for the original publication

Sample ID and Location fields. Each data entry is assigned a unique Sample ID to distinguish it from other data sources. Geographic location information includes the modern latitude and longitude (Latitude and Longitude) referencing WGS84 (World Geodetic System 1984), obtained directly from original literature or external sources whenever possible. For studies that do not provide exact coordinates, approximate locations are estimated based on geographic descriptions or accompanying maps, using tools such as Google Maps. Additionally, we record the broader sampling region (e.g., country or oceanic region) and specific sampling site details (such as province, county, or uplift names). The location fields also include the name of the drilling site or outcrop section (SiteName), which

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

identifies the precise drilling location or outcrop where samples were collected, providing valuable geographic context. Certain SiteNames are uniquely associated with major drilling projects (e.g., ODP, IODP), which is important for subsequent data supplementation and analysis. Some samples also have a SampleName, as designated by the original authors—typically a code or non-unique label reflecting the naming format in the primary literature. Although multiple samples in the database may share the same SampleName, each entry has a distinct Sample ID to ensure uniqueness across records.

We also provide paleolatitude and paleolongitude (PaleoLatitude and PaleoLongitude), calculated based on the geological age of each sample and using paleogeographic reconstruction tools such as PointTracker v7.0, built on the plate rotation model of Scotese and Wright (2018). Paleo-coordinate data are crucial for understanding the historical shifts in sample locations and their relationship to depositional environments (Percival et al., 2022; Li et al., 2025). To maintain consistency, all geographic coordinates are standardized to two decimal places.

Age fields. Each entry includes not only absolute age data but also a series of geologic age-related fields to provide precise temporal context. These fields enable targeted data retrieval across different geological time frames, facilitating comparisons with newly added data. The GeologicalAge field captures broad temporal frameworks, recorded as Epoch for the Phanerozoic (e.g., Early Triassic) and Era for the Precambrian (e.g., Neoproterozoic). For more refined stratigraphic resolution, the Stage field (e.g, Induan) is used, with the System as a substitute for Precambrian samples (e.g., Ediacaran). The Age field records the absolute age of each sample, following the International Chronostratigraphic Chart, GTS v202309. The Formation field notes the geological unit (formation or member) from which the sample was collected, aiding in understanding its depositional context and relation to surrounding strata (Murphy and Salvador, 1999). However, Formation data are generally limited to outcrop sections, as ocean drilling samples lack specific formation designations. The Unit field identifies particular stratigraphic units or geologic event layers, such as the Cretaceous pre-OAE2 or OAE2 (Jenkyns, 2010), which aids in correlating samples across recognized geological events. The RelativeDepth field records the sample's relative depth in the section or drill core, which is essential for high-resolution age analyses and sedimentation rate calculations.

Age data allocation follows these guidelines: when precise ages and geological age information for each sample were provided in the original source, these values are prioritized. For older studies with time boundaries that do not align with the GTS v202309 of the International Chronostratigraphic Chart, boundaries are adjusted to the latest standards. If exact ages are unavailable for all samples but

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

can be determined for two or more samples or layers, an age-depth model is constructed based on RelativeDepth to estimate the ages of individual samples. For samples with only one age constraint, the median age of the corresponding geologic interval is assigned.

Data fields. The dataset includes analyses of isotopic compositions, elemental concentrations, and specific components. To maintain consistency, all units were standardized during data collection, as original publications sometimes report these data in varying units. (1) Isotopic data include $\delta^{15}N$, $\delta^{13}C$, $\delta^{18}O$, and $\delta^{34}S$, all expressed in ‰ relative to international standards. Nitrogen isotopes are reported relative to atmospheric nitrogen (Air N_2), carbon and oxygen isotopes relative to the Vienna Pee Dee Belemnite (VPDB) standard, and sulfur isotopes relative to the Vienna Canyon Diablo Troilite (VCDT) standard (Hoefs, 2009). (2) Elemental concentrations include TN (Total Nitrogen), TOC (Total Organic Carbon), TS (Total Sulfur), $CaCO_3$, TC (Total Carbon), TIC (Total Inorganic Carbon), P, Al, K, Si, Ca, Ti, Na, Mg, Fe, as well as iron species data and LOI (Loss on Ignition), and they are reported in weight percent (wt %). Concentrations of other trace elements are standardized to parts per million (ppm). (3) Some data originally reported as oxide concentrations were converted to elemental concentrations based on stoichiometric ratios, such as P_2O_5. (4) Additional derived values include ratios of iron species, dry bulk density, and rock eval indices (Peters et al., 1986; Poulton and Canfield, 2005). These indices comprise Alkenone Content (C37, in nmol/g), Oxygen Index (OI, mg CO2/g TOC), Hydrogen Index (HI, mg HC/g TOC), maximum pyrolysis temperature (Tmax, °C), free hydrocarbons (S1, mg HC/g Rock), hydrocarbons generated from rock pyrolysis (S2, mg HC/g Rock), and CO_2 released from organic matter pyrolysis (S3, mg CO2/g Rock). Some inaccessible data points were visually extracted from figures using scatterplot recognition techniques, which are marked as "plot" in the Note field. Data with values exceeding detection limits or those erroneous (e.g., negative values for element concentration) were excluded from the dataset.

Petrological characteristic field. The petrological characteristic fields encompass information on lithology, depositional facies, and metamorphic grade, which provide essential contextual support for subsequent isotopic geochemistry analyses. (1) Lithology: The Lithology field records the original descriptions provided by authors, using terms such as "black shale" "mudstone" "limestone" and "breccia". The LithType field classifies these lithologies into broader categories, primarily as carbonate and siliciclastic (Tucker and Wright, 2009), with minor entries for phosphorite and iron formations. (2) Metamorphic grade: The metamorphic grade field reflects the extent of metamorphism the samples have undergone, based on original terminology whenever possible. Common terms include specific

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

metamorphic facies (e.g., amphibolite, greenschist) as well as general descriptors like "unmetamorphosed" and "low grade". For Cenozoic samples, which are generally assumed to have undergone minimal metamorphic alteration (Winter, 2014), any entries lacking detailed descriptions are uniformly designated as "unmetamorphosed". (3) Depositional setting: This field records the depositional environment of each sample, with terms like "neritic" "peritidal" "slope" and "abyssal" preserved from the original literature. For many ocean drilling samples, depositional settings are inferred from WaterDepth: depths of 500–2000 m are classified as "bathyal" and depths exceeding 2000 m are designated as "abyssal" ().

Data collection sources. Data in the database primarily originate from published literature and are traceable via DOI. Some data come from public databases such as PANGAEA, SGP, and NOAA. Each record includes multiple fields for source information, such as first author, publication year, article title, reference, DOI, and data source. Metadata fields have been standardized and cleaned via code to ensure consistency and machine readability, removing special characters while retaining complete citation formats. This structure allows users to trace data provenance, with DOI or Reference fields facilitating direct searches on Crossref for verification.

4 Technical validation

The DSMS-NI database has undergone meticulous curation and quality control (QC) to ensure data accuracy, consistency, and scientific value. Each record includes comprehensive metadata to support traceability and verification. While each entry and significant metadata contain a simple remarks field (excluded from the main database to prevent clutter), it notes the source or reason for inclusion, facilitating validation and cross-checking by the data management team. We implemented several QC measures to verify database accuracy.

Geographic coordinate verification. Latitude and longitude values were checked to confirm they fall within the valid ranges of -90 to 90 and -180 to 180, respectively. Sample coordinates were cross-referenced with country names and public national boundaries to ensure geographic accuracy. Modern sample coordinates were projected onto a global map with administrative boundaries (Figs. 2-3) to verify logical placements. If coordinates appeared on land or in other unexpected locations, each entry was manually reviewed and corrected as needed.

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

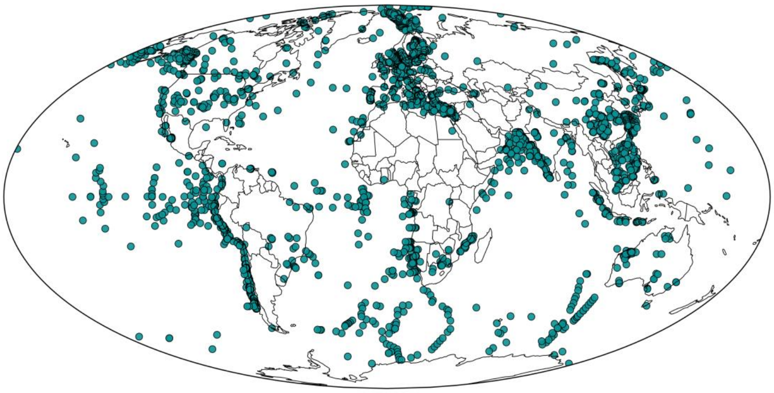

Figure 2. Distribution of sample sites on modern global map.

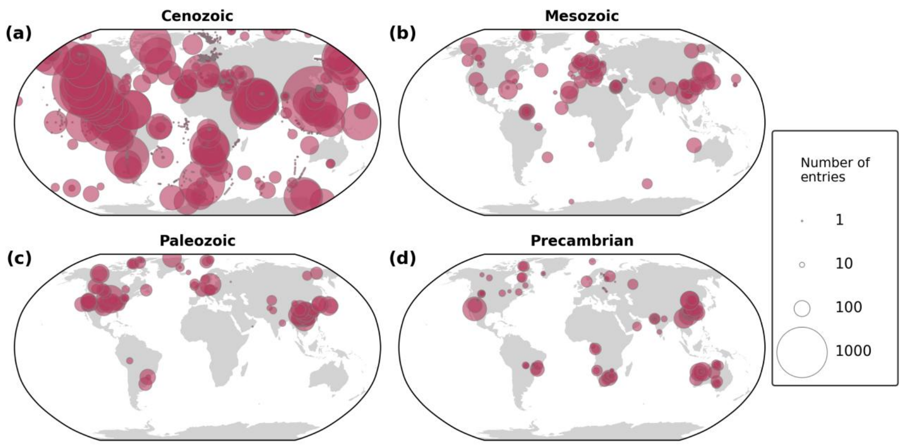

Figure 3. Spatial distribution of sampling sites and sample quantities by geological era in a modern geographic reference frame. The base map is adapted from Kocsis and Scotese (2021). The term "entries" refers to individual proxy values and their associated metadata (i.e., rows in the DSMS-NI database).

Paleocoordinate validation. Paleolatitudes and paleolongitudes were calculated using the G-Plates model (Scotese and Wright, 2018) and PointTracker v7.0 software, ensuring alignment with each sample's geological age and geographical context. Site locations were plotted on paleogeographic maps (Fig. 4) for further evaluation; any inconsistencies in paleocoordinates were flagged, reviewed, and adjusted accordingly.

Outlier detection. Frequency histograms and time-series scatter plots were generated to identify potential outliers in the dataset. Any extremely high or low values underwent secondary validation to confirm their accuracy. For instance, unusually high or low $\delta^{15}N$ values were rigorously checked against the original sources to verify the analytical methods and sample characteristics.

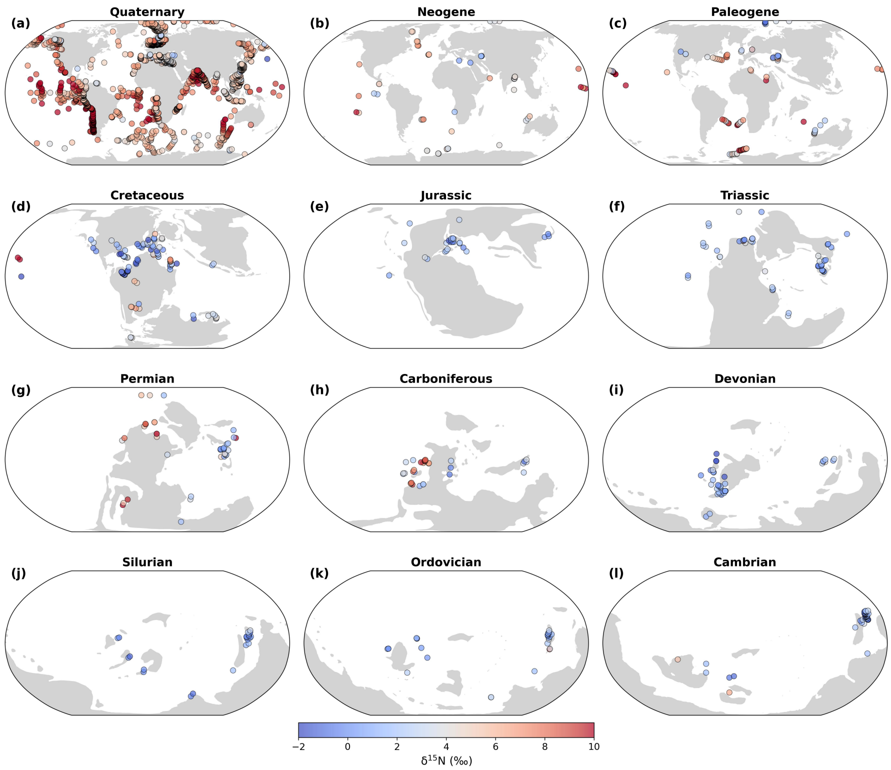

Figure 4. Paleogeographic distribution of $\delta^{15}N$ values by geological period. The base map is adapted from Kocsis and Scotese (2021).

Duplicate check. We conducted a comprehensive check for duplicate entries, especially for samples with similar GPS coordinates. All suspected duplicates were carefully compared, and necessary corrections were made to eliminate redundancy.

Age model calibration. For cases where geological time boundaries in older publications did not align with current standards, we calibrated these boundaries according to the latest Geological Time Scale (GTS v202309). For samples lacking precise ages, an age-depth model was constructed based on RelativeDepth information provided in the literature, allowing for high-resolution age estimation of individual samples. In instances where only a single known age was available for a sample point, the median age of the corresponding geological interval was assigned. To minimize errors, geological age data were entered using a standardized template to prevent typos, inconsistencies, or incorrect values. Automated analyses and cross-verification ensured that numerical ages corresponded accurately with designated eras and geological stages.

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

Data collection sources. Citation information within the reference field was obtained through automated methods from the CrossRef platform, ensuring uniformity in citation formatting (Hendricks et al., 2020). We used scripts to extract comprehensive bibliographic details for each publication, including author names, title, publication year, journal name, volume, page numbers, and DOI. This automation significantly reduced potential spelling errors and inconsistencies that may arise in manual entry. Extracted citation data were cross-checked against original entries in the database, and any discrepancies or errors were corrected manually by the data management team to maintain source accuracy and completeness.

5 General database statistics

The latest version of the DSMS-NI database comprises approximately 320,000 data entries, including 71 040 $\delta^{15}N$ records, spanning all geological periods from the Eoarchean (~3800 Ma) onward. These records originate from a diverse array of unique sampling sites, encompassing ocean drilling cores and outcrop sections. The $\delta^{15}N$ data are predominantly concentrated in the Phanerozoic, comprising 92.1 % of the total database, with further breakdowns showing 71.6 % in the Cenozoic, 8.3 % in the Mesozoic, and 12.2 % in the Paleozoic (Table 3 and Fig. 3). The following sections focus on first-order spatial and temporal trends in $\delta^{15}N$ data density, sampling locations, and values within DSMS-NI. The provided figures illustrate only a subset of the spatial-temporal patterns uniquely revealed by this extensive compilation, demonstrating the database's potential to advance research in paleoclimate, geochemistry, and paleoecology.

Table 3 The quantities and proportions for $\delta^{15}N$, $\delta^{13}C_{org}$, TN, and TOC of each geological era.

Proxy system	Cenozoic	Mesozoic	Paleozoic	Precambrian	Total
$\delta^{15}N$	50847	5894	8653	5646	71040
	71.6 %	8.3 %	12.2 %	7.9 %	
$\delta^{13}C_{org}$	10781	4902	6856	4493	27032
	40.0 %	18.1 %	25.4 %	16.6 %	
TN	31576	2855	6473	4942	45846
	68.9 %	6.2 %	14.1 %	10.8 %	
TOC	22613	5075	8597	5150	41435
	54.6 %	12.2 %	20.7 %	12.4 %	

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

5.1 Temporal density and evolution of $\delta^{15}N$

Given that the data are concentrated in the Phanerozoic, where ages are more precisely constrained, we performed a detailed stratigraphic breakdown of age distribution by stage within the Phanerozoic (Fig. 5). The distribution is uneven, with the highest data densities in recent periods, particularly the Holocene (0-12 ka), Late Pleistocene (12-129 ka), and Chibanian (129-770 ka). The high data density in the Quaternary primarily reflects the abundance of high-resolution records from ocean drilling projects, where individual cores contribute extensive and densely sampled datasets. In contrast, older geological periods exhibit data clusters around key events, such as biotic radiations, mass extinctions, and oceanic anoxic events (Bush and Payne, 2021). Notable gaps or low-density intervals occur from the mid-Cambrian to Early Ordovician, Silurian to Early Devonian, mid-Carboniferous to Early Permian, mid-Triassic to Early Jurassic, and Late Jurassic to Early Cretaceous.

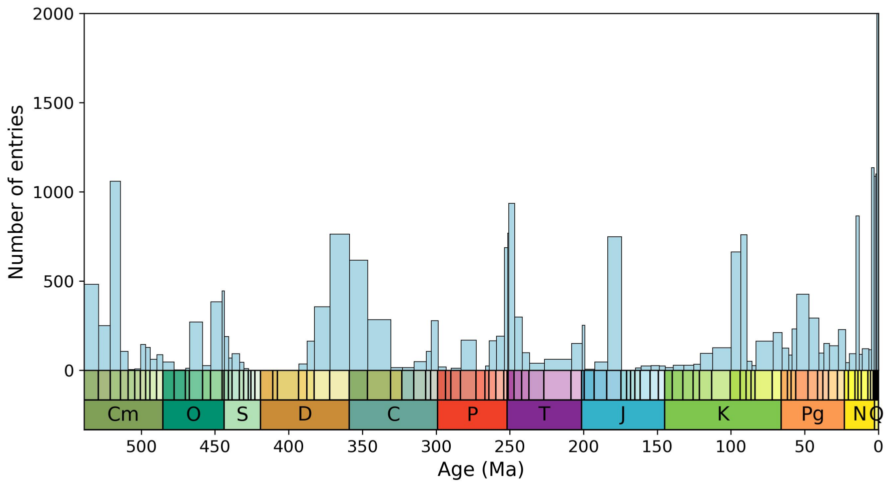

Figure 5. Number of data points binned by geologic stage. Data counts for the Holocene (0-12 ka), Late Pleistocene (12-129 ka), and Chibanian (129-770 ka) stages are 10 640, 21 754, and 8378, respectively; these counts are not displayed in the figure due to narrow column width. The Precambrian has only 5646 data points, accounting for 7.9%, and is not plotted. Cm: Cambrian; O: Ordovician; S: Silurian; D: Devonian; C: Carboniferous; P: Permian; T: Triassic; J: Jurassic; K: Cretaceous; Pg: Paleogene; N: Neogene; Q: Quaternary.

Overall, $\delta^{15}N$ values exhibit a unimodal distribution centered around +5 ‰, with a mean of 5.1 ± 9.1 ‰ (1σ; Fig. 6a). Examining data density by era, the Cenozoic has the highest overall peak, followed by the Precambrian, with lower peaks in the Paleozoic and Mesozoic (Fig. 6b). Plotting all

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

Open Access Earth System Science Data Discussions

$\delta^{15}N$ data on the Phanerozoic geological timescales reveals notable peaks in the Late Cretaceous, Carboniferous-Permian, mid-Triassic, Jurassic, and Early Cretaceous, which align with greenhouse-icehouse climate cycles (Fig. 7; Algeo et al., 2014). In the Precambrian, data distribution is more dispersed, potentially reflecting the instability of the nitrogen cycle or the influence of stronger metamorphic overprints (Fig. 8; Ader et al., 2016; Stüeken et al., 2024). Despite differences in paleolatitude, shifts in $\delta^{15}N$ exhibit consistent directional changes (increase or decrease) during key Phanerozoic transition events, such as the Permian-Triassic boundary (Knies et al., 2013; Du et al., 2021, 2023) and the Late Cretaceous (Meyers et al., 2009; Junium et al., 2018; Du et al., 2025b). LOWESS smoothing results reveal $\delta^{15}N$ peaks in the Neoarchean, Paleoproterozoic, and Ediacaran, i.e., periods closely associated with significant oxygenation events (Kipp et al., 2018; Koehler et al., 2019; Pellerin et al., 2024). These temporal patterns underscore the role of nitrogen isotopes in tracing the interplay between nitrogen cycling and shifts in Earth's oxygenation history.

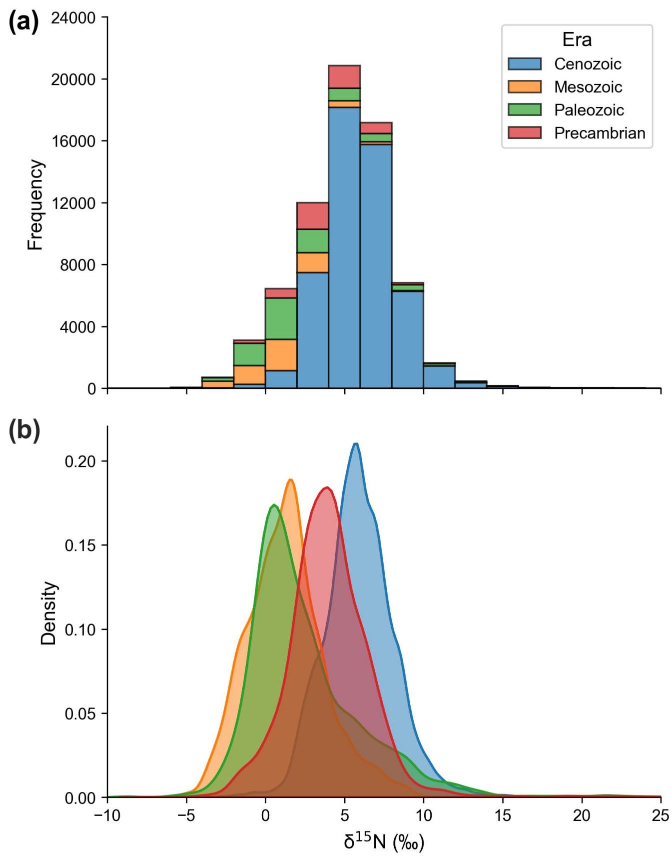

Figure 6. (a) Histogram and (b) density distribution of all $\delta^{15}N$ data.

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

Open Access Earth System Science Data Discussions

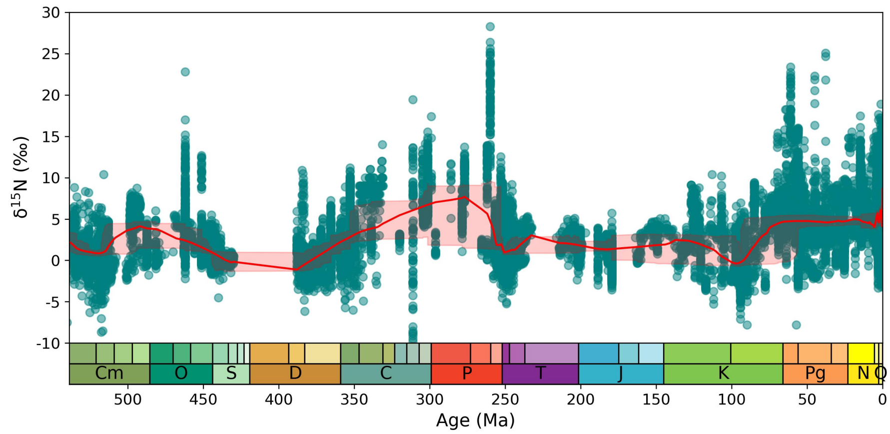

Figure 7. $\delta^{15}N$ data and LOWESS curve through Phanerozoic. A LOWESS factor of 0.03 and a confidence interval of 2.5–97.5 % were applied.

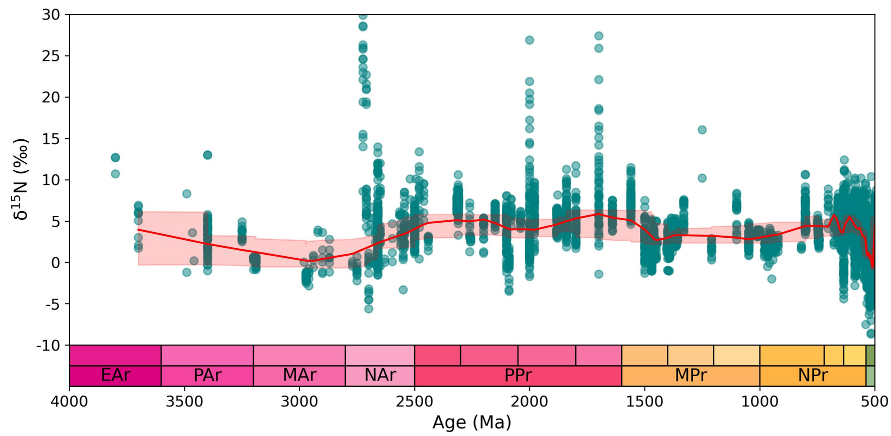

Figure 8. $\delta^{15}N$ data and LOWESS curve through Precambrian. A LOWESS factor of 0.01 and a confidence interval of 2.5–97.5 % were applied. EAr: Eoarchean; PAr: Paleoarchean; MAr: Mesoarchean; NAr: Neoarchean; PPr: Paleoproterozoic; MPr:Mesoproterozoic; NPr: Neoproterozoic.

5.2 Spatial density and characteristics of $\delta^{15}N$

Spatial trends in data density across the DSMS-NI database reveal substantial variability in both modern (Fig. 2) and paleogeographic distributions (Fig. 4). Ocean drilling sites are primarily concentrated along continental margins and deep-sea basins, with significant gaps in central oceanic regions (National Research Council, 2011). For older strata (pre-Cretaceous), sampling sites are clustered in North America, Europe, China, and South Africa (Fig. 2). In terms of latitude, $\delta^{15}N$

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

sampling in older strata is sparse in the modern equatorial region and the mid- to high-latitude areas of the Southern Hemisphere, aside from some Southern Hemisphere samples collected from Cenozoic ocean drilling sites (Fig. 3). When modern coordinates are converted to paleolatitudes and mapped onto paleogeographic reconstructions, the Quaternary shows the highest number of sites, followed by the Cretaceous. Only Cenozoic sites provide extensive latitudinal coverage (Fig. 9). In terms of marine spatial distribution, $\delta^{15}N$ data since the Cretaceous reflects global patterns to a certain degree (Fig. 4). However, pre-Jurassic data remain spatially concentrated, with Paleozoic sites limited to just two or three main areas. High-latitude sampling is generally scarce, with Paleozoic sites predominantly in the Southern Hemisphere and Mesozoic sites mainly in the Northern Hemisphere (Fig. 4).

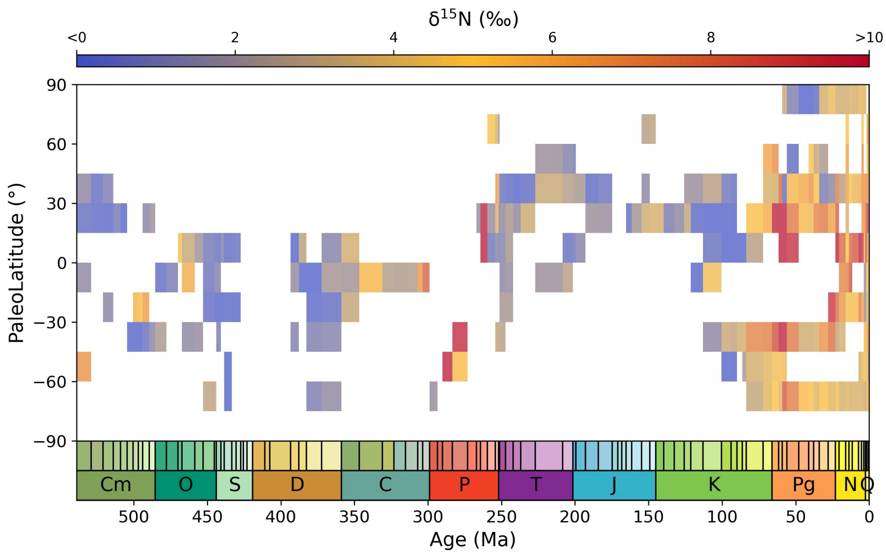

Figure 9. Spatio-temporal trends in $\delta^{15}N$ values through Phanerozoic, binned and averaged temporally by stage and spatially by 15° paleolatitudinal bins.

Significant spatial differences exist in $\delta^{15}N$ distribution across geological periods. Only a few periods allow for robust assessments of latitudinal gradients. For example, Cenozoic $\delta^{15}N$ values tend to decrease with increasing latitude globally, while Ediacaran $\delta^{15}N$ values exhibit the opposite trend (Fig. 9). To visualize spatial trends, average $\delta^{15}N$ values from each Phanerozoic period were mapped onto paleogeographic reconstructions for the respective period (Fig. 4). Quaternary records show the highest $\delta^{15}N$ values in areas associated with upwelling, such as the Arabian Sea, southeastern Indian Ocean, eastern equatorial Pacific, southwestern South America, and the western coast of Mexico (Fig. 4a; Altabet et al., 1999; Tesdal et al., 2013). For the Paleogene and Neogene, $\delta^{15}N$ hotspots are

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

generally located in deep-sea regions, potentially representing more positive $\delta^{15}N$ averages across ocean basins (Fig. 4b-c). In the Paleozoic and Mesozoic, $\delta^{15}N$ values are generally negative, lacking prominent hotspots except in the Carboniferous and Permian. Some periods exhibit slight trends of lower values at low latitudes and higher values at high latitudes, as well as more positive $\delta^{15}N$ values near coastlines compared to more distal marine settings (Fig. 4d-l). These observations suggest that increased temporal resolution is needed for in-depth analysis of spatial $\delta^{15}N$ variations. Given the current uneven distribution of sampling sites, further $\delta^{15}N$ studies across diverse regions are crucial for enhancing our understanding of the spatial characteristics of nitrogen cycle evolution.

6 Usage notes

6.1 Informed user notice

Each record (row) in the database includes detailed temporal and spatial metadata, along with lithology, metamorphic grade, and depositional facies information where available. These metadata are essential for evaluating the geological context and fidelity of nitrogen isotope data. However, this version of the database has certain limitations; for instance, it may not capture all possible geological age uncertainties or precise depositional environment details for some records. Despite our extensive efforts to accurately identify and quality-control each entry, given the vast dataset, some overlooked errors or data inconsistencies may remain. Users are encouraged to report any issues or omissions to the authors, as corrections will be incorporated into future database versions. We recommend that users carefully review metadata fields to ensure that the data aligns with their research needs.

In addition to $\delta^{15}N$ data, the database provides geochemical information such as TOC, total TN, $\delta^{13}C_{org}$, and major and trace element concentrations. These supplementary data are valuable for assessing factors that may influence nitrogen isotopes, such as organic matter preservation and redox conditions. Even when not directly paired with $\delta^{15}N$ values, we retain all relevant data to enable users to conduct correlation analyses via interpolation or other methods. Researchers are welcome to contribute additional geochemical data from the same sites or samples as they become available, allowing for updates and refinements in subsequent database releases.

6.2 Applying the database in deep time

When applying the database to deep-time studies, certain filtering criteria can be used. For instance, samples may be selected based on lithology, metamorphic grade, and other metadata to ensure that the data aligns with specific geological research contexts. Temporal, paleolatitude, and paleodepth information are critical for paleogeographic reconstructions and spatiotemporal distribution analyses, particularly when investigating paleoclimate change and global biogeochemical cycles. Further analysis of variations in latitude, basin characteristics, and water depth has the potential to yield significant insights. Given the rapid variability of nitrogen isotopes and their pronounced regional characteristics, filling temporal and spatial gaps and enhancing resolution are of great value—particularly for pivotal periods like the Ordovician-Silurian mass extinction, the Early Devonian terrestrial plant radiation, and the Late Jurassic-Early Cretaceous supercontinent breakup. The database is also especially suited for comparative studies of key geological periods, such as the Permian-Triassic boundary extinction and the Cretaceous OAE2. To support these applications, we have also provided a software tool on Zenodo, allowing users to generate heatmaps of $\delta^{15}N$ data distributions for specific time intervals. These heatmaps visualize the average spatial distribution of $\delta^{15}N$ for any selected geological interval, offering preliminary validation for user hypotheses and aiding in uncovering the evolution of the global nitrogen cycle.

7 Data availability

The DSMS-NI version 0.0.1 can be accessed via Zenodo at https://doi.org/10.5281/zenodo.15117375 (Du et al., 2025a) and via the GeoBiology website at https://geobiologydata.cug.edu.cn/ (last access: April 30 2025).

8 Code availability

The code used to validate the dataset and make the figures in this manuscript is available on Zenodo (https://doi.org/10.5281/zenodo.15758073). The paleocoordinates were estimated using the PointTracker v7 tool published by the PALEOMAP Project, which can be found at http://www.paleogis.com (last access: April 1 2025; https://doi.org/10.13140/RG.2.1.2011.4162, Scotese, 2008).

Author contributions. YD, HYS and HJS designed the study and secured funding. TJA, EES, SEG, JDC, YD, HZ, XKL, JP, YW, JK, XS, HS, DC and LT conducted data acquisition, curation and validation. YD, LW, JZ, QL, XCL and HY developed computational methodologies and provided technical support. YD prepared the paper with contribution from all co-authors.

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

Competing interests. The authors declare that they have no conflicts of interest.

Acknowledgements. This work has benefited from previous $\delta^{15}N$ datasets compiled by Jan-Erik Tesdal, Christophe Thomazo, Thomas J. Algeo, Eva E. Stüeken, Magali Ader, Xinqiang Wang, and Michael A. Kipp.

Financial support. This work has been funded by the State Key R&D Project of China (2023YFF0804000), the National Natural Science Foundation of China (42172032; 42325202; 42402316), the Postdoctoral Fellowship Program of CPSF (GZB20230679), the Postdoctor Project of Hubei Province (2024HBBHCXB087), the Natural Science Foundation of Hubei Province (2023AFA006).

References

Ader, M., Thomazo, C., Sansjofre, P., Busigny, V., Papineau, D., Laffont, R., Cartigny, P., and Halverson, G. P.: Interpretation of the nitrogen isotopic composition of Precambrian sedimentary rocks: Assumptions and perspectives, Chem. Geol., 429, 93–110, https://doi.org/10.1016/j.chemgeo.2016.02.010, 2016.

Algeo, T. J., Meyers, P. A., Robinson, R. S., Rowe, H., and Jiang, G. Q.: Icehouse–greenhouse variations in marine denitrification, Biogeosciences, 11, 1273–1295, https://doi.org/10.5194/bg-11-1273-2014, 2014.

Altabet, M. A., Murray, D. W., and Prell, W. L.: Climatically linked oscillations in Arabian Sea denitrification over the past 1 m.y.: Implications for the marine N cycle, Paleoceanography, 14, 732–743, https://doi.org/10.1029/1999PA900035, 1999.

Bebout, G. E., Cooper, D. C., Bradley, A. D., and Sadofsky, S. J.: Nitrogen-isotope record of fluid-rock interactions in the Skiddaw aureole and granite, English Lake District, Am. Mineral., 84, 1495–1505, https://doi.org/10.2138/am-1999-1002, 1999.

Bush, A. M. and Payne, J. L.: Biotic and Abiotic Controls on the Phanerozoic History of Marine Animal Biodiversity, Annu. Rev. Ecol. Evol. Syst., 52, 269–289, https://doi.org/10.1146/annurev-ecolsys-012021-035131, 2021.

Chase, B. M., Niedermeyer, E. M., Boom, A., Carr, A. S., Chevalier, M., He, F., Meadows, M. E., Ogle, N., and Reimer, P. J.: Orbital controls on Namib Desert hydroclimate over the past 50,000 years,

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

Geology, 47, 867–871, https://doi.org/10.1130/G46334.1, 2019.

Diepenbroek, M., Grobe, H., Reinke, M., Schindler, U., Schlitzer, R., Sieger, R., and Wefer, G.: PANGAEA—an information system for environmental sciences, Comput. Geosci., 28, 1201–1210, https://doi.org/10.1016/S0098-3004(02)00039-0, 2002.

Du, Y., Song, H., Tong, J., Algeo, T. J., Li, Z., Song, H., and Huang, J.: Changes in productivity associated with algal-microbial shifts during the Early Triassic recovery of marine ecosystems, Geol. Soc. Am. Bull., 133, 362–378, https://doi.org/10.1130/B35510.1, 2021.

Du, Y., Song, H., Grasby, S. E., Xing, T., Song, H., Tian, L., Chu, D., Wu, Y., Dal Corso, J., Algeo, T. J., and Tong, J.: Recovery from persistent nutrient-N limitation following the Permian–Triassic mass extinction, Earth Planet. Sci. Lett., 602, 117944, https://doi.org/10.1016/j.epsl.2022.117944, 2023.

Du, Y., Song, H., Stüeken, E. E., Grasby, S. E., Song, H., Tian, L., Chu, D., Dal Corso, J., Li, Z., and Tong, J.: Large nitrogen cycle perturbations during the Early Triassic hyperthermal, Geochim. Cosmochim. Acta, 382, 13–25, https://doi.org/10.1016/j.gca.2024.08.009, 2024.

Du, Y., Song, H., Algeo, T. J., Zhang, H., Peng, J., Wu, Y., Lai, J., Shu, X., Song, H., Wei, L., Zhang, J., Stüeken, E. E., Grasby, S. E., Dal Corso, J., Dai, X., Chu, D., Tian, L., Liang, Q., Li, X., Yao, H., and Song, H.: The global database of deep-time marine nitrogen isotope data [data set]. Zenodo. https://doi.org/10.5281/zenodo.15117375, 2025a.

Du, Y., Song, H., Algeo, T. J., Zhong, L., Li, J., and Song, H.: Tectonic controls on nitrogen cycling and ocean ventilation dynamics in the Late Cretaceous equatorial Atlantic, Earth Planet. Sci. Lett., 667, 119517, https://doi.org/10.1016/j.epsl.2025.119517, 2025b.

Farrell, Ú. C., Samawi, R., Anjanappa, S., Klykov, R., Adeboye, O. O., Agic, H., Ahm, A.-S. C., Boag, T. H., Bowyer, F., Brocks, J. J., Brunoir, T. N., Canfield, D. E., Chen, X., Cheng, M., Clarkson, M. O., Cole, D. B., Cordie, D. R., Crockford, P. W., Cui, H., Dahl, T. W., Mouro, L. D., Dewing, K., Dornbos, S. Q., Drabon, N., Dumoulin, J. A., Emmings, J. F., Endriga, C. R., Fraser, T. A., Gaines, R. R., Gaschnig, R. M., Gibson, T. M., Gilleaudeau, G. J., Gill, B. C., Goldberg, K., Guilbaud, R., Halverson, G. P., Hammarlund, E. U., Hantsoo, K. G., Henderson, M. A., Hodgskiss, M. S. W., Horner, T. J., Husson, J. M., Johnson, B., Kabanov, P., Brenhin Keller, C., Kimmig, J., Kipp, M. A., Knoll, A. H., Kreitsmann, T., Kunzmann, M., Kurzweil, F., LeRoy, M. A., Li, C., Lipp, A. G., Loydell, D. K., Lu, X., Macdonald, F. A., Magnall, J. M., Mänd, K., Mehra, A., Melchin, M. J., Miller, A. J., Mills, N. T., Mwinde, C. N., O'Connell, B., Och, L. M., Ossa Ossa, F., Pagès, A., Paiste, K., Partin, C. A., Peters, S. E., Petrov, P., Playter, T. L., Plaza-Torres, S., Porter, S. M.,

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

Poulton, S. W., Pruss, S. B., Richoz, S., Ritzer, S. R., Rooney, A. D., Sahoo, S. K., Schoepfer, S. D., Sclafani, J. A., Shen, Y., Shorttle, O., Slotznick, S. P., Smith, E. F., Spinks, S., Stockey, R. G., Strauss, J. V., Stüeken, E. E., Tecklenburg, S., Thomson, D., Tosca, N. J., Uhlein, G. J., Vizcaíno, M. N., Wang, H., White, T., Wilby, P. R., et al.: The Sedimentary Geochemistry and Paleoenvironments Project, Geobiology, 19, 545–556, https://doi.org/10.1111/gbi.12462, 2021.

Gard, M., Hasterok, D., and Halpin, J. A.: Global whole-rock geochemical database compilation, Earth Syst. Sci. Data, 11, 1553–1566, https://doi.org/10.5194/essd-11-1553-2019, 2019.

Hendricks, G., Tkaczyk, D., Lin, J., and Feeney, P.: Crossref: The sustainable source of community-owned scholarly metadata, Quant. Sci. Stud., 1, 414–427, https://doi.org/10.1162/qss_a_00022, 2020.

Hoefs, J.: Stable Isotope Geochemistry, Springer International Publishing, Berlin, https://doi.org/10.1007/978-3-030-77692-3, 2021.

Ishida, A., Hashizume, K., and Kakegawa, T.: Microbial nitrogen cycle enhanced by continental input recorded in the Gunflint Formation, Geochem. Persp. Let., 13–18, https://doi.org/10.7185/geochemlet.1729, 2017.

Jenkyns, H. C.: Geochemistry of oceanic anoxic events, Geochem. Geophys. Geosyst., 11, Q03004, https://doi.org/10.1029/2009GC002788, 2010.

Junium, C. K., Meyers, S. R., and Arthur, M. A.: Nitrogen cycle dynamics in the Late Cretaceous Greenhouse, Earth Planet. Sci. Lett., 481, 404–411, https://doi.org/10.1016/j.epsl.2017.10.006, 2018.

Kipp, M. A., Stüeken, E. E., Yun, M., Bekker, A., and Buick, R.: Pervasive aerobic nitrogen cycling in the surface ocean across the Paleoproterozoic Era, Earth Planet. Sci. Lett., 500, 117–126, https://doi.org/10.1016/j.epsl.2018.08.007, 2018.

Knies, J., Grasby, S. E., Beauchamp, B., and Schubert, C. J.: Water mass denitrification during the latest Permian extinction in the Sverdrup Basin, Arctic Canada, Geology, 41, 167–170, https://doi.org/10.1130/G33816.1, 2013.

Kocsis, Á. T. and Scotese, C. R.: Mapping paleocoastlines and continental flooding during the Phanerozoic, Earth Sci. Rev., 213, 103463, https://doi.org/10.1016/j.earscirev.2020.103463, 2021.

Koehler, M. C., Stüeken, E. E., Hillier, S., and Prave, A. R.: Limitation of fixed nitrogen and deepening of the carbonate-compensation depth through the Hirnantian at Dob's Linn, Scotland, Palaeogeogr. Palaeoclimatol. Palaeoecol., 534, 109321, https://doi.org/10.1016/j.palaeo.2019.109321, 2019.

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

Lai, J., Song, H., Chu, D., Dal Corso, J., Sperling, E. A., Wu, Y., Liu, X., Wei, L., Li, M., Song, H., Du, Y., Jia, E., Feng, Y., Song, H., Yu, W., Liang, Q., Li, X., and Yao, H.: Deep-Time Marine Sedimentary Element Database, Earth Syst. Sci. Data, 17, 1613–1626, https://doi.org/10.5194/essd-17-1613-2025, 2025.

Li, J., Song, H., Du, Y., Wignall, P. B., Bond, D. P. G., Grasby, S. E., Song, H., Dal Corso, J., Tian, L., and Chu, D.: Spatial and temporal heterogeneity of the marine nitrogen cycle during the end-Triassic mass extinction, Chem. Geol., 682, 122752, https://doi.org/10.1016/j.chemgeo.2025.122752, 2025.

Li, W., Zhou, L., Lin, Y., Zhang, H., Zhang, Y., Wu, X., Stevens, C., Yang, Y., Wang, H., Fang, Y., and Liang, F.: Interdisciplinary study on dietary complexity in Central China during the Longshan Period (4.5–3.8 kaBP): New isotopic evidence from Wadian and Haojiatai, Henan Province, The Holocene, 31, 258–270, https://doi.org/10.1177/0959683620970252, 2021.

Liu, Y., Magnall, J. M., Gleeson, S. A., Bowyer, F., Poulton, S. W., and Zhang, J.: Spatio-temporal evolution of ocean redox and nitrogen cycling in the early Cambrian Yangtze ocean, Chem. Geol., 554, 119803, https://doi.org/10.1016/j.chemgeo.2020.119803, 2020.

Liu, Z., Altabet, M. A., and Herbert, T. D.: Plio-Pleistocene denitrification in the eastern tropical North Pacific: Intensification at 2.1 Ma, Chem. Geol., 9, 2008GC002044, https://doi.org/10.1029/2008GC002044, 2008.

Meyers, P. A., Yum, J.-G., and Wise, S. W.: Origins and maturity of organic matter in mid-Cretaceous black shales from ODP Site 1138 on the Kerguelen Plateau, Mar. Pet. Geol., 26, 909–915, https://doi.org/10.1016/j.marpetgeo.2008.09.003, 2009.

Moretti, S., Auderset, A., Deutsch, C., Schmitz, R., Gerber, L., Thomas, E., Luciani, V., Petrizzo, M. R., Schiebel, R., Tripati, A., Sexton, P., Norris, R., D'Onofrio, R., Zachos, J., Sigman, D. M., Haug, G. H., and Martínez-García, A.: Oxygen rise in the tropical upper ocean during the Paleocene-Eocene Thermal Maximum, Science, 383, 727–731, https://doi.org/10.1126/science.adh4893, 2024.

Murphy, E. M. A. and Salvador, A.: International Subcommission on Stratigraphic Classification of IUGS International Commission on Stratigraphy, Episodes, 22, 1999.

National Research Council: Scientific Ocean Drilling: Accomplishments and Challenges, National Academies Press, Washington, D.C., https://doi.org/10.17226/13232, 2011.

Nichols, G.: Sedimentology and stratigraphy, John Wiley & Sons, 2009.

Pellerin, A., Thomazo, C., Ader, M., Rossignol, C., Rego, E. S., Busigny, V., and Philippot, P.:

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

Neoarchaean oxygen-based nitrogen cycle en route to the Great Oxidation Event, Nature, https://doi.org/10.1038/s41586-024-07842-x, 2024.

Percival, L. M. E., Marynowski, L., Baudin, F., Goderis, S., De Vleeschouwer, D., Rakociński, M., Narkiewicz, K., Corradini, C., Da Silva, A. -C., and Claeys, P.: Combined Nitrogen-Isotope and Cyclostratigraphy Evidence for Temporal and Spatial Variability in Frasnian–Famennian Environmental Change, Geochem. Geophys. Geosyst., 23, e2021GC010308, https://doi.org/10.1029/2021GC010308, 2022.

Peters, K. E.: Guidelines for Evaluating Petroleum Source Rock Using Programmed Pyrolysis, AAPG Bull., 70, 318–329, 1986.

Poulton, S. W. and Canfield, D. E.: Development of a sequential extraction procedure for iron: implications for iron partitioning in continentally derived particulates, Chem. Geol., 214, 209–221, https://doi.org/10.1016/j.chemgeo.2004.09.003, 2005.

Sahoo, S. K., Gilleaudeau, G. J., Wilson, K., Hart, B., Barnes, B. D., Faison, T., Bowman, A. R., Larson, T. E., and Kaufman, A. J.: Basin-scale reconstruction of euxinia and Late Devonian mass extinctions, Nature, 615, 640–645, https://doi.org/10.1038/s41586-023-05716-2, 2023.

Scotese, C.: The PALEOMAP Project PaleoAtlas for ArcGIS, version 1, 2, 16–31, ResearchGate [data set]. https://doi.org/10.13140/RG.2.1.2011.4162 , 2008.

Scotese, C. R. and Wright, N.: PALEOMAP Paleodigital Elevation Models (PaleoDEMS) for the Phanerozoic PALEOMAP Project, https://www.earthbyte.org/paleodem-resource-scotese-and-wright-2018/ (last access: 1 April 2025), 2018.

Song, H.Y., Xing, T., Stüeken, E. E., Du, Y., Zhu, Y., Tao, X., Ni, Q., and Song, H.J.: Isotopic differences and paleoenvironmental significance of nitrogen contained in bulk sedimentary rocks, decarbonated aliquots and kerogen extracts, Chem. Geol., 631, 121522, https://doi.org/10.1016/j.chemgeo.2023.121522, 2023.

Stüeken, E. E., Kipp, M. A., Koehler, M. C., and Buick, R.: The evolution of Earth's biogeochemical nitrogen cycle, Earth Sci. Rev., 160, 220–239, https://doi.org/10.1016/j.earscirev.2016.07.007, 2016.

Stüeken, E. E., Pellerin, A., Thomazo, C., Johnson, B. W., Duncanson, S., and Schoepfer, S. D.: Marine biogeochemical nitrogen cycling through Earth's history, Nat. Rev. Earth Environ., https://doi.org/10.1038/s43017-024-00591-5, 2024.

https://doi.org/10.5194/essd-2025-377
Preprint. Discussion started: 2 July 2025

Sun, Y.: Dynamics of nutrient cycles in the Permian–Triassic oceans, Earth Sci. Rev., 258, 104914, https://doi.org/10.1016/j.earscirev.2024.104914, 2024.

Tesdal, J.-E., Galbraith, E. D., and Kienast, M.: Nitrogen isotopes in bulk marine sediment: linking seafloor observations with subseafloor records, Biogeosciences, 10, 101–118, https://doi.org/10.5194/bg-10-101-2013, 2013.

Thomazo, C., Ader, M., and Philippot, P.: Extreme 15 N-enrichments in 2.72-Gyr-old sediments: evidence for a turning point in the nitrogen cycle, Geobiology, 9, 107–120, https://doi.org/10.1111/j.1472-4669.2011.00271.x, 2011.

Tribovillard, N., Algeo, T. J., Lyons, T., and Riboulleau, A.: Trace metals as paleoredox and paleoproductivity proxies: An update, Chem. Geol., 232, 12–32, https://doi.org/10.1016/j.chemgeo.2006.02.012, 2006.

Tucker, M. E. and Wright, V. P.: Carbonate sedimentology, John Wiley & Sons, 2009.

Uveges, B. T., Izon, G., Junium, C. K., Ono, S., and Summons, R. E.: Aerobic nitrogen cycle 100 My before permanent atmospheric oxygenation, Proc. Natl. Acad. Sci., 122, e2423481122, https://doi.org/10.1073/pnas.2423481122, 2025.

Wang, D., Ling, H.-F., Struck, U., Zhu, X.-K., Zhu, M., He, T., Yang, B., Gamper, A., and Shields, G. A.: Coupling of ocean redox and animal evolution during the Ediacaran-Cambrian transition, Nat. Commun., 9, 2575, https://doi.org/10.1038/s41467-018-04980-5, 2018.

Wilkinson, M. D., Dumontier, M., Sansone, S.-A., Bonino Da Silva Santos, L. O., Prieto, M., Batista, D., McQuilton, P., Kuhn, T., Rocca-Serra, P., Crosas, M., and Schultes, E.: Evaluating FAIR maturity through a scalable, automated, community-governed framework, Sci. Data, 6, 174, https://doi.org/10.1038/s41597-019-0184-5, 2019.

Winter, J. D.: Principles of igneous and metamorphic petrology, Pearson education Harlow, UK, 2014.

Xia, L., Cao, J., Stüeken, E. E., Hu, W., and Zhi, D.: Linkages between nitrogen cycling, nitrogen isotopes, and environmental properties in paleo-lake basins, GSA Bull., 134, 2359–2372, https://doi.org/10.1130/B36290.1, 2022.

Zhong, L., Peng, J., He, J., Du, Y., Xing, T., Li, J., Guo, W., Ni, Q., Hu, J., and Song, H.: Optimizations of the EA-IRMS system for $\delta^{15}N$ analysis of trace nitrogen, Appl. Geochem., 159, 105832, https://doi.org/10.1016/j.apgeochem.2023.105832, 2023.