PhanSST

Data entry instructions & description of fields

Document created: June 2022 (E. Judd)
Last updated: November 2022 (E. Judd)

General Guidelines

Thank you for volunteering to helping enter newly published or omitted datasets into PhanSST!

In this instruction manual, you will find descriptions and information regarding each of fields in the data entry template, parsed by subject. Before you begin entering data, please read through this document. Some important considerations to keep in mind:

- Please use a new data entry template for each unique sampling site (e.g., data from two different cores) and/or proxy type (e.g., Mg/Ca & TEX86 data).
- Required fields will be indicated below by an asterisk
- Fields without an asterisk are optional - if the information is available, please include it; however, not all fields will be applicable to all data
- If a field is not required and not applicable, please leave the cell blank; do not fill with '-' or 'N/A' values
- The data entry template is in a comma separated format (csv); to ensure the data entry forms are machine readable, please:
    - do **not** include any commas
    - do **not** include any special characters or accents (e.g., ß, ä, é)
- Once completed, please save the completed data entry template following the naming convention:
  PhanSST_<ProxyType>_<SiteName>_<EnteredBy>.csv
  For example, 'PhanSST_d18c_BassRiver_Judd.csv' or 'PhanSST_tex_ODP768_Judd.csv'
- Completed data entry templates can be uploaded on the PhanTASTIC website (paleo-temperature.org) or email directly to PhanSST@outlook.com

For more information, please refer to Judd et al., 2022 in *Scientific Data* (https://doi.org/10.1038/s41597-022-01826-0). If you have any questions, please don't hesitate to email us at PhanSST@outlook.com.

Sample site fields

SampleID: The unique identifier of each sample; applicable to both drill core and outcrop data

SiteName: The name of the drill core site or outcrop section

- If from drill core, use site name with all capital letters and no spaces (e.g., ODP768, IODPU1443)
- If from outcrop, please use section name when applicable (e.g., Bird Hill East, Meishan D)

SiteHole: The alphabetic (or in rare cases numeric) hole specifier; only applicable to drill core data

MBSF: Depth below the sea floor (m); applicable to drill core data only; always positive

MCD: The mean composite depth (m); applicable to drill core data only; always positive

SampleDepth: Stratigraphic height or depth of data not collected from an ocean drill core (m); can be positive or negative

Formation: The geologic formation name; please do not add the 'Formation' qualifier (e.g., Hungry Hollow, Bescie)

Country: The country in which the data were collected

ContinentOcean*: Two-letter code indicating the continent or ocean from which samples come (all lowercase). Options include:

- *af* – Africa
- *an* – Antarctica
- *ar* – Arctic Ocean
- *as* – Asia
- *at* – Atlantic Ocean
- *au* – Australasia
- *eu* – Europe
- *in* – Indian Ocean
- *me* – Mediterranean Sea
- *na* – North America
- *pa* – Pacific Ocean
- *sa* – South America
- *so* – Southern Ocean

ModLat*: modern latitude of the collection site, in decimal degrees

- Numeric value between -90 and 90
- Include negative signs when referencing Southern Hemisphere sites

ModLon*: modern longitude of the collection site, in decimal degrees

- Numeric value between -180 and 180
- Include negative signs when referencing Western Hemisphere sites

Age fields

Some sort of age constraint is required for all data. When filling in the age fields, you can opt to either enter a numeric age for each sample (i.e., *Age**) or to enter the relative age information (i.e., *Period**, *Stage**, and *StagePosition* if applicable). You do not need to enter both numeric and relative age information. The *Biozone* field is optional but is helpful when updating age information with new iterations of the GTS.

Age: Geologic age of data

- Single numeric value (in Ma)
- Calibrated to GTS2020

Period: The geochronologic period of the data (e.g., Jurassic)

Stage: The geochronologic age (i.e., stage) of the data (e.g., Bartonian, Hirnantian)

StagePosition: Further specification of relative age within its stage (early, middle, or late) if constrained

Biozone: The conodont, graptolite, and/or ammonite biozone; please only include accepted names as per GTS 2020

AgeFlag: Binary flag indicating if the *Age* field reflects: (0) a precise numeric assignment, (1) an estimate based on relative age information, or (2) that both numeric and relative age information were manually entered
(Note: this field is auto-generated and can be left blank)

Proxy fields

Note that the *DiagenesisFlag* field is only applicable to isotope and Mg/Ca data. This flag is expert-assigned; we're asking you to use your judgement when assigning this flag. In the case of foraminiferal data, the flag should refer to the preservation of the specimens themselves (i.e.,0 = glassy; 1 = frosted).

ProxyValue*: The proxy value as originally reported in publication (native proxy units)

ProxyType*: Reference to the proxy system from which the data derived. Options include:

- *d18c*: oxygen isotopes of calcite (per mille, VPDB)
- *d18a*: oxygen isotopes of aragonite (per mille, VPDB)
- *d18p*: oxygen isotopes of conodont phosphate (per mille, VSMOW)
- *mg*: magnesium to calcium ratios of planktonic foraminifera (mmol/mol)
- *tex*: TEX_{86}
- *uk*: alkenone unsaturation ratio

ValueType*: two-letter code indicating the averaging method of the data. Note that by definition, all TEX_{86} and alkenone data are population averages (i.e., *pa*). Options include,

- *im*: individual measurement from an individual specimen (e.g., single foram, single sample from brachiopod)
- *ia*: average of multiple measurements from an individual specimen (e.g., average of multiple different spot samples from a brachiopod)
- *pa*: population average (e.g., bulk foram measurement, average of measurements from a single stratigraphic horizon)

DiagenesisFlag: Binary expert-assigned flag indicating good (0) or questionable (1) preservation; this field is requisite for all isotope and Mg/Ca data but is not applicable to TEX_{86} nor alkenone data

Taxonomic, environmental, and ecological fields

Taxon1*: Two-letter (lowercase) code indicating the first-order taxonomic classification. Options include:

- *a*: Arthropod
- *br*: Brachiopod
- *m*: Mollusk
- *co*: Conodont
- *ha*: Haptophyte
- *pf*: Planktonic foraminifera
- *th*: Thaumarchaeota

Taxon2: Two-letter (lowercase) code indicating the second-order (class) taxonomic classification of data from mollusks. Options include:

- *bi*: Bivalve
- *ce*: Cephalopod
- *ga*: Gastropod
- *ot*: Other

Taxon3: Third-order (genus & species) classification following binomial nomenclature; data from organisms constrained to only to the genus level can be donated using 'spp.' (e.g., 'Globigerina spp.'); applicable – but not requisite – to all isotope and Mg/Ca data, but particularly useful with foraminiferal data

Environment: Depositional environment (e.g., mid-shelf, epeiric); applicable – but not requisite – to all data, but particularly useful with data collected in outcrop

Ecology: Ecological preference of the sampled taxon (e.g., surface, benthic); applicable – but not requisite –to all isotope and Mg/Ca data

Proxy specific fields

These are fields that are specific to a specific proxy system. If not entering data from that proxy system, then please leave these fields blank. If entering data from that specific proxy system, then the asterisk indicates a required field.

Oxygen isotopes of carbonate:

AnalyticalTechnique*: The analytical technique used to obtain the data (i.e., IRMS or SIMS)

CL: cathodoluminescence assessment (L = luminescent; SL = slightly luminescent; NL = not luminescent)

Elemental Suite: All reported elemental concentrations and ratios, including:

- *Fe*: Iron (ppm)
- *Mn*: Manganese (ppm)

- *Mg*: Magnesium (ppm)
- *Sr*: Strontium (ppm)
- *Ca*: Calcium (ppm)
- *Cawtp*: Calcium (weight percent)
- *Mg/Ca*: Magnesium to calcium (mmol/mol)
- *Sr/Ca*: Strontium to calcium (mmol/mol)

Oxygen isotopes of phosphate:

AnalyticalTechnique*: The analytical technique used to obtain the data (i.e., IRMS or SIMS)

NBS120c*: The NBS120c standard value used to correct the data (‰)

Durango: The Durango standard value used to correct the data (‰), required if data measured via SIMS

MaximumCAI: The maximum reported conodont color alteration index value for that sample or horizon

Mg/Ca of planktonic foraminifera:

ModWaterDepth*: The modern water depth of the sampling site (m)

CleaningMethod*: A binary flag to indicate either oxidative-only cleaning (0) or inclusion of a reductive cleaning step (1) of Mg/Ca data

TEX_{86}:

Fractional abundances*: Fields to indicate the fractional abundances of the GDGTs for each sample; requisite when reported

Index values*: Fields to indicate the values of the methane (*MI*), delta ring (*dRI*) and branched and isoprenoidal tetratether (*BIT*) indices for each sample; requisite when reported

Reference fields

LeadAuthor*: The last name of the first author of the original publication

Year*: The year of the original publication

PublicationDOI*: The DOI or other identifying hyperlink of the original publication
- Please ensure all DOIs are reported using the 'https://doi.org/…' convention

DataDOI: The DOI hyperlink to the data repository hosting the data (e.g., Pangaea, NOAA-NCEI)
- Please ensure all DOIs are reported using the 'https://doi.org/…' convention

EnteredBy*: Your last name (e.g., 'Judd'); this information is for internal use only – it will never be made publicly available