**1. Project**

**Title:** SISAL (Speleothem Isotopes Synthesis and AnaLysis Working Group) database Version 1.0

**Dates:** 2017-2018

**Funding organisations**: Past Global Changes (PAGES) programme, European Research Council (ERC 694481\_GC2.0), JPI-Belmont through the UK Natural Environmental Research Council (NE/P006752/1), Geological Survey Ireland Short Call 2017 (2017-SC-056), University College Dublin Seed Funding award (SF1428), European Geosciences Union (W2017/413), University of Reading, Irish Centre for Research in Applied Geosciences (iCRAG), European Association of Geochemistry (Early Career Ambassadors program 2017), Geological Survey Ireland, the Quaternary Research Association UK, Navarino Environmental Observatory, Stockholm University, Savillex, John Cantle.

**Grant number**: ERC 694481\_GC2.0, NE/P006752/1, 2017-SC-056, SF1428, W2017/413

**2. Dataset**

**Title:** SISAL (Speleothem Isotopes Synthesis and AnaLysis Working Group) database Version 1.0

**Summary description.** Stable isotope records from speleothems provide information on past climate changes, most particularly information that can be used to reconstruct past changes in precipitation and atmospheric circulation. SISAL (Speleothem Isotope Synthesis and Analysis) is an international working group of the Past Global Changes (PAGES) project. The working group aims to provide a comprehensive compilation of speleothem isotope records for climate reconstruction and model evaluation. The first version of the SISAL database contains oxygen and carbon isotope measurements from 371 individual speleothem records, and 10 composites from 174 cave systems worldwide, and metadata describing the cave settings and age models of these records.

**Publication year:** 2018

**Creators:** Kamolphat Atsawawaranunt,Sandy P. Harrison, Laia Comas Bru

**Organisation:** Geography and Environmental Science, University of Reading, UK and School of Earth Sciences, University College Dublin, Belfield. Dublin 4, Ireland

**Rights Holder:** University of Reading, University College Dublin

**3. Terms of use**

This dataset is licensed by the rights-holder(s) under a Creative Commons Attribution 4.0 International Licence: https://creativecommons.org/licenses/by/4.0/.

**4. Contents**

**Abstract:** Stable isotope records from speleothems provide information on past climate changes, most particularly information that can be used to reconstruct past changes in precipitation and atmospheric circulation. These records are increasingly being used to provide “out-of-sample” evaluations of isotope-enabled climate models. SISAL (Speleothem Isotope Synthesis and Analysis) is an international working group of the Past Global Changes (PAGES) project. The working group aims to provide a comprehensive compilation of speleothem isotope records for climate reconstruction and model evaluation. The SISAL database contains data for individual speleothems, grouped by cave system. Stable isotopes of oxygen and carbon (δ18O, δ13C) measurements are referenced by distance from the top or youngest part of the speleothem. Additional tables provide information on dating, including information on the dates used to construct the original age model and sufficient information to assess the quality of each data set and to create a standardized chronology across different speleothems. The metadata table provides location information, information about the full range of measurements carried out on each speleothem and information about the cave system that is relevant to the interpretation of the records, as well as citations for both publications and archived data.

There is a single MySQL database file (sisalv1.sql). Please check <https://dev.mysql.com/downloads/> to download and install MySQL. Once MySQL Community Server and MySQL Workbench are installed, the database can be imported and visualised. A schema must be created upon import. To import the SQL file, you follow:

* Open MySQL Workbench
* Connect to the connection you would like to store your database in. A connection is usually created during the installation process (usually root@localhost with the password defined during the installation process)
* Server>Data Import>Import from Self-contained file
* Browse to the SQL file you have downloaded
* Press New… next to the Default Target Schema to create a new schema (name this as appropriate, such as sisalv1)
* Press Import

Please note that once the database is imported, there are packages and modules in several programming languages which will allow you to connect to the database such as RMySQL in R, and MySQLdb in python.

There is a single compressed archive file (sisalv1.zip) comprising 14 CSV files corresponding to the 14 individual tables in the MySQL database. The CSV file names correspond to the table names. As these are flat CSV files, no relationships are defined here but the tables can be joined in different programming languages (R, Python, etc.) based on the foreign keys (shared column names between tables such as site\_id in the site and entity tables). The relationships are described in Figure 1 and the characteristics of each table are described in Table 1 – 14. Please note that CSV files are in UTF-8 characters, and special characters (such as Greek characters, and letters with accents which may appear in names and in citations) may not be reproduced correctly when open as default in Excel. Please see <https://scrapehero.freshdesk.com/support/solutions/articles/5000617795-how-to-open-csv-files-that-have-unicode-unprintable-or-weird-looking-characters-in-excel> to see how to open this in Excel.

There is a single compressed archive file (sisalv1\_codes.zip) comprising examples of codes and queries that can be used with the MySQL database, but also with the CSV files. Within this compressed archive file, there is:

* An html file (sisal\_intro\_example.html) which shows example SQL queries on the database
* A python file (sisal\_connect2db.py) demonstrating how to connect python to the database once the database has been uploaded into MySQL.
* An R file (sisal\_connect2db.R) demonstrating how to connect R to the database once the database has been uploaded into MySQL.
* An R file (sisal\_csv\_db.R) demonstrating how to load the CSV files into R and query these CSV files without the need to install MySQL.

**File structure:** The data is stored in a relational database (MySQL), which consists of 14 linked tables. Specifically: Site, entity, sample, dating, dating lamina, gap, hiatus, original chronology, d13C, d18O, entity link reference, references, composite link entity. Figure 1 shows the relationships between these tables. The structure and contents of each table is described below.



Figure 1: The structure of the SISAL database, showing individual tables (and their contents) and the nature of the relationships between them (where “many to one linkages” indicate that it is possible to have several entries in one table linked to a single entry in another table).

Table 1: Characteristics of the site table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field label** | **Definition** | **Format** | **Constraints** |
| site\_id | Unique identifier for each site, where a site is defined as a cave or cave system | Numeric | Positive integer |
| site\_name | Site name as given by original authors or as defined by us where there was no unique name given to the site | Text | None |
| latitude | Latitude of the cave site, given in decimal degrees, where N is positive and S is negative | Numeric | None |
| longitude | Longitude of the cave site, given in decimal degrees, E is positive and W is negative | Numeric | None |
| elevation | Elevation of the cave, in meters above sea level (where negative values indicate elevations below sea level) | Numeric | None |
| geology | Description of the rock type  | Text | selected from pre-defined list |
| rock\_age | Description of age of the rock | Text | selected from pre-defined list |
| monitoring | Indication of whether long-term monitoring of cave conditions have been carried out | Text | selected from pre-defined list |

Table 2: Characteristics of the entity table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Definition** | **Format** | **Constraints** |
| site\_id | Refers to the unique identifier for each site (as given in site table) | Numeric | Positive integer |
| entity\_id | Unique identifier for each entity, where an entity is defined as a speleothem or a speleothem composite | Numeric | Positive integer |
| entity\_name | Entity (speleothem) name as given by the author | Text | None |
| entity\_status | Status of the entity record, specifically whether it is current, is current but includes modifications or additional information, or it has been superseded by another record  | Text | selected from pre-defined list |
| corresponding\_current | This refers to the entity\_id of the record which replaces a superceded record or to the entity\_id of other current records that are linked to a current but partially modified record. | Numeric | Positive integer |
| depth\_ref | Indication of whether the reference point is the top or base of the speleothem | Text | selected from pre-defined list |
| cover\_thickness | Thickness of overlying bedrock above the speleothem (m) | Numeric | Positive decimal |
| distance\_entrance | Distance of the speleothem from the cave entrance (m) | Numeric | Positive decimal |
| speleothem\_type | Description of the speleothem type | Text | selected from pre-defined list |
| drip\_type | Description of the drip type | Text | selected from pre-defined list |
| d13C | Indication of whether δ13C measurements have been made of the speleothem  | Text | selected from pre-defined list |
| d18O | Indication of whether δ18O measurements have been measured | Text | selected from pre-defined list |
| d18O\_water\_equilibrium | Indication of whether studies assessing if the speleothem is precipitating in equilibrium with dripwaters have been done | Text | selected from pre-defined list |
| trace\_elements | Indication of whether trace elements have been measured  | Text | selected from pre-defined list |
| organics | Indication of whether organics have been measured  | Text | selected from pre-defined list |
| fluid\_inclusions | Indication of whether fluid inclusions have been measured  | Text | selected from pre-defined list |
| mineralogy\_petrology\_fabric | Indication of whether fabric measurements have been made  | Text | selected from pre-defined list |
| clumped\_isotopes | Indication of whether clumped isotopes have been measured | Text | selected from pre-defined list |
| noble\_gas\_temperatures | Indication of whether noble gases have been measured | Text | selected from pre-defined list |
| C14 | Indication of whether 14C measurements have been made  | Text | selected from pre-defined list |
| ODL | Indication of whether the optical density of luminosity measurements have been made  | Text | selected from pre-defined list |
| Mg\_Ca | Indication of whether the Mg/Ca ratio measurements have been made  | Text | selected from pre-defined list |
| contact | Name of the person who entered the data into the workbook | Text | None |
| data\_doi\_url | Digital Object Identifier (DOI) of the data or URL of the webpage from which the data can be obtained. | Text | None |

Table 3: Characteristics of the sample table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Format** | **Constraints** |
| entity\_id  | Refers to the unique identifier for each entity (as given in entity table) | Numeric | Positive integer |
| sample\_id  | Unique identifier for the sample | Numeric | Positive integer |
| sample\_thickness | Thickness of the sample analysed (mm) | Numeric  | Positive decimal |
| depth\_sample | Distance in mm from a reference point  | Numeric  | Positive decimal |
| mineralogy | Description of the mineralogy of the sample | Text  | selected from pre-defined list |
| arag\_corr | Indication of whether the isotope measurements have been corrected in aragonite samples | Text | selected from pre-defined list |

Table 4: Characteristics of the dating information table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Format** | **Constraints** |
| dating\_id | Unique identifier for each date  | Numeric  | Positive integer |
| entity\_id | Refers to the unique identifier for each entity (as given in entity table) | Numeric  | Positive integer |
| date\_type | Description of the dating method used  | Text | selected from pre-defined list |
| depth\_dating | Distance in mm from a reference point  | Numeric  | Positive decimal |
| dating\_thickness | Thickness of dated sample in mm  | Numeric  | Positive decimal |
| lab\_num | The laboratory number of the dated sample | Text | None |
| material\_dated | Mineralogy of the dated sample | Text | selected from pre-defined list |
| min\_weight | Minimum weight of the dated sample in mg | Numeric  | Positive decimal |
| max\_weight | Maximum weight of the dated sample in mg | Numeric  | Positive decimal |
| uncorr\_age | Uncorrected age of the dated sample in years | Numeric  | None |
| uncorr\_age\_uncert\_pos | Positive uncertainty of the uncorrected age of the dated sample in years | Numeric  | Positive decimal |
| uncorr\_age\_uncert\_neg | Negative uncertainty of the uncorrected age of the dated sample in years | Numeric  | Positive decimal |
| 14C\_correction | Percentage dead carbon present in dated sample | Numeric  | Positive decimal  |
| calib\_used | Calibration method used to convert C14 dates to calendar years  | Text | selected from pre-defined list |
| date\_used | Indication of whether the date is used in the original age model  | Text | selected from pre-defined list |
| 238U\_content | 238U content of the dated sample in ppb | Numeric  | None |
| 238U\_uncertainty | 238U 2-sigma uncertainty of dated sample in ppb | Numeric  | None |
| 232Th\_content | 232Th content of the dated sample in ppt | Numeric  | None |
| 232Th\_uncertainty | 232Th 2-sigma uncertainty of the dated sample in ppt | Numeric  | None |
| 230Th\_content | 230Th content of the dated sample in ppt | Numeric  | None |
| 230Th\_uncertainty | 230Th 2-sigma uncertainty of the dated sample in ppt | Numeric  | None |
| 230Th\_232Th\_ratio | 230Th/232Th activity ratio of the dated sample | Numeric  | None |
| 230Th\_232Th\_ratio\_uncertainty | 230Th/232Th activity ratio 2-sigma uncertainty of the dated sample | Numeric  | None |
| 230Th\_238U\_activity | 230Th/238U activity ratio of the dated sample | Numeric  | None |
| 230Th\_238U\_activity\_uncertainty | 230Th/238U activity ratio 2-sigma uncertainty of the dated sample | Numeric  | None |
| 234U\_238U\_activity | 234U/238U activity ratio of the dated sample | Numeric  | None |
| 234U\_238U\_activity\_uncertainty | 234U/238U activity ratio 2-sigma uncertainty of the dated sample | Numeric  | None |
| ini\_230Th\_232Th\_ratio | Initial 230Th/232Th activity ratio for the detrital correction | Numeric  | None |
| ini\_230Th\_232Th\_ratio\_uncertainty | Initial 230Th/232Th activity ratio uncertainty for the detrital correction | Numeric  | None |
| decay\_constant | Description of the half-life used for 230Th and 234U for U/Th samples | Text | selected from pre-defined list |
| corr\_age | Corrected age of the dated sample in years | Numeric  | None |
| corr\_age\_uncert\_pos | Positive uncertainty of corrected age of the dated sample in years | Numeric  | Positive decimal |
| corr\_age\_uncert\_neg | Negative uncertainty of corrected age of the dated sample in years | Numeric  | Positive decimal |

Table 5: Characteristics of the lamina dating table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Format** | **Constraints** |
| dating\_lamina\_id | Unique identifier for each lamina  | Numeric | Positive integer |
| entity\_id | Refers to the unique identifier for each entity (as given in entity table) | Numeric | Positive integer |
| depth\_lam | Depth of the midpoint of the lamina in mm from a reference point  | Numeric | Positive integer |
| lam\_thickness | Thickness in mm of the sample dated | Numeric  | Positive decimal |
| lam\_age | Age in years of individual lamina | Numeric | None |
| lam\_age\_uncert\_pos | Positive counting uncertainty of individual lamina in years | Numeric | Positive decimal |
| lam\_age\_uncert\_neg | Negative counting uncertainty of individual lamina in years | Numeric | Positive decimal |

Table 6: Characteristics of the hiatus place mark information table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Format** | **Constraints** |
| sample\_id | Refers to the unique identifier for each sample (as given in sample table) | Numeric | Positive integer |
| hiatus | Indication of an hiatus | Text | selected from pre-defined list |

Table 7: Characteristics of the gap place mark information table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Format** | **Constraints** |
| sample\_id | Refers to the unique identifier for the sample (as given in sample table)  | Numeric | Positive integer |
| gap | Indication of a gap | Text | selected from pre-defined list |

Table 8: Characteristics of the original chronology information table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Format** | **Constraints** |
| sample\_id  | Refers to the unique identifier for the sample (as given in sample table) | Numeric | Positive integer |
| interp\_age | Calendar age of the sample in years | Numeric  | Positive decimal |
| interp\_age\_uncert\_pos | Positive uncertainty on the age of the sample in years | Numeric  | Positive decimal |
| interp\_age\_uncert\_neg | Negative uncertainty on the age of the sample in years | Numeric  | Positive decimal |
| age\_model\_type | Description of the age model used in the original publication | Text  | selected from pre-defined list |
| ann\_lam\_check | Indication that verification that laminae are annual, in cases where lamina counting is used to construct the age model | Text  | selected from pre-defined list |
| dep\_rate\_check | Indication that verification of the deposition rate has been made | Text  | selected from pre-defined list |

Table 9: Characteristics of the carbon isotope data table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Format** | **Constraints** |
| sample\_id  | Refers to the unique identifier for the sample (as given in sample table) | Numeric | Positive integer |
| d13C\_measurement | Original δ13C measurement | Numeric  | None |
| d13C\_precision | Laboratory precision on the δ13C measurement | Numeric  | None |
| iso\_std | Description of isotopic standard used  | Text  | selected from pre-defined list |

Table 10: Characteristics of the oxygen isotope data table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Format** | **Constraints** |
| sample\_id  | Refers to the unique identifier for the sample (as given in sample table) | Numeric | Positive integer |
| d18O\_measurement | Original δ18O measurement | Numeric  | None |
| d18O\_precision | Laboratory precision on the δ18O measurement | Numeric  | None |
| iso\_std | Description of isotopic standard used  | Text  | selected from pre-defined list |

Table 11: Characteristics of the publication information table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Type** | **Constraints** |
| ref\_id | Unique identifier for the reference | Numeric | Positive integer |
| Citation | Full citation for the original publication  | Text | None |
| publication\_DOI | Digital Object Identifier (DOI) of publication | Text | None |

Table 12: Characteristics of the link entity and publication information table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Format** | **Constraints** |
| entity\_id  | Refers to the unique identifier for the entity (as given in entity table) | Numeric | Positive integer |
| ref\_id  | Refers to the unique identifier for the publication (as given in the publication information table) | Numeric | Positive integer |

Table 13: Characteristics of the link composite and entity table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Format** | **Constraints** |
| composite\_entity\_id  | Refers to the unique identifier for a composite entity (as given in entity table) | Numeric | Positive integer |
| single\_entity\_id | Refers to the unique identifier for each single entity in the composite (as given in entity table) | Numeric | Positive integer |

Table 14: Characteristics of the notes table

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Format** | **Constraints** |
| site\_id  | Refers to the unique identifier for each site (as given in site table) | Numeric | Positive integer |
| notes | Notes and additional information about the site | Text | None |