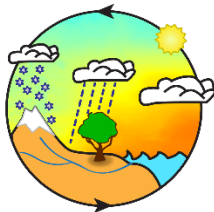


# Data Repositories

Gabriel Bowen  
Geology & Geophysics  
University of Utah



# Data Users

- ⊕ Reference data
- ⊕ Meta-analyses
- ⊕ Model benchmarks
  
- ⊕ Want...
  - ⊕ Comprehensive
  - ⊕ Easy to find
  - ⊕ Easy to use



# Data Producers

- ⊕ Funding agencies (OSTP Expanding Access Mandate)
- ⊕ Publication/journals
- ⊕ Cultural push/pull
- ⊕ Want...
  - ⊕ No extra work
  - ⊕ Exposure/credit



# The Ideal Repository?

- ⊕ Agency- vs. community-driven
- ⊕ Structured vs. flat
- ⊕ Retrospective vs. active
- ⊕ Access models (FAIR, structured/standardized metadata, APIs)

# IAEA - WISER

IAEA NUCLEUS WISER Contact Us Welcome Bowen, Gabriel

 **IAEA** | **WISER** Water Isotope System for data analysis  
visualization and Electronic Retrieval

About Datasets

[Home](#) > [NUCLEUS](#) > [Wiser](#) > [About](#)

## WISER – Water Isotope System for data analysis, visualization and Electronic Retrieval

Isotopic and ancillary data available in this database are generated under the IAEA Water Resources Programme's monitoring networks: the Global Network of Isotopes in Precipitation (GNIP), which is jointly operated with the World Meteorological Organization (WMO), and of the Global Network of Isotopes in Rivers (GNIR). These data sets are available in tabulated form through the Water Isotope System for Data Analysis, Visualization, and Electronic Retrieval (WISER). From October 2017 on, the WISER portal version 3 will be used for data dissemination.

A number of attractive features are retained from the previous version (such as using the IAEA-NUCLEUS-Single-Sign-On, which allows to use existing user accounts), to restore others (such as the map-based selection and the full functionality of the selection by country list), but also to skip a number of poorly-used functions (such as the extensive plotting engine). Furthermore, this version improves some elements of the previous interface, including the statistic treatment of the GNIP-Monthly data category into exported files, and provides with a full list of references and acknowledgements of the data to be downloaded.

### Data structure

Data categories have been rearranged to better reflect the varieties within different data categories:

- Data collected under the GNIP programme is now grouped into the following main categories: Daily/high-frequency, Monthly, Monitoring and Synoptic. Special categories include GNIP-Vapour and GNIP-Cryo (Antarctic and other datasets related to the cryosphere).
- Data collected under the GNIR programme is now grouped into two categories: Monitoring and Synoptic data.

Additional Details on the data categories can be found on the [GNIP](#) and [GNIR](#) web sites. The [NUCLEUS FAQs](#) are available for questions related to account management. For questions not answered there, pls. contact the WISER contact point at [WISER.Contact-Point@iaea.org](mailto:WISER.Contact-Point@iaea.org).

### Obtaining and citing GNIP/GNIR data

GNIP and GNIR data and summary statistics can be obtained free of charge from the WISER application. Download of WISER data is available in Microsoft Excel® format and is limited to 5,000 records per file. For the sake of continuity, the current CSV export format with fixed column heads is still available but only with the attributes used for GNIP-Monthly.

<https://nucleus.iaea.org/Pages/GNIPR.aspx>

# IAEA - WISER

- ⊕ Long-term international coverage (legacy + ongoing)
- ⊕ Pretty good QC
- ⊕ Growing support for non-traditional (e.g., rivers, vapor) data
  
- ⊕ Agency-controlled
  - ⊕ Not very agile
  - ⊕ Branding agency-focused (e.g., GNIP)
- ⊕ Data access and use restrictions
- ⊕ Search (currently) pretty limited



# SWVID

Yale University

## Stable Water Vapor Isotopes Database

[Home](#) [Data Policy](#) [Data Download](#) [Publications](#) [Contact Us](#)

Hosted by Yale University and sponsored by the U. S. National Science Foundation, the Stable Water Vapor Isotope Database (SWVID) website archives high-frequency vapor isotope data collected with instruments based on infrared isotopic spectroscopy. The goal of this centralized data depository is to facilitate investigation that transcends disciplinary and geographic boundaries.

Last update (2018/11/12)

Click on the interactive map to view individual site information




<https://vapor-isotope.yale.edu/>

# SWVID

- ⊕ Indexed catalog of vapor monitoring records (currently ~40 sites)
- ⊕ Functional & uniform metadata, format
- ⊕ Easy search, open access
  
- ⊕ No metadata standard
- ⊕ Flat data
- ⊕ Lots of complimentary data could be linked (e.g., flux tower)



# Oceans

 National Aeronautics and Space Administration  
Goddard Institute for Space Studies

Goddard Space Flight Center  
Sciences and Exploration Directorate  
Earth Sciences Division

[GISS Home](#)  
[News & Features](#)  
[Projects & Groups](#)  
**[Datasets](#)**  
[Publications](#)  
[Software](#)  
[Education](#)  
[Events](#)  
[About GISS](#)

## Global Seawater Oxygen-18 Database

This database is a collection of over 26,000 seawater O-18 values made since about 1950. We hope that other researchers can benefit from the not-inconsiderable pain it took to collect, verify and put together all these different sources.

This is an amalgamation of independent efforts by Gavin Schmidt, Grant Bigg and Eelco Rohling which combines our separately evolved interests. Partial versions of these databases appeared in [Schmidt \(1999\)](#) and [Bigg and Rohling \(2000\)](#). These publications can (should!) be referenced, but references to the whole database (currently version 1.22) should include the web address as follows:

Schmidt, G.A., G. R. Bigg and E. J. Rohling. 1999. "Global Seawater Oxygen-18 Database - v1.22"  
<https://data.giss.nasa.gov/o18data/>

Since this data comes from different laboratories, different standards have been used and it is not always possible to make everything consistent. If you find that strange phenomena occur across different data sets, please don't blame us! Make sure to read the notes that go along with each reference, and try and go back to the original articles (where possible) for clarification. If there are still unresolved issues, please let us know and we will try to investigate. If you know of more pertinent data that could usefully be included here, please let us know and it will happen. Nominally all delta values are with respect to V-SMOW. Published values that used an older standard (original SMOW or PDB) have been corrected (and there is a note in the data reflecting this).

**Note added 27 Dec 2006:** Temperature measurements can be in-situ or potential temperature (theta). If this matters for you, please check with the original reference.

<https://data.giss.nasa.gov/o18data/>

# Paleo Data

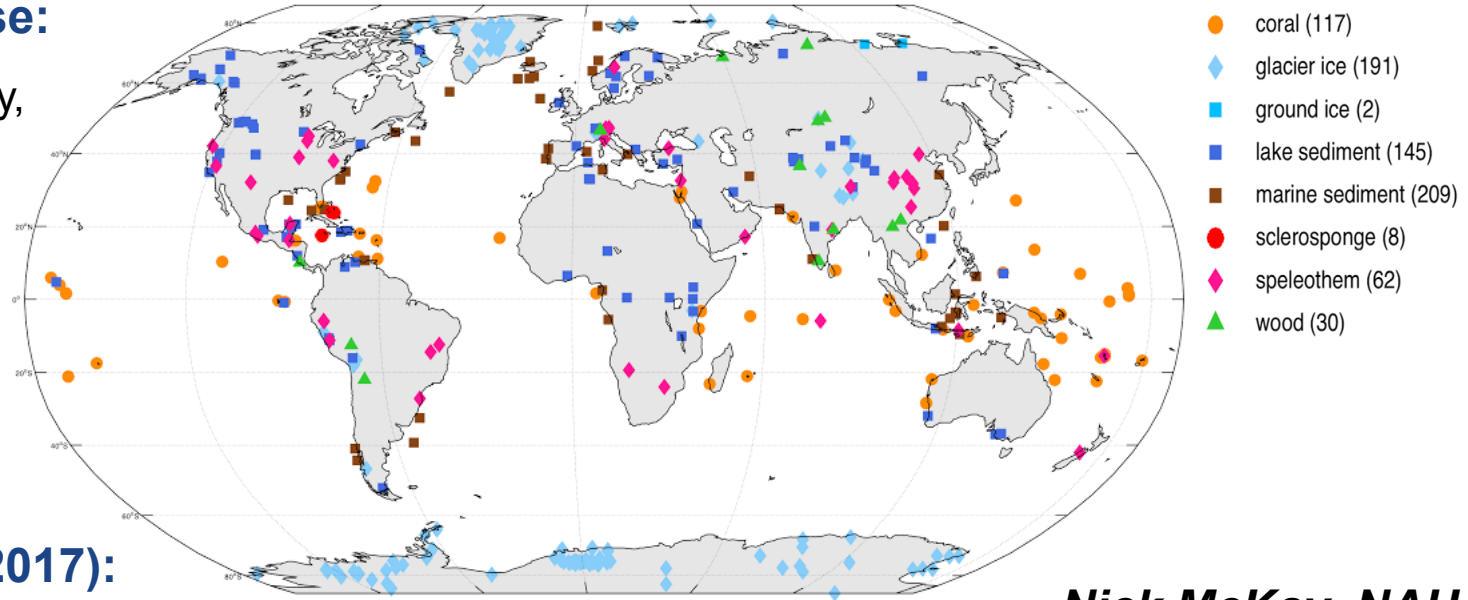
- ⊕ Some options here...
- ⊕ Classic & emerging 'catch all' repositories (NCDC, Pangea, NEOTOMA)
  - ⊕ Various levels of structure, accessibility
  - ⊕ Various levels of PITA to deposit
  - ⊕ Various levels of isotope-specific support
- ⊕ Synthesis-driven effort – Iso2k (Konecky + collabs)

# Iso2k

## The Iso2k database:

- Multi-archive, proxy, resolution, realm
- Tons of metadata

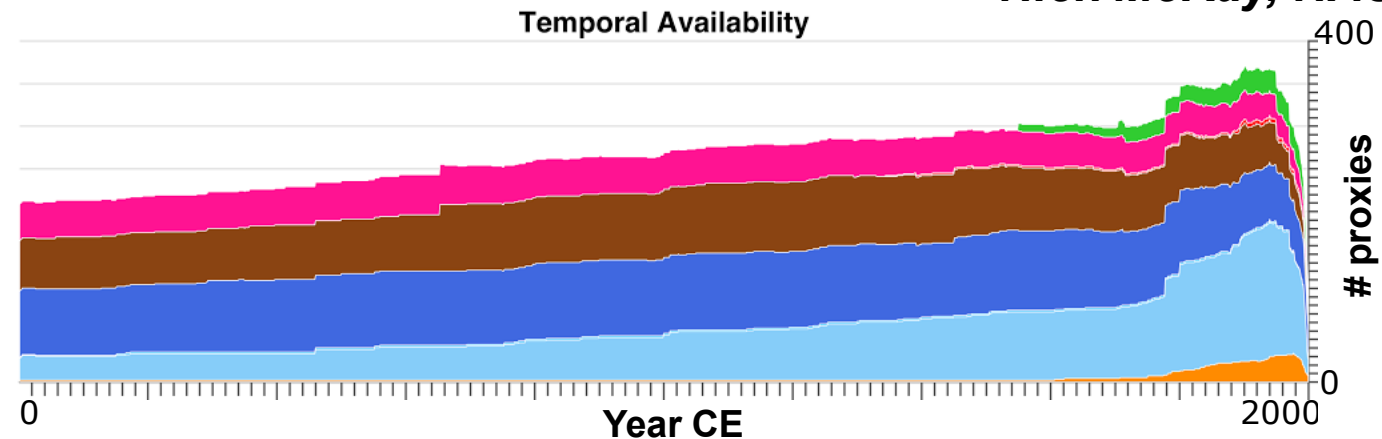
Iso2k network version 0.6 (629 records from 432 sites)



## “Science beta” version 0.6 (Nov. 2017):

- Major QC done
- Chronological data mostly done (95% of marine sed, 75% of lake sed)
- LiPD format
- MATLAB, R, Python

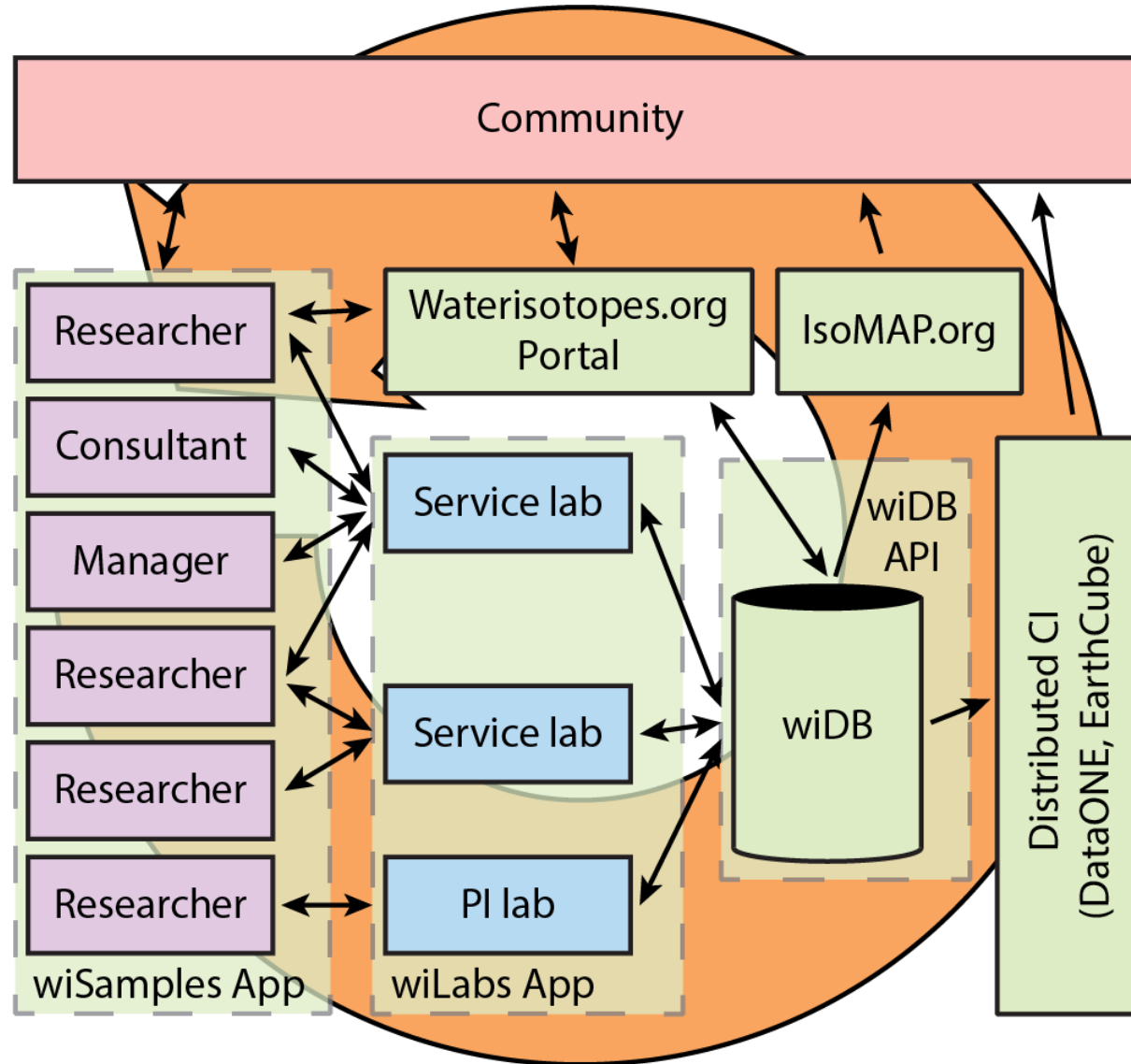
*Nick McKay, NAU*



# wiDB/IsoBank

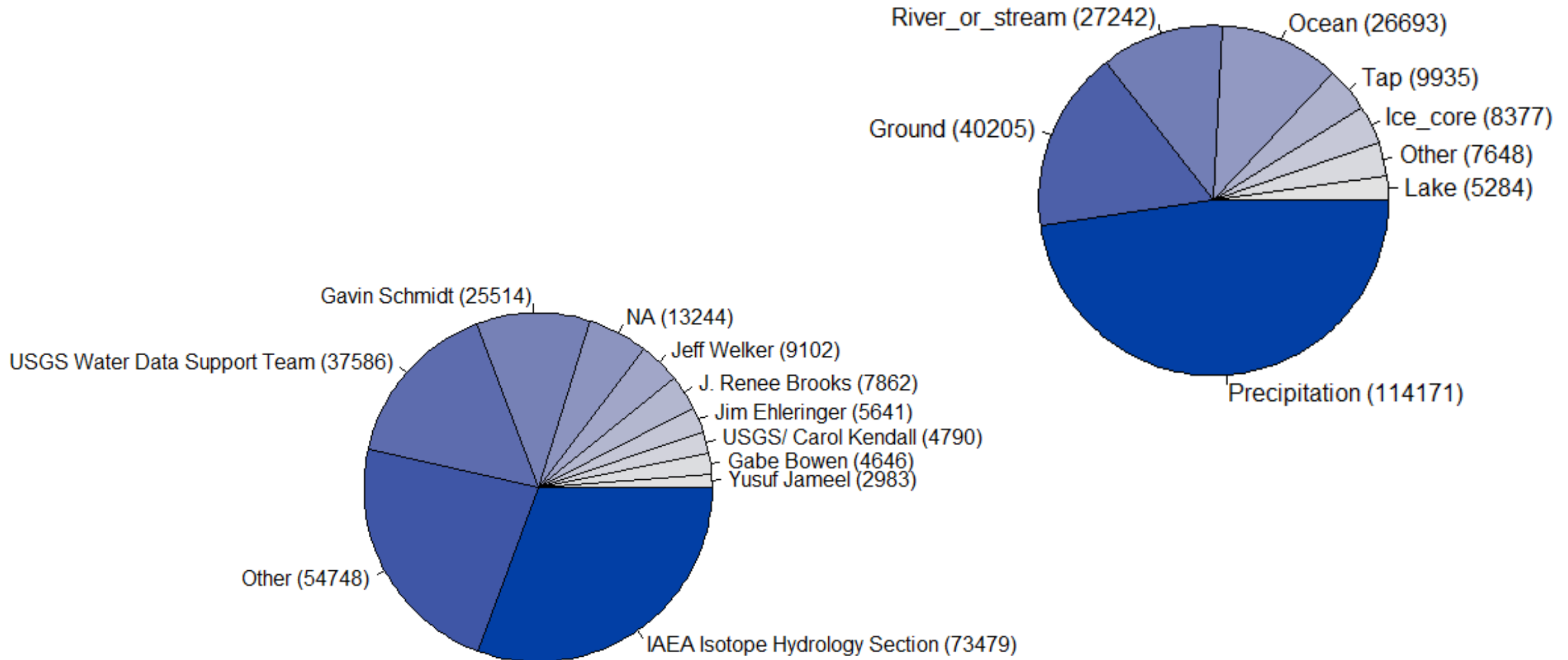
- ⊕ Waterisotopes Database
  - ⊕ Structured relational DB hosting all types of water isotope data
  - ⊕ Embedded metadata model (simple and uniform but not based on community standards)
  - ⊕ Grown from meta-analysis project, growing community involvement
- ⊕ The bigger picture – a new model of how we archive our data?

# wiDB – The Big Picture



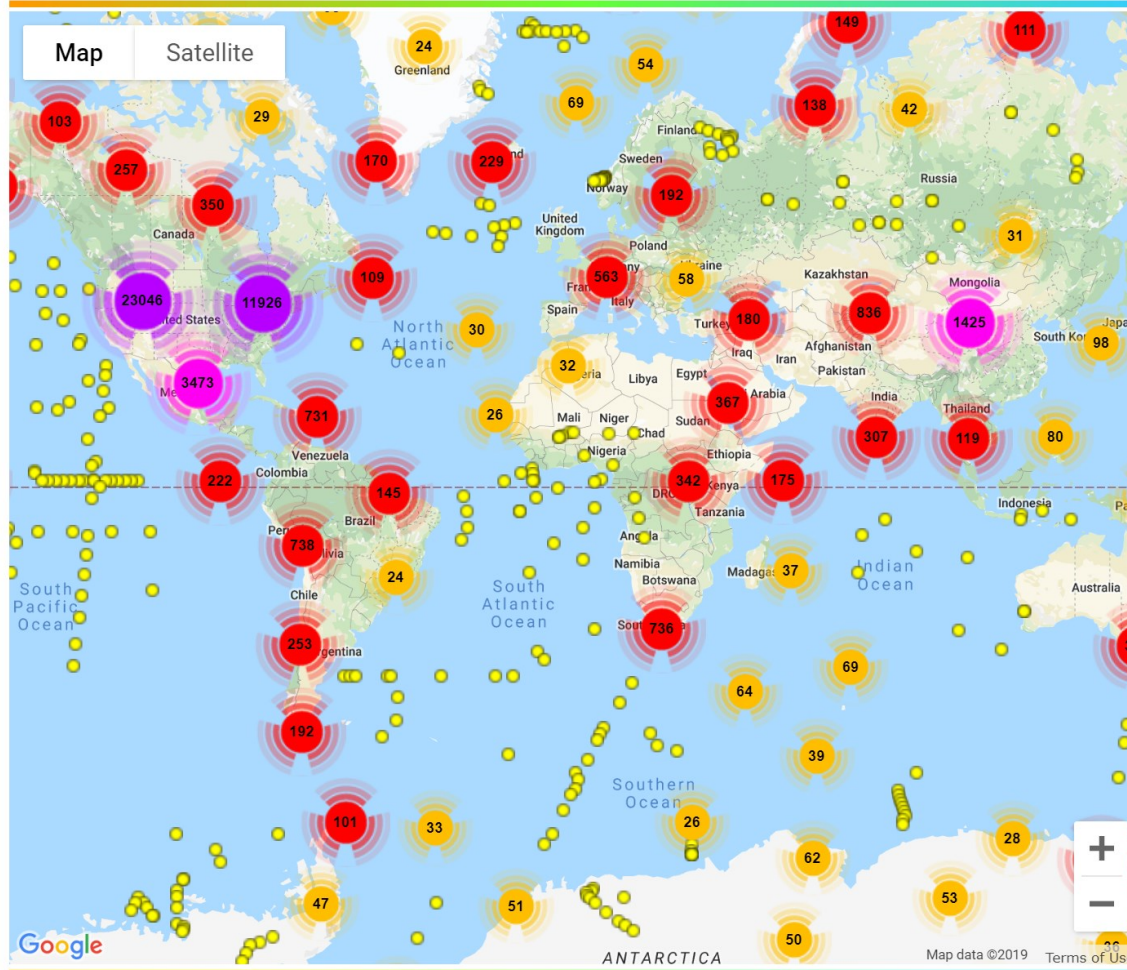
**Poster!**

# wiDB





# wiDB Access



**Recent datasets added:**  
00229 - Yang daily precip - Xiaoxin Yang

North Latitude   
West Longitude  East Longitude   
South Latitude

Country   
State/Province

Type   
Project ID

Collection Date  To

Elevation  To

$\delta^2\text{H}$    $\delta^{18}\text{O}$

- Web portal
- API

<http://waterisotopes.org>

[Data use policy](#) || [Contribute data](#)



# wiDB -> IsoBank

- ✦ IsoBank is a bottom-up community data initiative for isotope science
  - ✦ NSF-BIO \$, covers Earth/environmental, biological, anthropological sciences
  - ✦ 4 committees on metadata standards, procedures & community engagement (Adriana Bailey as rep for atmospheric water/vapor)



Why we need a centralized repository for isotopic data

Jonathan N. Pauli<sup>a,1</sup>, Seth D. Newsome<sup>b</sup>, Joseph A. Cook<sup>c</sup>, Chris Harrod<sup>d</sup>, Shawn A. Steffan<sup>e,f</sup>, Christopher J. O. Baker<sup>g</sup>, Merav Ben-David<sup>h</sup>, David Bloom<sup>i</sup>, Gabriel J. Bowen<sup>j</sup>, Thure E. Cerling<sup>j</sup>, Carla Cicero<sup>k</sup>, Craig Cook<sup>h</sup>, Michelle Dohm<sup>l</sup>, Prarthana S. Dharampal<sup>f</sup>, Gary Graves<sup>m,n</sup>, Robert Gropp<sup>o</sup>, Keith A. Hobson<sup>p</sup>, Chris Jordan<sup>q</sup>, Bruce MacFadden<sup>r</sup>, Suzanne Pilaar Birch<sup>s,t</sup>, Jorrit Poelen<sup>u</sup>, Sujeewan Ratnasingham<sup>v</sup>, Laura Russell<sup>i</sup>, Craig A. Stricker<sup>w</sup>, Mark D. Uhen<sup>x</sup>, Christopher T. Yarnes<sup>y</sup>, and Brian Hayden<sup>z</sup>

# Thoughts...

- ⊕ Plenty of places to stick your data
- ⊕ What we lack
  - ⊕ Community supported standards that make data and metadata accessible and interoperable
  - ⊕ Repositories that prioritize and support FAIR principles
  - ⊕ Consensus on best practices for archiving water isotope data
  - ⊕ Tools for discovery and retrieval across resources