On the PAGES2k project, paleoclimate cyberinfrastructure and integrating paleoclimatology and paleoseismology

Nick McKay, Julien Emile-Geay & the PAGES 2k Consortium
Acknowledgments

PAGES2K Consortium (in revision), A global multiproxy database for temperature reconstructions of the Common Era, *Scientific Data*.

Julien Emile-Geay (JEG), Nicholas P. McKay (NPM), Darrell S. Kaufman (DSK), Lucien von Gunten (LvG), Jianghao Wang (JW), Kevin J. Anchukaitis (KJA)

Nerilie J. Abram (NJA), Jason A. Addison (JAA), Mark A. J. Curran (MAC), Michael N. Evans (MNE), Benjamin J. Henley (BJH), Zhixin Hao (ZH), Belen Martrat (BM), Helen V. McGregor (HVM), Raphael Neukom (RN), Gregory T. Pederson (GTP), Barbara Stenni (BS), Kaustubh Thirumalai (KT), Johannes P. Werner (JPW), Chenxi Xu (CX)

Dmitry V. Divine (DVD), Bronwyn C. Dixon (BCD), Joelle Gergis (JG), Ignacio A. Mundo (IAM), Takeshi Nakatsuka (TN), Steven J. Phipps (SJP), Cody C. Routson (CCR), Eric J. Steig (EJS), Jessica E. Tierney (JET), Jonathan J. Tyler (JJT)

Kathryn J. Allen (KatJA), Nancy A.N. Bertler (NAB), Jesper Björklund (JB), Brian M. Chase (BMC), Min-Te Chen (MTC), Edward Cook (EC), Rixt de Jong (RdJ), Kristine L. DeLong (KLD), Daniel A. Dixon (DAD), Alexey A. Ekaykin (AAE), Vasile Ersek (VE), Helena L. Filipsson (HLF), Pierre Francus (PF), Mandy B. Freund (MBF), Massimo Frezzotti (MF), Narayan P. Gaire (NPG), Konrad Gajewski (KG), Quansheng Ge (QG), Hugues Goosse (HG), Anastasia Gornostaeva (AG), Martin Grosjean (MG), Kazuho Horiuchi (KH), Anne Hormes (AH), Katrine Husum (KatH), Elisabeth Isaksson (EI), Selvaraj Kandasamy (SK), Kenji Kawamura (KK), K. Halimeda Kilbourne (KHK), Nalan Ko (NC), Guillaume Leduc (GL), Hans W. Linderholm (HWL), Andrew M. Lorrey (AML), Vladimir Mikhalenko (VM), P. Graham Mortyn (PGM), Hideaki Motoyama (HM), Andrew D. Moy (ADM), Robert Mulvaney (RM), Philipp M. Munz (PMM), David J. Nash (DJN), Thomas Opel (TO), Hans Oerter (HO), Anais J. Orsi (AJO), Dmitriy V. Ovchinnikov (DO), Trevor J. Porter (TJP), Heidi A. Roop (HAR), Casey Saenger (CS), Masaki Sano (MasS), David Sauchyn (DJS), Krystyna M. Saunders (KMS), Kate Sinclair KS), Marit-Solveig Seidenkrantz (MSS), Mirko Severi (MirS), Xuemei Shao (XS), Marie-Alexandrine Sicre (MAS), Michael Sigl (MicS), Scott St. George (SSG), Jeanine-Marie St. Jacques (JSJ), Meloth Thamban (MT), Udya Kuwar Thapa (UKT), Elizabeth R. Thomas (ERT), Chris Turney (CT), Ryu Uemura (RU), Andre E. Viau (AEV), Diana O. Vladimirova (DOV), Eugene R. Wahl (ERW), James W.C. White (JWW), Zicheng Yu (ZY), Jens Zinke (JZ)

➡ latest temperature db: 98 scientists from 22 countries
PAGES in FutureEarth

“The farther backward you can look, the farther forward you are likely to see”
Winston Churchill

• PAGES = Past Global Changes
  ➡ PAGES = core project of Future Earth
• Distributed collaboration
  ➡ Unique data stewardship challenges (and opportunities)
  ➡ Good laboratory for next generation paleoscience
  ➡ We’ve made a lot of mistakes, so you don’t have to
Why PAGES 2k?

1. Hughes’ First Law of Paleoclimatology
Why PAGES 2k?

- Archive types:
  - bivalve
  - borehole
  - coral
  - documents
  - glacier ice
  - hybrid
  - lake sediment
  - marine sediment
  - sclerosponge
  - speleothem
  - tree

- Resolution (years):
  - 1/12
  - 1/6
  - 1/2
  - 1
  - 2
  - 5
  - 10
  - 20
  - 50
  - 100
  - 200
  - 400
  - 600
  - 800
  - 1000
  - 1200
  - 1400
  - 1600
  - 1800
  - 2000

- Temporal Availability:
  - # proxies

- Year (CE):
  - 0
  - 50
  - 100
  - 150
  - First Millennium

Why PAGES 2k?
Why PAGES 2k?

2. Engage regional expertise
3. Create an open, common, evolving community resource
A brief history... phase 1
A brief history… phase 1

Database proxy data types used for each regional reconstruction. Supplementary regions used in this study. The pie charts represent the fraction of

Figure 1

... in regional pollen-based reconstructions (0.334) multiplied by those that capture the serial correlation structure of the time series (ref.

... period, domains and gridded instrumental products specified for each region in the Supplementary Information.

... regression; CPS, composite plus scaling; PPR, point-by-point regression.

... for the warm season (Sept–Feb).

... values were calculated for the correlation between reconstructed values and target instrumental data using the seasons, calibration

... of the decadal tree-ring-based calibration itself (0.929); Correlation for decadal mean reconstruction.

... values for calibrations are all <0.001 as estimated by phase-randomization Monte-Carlo methods

... of the 30-year-averaged PAGES 2k Network reconstructions

... exhibit a significant change in the time series. Generally, the longer the proxy record, the more likely it is to

... trends in the 30-year-averaged PAGES

... term cooling trend followed by recent warming during the twentieth century, except Antarctica. Before

... personal biases in some proxies.

... records were also analysed to deter

... values in the alternative reconstructions (Supplementary

... agree with those in the alternative reconstructions (Supplementary

... of the decadal tree-ring-based calibration itself (0.929); Correlation for decadal mean reconstruction.

... values for calibrations are all <0.001 as estimated by phase-randomization Monte-Carlo methods

... of the 30-year-averaged PAGES 2k Network reconstructions

... exhibit a significant change in the time series. Generally, the longer the proxy record, the more likely it is to

... trends in the 30-year-averaged PAGES

... term cooling trend followed by recent warming during the twentieth century, except Antarctica. Before

... personal biases in some proxies.

... records were also analysed to deter

... values in the alternative reconstructions (Supplementary

... agree with those in the alternative reconstructions (Supplementary

... of the decadal tree-ring-based calibration itself (0.929); Correlation for decadal mean reconstruction.

... values for calibrations are all <0.001 as estimated by phase-randomization Monte-Carlo methods

... of the 30-year-averaged PAGES 2k Network reconstructions

... exhibit a significant change in the time series. Generally, the longer the proxy record, the more likely it is to

... trends in the 30-year-averaged PAGES

... term cooling trend followed by recent warming during the twentieth century, except Antarctica. Before

... personal biases in some proxies.
A brief history... phase 1

Several studies have investigated the cause of cooling between the tenth and sixteenth centuries, suggesting a potential role of solar irradiance, volcanic activity, land-cover changes and orbitally driven atmospheric variability (Supplementary Fig. S4). An ensemble of simulations performed with a climate model of intermediate complexity (Supplementary Note 341) simulates a multi-millennial cooling in summer as a delayed response to three longest volcanic-solar downturns (1160 to 1370). In the Arctic and Europe, temperatures were relatively high during the thirteenth and northern Scandinavia, a sustained warm period occurred later, from around 1200 and 1300 in the Arctic, 1200 and 1300 in South America and 1200 and 1300 in Australasia, a sustained warm period occurred later, from around 1200 and 1300 in South America. Temperatures did not reach a peak before the 1400-1500 interval in all four Northern Hemisphere regions. In South America, a climate model of intermediate complexity (Supplementary Note 341) simulates a multi-millennial cooling in winter as a delayed response to an orbitally driven decrease in local summer insolation, as suggested for the Arctic and northern Scandinavia (Mann-Kendall trend test, 1990–1970). North America includes a shorter tree-ring-records set for the pre-twentieth centuries of the second millennium compared with those of the late (pre-twentieth) centuries of the second millennium. Commensurate cooling occurred during the last millennium; for example, some records of volcanic activities (Fig. S6), which is extremely unlikely to occur by chance (one-tailed probability test, 0.001). Of these, three-quarters of the slopes are statistically significant (P < 0.05).}

**Volcanic-solar downturns**

- **Arctic**
- **Europe**
- **Asia**
- **North America**
- **South America**
- **Australasia**

**Count**

- **Arctic**
- **Europe**
- **North America - trees**
- **North America - pollen**
- **Australasia**
- **South America**
- **Antarctica**

**Year AD**

- 1
- 200
- 400
- 600
- 800
- 1000
- 1200
- 1400
- 1600
- 1800
- 2000

**Standardized rel. to AD 1990–1970 (SD)**

- 3
- 2
- 1
- 0
- -1
- -2
- -3

**Temperature evolution at multi-centennial scales**

- The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighteenth century, is consistent with the underlying millennial-scale cooling trend. The relative regional warmth, especially during the eighth...
A brief history...
A brief history... phase 1

Shortcomings:
1. Non-uniform data collection
2. Relatively metadata-poor
3. No oceans
4. Global reconstruction
Phase 2
Criteria

1. Length
2. Resolution
3. Thermal sensitivity
The data

The data entered into a spreadsheet. The table contains various columns labeled from A to W, with data entries for years from 2007 to 2023. The table includes columns such as Year, Value, and Composite of WAS 2005A (90% Variance Matched) and WAS 2006A. The data seems to be numerical values with some text notes at the bottom of the table.
Why data standards?
Why data standards?

How have we for so long managed without data standards?

Going from “long tail” field investigations to Big Science requires unification (not sufficient, but necessary)

Examples: genomics, astronomy, particle physics
Premise behind LiPD:

1. All paleorecords share common features, and allow for a common structure

3. That structure must be flexible

5. The data must be accessible
LiPD Dataset v1.2

- geospatial metadata
- publication metadata
- funding metadata

- paleoData
  - paleoMeasurement Table
    - ensemble Table
    - methods
    - distribution Table
    - summary Table
  - paleoModel

- chronData
  - chronMeasurement Table
    - ensemble Table
    - methods
    - distribution Table
  - chronModel
    - summary Table
integration

Google Sheets

LiPD
Measured Variable: trsgi
Units: NA
climateInterpretation:
  climateParameter: T
  climateParameterDetail: air
  Seasonality: Summer
  Direction: positive
useInGlobalTemperatureAnalysis: TRUE

Proxy vs. MAT ($\rho_{loc} = 0.42$)
Proxy vs. JJA ($\rho_{loc} = 0.69$)
Proxy vs. DJF ($\rho_{loc} = 0.11$)
archiveType: marine sediment
PAGES2k region: O2kLR
Authors: Nieto-Moreno, V.; Martínez-Ruiz, F.; Willmott, V.; García-Orellana, J.; Masqué, P.; Sinninghe Damsté, J. S.
Year: , DOI: 10.1016/j.orggeochem.2012.11.001
parameter: temperature
units: deg C
climateInterpretation:
  climateParameter: T
  climateParameterDetail: sea_surface
  Seasonality: Annual
  Direction: positive
  QC Notes: ?

Calibration Information
  calibration_equation: est. SST = (UK'37 - 0.044) / 0.033
  calibration_reference: doi:10.1016/S0016-7037(98)00097-0
  calibration_uncertainty: 1.5 deg C
  calibration_uncertaintyType: RMSE
  calibration_notes: MUE1998: T(sediments) = (Uk’37-0.044)/0.033, (Müller et al., (1998); global (60S-60N), 0-29°C, annual at 0 m, n=370, R2=0.958)
WHAT IS LINKED EARTH?

EarthCube “Integrated Activity”

- paleoclimate data crowd-curation platform
- social enterprise to standardize paleoclimate data
- enable (semantic) links between disciplines

LinkedEarth: springboard to better, easier paleoscience

**GOALS**

1. Crowdsource data curation through a social-technical system to develop linked-data compliant standard for paleoclimate observations.
2. Develop social codes to process these datasets.

**CONCEPT**

- Data
- Ontology
- Entities
- Workflows
- Software

Crowdsource data curation through a social-technical system to develop linked-data compliant standard for paleoclimate observations.

Develop social codes to process these datasets.
### Arc-HalletLake.McKay.2008

**Standard Properties**

<table>
<thead>
<tr>
<th>Dataset Name</th>
<th>Arc-HalletLake.McKay.2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archive Type</td>
<td>Lake sediment</td>
</tr>
<tr>
<td>Collection Name</td>
<td>HT01</td>
</tr>
<tr>
<td>Investigators</td>
<td>Not defined!</td>
</tr>
<tr>
<td>Publication</td>
<td>Publication.10.1029/2007GL032876</td>
</tr>
<tr>
<td>GeoLocation</td>
<td>Arc-HalletLake.McKay.2008.Location</td>
</tr>
<tr>
<td>Chron Data</td>
<td>Not defined!</td>
</tr>
</tbody>
</table>

**Properties**

<table>
<thead>
<tr>
<th>Age Units</th>
<th>BP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Metadata Worksheet</td>
<td>ocqumc6</td>
</tr>
<tr>
<td>Google Spreadsheet</td>
<td><a href="https://docs.google.com/spreadsheets/d/1oQpylJSn">https://docs.google.com/spreadsheets/d/1oQpylJSn</a></td>
</tr>
<tr>
<td>Year Units</td>
<td>AD</td>
</tr>
</tbody>
</table>
Stable carbon isotopes in trees

The source of carbon in trees is atmospheric CO$_2$. The CO$_2$ enters the tree though the leaf stomata. An isotopic fractionation or discrimination against the heavy isotope $^{13}$C occurs during the diffusion through the stomata. Generally, the heavier isotopic species diffuse more slowly than the lighter isotopic species. The diffusivity of $^{13}$CO$_2$ in air has been calculated to be 4.4% less than that of $^{12}$CO$_2$ [2]. This simplistic value has been adopted to describe the fractionation due to diffusion through the leaf stomata. However, in reality, the situation is more complex since isotopes fractionation also occurs in the laminar boundary layer surrounding the leaf, during collision with guard cells, in the intracellular air spaces, and with water vapor [3].

Another fractionation occurs during carboxylation by the enzyme used to fix carbon, which, for C3 plants, is ribulose-1,5-biphosphate (RuBisCO). The magnitude of the fractionation depends on the photosynthetic pathway (C3, C4, or CAM), the environment, the genotype and many other factors [4]. For C3 plants, this fractionation has been measured and is on the order of -27% [5].

The $^{13}$C of tree ring cellulose reflects the $^{13}$C of internal CO$_2$, whose isotopic composition is affected by the rate at which CO$_2$ enters the leaf and the rate at which CO$_2$ is taken up for photosynthesis, each step being associated with discrimination of $^{13}$CO$_2$. The $^{13}$C of cellulose of therefore a measure of the stomatal conductance, which is mainly controlled by moisture stress and humidity, and the photosynthetic rate, which depends on temperature and sunlight hours (Fig. 1) [1] [6] [7].

Quantitatively, carbon isotope fractionation in response to the atmospheric isotopic signature, diffusion through the stomata, carboxylation, and environmental factors can be described by the following equation [2]:

$$\delta^{13}C_{\text{cellulose}} = \delta^{13}C_{\text{atm}} - a - (b - a) \left( \frac{c_i}{c_a} \right)$$

where:
- $\delta^{13}C_{\text{cellulose}}$ is the measured carbon isotopic composition of the cellulose
- $\delta^{13}C_{\text{atm}}$ is the carbon isotopic signature of the atmosphere
- $a$ is the discrimination associated with the transport of CO$_2$ though the stomata -4.4%
- $b$ is the fractionation associated with carboxylation, -27%
Data and code for the second phase of the PAGES2k synthesis of temperature-sensitive proxies

<table>
<thead>
<tr>
<th>Branch:</th>
<th>master</th>
<th>New pull request</th>
<th>New file</th>
<th>Find file</th>
<th>HTTPS</th>
<th>Download ZIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **CommonClimate** added verif_stats.m

  - `data`: Made compliant with version 1.7.1
  - `figs`: Made compliant with version 1.7.1
  - `utilities`: Made compliant with version 1.7.1
  - `.gitignore`: added verif_stats.m
  - `LICENSE`: Initial commit
  - `README.md`: updated README
  - `pages2k_compositeByArchive.m`: Made compliant with version 1.7.1
  - `pages2k_compositeByLatitudeBand.m`: Made compliant with version 1.7.1
  - `pages2k_composite_prep.m`: Made compliant with version 1.7.1
  - `pages2k_composite_recordLength.m`: Updated README; added "recordLength" analysis; streamlined code
  - `pages2k_composite_regression.m`: Made compliant with version 1.7.1
  - `pages2k_composite_sensitivity.m`: Made compliant with version 1.7.1
  - `pages2k_composite_workflow.m`: Made compliant with version 1.7.1

Latest commit 959124a 5 days ago
The current state of the database (v 1.12.0)
b) Resolution

- **Archive types**:
  - Bivalve
  - Borehole
  - Coral
  - Documents
  - Glacier ice
  - Hybrid
  - Lake sediment
  - Marine sediment
  - Sclerosponge
  - Speleothem
  - Tree

- **Resolution** (years):
  - 1/12
  - 1/6
  - 1/2
  - 1
  - 2
  - 5
  - 10
  - 20
  - 50
  - 100
  - 200
  - 400
  - 600
  - 800
  - 1000
  - 1200
  - 1400
  - 1600
  - 1800
  - 2000

- **Temporal Availability** (# proxies)

- **Year (CE)**:
  - 0
  - 50
  - 100
  - 150
  - First Millennium
Is it useful for understanding past temperature?
Does it capture the 20th century?
How about low-resolution records?
High vs low resolution

b) HR screen = none; LR screen = none
50-yr bins, Δt 5-yr cutoff

Temperature (°C)

nHR= 570
nLR= 104
HadCRUT4
Lessons Learned

1. Identify metadata
2. Structure metadata hierarchically
3. Selection criteria
4. Iterative data and metadata assimilation
5. Flexible scientific workflow
Thanks!

Write us!
nick@nau.edu
julieneg@usc.edu