

**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 Bjrklund (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 Mikhaleenko (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), Jeannine-Marie St. Jacques (JSJ), Meloth Thamban (MT), Udyia 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)

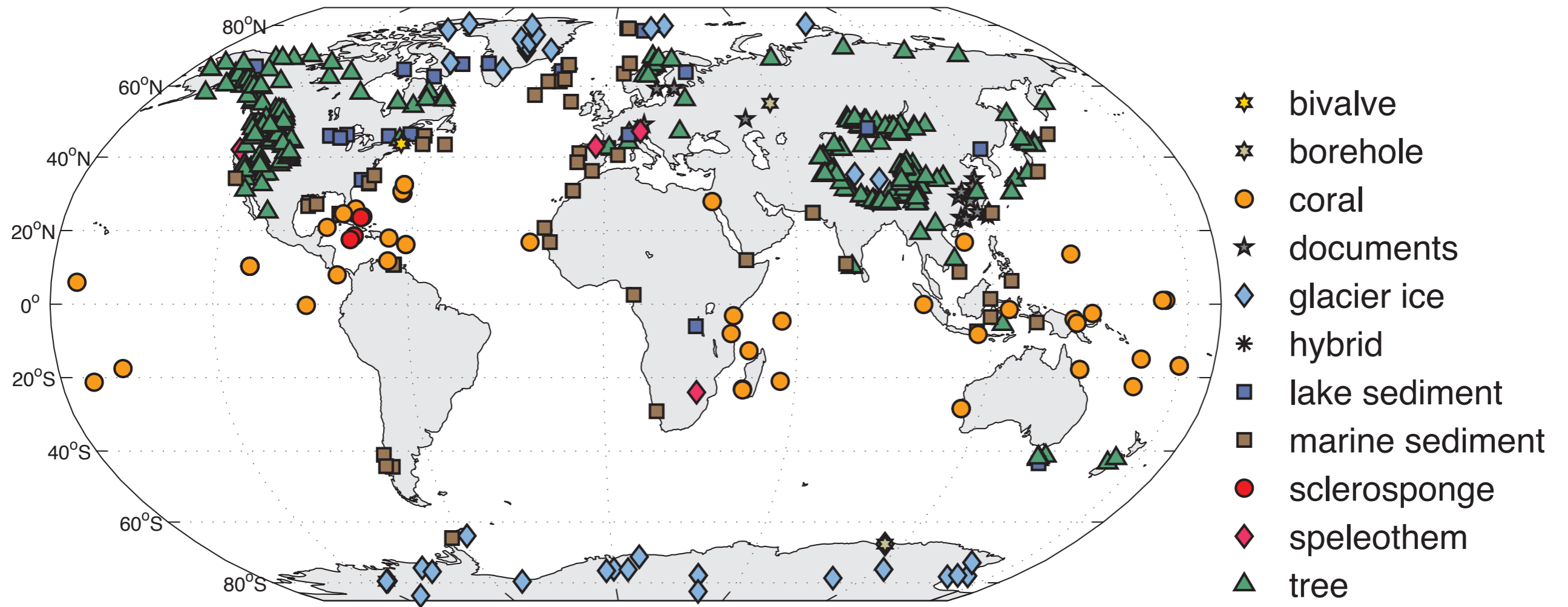
- ➔ Good laboratory for next generation paleoscience

- ➔ 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?



Why PAGES 2k?

2. Engage regional expertise

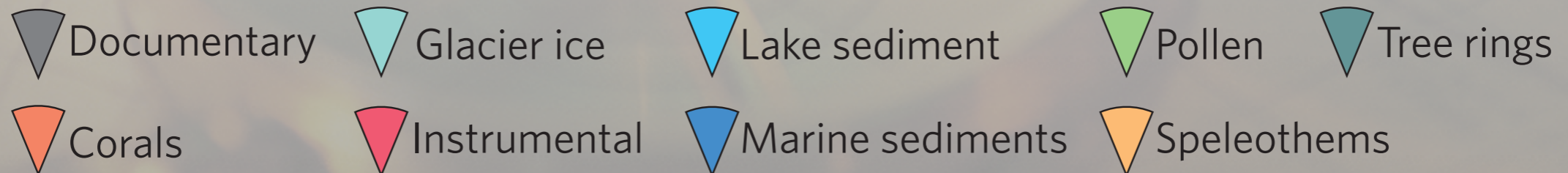
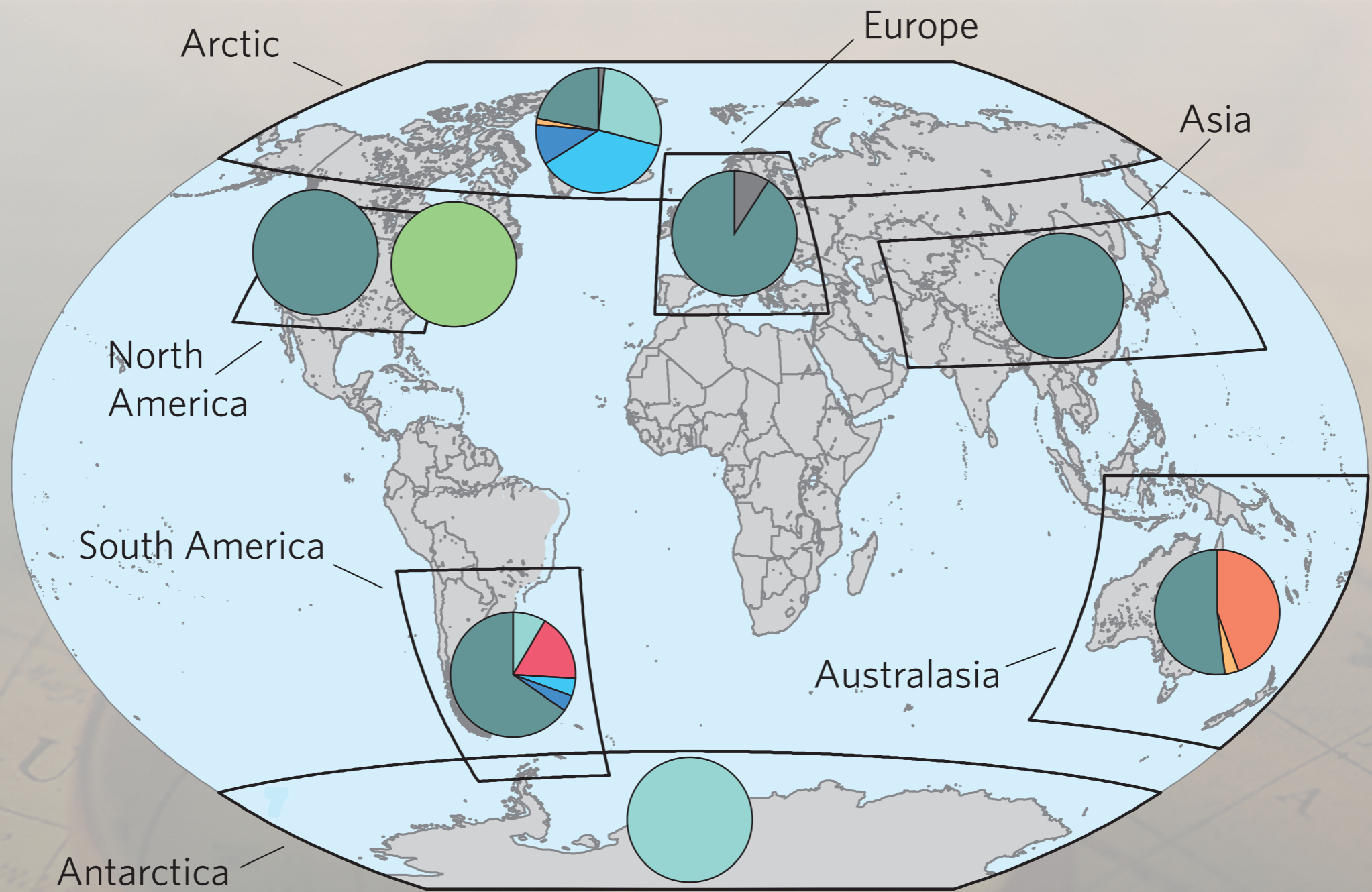
Why PAGES 2k?

3. Create an open, common, evolving
community resource

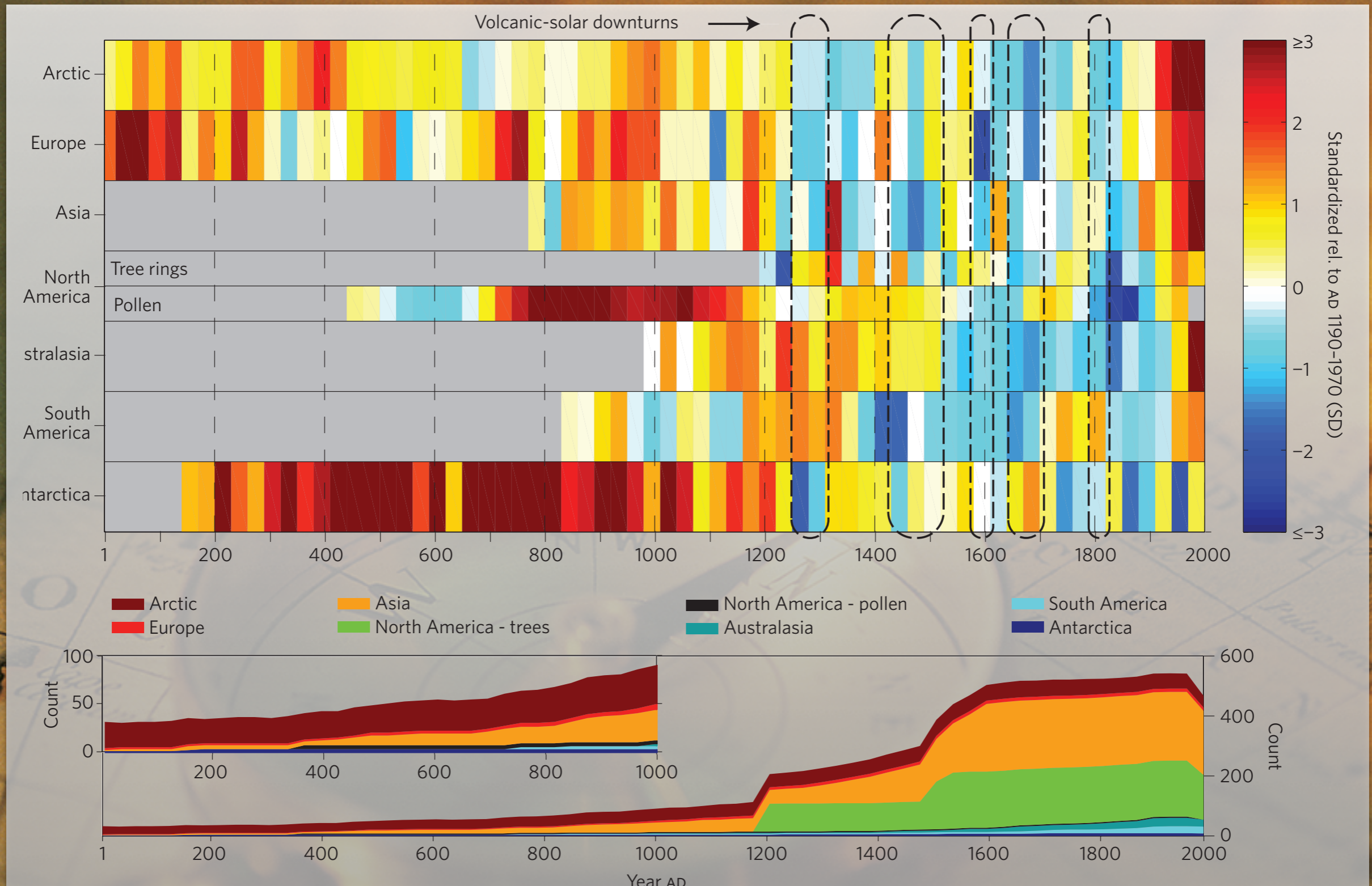
A brief history... phase 1



A brief history... phase 1



A brief history... phase 1



A bri

home > archive > issue > progress article > article > Article metrics

Article metrics for:

Continental-scale temperature variability during the past two millennia

PAGES 2k Consortium

Nature Geoscience 6, 339–346 (2013) | doi:10.1038/ngeo1797

Last updated: 6 September 2016 17:39:46 EDT

Total citations



Online attention

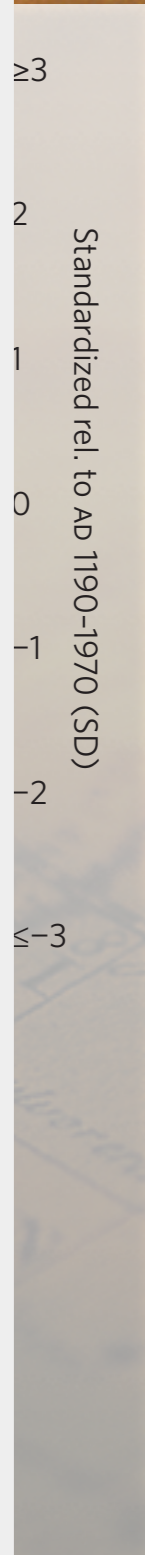
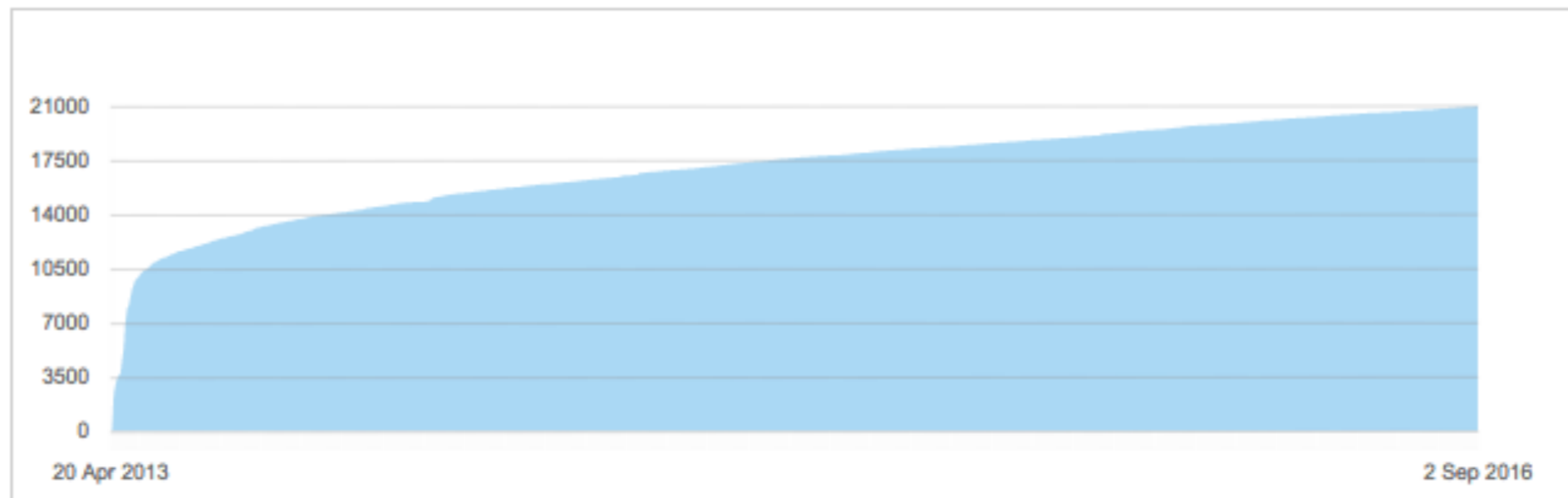


This Altmeter score means that the article is:

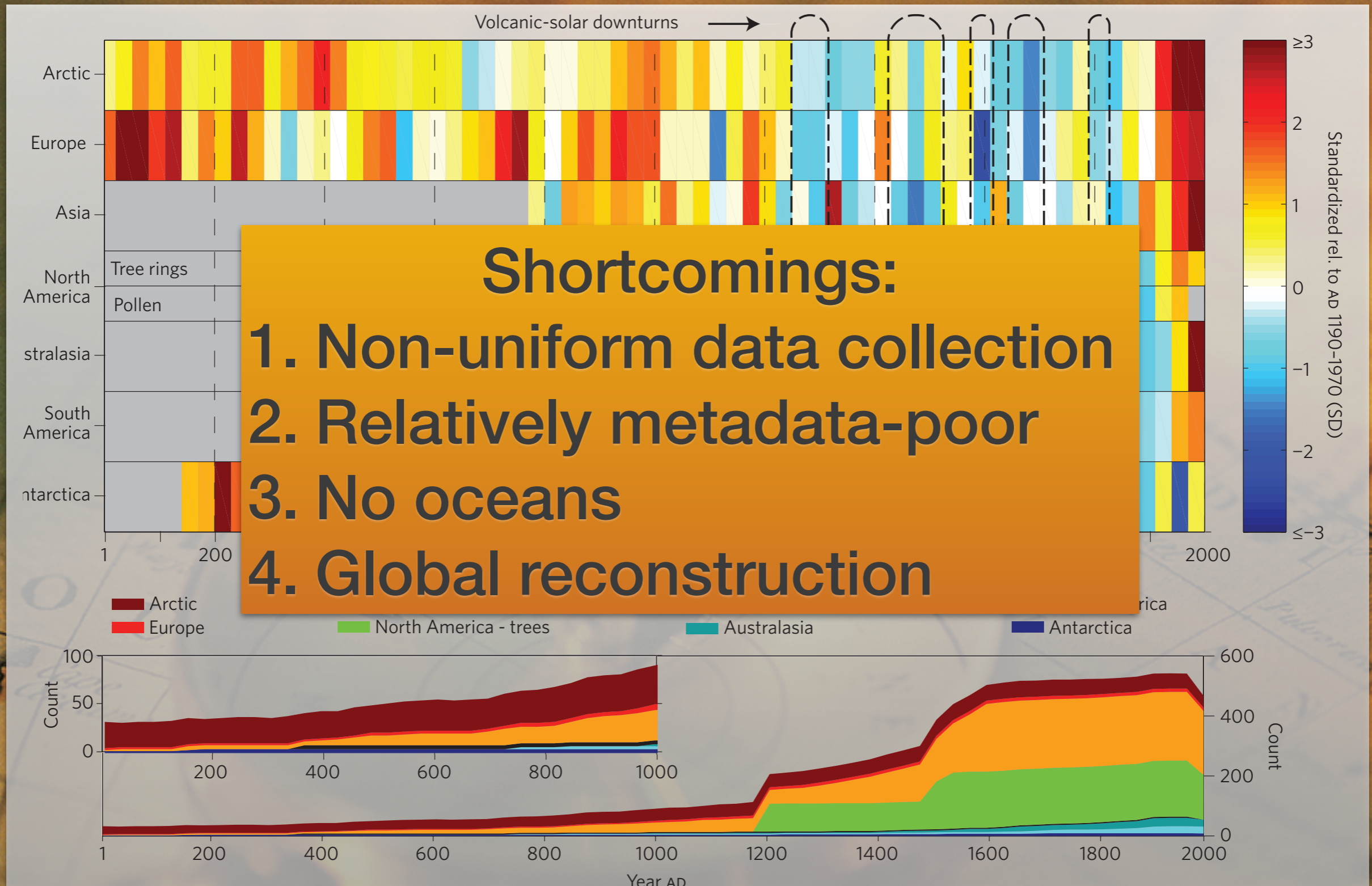
- in the 99 percentile (ranked 84th) of the 96,180 tracked articles of a similar age in all journals
- in the 98 percentile (ranked 1st) of the 74 tracked articles of a similar age in Nature Geoscience

Page views

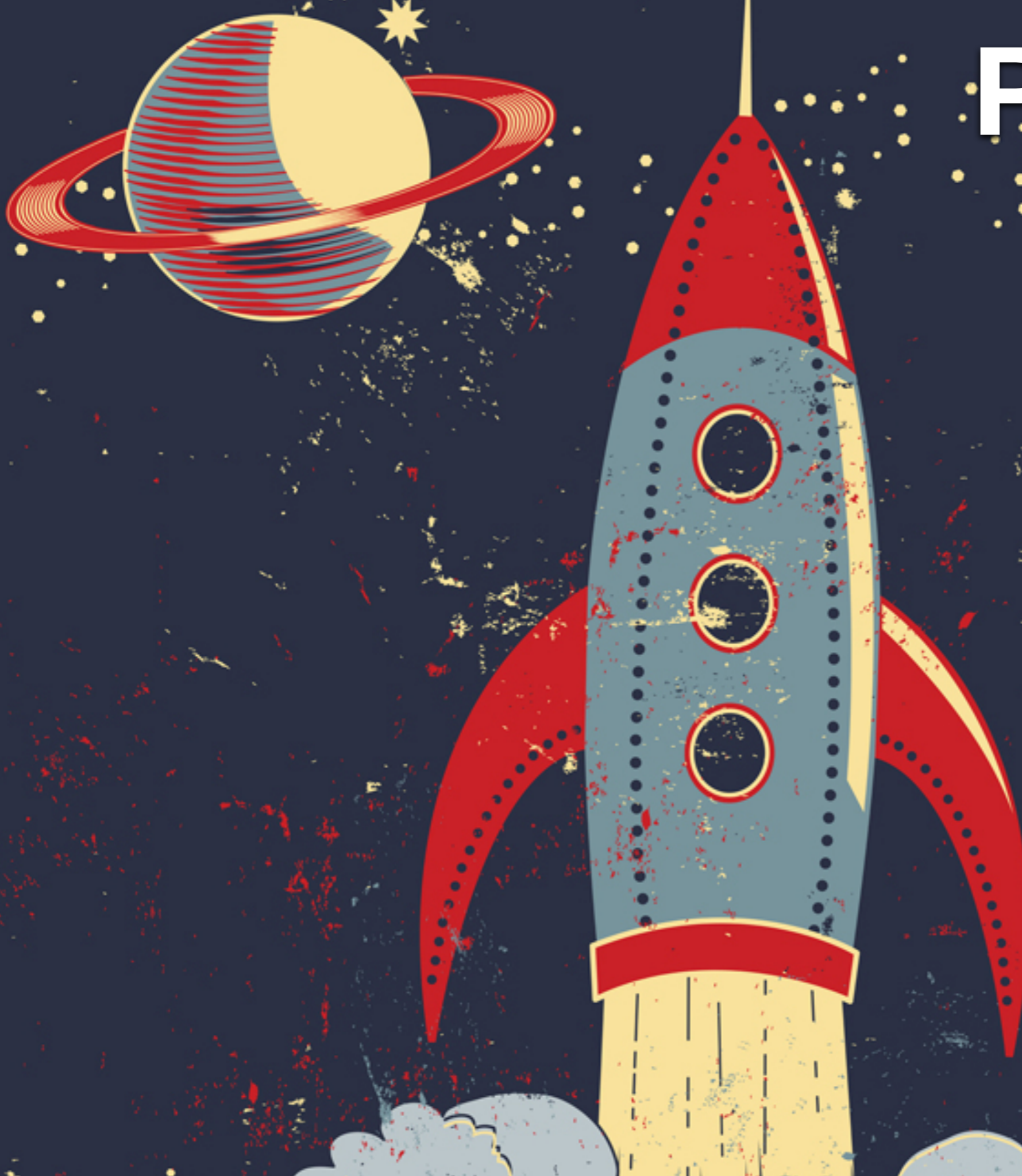
20,722



A brief history... phase 1



Phase 2



Criteria

1. Length
2. Resolution
3. Thermal sensitivity

The data

DatabaseS1-All-proxy-records.xlsx

80%

Search in Sheet

Home Layout Tables Charts SmartArt Formulas Data Review

Font: Calibri, 12, Bold, Italic, Underline, Text Color, Background Color

Alignment: Left, Center, Right, Wrap Text, Merge

Number: General, Percentage, Decimals, Fractions

Format: TableStyleLig..., Normal

Cells: Insert, Delete, Format

Themes: Themes, Aa

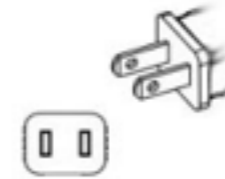
fx: -33.05272727

PAGES ID	Ant_1	Ant_2	Ant_3	Ant_4	Ant_5	Ant_6	Ant_7	Ant_8	Ant_9	Ant_10	Ant_11	M	N	O	P	Q	R	S	Y	U	V	W	
Year C.E.	Talos Dome	DSS d180	Plateau Rem	IND-22 B4 d	EDML d180	WAIS 2005A	WAIS 2006A	ITASE 00-1 d	ITASE 00-5 d	Ant_10	Ant_11		Composite of WAIS 2005A (d180 Variance Matched) and WAIS2006A Used in alternate reconstructions and site-level analyses instead of Ant_6 and Ant_7. Details in Supplementary Information										
2007		-21.66811									-158.1												
2006		-22.66469									-135.7												
2005		-22.37433				-33.32873					-146.1		-34.0948										
2004		-23.49186				-35.67324					-145.6		-34.4138										
2003		-22.9755				-33.15743					-150.6		-34.081										
2002		-22.04853				-34.2854					-152.2		-34.0435										
2001		-23.13156				-34.40308					-130		-33.9937										
2000		-23.15101				-32.89417		-31.85171	-278.8096		-152.6		-33.7916										
1999		-23.64831				-35.09511		-32.75071	-289.3724		-138.8		-33.6691										
1998		-22.59063				-31.76437		-30.83333	-277.3247		-138.5		-32.9271										
1997		-23.01687				-32.97733		-30.8615	-269.3756		-147.3		-32.7303										
1996		-21.36637				-32.70684		-31.54231	-257.6471		-142		-32.9998										
1995	-309.5081	-21.55999			-43.68122	-32.70684		-31.54231	-257.6471		-142		-32.9998										
1994	-303.9172	-22.82129		-25.06944	-45.05705	-32.51447		-31.35167	-265.7514		-143.7		-33.1723										
1993	-302.4587	-23.25602		-25.20537	-47.01522	-32.61222		-32.3935	-276.5133		-141.2		-32.6221										
1992	-294.844	-21.68464		-24.60958	-45.72084	-32.52732		-31.80222	-255.1136		-138		-32.6817										
1991	-285.526	-22.26711		-24.0049	-41.36889	-32.86738		-30.98476	-258.3692		-143.4		-33.0954										
1990	-301.262	-23.25961		-23.40022	-41.10128	-33.99008		-32.33588	-267.9857		-156		-33.69										
1989	-275.12	-22.03031		-23.2743	-40.83366	-34.24384		-32.24462	-266.5433		-126.6		-33.7703										
1988	-280.249	-21.44053		-23.66289	-43.48982	-32.91622		-33.67417	-269.7013		-149		-33.4473										
1987	-285.474	-23.1244		-24.05147	-44.54067	-33.00252		-33.87846	-273.8571		-153.5		-34.1016										
1986	-285.738	-23.14153	-53.85	-24.44006	-45.41691	-37.14629		-35.61417	-278.4777		-162.3		-35.4659										
1985	-271.042	-22.76242	-54.38354	-24.64713	-45.32409	-35.28455		-33.20688	-277.4138		-140		-35.4313										
1984	-263.499	-21.06345	-54.58971	-24.84993	-45.77673	-34.68641		-31.46304	-261.192		-144.2		-34.5241										
1983	-277.158	-21.88236	-56.24661	-25.05272	-43.09658	-33.28204		-32.85682	-272.7843		-142.8	-29.661	-33.6941										
1982	-298.116	-21.8581	-54.53449	-25.1823	-44.19635	-33.06917		-32.86167	-277.4522		-151.7	-29.436	-33.5716										
1981	-288.047	-21.59147	-55.29418	-25.22965	-45.35422	-34.77311		-31.97813	-278.3922		-142.9	-28.767	-33.8511										
1980	-290.025	-22.08254	-51.87163	-25.277	-44.86582	-32.65334		-30.944	-242.3518		-156.5	-28.24	-34.1225										
1979	-291.552	-22.69513	-52.3945	-25.32434	-43.40654	-36.07419		-33.12875	-283.925		-150.6	-30.006	-35.225										
1978	-301.62	-23.03467	-55.83737	-25.29352	-40.79817	-36.17505		-33.16125	-287.4414		-146.3	-29.999	-35.8477										
1977	-285.397	-21.20442	-57.12225	-25.25986	-47.69065	-35.87187		-34.63692	-287.77		-136.3	-30.657	-35.3464										
1976	-294.468	-22.78758	-54.84587	-25.22621	-46.36438	-33.3487		-31.97	-279.6325		-155.2	-30.273	-34.7217										
1975	-301.022	-20.64793	-56.72503	-25.25068	-44.51183	-35.83		-33.65308	-267.4		-145.9	-28.786	-35.0994										
1974	-322.219	-21.50584	-56.99115	-25.34366	-42.60509	-35.34615		-33.06545	-255.6638		-139.7	-29.964	-35.3419										
1973	-286.54	-23.3331	-56.35583	-25.43665	-43.55332	-35.21332		-33.05273	-268.665		-134.9	-28.837	-35.1179										
1972	-268.81	-22.0923	-54.87197	-25.52963	-44.38873	-34.91282		-32.66889	-260.625		-153.8	-29.435	-34.4962										
1971	-295.168	-21.92227	-55.05567	-25.28012	-43.83207	-33.07501		-32.224	-281.3571		-140.9	-29.383	-33.7062										
1970	-291.199	-22.61669	-54.83099	-25.01372	-46.20849	-33.14026		-33.01429	-266.899		-156.5	-29.305	-33.8925										
1969	-311.147	-22.91482	-54.38783	-24.74732	-44.44358	-35.77421		-34.01273	-260.3967		-142.9	-30.063	-34.8625										
1968	-284.006	-22.50173	-53.94729	-24.64786	-44.29321	-35.15153		-33.5925	-248.9556		-158.3	-30.459	-35.0388										
1967	-281.633	-22.04655	-54.15811	-24.75492	-43.47159	-34.83352		-33.3725	-268.8329		-156.2	-29.477	-34.2293										
1966	-299.281	-23.90001	-54.32483	-24.86198	-45.79374	-32.05196		-33.25214	-286.4613		-146.3	-29.927	-33.239										
1965	-304.242	-22.27008	-56.37793	-24.96904	-43.68138	-33.49183		-32.30357	-263.4743		-150.7	-30.84	-33.3389										
1964	-290.474	-22.22099	-57.20431	-24.33634	-43.4143	-33.81644		-32.09667	-261.5429		-157.7	-29.436	-34.0963										
1963	-304.006	-22.15181	-55.68049	-23.67521	-44.7377	-35.28879		-33.4025	-276.799		-143.3	-30.327	-34.9108										
1962	-309.525	-22.59197	-55.32304	-23.59355	-44.85414	-35.72318		-34.891	-275.8056		-152.2	-30.054	-34.9762										
1961	-299.887	-22.76725	-52.96992	-23.71759	-43.68287	-33.52018		-32.4275	-273.9229		-150.9	-29.281	-34.3697										

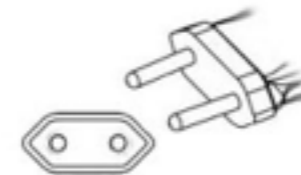
ReadMe Metadata AntProxies ArcProxies AsiaProxies AusProxies EurProxies NAMPol NAMTR SAMProxies

Normal View Ready Sum=-33.05272727

Why data standards?



US type



EU type



UK type



Euro plug

~~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

The data



Linked PaleoData

The screenshot shows a journal article page with a purple header and navigation menu. The article title is "Technical note: The Linked Paleo Data framework – a common tongue for paleoclimatology" by Nicholas P. McKay and Julien Emile-Geay. The page includes a search bar, download options for PDF and XML, and a short summary. The EGU logo is visible in the top right corner.

Climate of the Past
An interactive open access journal

EGU.eu | EGU Journals

Submit a manuscript | Manuscript tracking

Clim. Past, 12, 109–114, 2016
http://www.clim-past.net/12/109/2016/doi:10.5194/cp-12-109-2016
© Author(s) 2016. This article is distributed under the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>)

Volume 12, issue 4

Copernicus Publications
The Innovative Open Access Publisher

Metrics | Related articles

29 Apr 2016

Search articles

Search
Author

Download

PDF | FULL-TEXT XML

• Supplement (15 KB)

Short summary
The lack of accepted data

Technical note

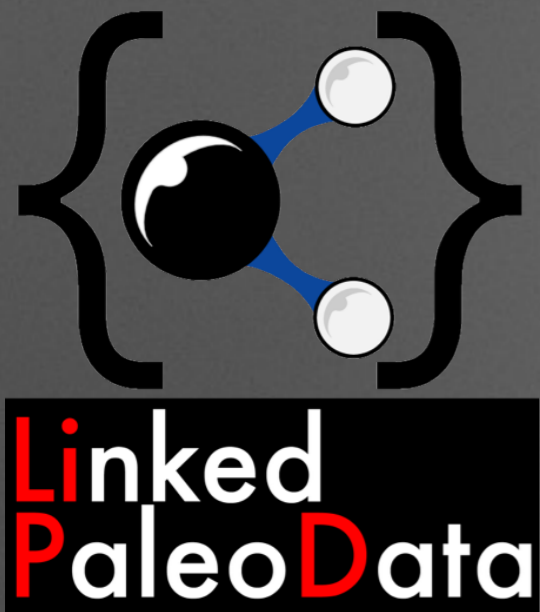
Technical note: The Linked Paleo Data framework – a common tongue for paleoclimatology

Nicholas P. McKay¹ and Julien Emile-Geay²

¹School of Earth Sciences and Environmental Sustainability, Northern Arizona University, Flagstaff, AZ 86011, USA
²Department of Earth Sciences & Center for Applied Mathematical Sciences, University of Southern California, Los Angeles, CA 90089, USA

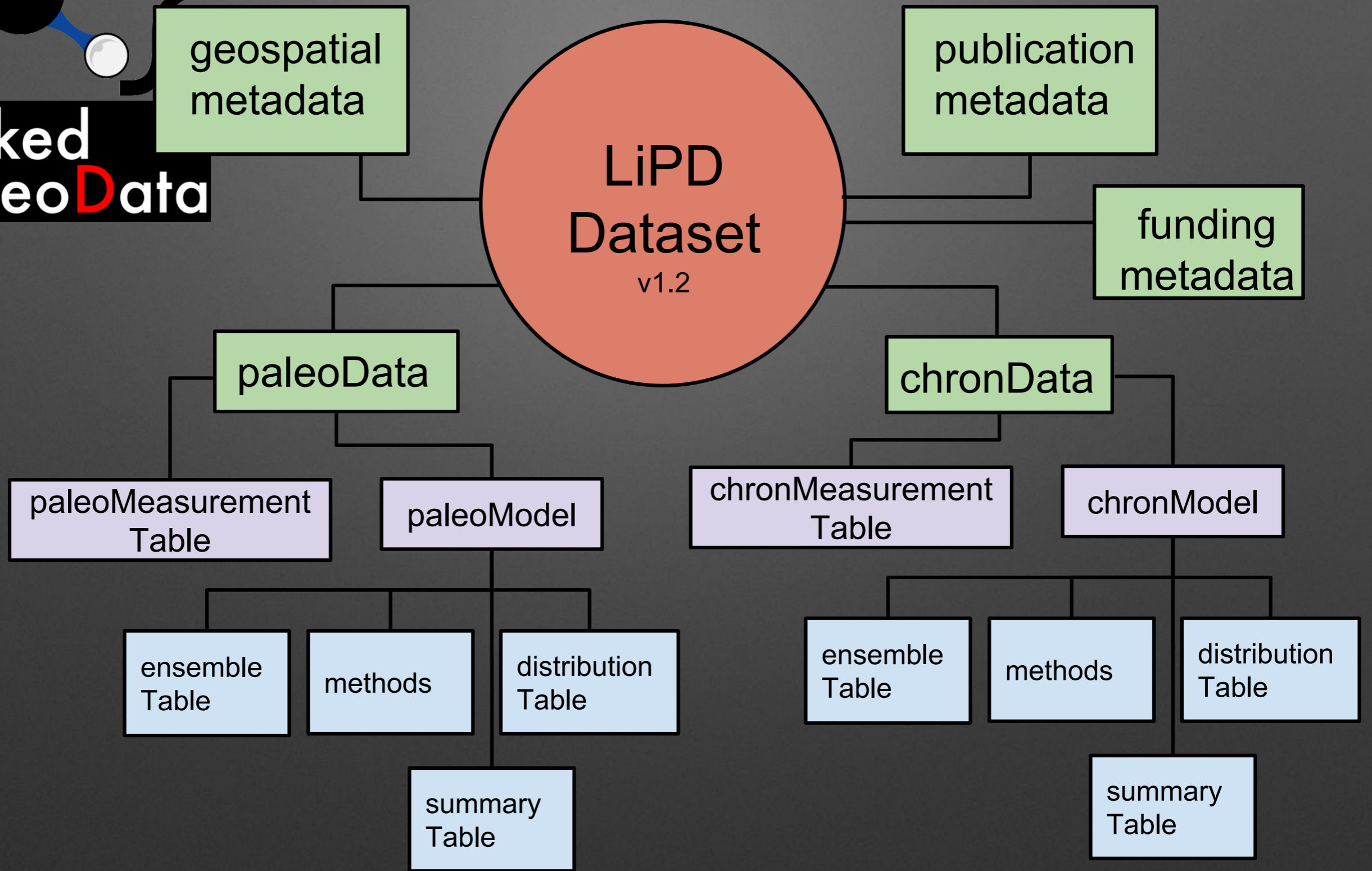
Received: 02 Jul 2015 – Published in Clim. Past Discuss.: 11 Sep 2015
Revised: 10 Mar 2016 – Accepted: 24 Mar 2016 – Published: 29 Apr 2016

Abstract. Paleoclimatology is a highly collaborative scientific endeavor, increasingly reliant on online databases for data sharing. Yet there is



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



PAGES NETWORK 2k



integration





integration



record #337: Eur-CentralEu.Buntgen.2011

archiveType:tree

PAGES2k region: Eur

Authors: B ntgen

Year: 2011, DOI: 10.1126/science.1197175

Measured Variable:trsgj

Units: NA

climateIntepretation:

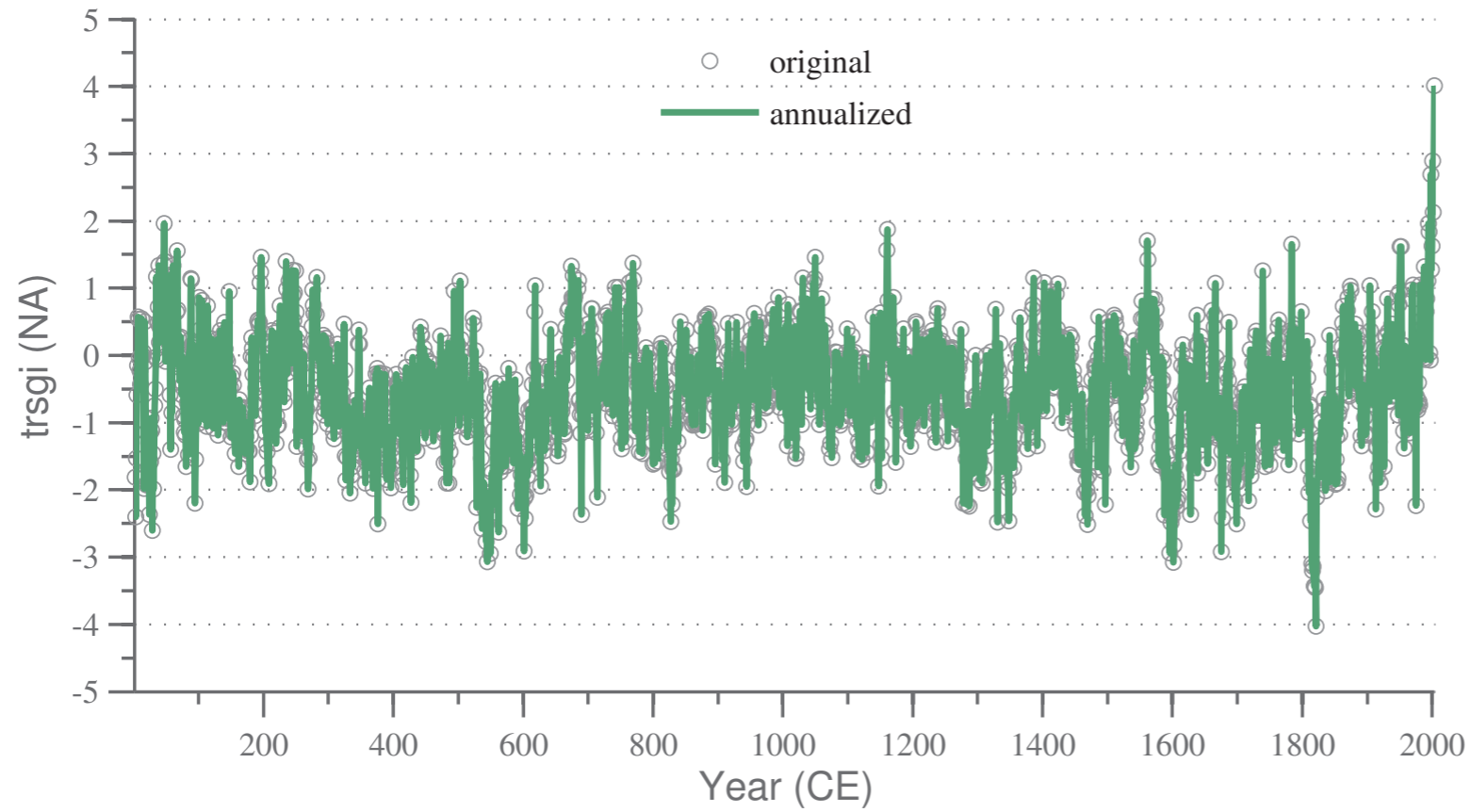
climateParameter: T

climateParameterDetail: air

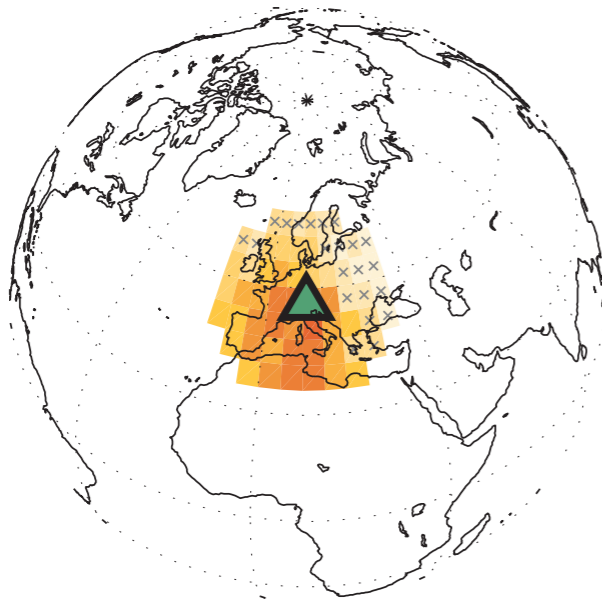
Seasonality: Summer

Direction: positive

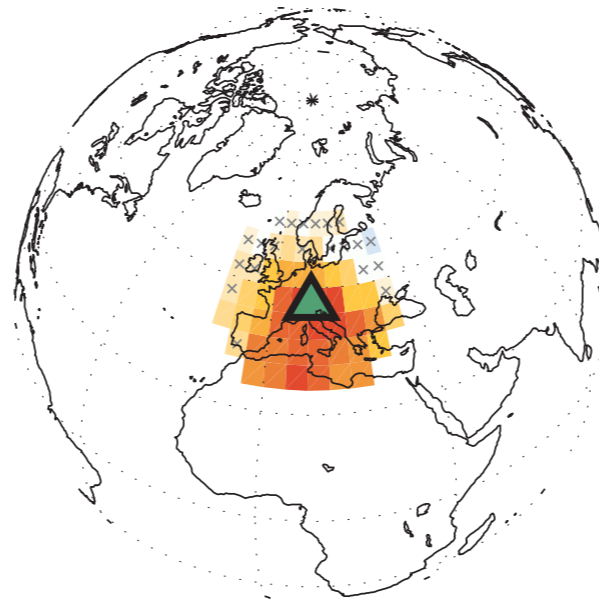
useInGlobalTemperatureAnalysis: TRUE



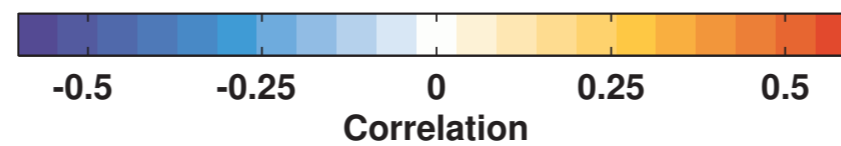
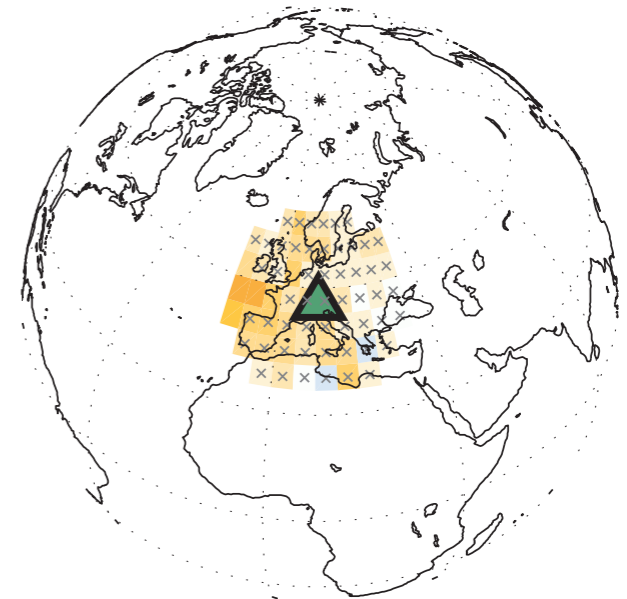
Proxy vs. MAT ($\rho_{loc} = 0.42$)



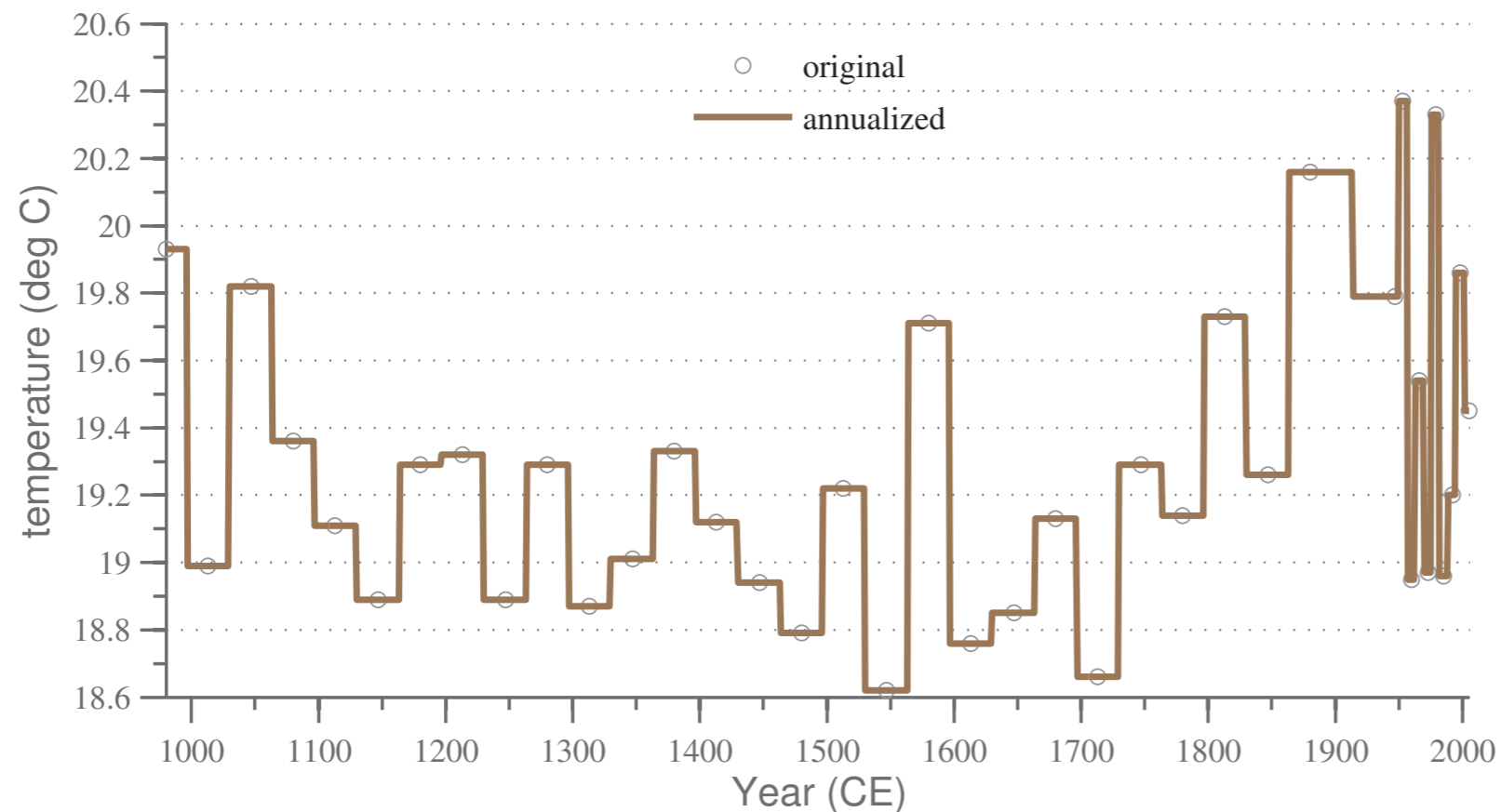
Proxy vs. JJA ($\rho_{loc} = 0.69$)



Proxy vs. DJF ($\rho_{loc} = 0.11$)



record #499: O2kLR-AlboranSea.Nieto-Moreno.2013



archiveType: **marine sediment**

PAGES2k region: O2kLR

Authors: Nieto-Moreno, V.; Mart\u00ednez-Ruiz, F.; W

Year: , DOI: 10.1016/j.orggeochem.2012.11.001

parameter: **temperature**

units: deg C

climateIntepretation:

climateParameter: T

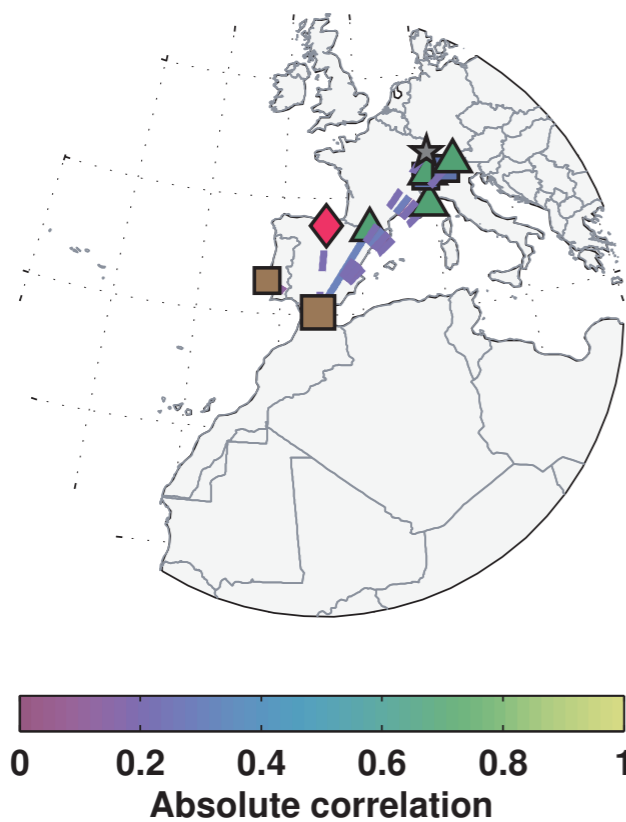
climateParameterDetail: sea_surface

Seasonality: Annual

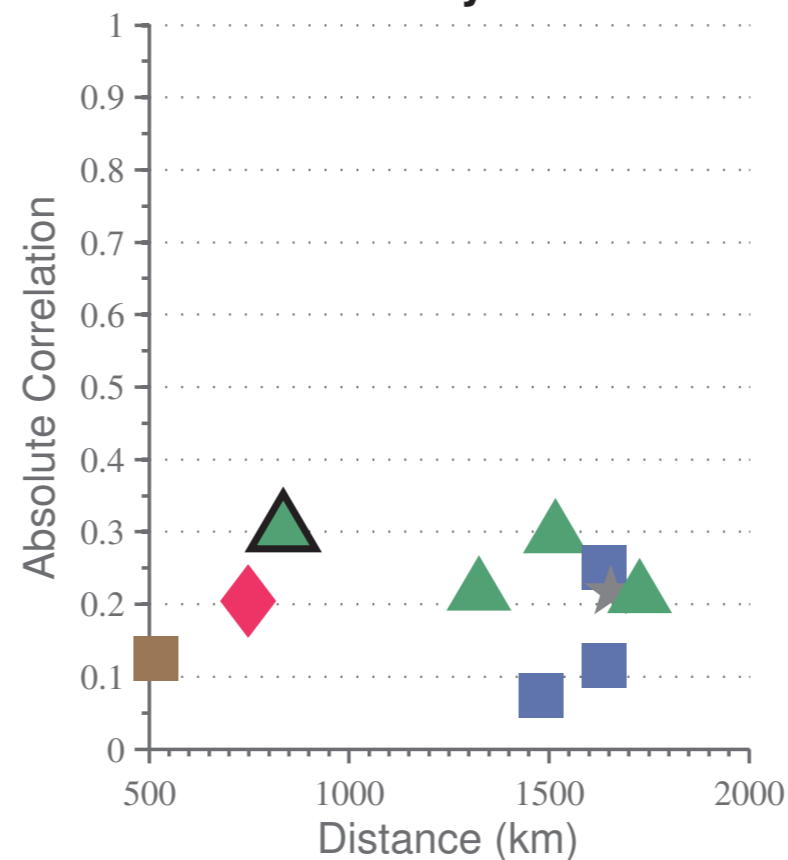
Direction: positive

QC Notes: ?

Correlation to HR neighbors



Proximity Effect



Calibration Information

calibration_equation: est. SST = (UK'37 - 0.044

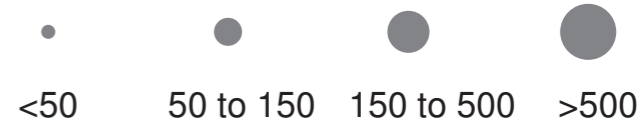
calibration_reference: doi:10.1016/S0016-7037

calibration_uncertainty: 1.5 deg C

calibration_uncertaintyType: RMSE

calibration_notes: MUE1998: T(sediments) = (U

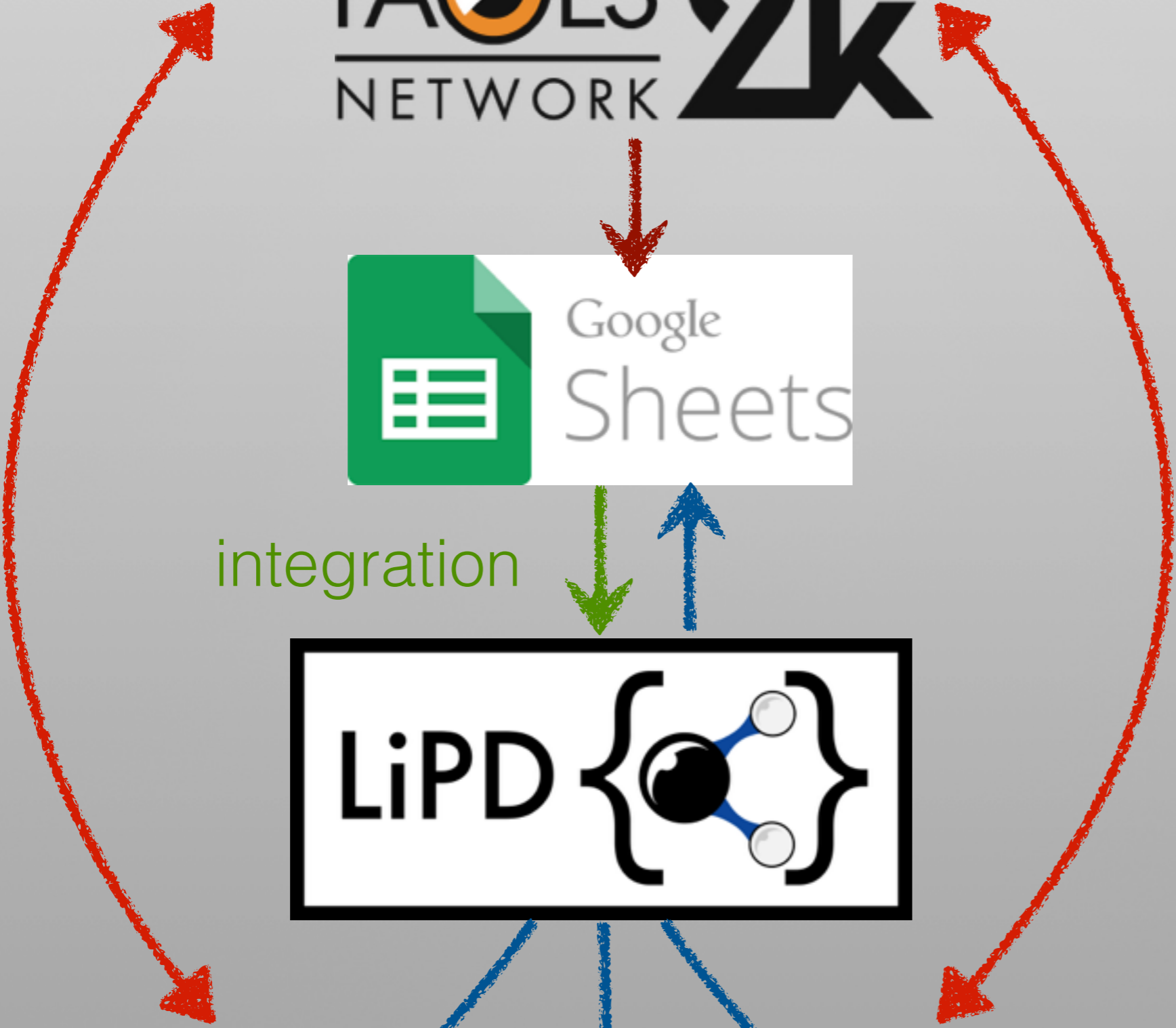
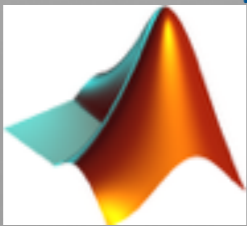
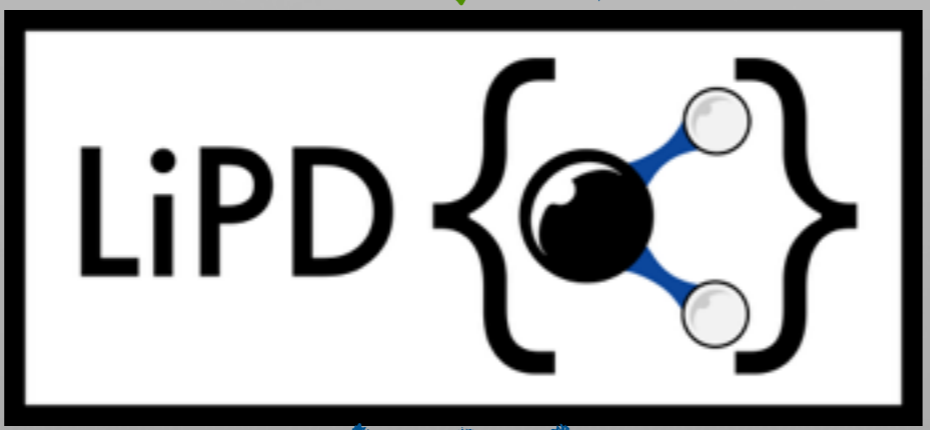
years of overlap



PAGES NETWORK 2k

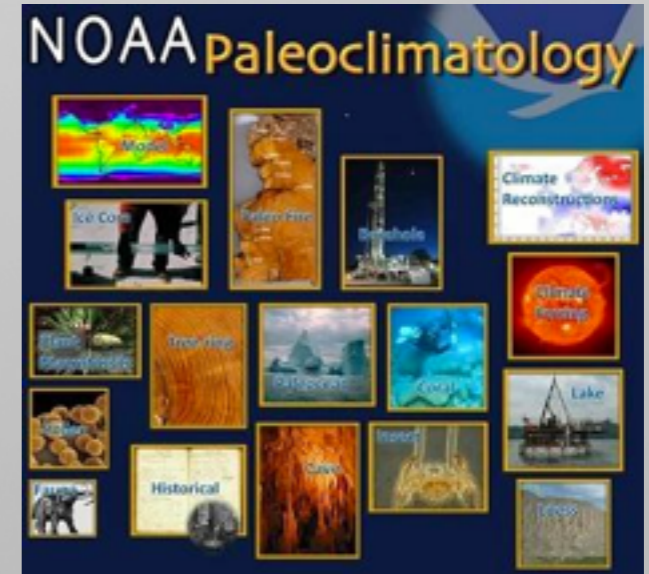


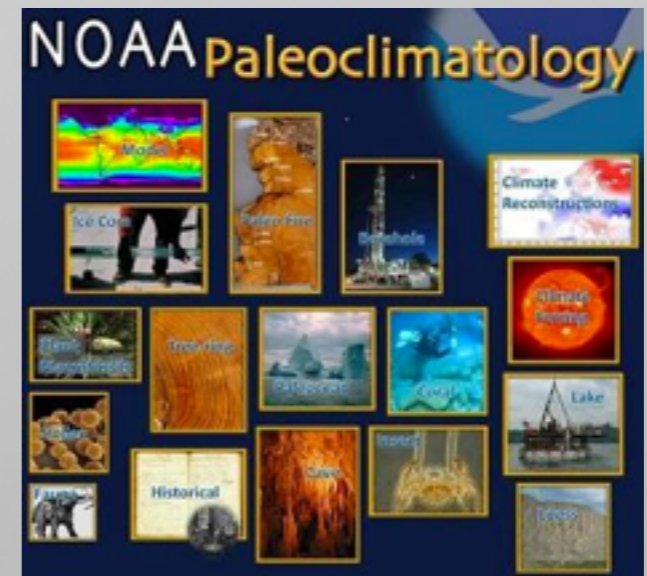
integration





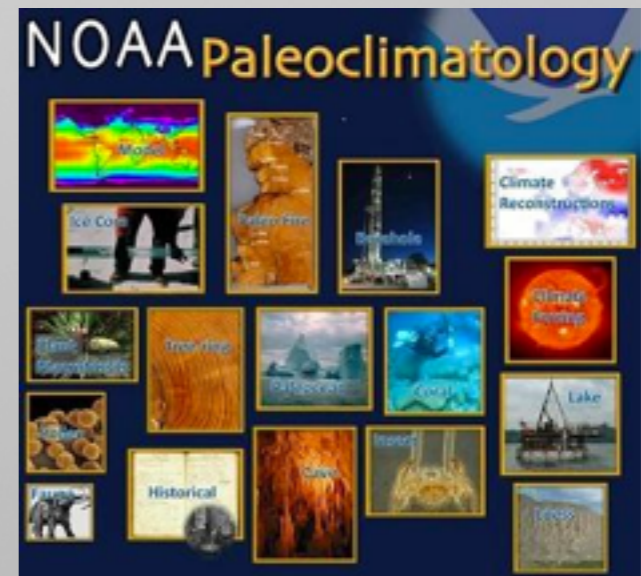
integration





integration





integration

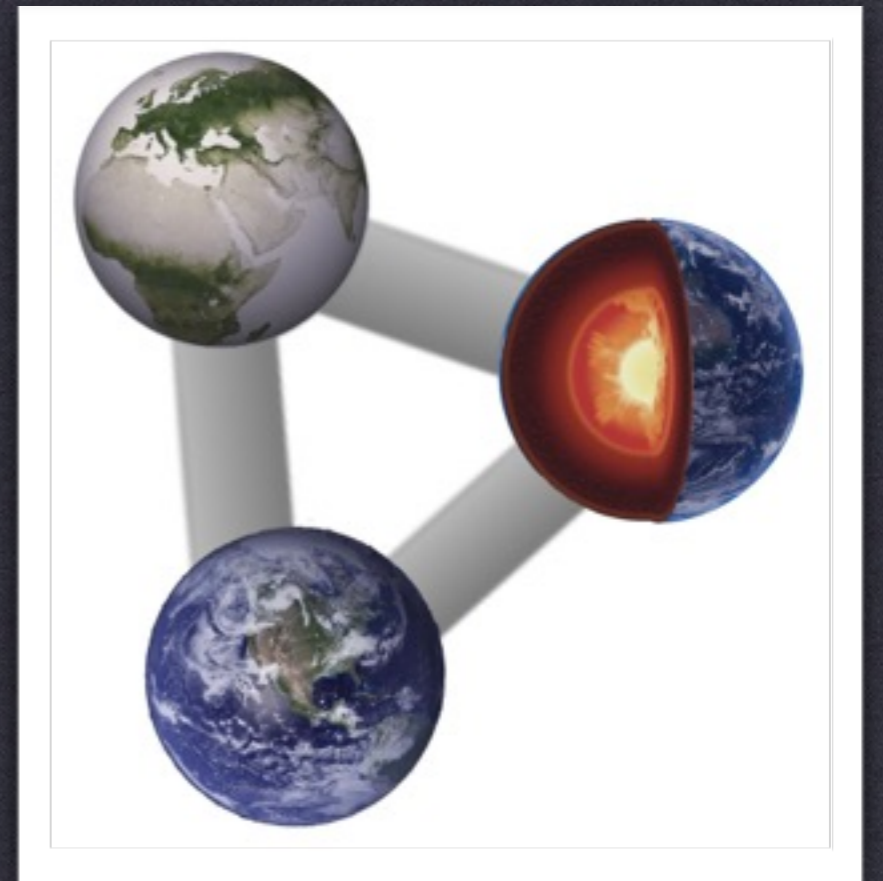


WHAT IS LINKED EARTH?

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



**Nick McKay
(NAU Earth)**



**Julien Emile-Geay
(USC Earth)**



**Yolanda Gil
(USC ISI)**



Deborah Khider (USC Earth)



(Community Liaison)

Daniel Garijo (USC ISI)



Varun Ratnakar (USC ISI)



THE TEAM

<http://earthcube.org/group/linkedeearth>



LinkedEarth

The Future of Past Climates: Linked Earth and 21st century Paleoclimatology.

1

GOALS



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



Develop social codes to process these datasets.

2

CONCEPT





Page [Discussion](#)


Arc-HalletLake.McKay.2008

- [Main page](#)
- [Recent changes](#)
- [Random page](#)
- [Help](#)

Tools





- [What links here](#)
- [Related changes](#)
- [Special pages](#)
- [Printable version](#)
- [Permanent link](#)
- [Page information](#)
- [Browse properties](#)

Standard Properties

 Dataset Name	Arc-HalletLake.McKay.2008
 Archive Type	Lake sediment
 Collection Name	HT01
 Investigators	Not defined!
 Publication	Publication.10.1029/2007GL032876
 GeoLocation	Arc-HalletLake.McKay.2008.Location
 Paleo Data	Arc-HalletLake.McKay.2008.Data.PaleoData
 Chron Data	Not defined!

Legend:  Not defined,  Valid,

Properties

 Age Units	BP
 Google Metadata Worksheet	ocqumc6
 Google Spreadsheet	https://docs.google.com/spreadsheets/d/1oQpyIJSn
 Year Units	AD

Credits

Stable carbon isotopes in trees

The source of carbon in trees is atmospheric CO_2 . The CO_2 enters the tree through 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 $^{13}\text{CO}_2$ in air has been calculated to be 4.4% less than that of $^{12}\text{CO}_2$. 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.

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

The $\delta^{13}\text{C}$ of tree ring cellulose reflects the $\delta^{13}\text{C}$ 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 $^{13}\text{CO}_2$. The $\delta^{13}\text{C}$ of cellulose is 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).

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:

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

where:

- $\delta^{13}\text{C}_{\text{cellulose}}$ is the measured carbon isotopic composition of the cellulose
- $\delta^{13}\text{C}_{\text{atm}}$ is the carbon isotopic signature of the atmosphere
- a is the diffusion fractionation associated with the transport of CO_2 through the stomata -4.4‰
- b is the fractionation associated with carboxylation, -27‰

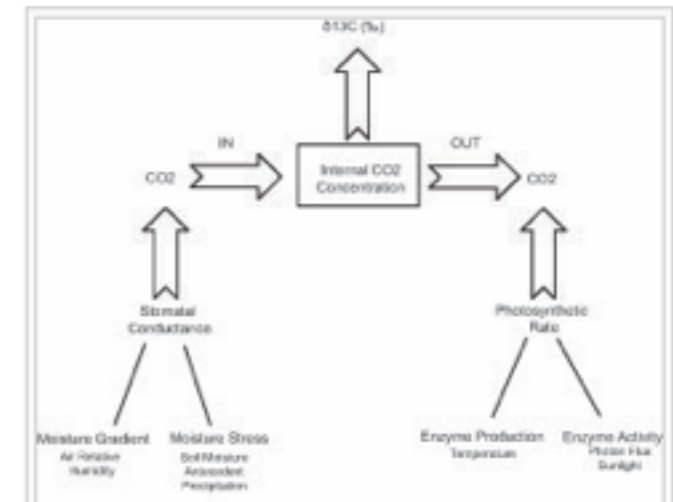
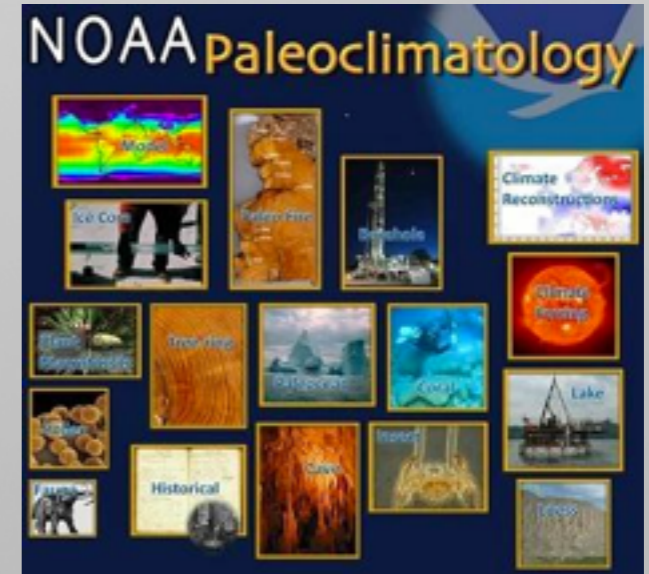
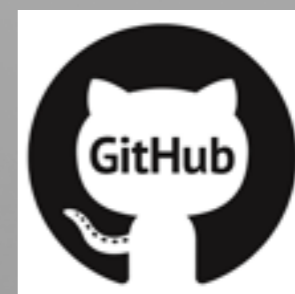


Fig. 1: Model of the influence of climate variables on the carbon isotopic composition of plant cellulose.



integration





This repository Search

Pull requests Issues Gist



CommonClimate / PAGES2k_phase2 PRIVATE

Unwatch 11 Star 0 Fork 2

Code Issues 0 Pull requests 0 Wiki Pulse Graphs

Data and code for the second phase of the PAGES2k synthesis of temperature-sensitive proxies

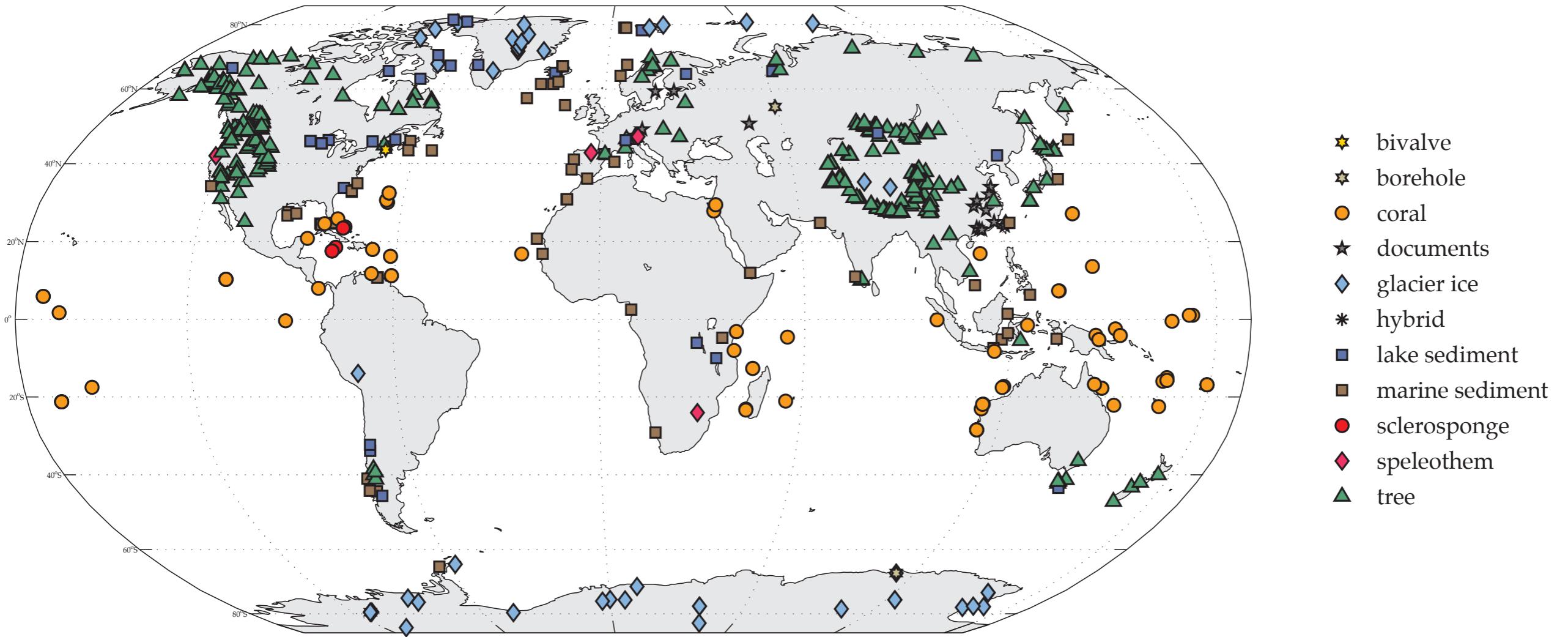
22 commits 2 branches 0 releases 2 contributors

Branch: master New pull request New file Find file HTTPS https://github.com/CommonC... Download ZIP

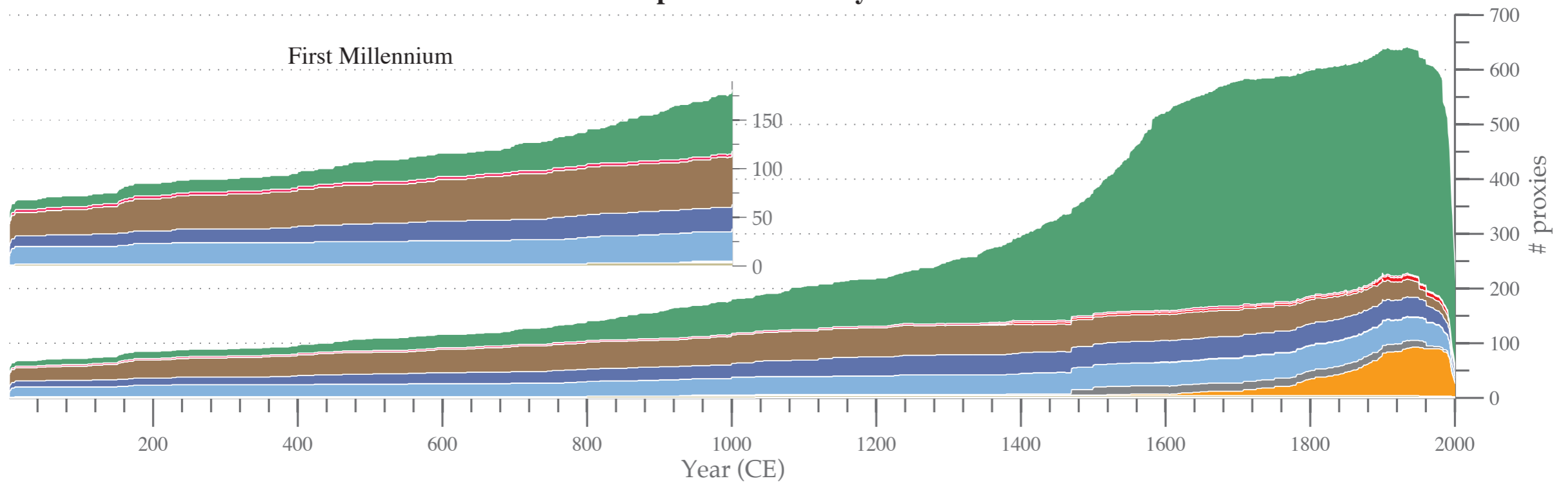
CommonClimate	added verif_stats.m	Latest commit 959124a 5 days ago
data	Made compliant with version 1.7.1	5 days ago
figs	Made compliant with version 1.7.1	5 days ago
utilities	added verif_stats.m	5 days ago
.gitignore	test of branching by Kevin	a month ago
LICENSE	Initial commit	a month ago
README.md	updated README	22 days ago
pages2k_compositeByArchive.m	Made compliant with version 1.7.1	5 days ago
pages2k_compositeByLatitudeBand.m	Made compliant with version 1.7.1	5 days ago
pages2k_composite_prep.m	Made compliant with version 1.7.1	5 days ago
pages2k_composite_recordLength.m	Updated README; added "recordLength" analysis; streamlined code	28 days ago
pages2k_composite_regression.m	Made compliant with version 1.7.1	5 days ago
pages2k_composite_sensitivity.m	Made compliant with version 1.7.1	5 days ago
pages2k_composite_workflow.m	Made compliant with version 1.7.1	5 days ago

The current state of the database (v 1.12.0)

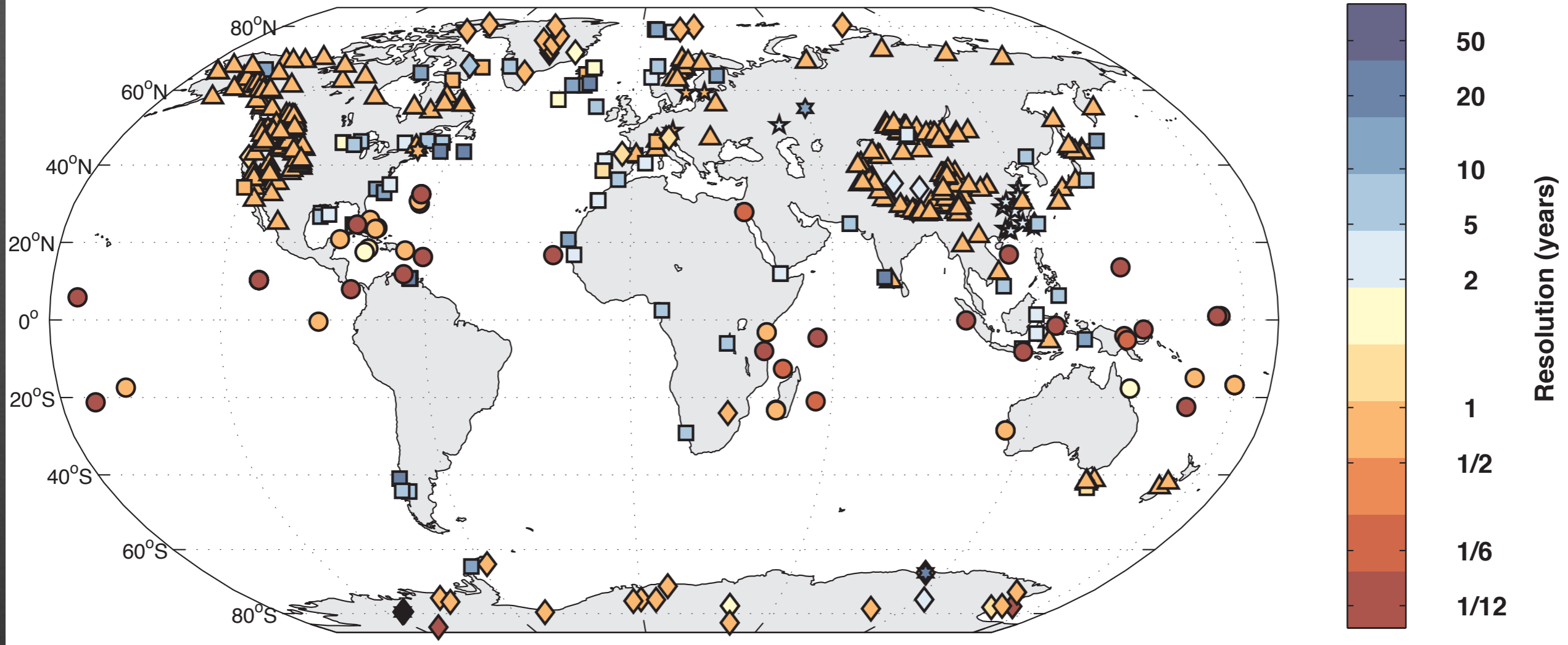
PAGES 2K network version 1.12.0 (683 records from 642 sites)



Temporal Availability

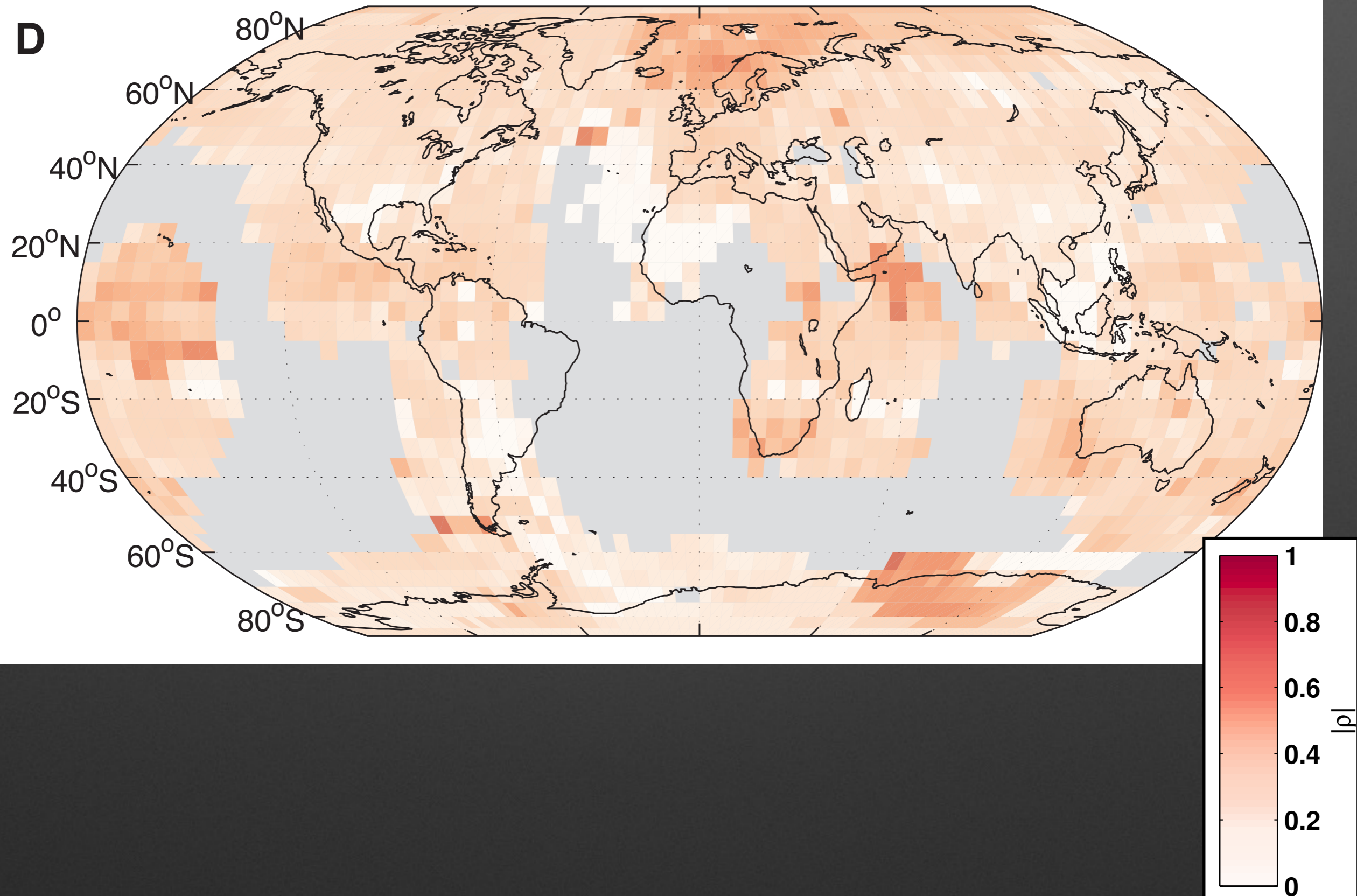


b) Resolution

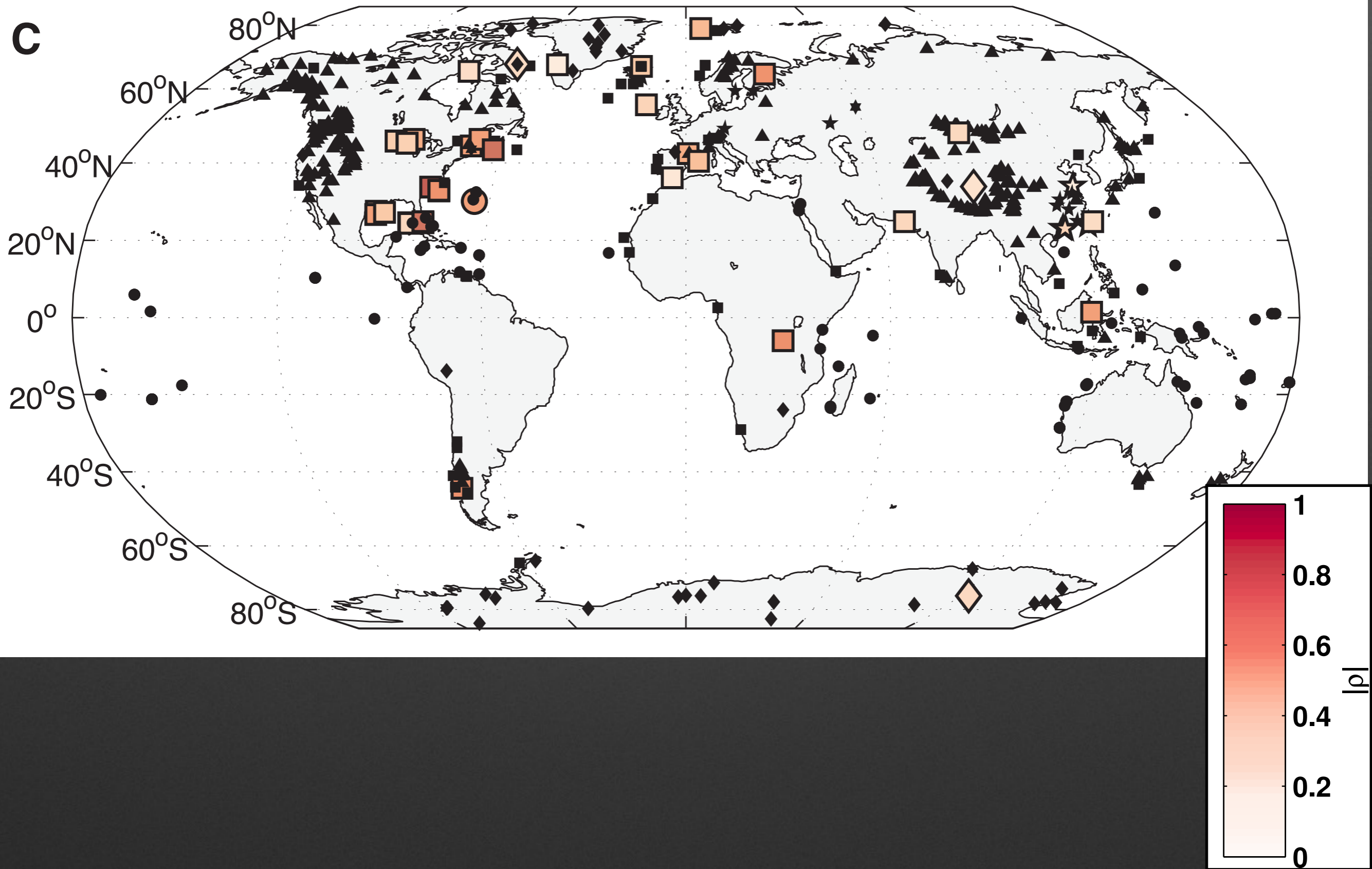


**Is it useful for understanding
past temperature?**

Does it capture the 20th century?

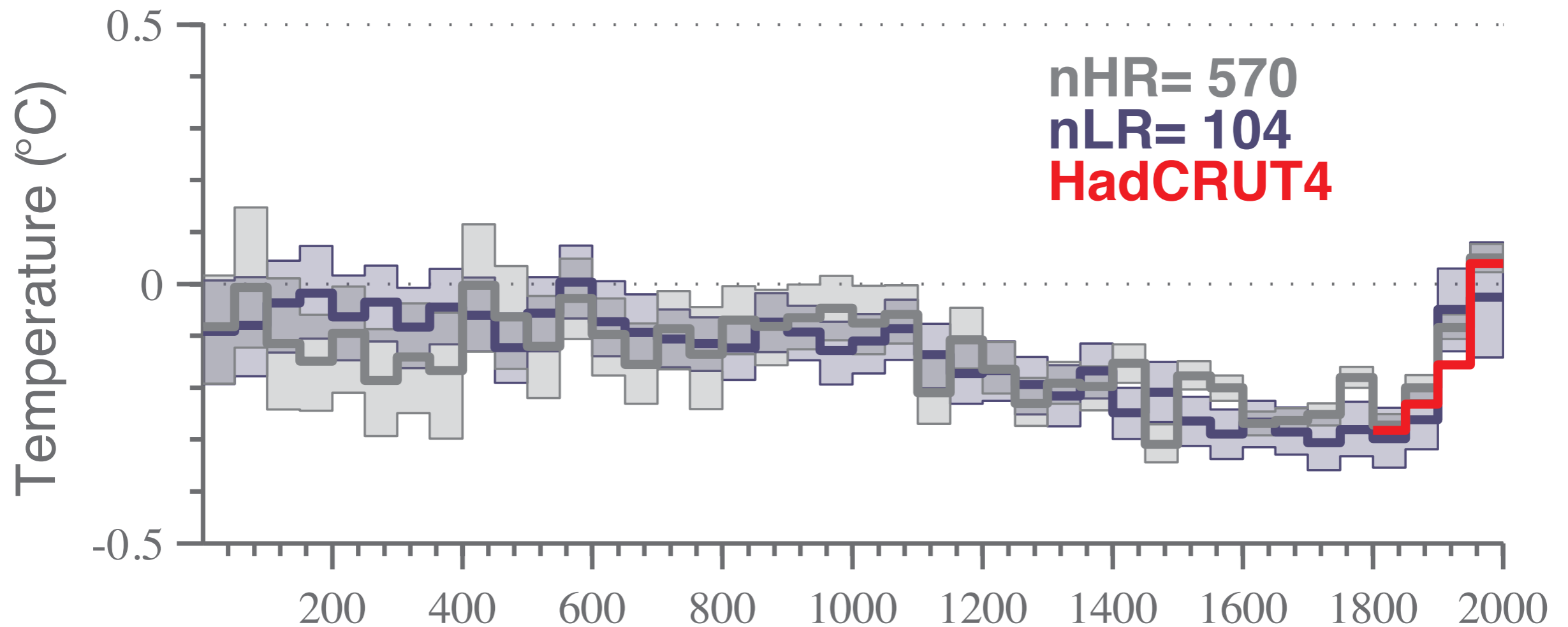


How about low-resolution records?



High vs low resolution

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



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