DOCU-CLIM: A global documentary climate data set for climate reconstructions

Angela-Maria Burgdorf[1,2], Stefan Brönnimann[1,2] and coauthors

1 Institute of Geography, University of Bern, Switzerland

2 Oeschger Centre for Climate Change Research, University of Bern, Switzerland

corresponding author: Angela-Maria Burgdorf (angela-maria.burgdorf@giub.unibe.ch)

Abstract

Documentary climate data describe evidence of past climate arising from predominantly written historical documents such as (weather) diaries, chronicles, newspapers, or logbooks. Over the past decades, historians and climatologists have generated numerous document-based time series and have used them for reconstructing local and regional climate. However, a global data set of documentary climate time series has never been compiled, and consequently documentary data are rarely used in large-scale climate reconstructions. Here, we present the first global multi-variable collection of documentary climate records. The data set DOCU-CLIM comprises 623 time series (both previously published and hitherto unpublished) providing information on temperature, precipitation, and wind. The series are evaluated by formulating forward models from monthly climate fields to reproduce the documentary time series in an overlapping period. Results show a strong climate signal particularly in the temperature series. Correlations are somewhat smaller for precipitation, which may be due to worse instrumental data, worse documentary data, or a more locally confined signal. Overall, we ascribe considerable potential to documentary records as climate proxies, especially in regions and seasons not well represented by natural proxies.

1. Background & Summary

Information on past climate has played an important role in climate science. While the main research focus was long on past annual temperature regimes, the questions that are addressed today with the help of palaeoclimatological data are multifaceted, comprising changes in the water cycle, in extremes, or in atmospheric dynamics. This in turn poses a challenge for the production of palaeoclimatic data sets. New approaches such as off-line palaeodata assimilation[1–4] provide past climate fields in increasing spatial and temporal resolution. However, all reconstructions essentially depend on sufficient input data.

Present climate field reconstructions are largely based on natural proxies such as tree rings, corals, bivalves, or ice cores for which extensive compilations of temperature sensitive proxies exist [5,6]. In particular, tree rings are widely used, among others due to their extensive spatial distribution across the globe. However, like most natural proxies, their resolution is annual, and the climate signal mainly limited to the growing season. Documentary proxies, i.e., climate data originating from historical documents, could provide an essential contribution since they cover seasons (boreal winter) and regions (e.g., East Asia) that are otherwise not well covered with natural proxies. Furthermore, documentary data are often perfectly dated, and they can have a high temporal resolution. Despite these advantages, they are largely overlooked and only marginally used in the field of large-scale climate reconstruction, among others since they are not readily available on the main archives used by climate scientists. The PAGES 2k database[7] for instance only includes a handful of documentary proxy series in their compilation, most of which have a decadal resolution. In recent years, there has been a great international effort to advance the role of proxies from the archives of society in the field of climate reconstructions. The PAGES CIRAS working group (Climate Reconstruction and Impacts from the Archives of Societies) was founded in 2018 and is working towards that goal. The Palgrave handbook of climate history[8] provided a first global overview of documentary climate data by publishing regional chapters. Based on this and other many other sources, Burgdorf [9] recently inventoried documentary record series from a literature research and a data base search. The inventory contains 688 entries, not all of which are, however, publicly available, and some have not yet been digitized. Here we publish a subset of the data inventoried in Burgdorf [9]. The DOCU-CLIM data set was initially compiled for a project with the aim of producing a novel global palaeo-reanalysis starting in 1420. For that reason, the focus is on series that provide information in the window 1400-1880 (in later periods, sufficient instrumental information is available). However, the data set could also be used for other quantitative analyses. In this paper we present the DOCU-CLIM data set (see Supplement Table 1 for an overview of all records) and evaluate the usefulness for climate reconstruction by using forward models.

2. Methods

Compilation and Data rescue

Over the past decades, historians have produced numerous data sets in which documentary data have been translated to quantitative climate information that could be useful for climate science. However, as their focus is commonly regional or local, these data are often not submitted to global data repositories such as the NOAA Palaeoclimatology website (https://www.ncei.noaa.gov/access/paleo-search/), but rather published on project or personal websites, or, unfortunately still very often, not at all. Even if data sets are incorporated into data bases, they may not always be organized in a manner suitable for climate scientists, particularly when working with time series. In this work we focus exclusively on document-based time series data, which represents a small subset of the body of documentary climate data.

We compiled documentary data from various sources, performed data rescue for others, combined the data in a common format and evaluated the data using forward models. Figure 1 illustrates the general work flow followed in this project. Methodological details on each step are given in the following.

For our work we have searched a number of compilations as listed in Table 1 (from Burgdorf [9]). The cataloged data followed a set of criteria, some of which were dictated by our own intended use (e.g., only material overlapping the period 1400-1880; a minimum record length of 30 years of which 20 must be before 1880 and therefore allowing calibration; only information on temperature, precipitation, or wind). As detailed in Burgdorf[9], the focus was predominantly on the English literature that is accessible electronically and that show significant potential for climate reconstructions. We predominantly used secondary material (i.e., derived time series or indices) to ensure the inclusion of expert source interpretation. Only for phenological data, we also searched (and digitized) original values.

After cataloguing, for obtaining the actual data, we also contacted many authors directly in cases when a data set was not available in a repository. However, we only compiled data series that are open access and allow us to redistribute the data.

Table 1. Overview of available global and national repositories and databases containing documentary evidence. N_{all} is the total number of series available on the platform, and N_{docu} is the number of series or databases based on documentary evidence available on the platform available prior to 1880 CE (from Burgdorf [9]).

Name of repository or database	Abbreviation	Region	N_{all}	N_{docu}	Reference	URL
PAGES2k Global 2,000 Year Multiproxy Database	PAGES2k	Global	692	14	Emile-Geay et al. (2017)	https://www.ncei.noaa.gov/access/paleo-search/study/21171 (last access: 30 May 2022)
NOAA/World Data Service for Paleoclimatology archives	NOAA Paleo	Global	>10000	61		https://www.ncei.noaa.gov/access/paleo-search/ (last access: 30 May 2022)
Euro-Climhist	Euro-Climhist	Switzerland/ central Europe	65	27	Pfister et al. (2017)	https://www.euroclimhist.unibe.ch/en/ (last access: 30 May 2022)
Tambora.org	Tambora.org	Germany	4	4	Riemann et al. (2015)	http://www.tambora.org (last access: 30 May 2022)
National Snow and Ice Database: Global Lake and River Ice Phenology	NSIDC	Northern Hemisphere	865	39	Benson et al. (2000), updated 2020	https://nsidc.org/data/g01377/versions/1 (last access: 30 May 2022)
Japan Climate Data Project	JCDP	Japan	14	3		https://jcdp.jp (last access: 30 May 2022)
Climatological Database for the World's Oceans	CLIWOC	Global	1624		García-Herrera et al. (2005)	https://www.historicalclimatology.com/cliwoc.html (last access: 30 May 2022)
Institute for Ocean Technology Ice Database	Ice Data	Canada	4	4		http://www.icedata.ca (last access: 30 May 2022)
KNMI Climate Explorer	Climate Explorer	Global	>200	~10		https://climexp.knmi.nl/start.cgi?id=someone@

						somewhere (last access: 30 May 2022)
Red Española de Reconstrucción Climática a Partir de Fuentes Documentales	RECLIDO	Spain	7	7		http://stream- ucm.es/ RECLIDO/es/home-es.htm (last access: 30 May 2022)
Salvá Sinobas	Salvá Sinobas	Iberian Peninsula	18	5		http://salva- sinobas.uvigo. es/index.php (last access: 30 May 2022)
Variabilidad y Reconstrucción del Clima	Vareclim	Global	5	5		https://www.upo.es/ vareclim/index.php (last access: 30 May 2022)
Reconstructed East Asian Climate Historical Encoded Series	REACHES	China	1	1	Wang et al. (2018)	https://www.ncdc.noaa. gov/paleo- search/study/ 23410 (last access: 30 May 2022)
Tracking Extremes of Meteorological Phenomena Experienced in Space and Time	TEMPEST	United Kingdom	5	5	Veale et al. (2017)	https://www.nottingham. ac.uk/research/groups/ weather-extremes/research/ tempest-database.aspx (last access: 30 May 2022)

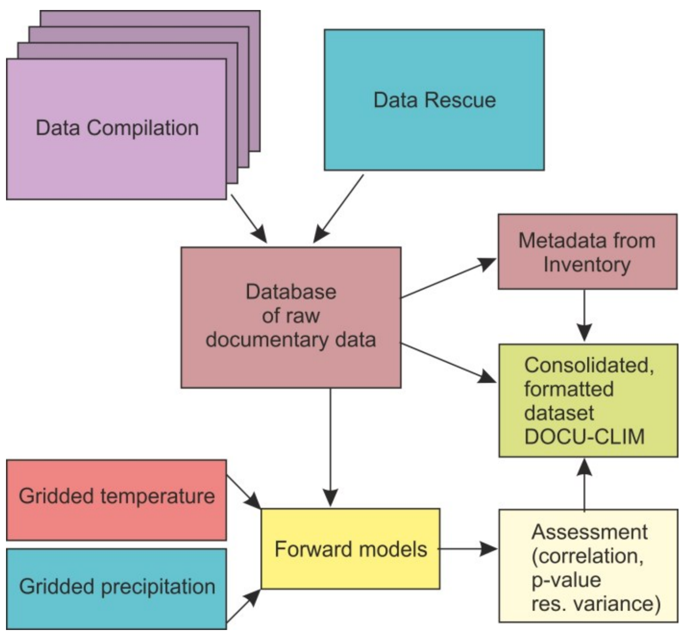
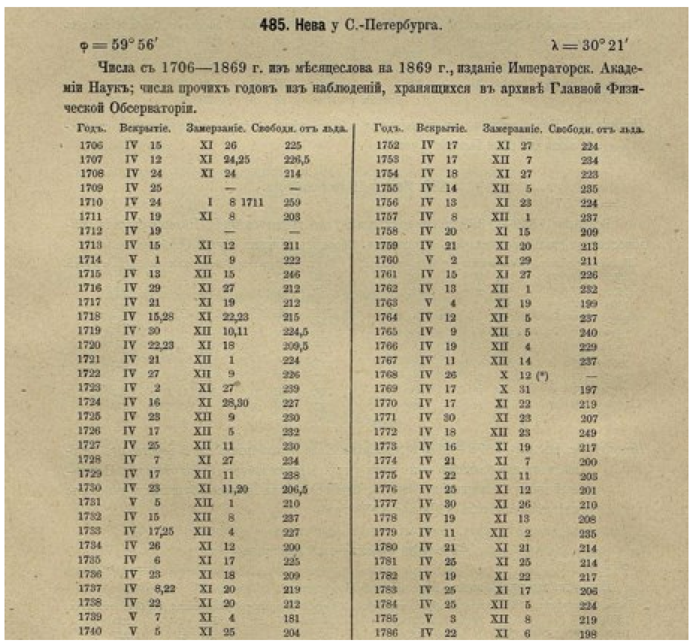


Figure 1. Flow chart for the generation of the documentary data set.

In addition to compiling existing data series, data rescue work was undertaken. Some series were digitized from original or secondary sources such as yearbooks. Others were measured from graphs published in the 1970s in cases in which the underlying data were not available electronically. Many data series digitized in the 1970s or even 1980s have not made the transition to the era of electronic publishing and open data policies. A map of the 195 time series rescued is given in Figure 2 (some of the new ice phenology records were used and described by Reichen et al. [10]).

Figure 2. Map of rescued data.



485. Нева у С.-Петербурга.

φ = 59° 56′ λ = 30° 21′

Числа съ 1706—1869 г. изъ мѣсяцеслова на 1869 г., изданіе Императорск. Академіи Наукъ; числа прочихъ годовъ изъ наблюденій, хранящихся въ архивѣ Главной Физической Обсерваторіи.

Годъ.	Вскрытіе.	Замерзаніе.	Свободн. отъ льда.	Годъ.	Вскрытіе.	Замерзаніе.	Свободн. отъ льда.
1706	IV 15	XI 26	225	1752	IV 17	XI 27	224
1707	IV 12	XI 24,25	226,5	1753	IV 17	XII 7	234
1708	IV 24	XI 24	214	1754	IV 18	XI 27	223
1709	IV 25	—	—	1755	IV 14	XII 5	235
1710	IV 24	I 8 1711	259	1756	IV 13	XI 23	224
1711	IV 19	XI 8	203	1757	IV 8	XII 1	237
1712	IV 19	—	—	1758	IV 20	XI 15	209
1713	IV 15	XI 12	211	1759	IV 21	XI 20	213
1714	V 1	XII 9	222	1760	V 2	XI 29	211
1715	IV 13	XII 15	246	1761	IV 15	XI 27	226
1716	IV 29	XI 27	212	1762	IV 13	XII 1	232
1717	IV 21	XI 19	212	1763	V 4	XI 19	199
1718	IV 15,28	XI 22,23	215	1764	IV 12	XII 5	237
1719	IV 30	XII 10,11	224,5	1765	IV 9	XII 5	240
1720	IV 22,23	XI 18	209,5	1766	IV 19	XII 4	229
1721	IV 21	XII 1	224	1767	IV 11	XII 14	237
1722	IV 27	XII 9	226	1768	IV 26	X 12 (*)	—
1723	IV 2	XI 27*	239	1769	IV 17	X 31	197
1724	IV 16	XI 28,30	227	1770	IV 17	XI 22	219
1725	IV 23	XII 9	230	1771	IV 30	XI 23	207
1726	IV 17	XII 5	232	1772	IV 18	XII 23	249
1727	IV 25	XII 11	230	1773	IV 16	XI 19	217
1728	IV 7	XI 27	234	1774	IV 21	XI 7	200
1729	IV 17	XII 11	238	1775	IV 22	XI 11	203
1730	IV 23	XI 11,20	206,5	1776	IV 25	XI 12	201
1731	V 5	XII 1	210	1777	IV 30	XI 26	210
1732	IV 15	XII 8	237	1778	IV 19	XI 13	208
1733	IV 17,25	XII 4	227	1779	IV 11	XII 2	235
1734	IV 26	XI 12	200	1780	IV 21	XI 21	214
1735	IV 6	XI 17	225	1781	IV 25	XI 25	214
1736	IV 23	XI 18	209	1782	IV 19	XI 22	217
1737	IV 8,22	XI 20	219	1783	IV 25	XI 17	206
1738	IV 22	XI 20	212	1784	IV 25	XII 5	224
1739	V 7	XI 4	181	1785	V 3	XII 8	219
1740	V 5	XI 25	204	1786	IV 22	XI 6	198

Figure 3. Excerpt from Rykachev[11] showing freezing and thawing dates for the Neva River in St. Petersburg/Leningrad.

Formatting

The data were formatted in a way that allows for straightforward ingestion into data assimilation approaches; the format is shown in Figure 4. In each file, one line covers one year. As a consequence, monthly data are stored in 12 files, one for each calendar month. The first seven entries contain information on the series, the version, and the location. These are identical for each line in the file. Then come the year, the month (only given if record resolution is monthly, then the specific month is indicated, e.g., a monthly index). Where documentary information refers to a varying time (e.g., date of freezing), the month is set to NA. The column "STATISTIC" indicates whether the observation is a state (such as a date of freezing), or a mean value (e.g., a seasonal mean index, in which case the indicated month gives the last month of the season and the column "WINDOW" the number of months averaged. The column "BOREALSEASON" indicates the closest match to a season (annual, winter (Dec-Feb), spring (Mar-May), summer (Jun-Aug), fall (Sep-Nov); for ice freezing and thawing series we additionally used "earlywinter" and "latespring"). The next columns indicate variable name, unit, and type (for all series in this paper, the type is "DOCU"), then follows the column "VALUE" that contains the actual time series values. The next seven columns indicate our attempt to fit a forward model, as is described in the next section. Finally, the last column "META" contains further information (if several, then separated by "|"). In all cases it contains the original value (which is often the same as the value itself). In the example in Fig. 4, this is the freezing date given in yr-mon-day. For further metadata on the series and collections, the reader is referred to the inventory by Burgdorf [9]. The corresponding ID is given in the column "ID_INVENTORY".

General information	ID	1213
	VERS	V1.0
	NAME	Vyatka_River_Slobodskoy_ice_phenology
	LOCATION	Slobodskoy
	LON	50.2
	LAT	58.7333
	ALT	NA
Specific time series information	YEAR	1829
	MONTH	NA
	STATISTIC	state
	WINDOW	NA
	BOREALSEASON	latespring
	VARIABLE	break_up_date
	UNIT	DOY
	TYPE	DOCU
	VALUE	122
Model evaluation	REF_PERIOD_START	1829
	REF_PERIOD_END	1879
	REFDATA	BEST
	MODEL	MarApr
	CORREL	0.486974747
	PVAL	0.001508746
	OBS_ERR_VARIANCE	36.95332339
	META	orig.value=1829-5-2

Figure 4. Data format using the example of the first line of series 1213 (transposed for better visualization).

Evaluation

To evaluate the climate signal in the documentary climate data, we formulated forward models based on global monthly climate data fields. Forward models are also directly relevant for climate reconstruction approaches such as data assimilation. We used temperature fields from BEST [12] and HadCRUT5 [13] and precipitation fields from GPCC [14] and extracted the time series from the closest grid point. The number of overlapping years must be more than 20 years for the following evaluation (for African dryness indices, as an exception, we accepted 10 overlapping years as otherwise no evaluation would have been possible on the entire data set). The forward model took the form of a multiple regression model of the corresponding variable in different months. If the season or month was specified in the source (e.g., monthly, seasonal, or annual indices), these months were taken. If no season was specified, we used the annual mean value. In case of events that were indicated as date (e.g., phenological data), the variables were determined in a backward selection approach. In this case, the models initially included 6 months prior to, and including, the event in question (defined as the 90th percentile). Then the selection proceeded backward. Variables were only kept if they were significant at $p<0.1$. Insignificant months in between two significant months were also kept. If no significant months remained, no model was calculated. Strongly skewed variables were transformed logarithmically.

We then fit regression models with a least squares estimator and used the correlation along with the p-value, as the number of data pairs varies, as a measure of fit. These two measures are plotted as global maps in the following (Fig. 7). All information on the evaluation is indicated in the example data file (Fig. 4), including the reference period used, the reference data set (BEST in this case), the model (monthly mean temperature of October and November were included in a multiple regression model; more complex models such as transformed variables would be indicated here), correlation, p-value, and error variance of the residuals. It should be kept in mind that, first, this error is not a measure of the error of the documentary proxy, but of the difference between the observation and the forward modelled observation. It thus also contains the error in the instrumental climate data and in the interpolation. However, in data assimilation approaches, this is often what is required. Second, this evaluation measures the error only in recent times when climate data are available. The quality of the documentary data in the recent period may not be comparable to that in earlier parts of the series. For some of the ice phenological records, we also have digitized nearby temperature records as the existing global data bases had no data in close vicinity. These new data have been published in Lundstad et al. (submitted) and Reichen et al.[10], and in these cases “REFDATA” is “station”.

For all proxies that do not overlap with instrumental climate fields or no long station record nearby exist, independent evaluation was not possible. This is indicated with “NA” throughout the evaluation section of the data file. It would be possible to compare these cases with reconstructions such as EKF400v2 [4]. This is a global, monthly three-dimensional climate reconstruction covering 1600-2003 and based on an off-line assimilation approach of proxy data (e.g., tree-ring width, maximum late wood density), documentary data, and early instrumental data into an ensemble of atmospheric model simulation. However, in many cases the documentary data under question were assimilated (and hence are not independent), while in cases where no information is available in the vicinity, EKF400v2 basically represents a model simulation, and no correlation is expected. We therefore use EKF400v2 only in the last Section for case studies.

Other proxies (e.g., documentary indices) continue into the instrumental era as the corresponding authors have complemented them, e.g., with degraded instrumental data. Since this data set was compiled for a specific purpose, namely the assimilation, together with instrumental data and other proxies, into a set of climate model simulations, these values were not removed in. However, for the evaluation in this paper, the calibration period in such cases is limited to years before 1900.

3. Data records

The data set DOCU-CLIM comprises 623 records, totaling more than 90,000 values. The spatial distribution is shown in Figure 5. Most of the document-based climate records are from Europe, which is partly due to our selection criteria. However, there are also records from Asia and North America. Data for Africa mostly concern precipitation[15]. We have only few documentary records from South America [16,17] and only one from Australia[18].

One of the advantages of documentary proxy data is that all seasons are covered. Plant phenology indicates temperature in spring and summer (sometimes fall), ice phenology in early winter or late spring. Many of the indices are seasonal or even monthly. Some documentary data indicate annual conditions (interpreted as annual means).

The oldest records that reach back to the 15th century are mostly from Europe, but old records are also available for China and Japan. Some records from South America reach back to the 16th century, those from North America to the 18th century, while the bulk of the African records starts in 1800. This is also the approximate starting date for many ice phenological records. Oldest records are typically also the longest, as many continue to the start of instrumental observations or even beyond.

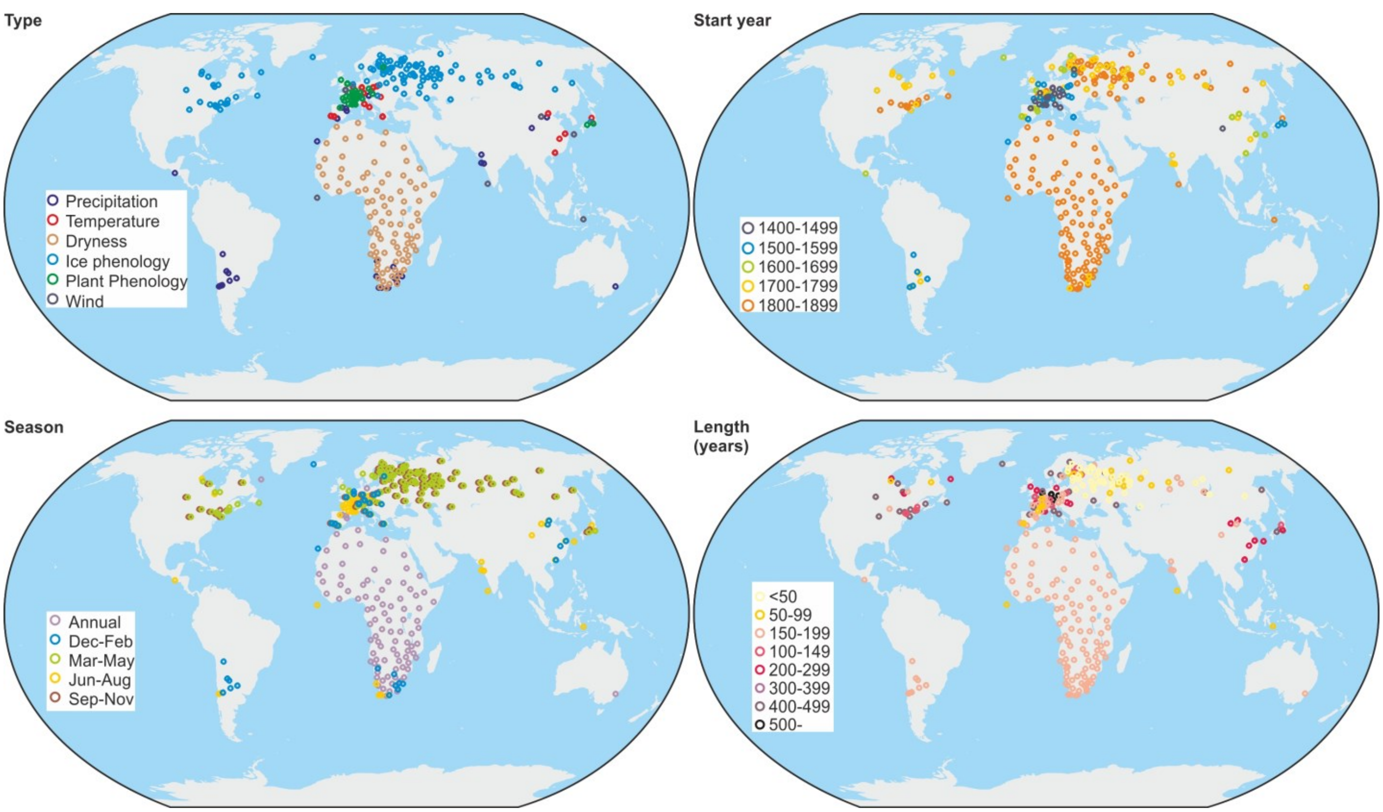


Figure 5. Map of all documentary data categorized according to (top left) the variable, (top right) the start year, (bottom left) the season covered, and (bottom right) the length.

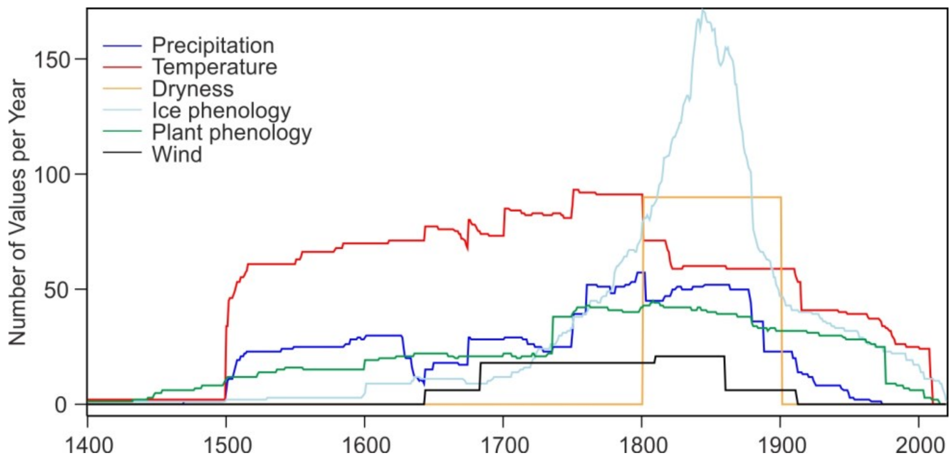


Figure 6. Number of values in the DOCU-CLIM data set as a function of year and type of proxy.

In terms of temporal coverage, Fig. 6 shows that many of the records start after 1500, the peak is in the late 19th century (note that from 1880 onward, no new records were added, although a large volume of phenological series that start later would exist). Temperature and precipitation indices are the most frequent types of record, but except for wind [19–22], all proxy types are relatively frequent. In the 19th century, when many weather stations were already measuring temperature in Europe, the number of documentary temperature series decreases and ice phenological records dominate.

4. Technical validation

For many of the series, a technical validation was performed by the original authors. These series include arguably the most specific, most expert information, combining the local knowledge both on the historical sources and on the local climate characteristics. Authors are encouraged to consult the original publication (references are always indicated, see Supplementary material for extensive information) for specific details.

In the following, we report on the results of our own technical validation, as described in Section 2 above. In Figure 7 we show correlations for forward models that we calibrated in gridded instrumental-based data sets. All records that have no overlap with observations (and thus evaluation was not possible) are shown here as white circles. This concerns 151 records.

High to very high correlations are found over Europe, North America and Asia. Many of them concern indices, or plant or ice phenology. Somewhat lower but still highly significant correlations are found over South America. Mixed correlations are found over Africa, where most of the series were evaluated based on only 10 years of overlap. Moreover, the precipitation data set[14] with which the data were compared may have large errors in these pre-1900 years. Still, several significant correlations can be found. Over Southern Africa, numerous precipitation indices show significant correlations.

A histogram of the correlations shows that the overwhelming number of series exhibits correlations above 0.5, the peak in the distribution is at correlations between 0.7 and 0.9. This is higher than the correlations found for (more complex) forward modelling of tree rings[6]. The highest correlations are typically found for temperature and precipitation indices, but also ice phenological data and grape harvest dates typically have high correlations.

In all, the evaluation of the 446 records with valid models demonstrates that many of the documentary series have great potential for science applications. However, the series that were not subject to evaluation (due to missing instrumental series in proximity or missing overlap altogether) might require further examination before usages in a climate reconstruction context.

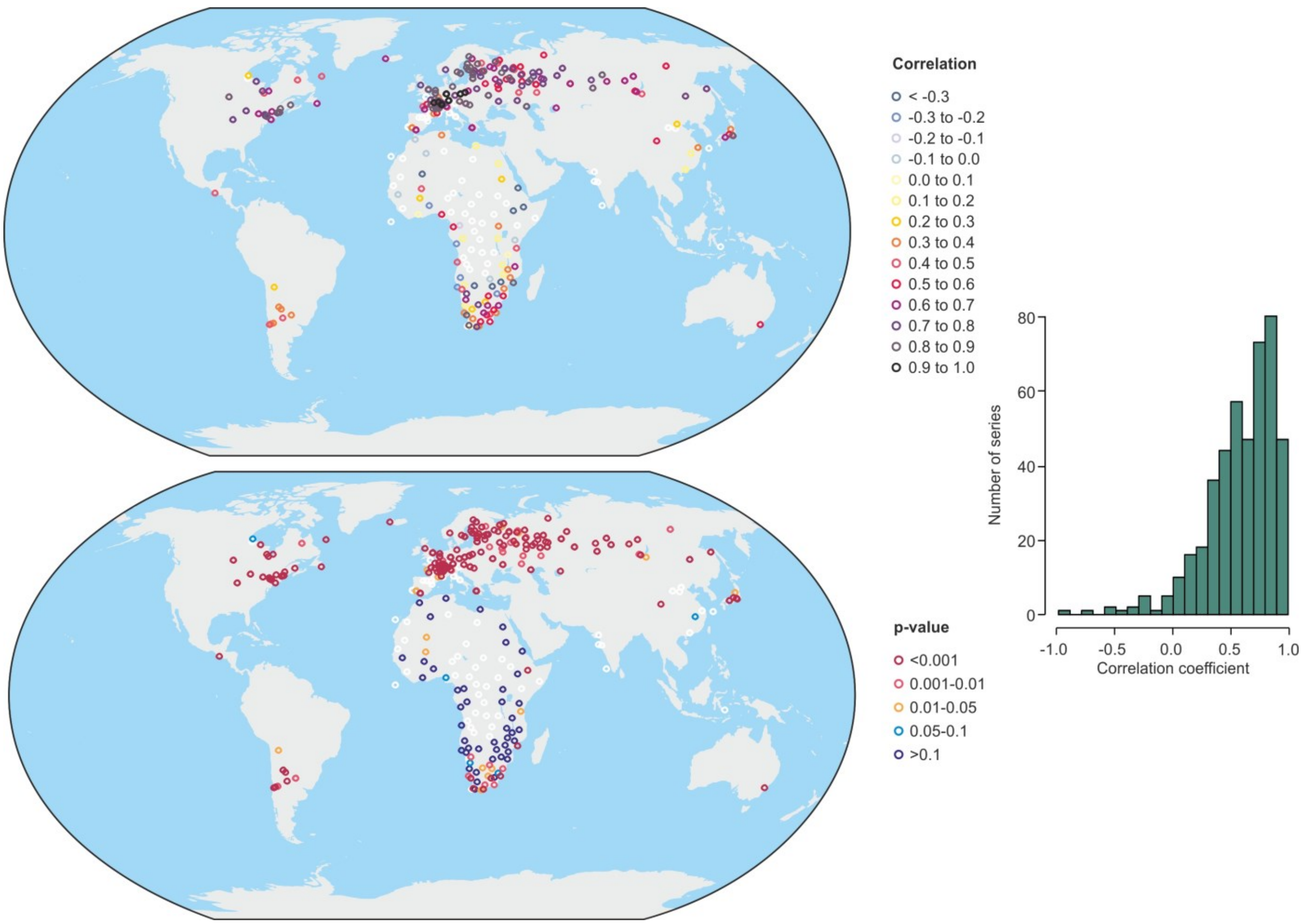


Figure 7. Map of correlations coefficients (top), p-values (bottom) and histogram showing the distribution of the correlation coefficients. White circles indicate series where no evaluation was performed.

Finally, a small case study of the year 1835 is presented in order to elucidate the potential of this collection for climate studies. In January 1835, the volcano Cosigüina in Nicaragua erupted and released massive amounts of sulfuric aerosols into the atmosphere. It is considered one of the largest historical volcanic eruption in the Americas and lead to wide-reaching environmental impacts[23]. We therefore investigate temperature-related series as well as precipitation or dryness-related series for the year 1835. For temperature, we distinguished two seasons: spring to summer (March to July), and fall and early winter (August to December), as this fits best with what the documentary data cover (thawing dates and spring/summer phenology, freezing dates and fall phenology). Monthly series were averaged to these seasons. For precipitation and dryness indices, we considered all season (sub-annual seasons were averaged to annual means). As a reference we chose the period 1841-1870 (more than 20 years must have data) and standardized the series with respect to this reference. Finally, the sign of series was adjusted such that, for temperature, positive means warmer (e.g., spring flowering or thawing dates were multiplied with -1 as earlier dates indicate warming; the sign of freezing dates was kept as early freezing means low temperatures). We then compare these anomalies to the EKF400v2 reanalysis where we performed the same procedure with global monthly fields. Seasonal and annual averages were formed, and the fields for 1835 were expressed as standardized anomalies from the period 1841-1870. Figure 8 shows the standardized anomalies of the documentary proxies (top row) and EKF400v2 (bottom row) for the year 1835. The two sets of data are entirely independent, and they are plotted on the same scale.

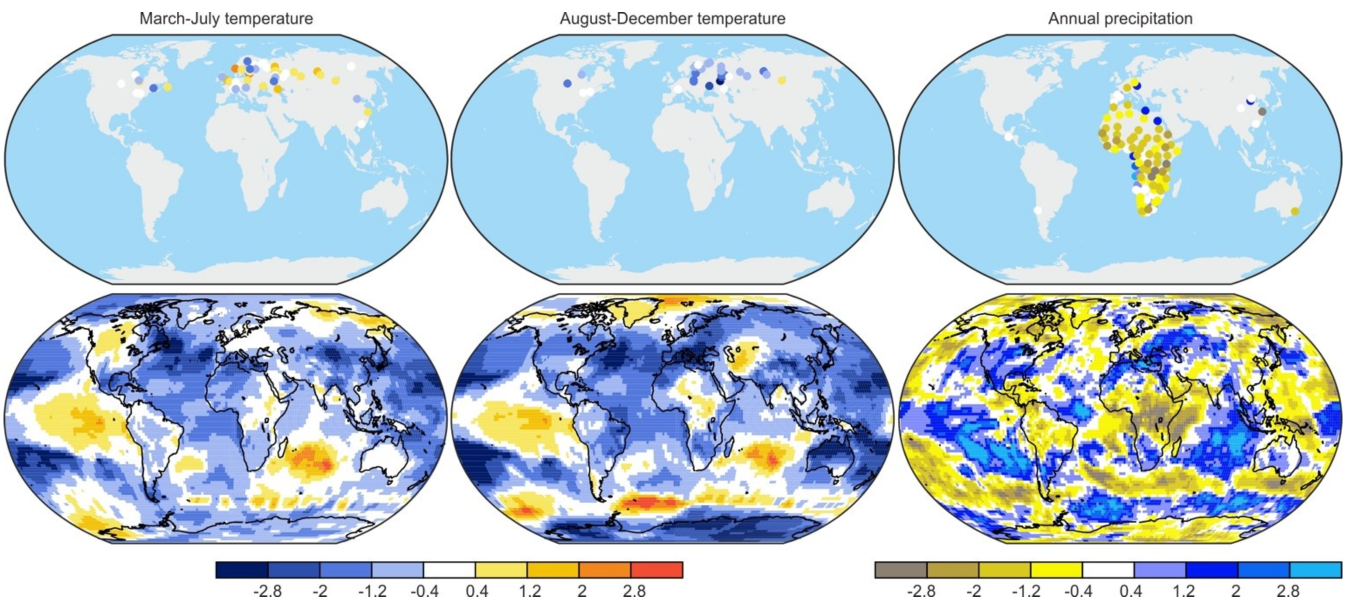


Figure 8. Standardized climatic anomalies in 1835 (with respect to the period 1841-1870) for (top) documentary data and (bottom) EKF400v2 for temperature in (left) spring to summer and (middle) fall to early winter and (right) for precipitation (all seasons or annual means considered).

EKF400v2 shows a general cooling that is arguably related to the volcanic eruption. However, not all regions cool in all seasons. In spring and summer, Europe and northern Eurasia have standardized anomalies around zero, in some regions even positive. Although EKF400v2 has no information from Siberian rivers except for the Angara in Irkutsk, these also show neutral or slightly warm conditions. In fall and early winter, the temperature proxies suggest a general cooling across the northern midlatitudes, particularly the ice phenological data (early freezing of rivers). This corresponds well with the EKF400v2 anomaly fields. Finally, precipitation in EKF400v2 indicates drying in most parts of Africa, which is also found in the documentary data in Africa (none of which were assimilated into EKF400v2), and wetting around the Mediterranean, which is in agreement with some documentary series but not all. In all, our analysis shows that spatiotemporal climatic signals are captured in the data set of documentary data (DOCU-CLIM). Furthermore, the large-scale cooling as well as the drying of African monsoon regions is in agreement with the expected effects of a tropical volcanic eruption[24].

Usage note

The data set DOCU-CLIM can be used for climate reconstruction, particularly for data assimilation (which can make full use of the data and the metadata provided on the forward modeling, although it is recommended to recalibrate the forward models). Some series could not be validated or only with EKF400 and should be further analyzed. DOCU-CLIM provides a global view and can be combined with other proxies, e.g., from PAGES 2k[5]. Care should be taken when evaluating the series for trends. We have not analyzed the suitability of the records for trend analyses and advise to test this further before using the data set for trend analyses.

Code availability

R code attached

Acknowledgements

This work was supported by the European Commission (ERC Grant PALAEO-RA, 787574) and by Swiss National Science Foundation project WeaR (188701). Simulations underlying EKF400v2 were performed at the Swiss National Supercomputing Centre CSCS.

References

1. Bhend, J., Franke, J., Folini, D., Wild, M. & Brönnimann, S. An ensemble-based approach to climate reconstructions. *Clim. Past* **8**, 963–976 (2012).
2. Hakim, G. J. *et al.* The last millennium climate reanalysis project: Framework and first results. *J. Geophys. Res. Atmos.* **121**, 6745–6764 (2016).
3. Tardif, R. *et al.* Last Millennium Reanalysis with an expanded proxy database and seasonal proxy modeling. *Clim. Past* **15**, 1251–1273 (2019).
4. Valler, V., Franke, J., Brugnara, Y. & Brönnimann, S. An updated global atmospheric paleo-reanalysis covering the last 400 years. *Geosci. Data J.* **00**, 1–19 (2021).
5. Emile-Geay, J. *et al.* A global multiproxy database for temperature reconstructions of the Common Era.

Sci. Data **4**, 170088 (2017).

6. Breitenmoser, P., Brönnimann, S. & Frank, D. Forward modelling of tree-ring width and comparison with a global network of tree-ring chronologies. *Clim. Past* **10**, 437–449 (2014).
7. A global multiproxy database for temperature reconstructions of the Common Era. *Sci. Data* **4**, 170088 (2017).
8. *The Palgrave Handbook of Climate History*. (Palgrave Macmillan UK, 2018). doi:10.1057/978-1-137-43020-5.
9. Burgdorf, A.-M. A global inventory of quantitative documentary evidence related to climate since the 15th century. *Clim. Past* **18**, 1407–1428 (2022).
10. Reichen, L. *et al.* A decade of cold Eurasian winters reconstructed for the early 19th century. *Nat. Commun.* **13**, 1–9 (2022).
11. Rykachev, M. *Openings and freezings of rivers in the Russian Empire*. (тип. Имп. Акад. наук, 1886).
12. Rohde, R. *et al.* A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011. *Geoinformatics Geostatistics An Overv.* **01**, 1–7 (2013).
13. Morice, C. P. *et al.* An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set. *J. Geophys. Res. Atmos.* **126**, e2019JD032361 (2021).
14. Schneider, U. *et al.* Evaluating the Hydrological Cycle over Land Using the Newly-Corrected Precipitation Climatology from the Global Precipitation Climatology Centre (GPCC). *Atmos. 2017, Vol. 8, Page 52* **8**, 52 (2017).
15. Nicholson, S. E., Klotter, D. & Dezfuli, A. K. Spatial reconstruction of semi-quantitative precipitation fields over Africa during the nineteenth century from documentary evidence and gauge data. *Quat. Res. (United States)* **78**, 13–23 (2012).
16. Prieto, M. del R. & García Herrera, R. Documentary sources from South America: Potential for climate reconstruction. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **281**, 196–209 (2009).
17. Neukom, R. *et al.* An extended network of documentary data from South America and its potential for quantitative precipitation reconstructions back to the 16th century. *Geophys. Res. Lett.* **36**, L12703 (2009).
18. Fenby, C. & Gergis, J. L. Rainfall variations in south-eastern Australia part 1: Consolidating evidence from pre-instrumental documentary sources, 1788-1860. *Int. J. Climatol.* **33**, 2956–2972 (2013).
19. Barriopedro, D. *et al.* Witnessing North Atlantic westerlies variability from ships' logbooks (1685-2008). *Clim. Dyn.* **43**, 939–955 (2014).
20. Gallego, D., Ordóñez, P., Ribera, P., Peña-Ortiz, C. & García-Herrera, R. An instrumental index of the West African Monsoon back to the nineteenth century. *Q. J. R. Meteorol. Soc.* **141**, 3166–3176 (2015).
21. Gallego, D., García-Herrera, R., Peña-Ortiz, C. & Ribera, P. The steady enhancement of the Australian Summer Monsoon in the last 200 years. *Sci. Rep.* **7**, 1–7 (2017).

22. Vega, I. *et al.* Reconstructing the western North Pacific summer monsoon since the late nineteenth century. *J. Clim.* **31**, 355–368 (2018).

23. Longpré, M. A., Stix, J., Burkert, C., Hansteen, T. & Kutterolf, S. Sulfur budget and global climate impact of the A.D. 1835 eruption of Cosigüina volcano, Nicaragua. *Geophys. Res. Lett.* **41**, 6667–6675 (2014).

24. Raible, C. C. *et al.* Tambora 1815 as a test case for high impact volcanic eruptions: Earth system effects. (2016) doi:10.1002/wcc.407.