ADVANCED REVIEW

WIREs CLIMATE CHANGE WILEY

Climate and society in long-term perspective: Opportunities and pitfalls in the use of historical datasets

Bas J. P. van Bavel[1] | **Daniel R. Curtis[2]** | **Matthew J. Hannaford[3]** |
Michail Moatsos[1] | **Joris Roosen[1]** | **Tim Soens[4]**

[1]Department of History and Art History, Utrecht University, Utrecht, Netherlands

[2]Erasmus School of History, Culture and Communication, Erasmus University Rotterdam, Rotterdam, Netherlands

[3]School of Geography, University of Lincoln, Lincoln, UK

[4]Department of History, University of Antwerp, Antwerp, Belgium

Correspondence
Bas J. P. van Bavel, Department of History and Art History, Utrecht University, Utrecht 3512 JE, Netherlands.
Email: b.j.p.vanbavel@uu.nl

Edited by Matthias Heymann, Domain Editor, and Mike Hulme, Editor-in-Chief

Recent advances in paleoclimatology and the growing digital availability of large historical datasets on human activity have created new opportunities to investigate long-term interactions between climate and society. However, noncritical use of historical datasets can create pitfalls, resulting in misleading findings that may become entrenched as accepted knowledge. We demonstrate pitfalls in the content, use and interpretation of historical datasets in research into climate and society interaction through a systematic review of recent studies on the link between climate and (a) conflict incidence, (b) plague outbreaks and (c) agricultural productivity changes. We propose three sets of interventions to overcome these pitfalls, which involve a more critical and multidisciplinary collection and construction of historical datasets, increased specificity and transparency about uncertainty or biases, and replacing inductive with deductive approaches to causality. This will improve the validity and robustness of interpretations on the long-term relationship between climate and society.

This article is categorized under:
   Climate, History, Society, Culture > Disciplinary Perspectives

KEYWORDS
climate and society, conflict, historical datasets, long-term, plague

1 | NEW OPPORTUNITIES

Concerns over the implications of global climate change have given new impetus to the need to understand the effects of climatic changes on society over the long term (Adamson, Hannaford, & Rohland, 2018; Haldon, Elton, et al., 2018). Recent advances in paleoclimatology and historical climate reconstruction have resulted in increased numbers of highly-resolved series of long-term climatic changes (Emile-Geay et al., 2017). Similarly, the digital age has enhanced our capacity to produce bigger datasets on historical variations in human activity than ever before. These new opportunities have led to an increasing use of historical data to untangle climate-society linkages.

Although the uncertainties and pitfalls associated with paleoclimate records are now well-known (Brádzil, Pfister, Wanner, Von Storch, & Luterbacher, 2005), the use of large historical datasets on human activity creates potential dangers. Even readily observable and quantifiable "hard" data from the natural sciences have their uncertainties and limitations; this applies to a greater degree for "soft" data from the social sciences, while "historical soft" data are even more problematic. In a sense, historical data are the result of uncontrolled experiments from which a biased subset of results survives. Using historical data noncritically, therefore, can lead to misguided or even spurious conclusions on long-term climate-society linkages. A critical investigation of historical big data, its gathering, processing and analysis, is therefore needed to remove its aura of scientific objectivity and to avoid the proliferation of "bad history" in scientific journals (as observed for medical history) (King & Green, 2018).

In this article, we present a systematic critique of historical datasets and the pitfalls that can ensue from their noncritical use. We do so through both a systematic review and a qualitative analysis, scrutinizing the three domains where climate research has made use of quantitative historical datasets very clearly: the links between climate and (a) conflict incidence, (b) plague outbreaks, and (c) agricultural productivity changes. While historical data have also been used to explore related questions such as the link between climate variability and migration (see Mauelshagen, 2018), we have selected these three areas because of the extensive use of quantitative rather than qualitative methodologies.

Previous critiques of this literature are mainly limited to assessing methodological flaws or sampling bias in studies on climate-conflict relationships from the mid-twentieth century onward (Adams, Ide, Barnett, & Detges, 2018; Buhaug, 2010; Buhaug et al., 2014; Gleditsch, 2012; Klomp & Bulte, 2013), while reviews of the literature linking climate and economy have focused entirely on the period after 1950 (Carleton & Hsiang, 2016). Some literature has reviewed the connections between climate and conflict, disease, and food production, going much further back in time (Degroot, 2018; Degroot, 2019; Webb, 2018; White, 2018), but this article represents a first systematic attempt to specifically critique the use and interpretation of historical datasets on human activity in research on the climate-society nexus across a broad timespan. We first discuss the methods and results of the main studies in the three subdomains, followed by a bibliometric analysis of relevant articles published in highly-ranked scientific journals. Thereafter, we analyze the main pitfalls posed by the content of historical datasets, their use, and the interpretation of the results. We conclude by proposing three sets of interventions to overcome these pitfalls.

2 | STUDIES ON LONG-TERM CLIMATE-SOCIETY LINKAGES

In recent years a new body of quantitative research has emerged on the causal relationship between climate and human activity. In many cases, paleoclimatic data have been correlated with historical data to "explain" human phenomena as an outcome of climatic change (Zhang, Brecke, Lee, He, & Zhang, 2007; Zhang et al., 2011; Burke, Miguel, Satyanath, Dykema, & Lobell, 2009). Analyses have focused on three time periods: the post-1950 period using modern big data, the historical period using datasets assembled from written documents, and earlier periods using the archeological record. Here, we focus on studies using historical data spanning previous centuries, including large datasets on conflicts (Brecke, 1999; Kohn, 1999; Luard, 1986; Wright, 1942), plagues (Biraben, 1975/6; Büntgen, Ginzler, Esper, Tegel, & McMichael, 2012), grain prices and wages (Allen, 2001; Beveridge, 1922), agricultural productivity (Daux et al., 2012; Slicher van Bath, 1963), population (McEvedy & Jones, 1978), and technological innovations (Simonton, 1980). Some of these datasets are simply digitized versions of older datasets, while others have incorporated older datasets into new efforts.

The focus of historical data collection and digitization on conflicts, disease and agricultural productivity has led to three corresponding subdomains on climate-society linkages. We briefly introduce each subdomain.

2.1 | Conflict

The availability of conflict datasets that span several centuries has led scientists to use history as a testing ground to analyze the possible causal relationship between climatic change and conflict, warfare and unrest (Zhang, Zhang, Lee, & He, 2007; Zhang et al., 2011; Burke et al., 2009; Zhang, Brecke, et al., 2007; Tol & Wagner, 2010; Büntgen et al., 2011; Hsiang, Burke, & Miguel, 2013; Lee & Zhang, 2015; Yin, Su, & Fang, 2016). Of the analyses that focus on long (multi-centennial) timescales, most find that long-term fluctuations in war frequency correlated with cold periods, in particular those associated with the Little Ice Age, both in Europe (Tol & Wagner, 2010; Zhang et al., 2011; Zhang, Brecke, et al., 2007) and China (Yin et al., 2016; Zhang, Zhang, et al., 2007). In some cases, historical analyses are also viewed as direct analogies for future climate-conflict relationships, insofar as temperature increases could lead to a decrease in conflict in Europe (Tol & Wagner,

2010), or an increase in sub-Saharan Africa (Burke et al., 2009). Such claims make it crucial to scrutinize the validity of these studies.

2.2 | Plague

Given that reservoirs of contemporary plague in a variety of ecosystems are sensitive to climatic fluctuations (Stenseth et al., 2006), growing attention is paid to the relationship between climate and historical plagues. Analysis of historical data on the occurrence of plague outbreaks going back to the fourteenth century has suggested that a variety of short-term climatic variations and extreme weather events are correlated with the most significant outbreaks of Second Pandemic plagues in Europe (Schmid et al., 2015; Yue & Lee, 2018), Third Pandemic plagues in China (Xu et al., 2011; Xu et al., 2014;) and India (Lewnard & Townsend, 2016), twentieth-century plagues in the United States and Madagascar (Ben-Ari et al., 2008; Kreppel et al., 2014), and potential plague epidemics in Ming and Qing China (Lee et al., 2017; Tian et al., 2017). Given that the World Health Organization has recently categorized plague as a re-emerging global health threat (World Health Organization, 2009), epidemiologists are concerned that future climatic change might increase the global occurrence of plague (Stenseth et al., 2008). Accordingly, it is crucial to establish the nature of climate-plague connections.

2.3 | Agricultural productivity

Agricultural output is directly influenced by temperature and/or precipitation, although the precise nature of the interaction is highly dependent on the type of crop, geography, production methods and technology (Federico, 2005). Combining paleoclimatic evidence with historical data on cereal yields, harvest dates and prices, it becomes possible to question the impact of climatic changes on patterns of agricultural production (Cook & Wolkovich, 2016; Olmstead & Rhode, 2011). Furthermore, starting from the nexus between climatic change and agricultural productivity, wider questions have been addressed on the impact of climate on macro-economic performance (Pei, Zhang, Lee, & Li, 2014; Pei, Zhang, Li, & Lee, 2013; Pei, Zhang, Li, & Lee, 2015), migration (Büntgen et al., 2011), technological innovations (De Dreu & van Dijk, 2018), and population fluctuations and movements (Zhang et al., 2011; Zhang, Brecke, et al., 2007). Most of these studies see climatic change—via stress on food production—as the most important agent behind these phenomena. These findings are significant given current concerns over the impact of climate change on food security (Battisti & Naylor, 2009; Nelson et al., 2016), and are therefore in need of scrutiny.

3 | BIBLIOMETRIC ANALYSIS

We now proceed to identify a concise set of pitfalls within the literature concerning the three subdomains. In order to avert potential issues of bias, our identification of common pitfalls begins with a systematic bibliometric analysis. The purpose of this is to more objectively obtain a sample of relevant studies and to guard against any accusation of arbitrary choices or the use of "strawmen": accordingly, this article should not be viewed as a comprehensive critique of the whole literature. The aim of this analysis is to identify how typical these pitfalls are among those articles published within the highest-ranking peer-reviewed journals. It is not meant to be the only source of studies to be discussed, but is a systematically collected sample of studies in the top journals of their respective fields which allows us to avoid drawing our conclusions on a set of arbitrarily chosen articles.

3.1 | Systematic search for relevant publications

To obtain a systematic identification of relevant publications we operationalize a set of broad but relevant keywords for each of the subdomains.[1] In turn, we apply a keyword-based full-text search within the five leading "general purpose" scientific journals, plus the three highest-ranked academic journals in the most relevant fields to these subdomains as classified by Scimago Journal Rankings (SJR). These SJR categories are: safety research (conflict studies), infectious diseases (plague), food science (agricultural productivity), and atmospheric science (climate research), surveyed for the period from January 2007 to December 2018 (see S2 section in the Supporting Information for more information). As detailed in Table S1 (in Supporting Information), our keyword-based full-text search, performed using the Scopus service, produced a total of 21,300 articles, the vast majority of which were clear false positives, referring to articles that were from very different research topics. Two of the co-authors were assigned to each subdomain, and each of them independently filtered the abstracts to relevant and

irrelevant. Overall, 33 unique articles were identified as relevant for the fields under investigation (with some of them falling into more than one field), 23 of which made use of historical datasets in a quantitative analysis and focused on the pre-1950 period.[2]

The ratio of the 33 relevant articles over the total keyword-matched articles implies that the search terms, which were also applied with word stemming, are generic enough to produce numerous false positive results. Having a large pool of articles to begin our manual screening with, allows us to reduce as much as possible the false negatives, which we think would be a more urgent problem in our analysis. We simply want to avoid missing potential articles that fit our requirements. This necessitates the use of generic keywords (along with the stemming terms) that would in all likelihood be used in an article that fits each subdomain, which in turn explodes the number of false positives.

At the same time, a high share of the relevant articles made quantitative use of historical datasets (23 out of 33). This high share implies that quantitative use of historical datasets is common practice in the research on long-term climate-society linkages, or at least in high ranking journals. Also, this indicates that the methodological problems discussed below are highly relevant within the entire subdomains. Due to their bold claims, some of these publications are taken up by mainstream media, adding to their visibility in the public eye. Such claims about the impact of future climate change, for example, in terms of conflict, may turn public discourse in unwarranted directions when based on noncritical use of historical data.

3.2 | Analysis of relevant publications

To identify possible pitfalls, we analyzed the use and interpretation of historical data through eight closed-ended (yes/no) questions. Table 1 provides the detailed definition of each pitfall type. While the first question is designed to filter out those articles that refer to historical data but do not contain quantitative analysis, the seven remaining questions correspond to the pitfalls. Those possible pitfalls represent common concerns in the use of historical datasets that can undermine the validity of the relevant findings. As Table 1 shows, each of our pitfalls corresponds to a specific closed-ended question, which enabled us to provide a sharp categorization of each article. For each of the 23 articles that make quantitative use of historical data, the two co-authors assigned to the particular subdomain independently scored the articles, while a third co-author oversaw all scoring.[3]

The results of the scoring are summarized in Table 2. As the table makes clear, in columns 4–10, the majority of the studies that employ historical data do not avoid most of the pitfalls. Moreover, for 5 out of 7 pitfalls the articles that do not avoid them also have a statistically significant higher "field-weighted cited impact" factor (FWCI, as provided by Scopus service).[4] For the other 2, the articles that fall into a pitfall have a 76% higher FWCI on average, but the difference is not statistically significant.

Using this bibliometric search as a systematic starting point, we next review the content, use, and interpretation of historical data employed in research on long-term climate-society linkages, and identify a set of common problems. The parts of this

TABLE 1 Criteria used for the bibliometric analysis

Criteria	Corresponding closed-ended question
Quantitative analysis	Is quantitative analysis (e.g., statistical testing) employed using historical data (before 1950)? Only those articles with a yes in this question are used in the subsequent scoring and analysis.
Data critique	Is historical source critique present that moves beyond a mere description of the database, by discussing the nature/type of the historical documents behind the data as well as their limitations?
Temporal critique	Is there critical reflection on how a lack of full temporal coverage of the data employed might influence the results?
Geographical critique	Is there critical reflection on how a lack of full geographical coverage of the data employed might influence the results?
Avoid false uniformity	Do the authors guard their analysis against the creation of false uniformity across space and time? For instance, by not giving equal weight in the statistical analysis to sources of uneven quality (or varying source types) or availability?
Societal contextualization	Are the specific context and characteristics of societies, regions and localities where the data derive from, taken into account in the analysis?
Avoid causal claim	Do the authors avoid claiming causality on the basis of results derived from analysis of historical data which fails to meet one of the criteria above?
Historiography	Do the authors use prevailing, up-to-date historical ideas and theory to discuss and analyze their results?

Note: These questions were constructed to be closed-ended, allowing only a yes/no answer.

TABLE 2 Scoring results of bibliometric analysis

Study	Scopus Cit. (FWCI)	Quant. analysis	Data crit.	Temp. crit.	Geo. crit.	Avoid false uniformity	Context society	Avoid causal. claim	Hist.
Climate-conflict (14 out of 1,095 keyword-based matched articles)									
Büntgen et al., 2011	581 (11.33)	Y	N	N	N	N	N	Y	N
Drake, 2017	2 (0.28)	Y	N	N	N	N	Y	Y	N
Endfield, 2012		N							
Gartzke, 2012	34 (7.12)	Y	N	N	N	N	N	N	Y
Haldon, Elton, et al., 2018	12 (2.52)	Y	Y	Y	Y	Y	Y	Y	Y
Hsiang et al., 2013	436 (12.55)	Y	N	N	N	N	N	N	N
Kaniewski et al., 2017		N							
Manning et al., 2017	12 (2.08)	Y	Y	Y	Y	Y	Y	Y	Y
McMichael, 2012		N							
Tan et al., 2015		N							
Tian et al., 2017	5 (0.57)	Y	N	Y	Y	N	N	N	N
Wig, 2016	7 (2.11)	Y	Y	Y	N	N	N	N	N
Zhang, Brecke, et al., 2007	241 (1.58)	Y	N	N	N	N	N	N	N
Zhang et al., 2011	187 (3.11)	Y	N	N	N	N	N	N	N
Climate-plague (11 out of 6,245 keyword-based matched articles)									
Ben-Ari et al., 2012	11 (0.3)	Y	Y	Y	Y	Y	Y	Y	Y
Brook, 2017	0 (0)	Y	Y	Y	Y	Y	Y	Y	Y
Helama et al., 2018		N							
Lewnard & Townsend, 2016	4 (0.32)	Y	N	Y	Y	Y	Y	Y	N
McMichael, 2012		N							
Schmid et al., 2015	74 (3.52)	Y	N	N	N	N	N	Y	Y
Streeter, Dugmore, & Vésteinsson, 2012		N							
Tian et al., 2017	5 (0.57)	Y	N	Y	Y	N	N	N	N
Welford & Bossak, 2009	15 (0.59)	Y	N	N	Y	N	N	Y	Y
Xu et al., 2011	36 (0.78)	Y	N	N	N	N	N	Y	N
Yue, Lee, & Wu, 2017	8 (1.14)	Y	N	N	N	N	N	Y	N
Climate-productivity (16 out of 13,960 keyword-based matched articles)									
Battisti & Naylor, 2009		N							
Büntgen et al., 2011	581 (11.33)	Y	N	N	N	N	N	Y	N
Cook & Wolkovich, 2016	22 (3.92)	Y	N	N	Y	N	N	N	N
De Dreu & van Dijk, 2018	1 (1)	Y	N	N	N	N	N	N	N
Drake, 2017	2 (0.28)	Y	N	N	N	N	Y	Y	N

(Continues)

TABLE 2 (Continued)

Study	Scopus Cit. (FWCI)	Quant. analysis	Data crit.	Temp. crit.	Geo. crit.	Avoid false uniformity	Context society	Avoid causal. claim	Hist.
Helama et al., 2018		N							
Kaniewski et al., 2017		N							
Kukal & Irmak, 2018		N							
Nelson et al., 2016		N							
Olmstead & Rhode, 2011	36 (0.31)	Y	Y	Y	Y	Y	Y	Y	Y
Paprotny, Sebastian, Morales-Nápoles, & Jonkman, 2018	11 (5.23)	Y	Y	Y	Y	N	N	Y	N
Pei et al., 2014	26 (1.65)	Y	N	N	N	N	N	N	Y
Pei et al., 2015	14 (1.25)	Y	N	N	N	N	N	N	Y
Shennan et al., 2013		N							
Zhang, Brecke, et al., 2007	241 (1.58)	Y	N	N	N	N	N	N	N
Zhang et al., 2011	187 (3.11)	Y	N	N	N	N	N	N	N
		23 Y	16 N vs. 7 Y	14 N vs. 9 Y	13 N vs. 10 Y	17 N vs. 6 Y	16 N vs. 7 Y	11 N vs. 12 Y	13 N vs. 10 Y
FWCI per article[10]			3.17 N vs. 1.79 Y	3.56 N vs. 1.49 Y*	3.65 N vs. 1.58 Y*	3.4 N vs. 0.92 Y**	3.59 N vs. 0.83 Y***	4.2 N vs. 1.42 Y**	3.38 N vs. 1.93 Y

Note: see table 1 for the explanation of the criteria employed; Y for Yes stands for pitfall avoidance, while N for No stands for a failure in avoiding a pitfall.

* Statistically significant at 10%, ** statistically significant at 5%, and *** statistically significant at 1% using independent 2-group two-sided Mann–Whitney *U*-test, as the data fail normality tests.

review mainly focus on articles captured by the bibliometric search, but, as discussed below, also incorporate other examples of additional articles from the relevant literature published outside the aforementioned high-ranked journals to substantiate our arguments. This suggests that the identified pitfalls might be a part of a phenomenon not exclusive to a few selected high-ranking journals.

4 | CONTENT OF HISTORICAL DATASETS

Although historical climatologists have made methodological allowances for the limitations of historical documentary sources to reconstruct climate in the pre-instrumental period (Brádzil et al., 2005), the majority of recent research that links this climatic evidence to historical datasets on human activity has shown a lack of appreciation for the original source material. Indeed, while source and data critique are the cornerstone of the discipline of history, the majority of studies in our bibliometric analysis offer little in this regard (Table 2). Instead, justification of the chosen dataset follows a similar pattern: the dataset is introduced in a paragraph (usually in Supporting Information), its characteristics—typically that the dataset is the largest and most "representative" of its kind (Zhang et al., 2011; Zhang, Brecke, et al., 2007)—are described, and the data extracted for a given region are then used as a dependent variable. We argue here that this lack of critique has obscured unevenness in the temporal and spatial representativeness of the datasets in question, as well as uncertainty over the representativeness of the specific historical variables employed.

4.1 | Unevenness in temporal and spatial representativeness

Since the digitization of a historical database of plagues in Europe and Northern Africa in the period 1347–1900 (Büntgen et al., 2012), a number of scholars have used this dataset in combination with climate reconstructions to account for the spatial

and temporal distribution of plague outbreaks (Büntgen et al., 2013; Schmid et al., 2015; Yue & Lee, 2018). This database is nevertheless not newly-collected information, but appendices from the work of one historian in the 1970s (Biraben, 1975/6). Similarly, data collected for Third Pandemic plague cases in China in the period 1772–1964 (although mainly corresponding only to Yunnan before 1850) and used to establish a relationship with climate (Xu et al., 2014; Xu et al., 2011) are not newly collected but were compiled by a Chinese Academy of Sciences team in the period 1963–80. Few of the works using these datasets reflect on the context behind the initial data collection, but simply take the digitized version as a state-of-the-art dataset rather than a product of the uneven spatial and temporal distribution and accessibility of historical source material.

Such problems of chronological or geographical representativeness often pass without even cursory recognition. Some have used historical plague datasets, for example, to give an impression of the chronological development of plague activity over time (Büntgen et al., 2012; Schmid et al., 2015; Yue, Lee, & Wu, 2016). However, what these timeseries show is not numbers of plague "incidences" or 'outbreaks' but simply uneven distribution of surviving and consulted source material over time, exacerbated by uneven scholarly interest in specific periods or outbreaks such as the Black Death of 1347–1352. In the study on the plague-climate interaction in Ming-Qing China, the supporting documents even show that the more numerous spikes in recorded epidemics from the nineteenth century onward are simply connected to a much fuller documentary record in general (Tian et al., 2017).

Similar problems exist within the historical climate-conflict literature. One of the most frequently used conflict datasets, for example, is Brecke's "Conflict Catalog" (1999). This dataset is partly based on a compilation by Luard (1986), and partly on later secondary works, which have their own foci and biases. Indeed, by Brecke's own admission, the catalog is an "unfinished product," with errors "especially as we go back in time and into particular regions of the world" (Brecke, 2012, p. 1). Coverage in the Southern Hemisphere—where data going back to 1400 are used by Zhang, Brecke, et al. (2007) and Lee and Zhang (2015)—is seriously deficient before 1800, with most entries skewed toward larger conflicts or those between colonial powers and indigenous populations. Even in the nineteenth century there are vast oversimplifications. The multitude of inter- and intra-group conflicts of the 1820s in southeast Africa, for example, are grouped into one decade-long conflict of the "Zulu tribes," a notion that dates back to early colonial writings which routinely exaggerated the effects of conflict (and in some cases fabricated its existence) (Hannaford & Nash, 2016). Lack of accessible information on how and why these original judgments were arrived at makes it difficult for researchers to constructively engage with or refine these datasets, yet a minority of cases in the climate-conflict literature acknowledge such issues (Haldon, Elton, et al., 2018; Tian et al., 2017; Manning et al., 2017; Wig, 2016).

Lack of representativeness is also a frequent problem in studies on agricultural productivity. For instance, European yield ratios—used as proxy for productivity in a number of studies (Pei et al., 2013; Pei et al., 2014; Pei et al., 2015; Zhang et al., 2011)—are mainly derived from the dataset of Slicher van Bath (1963), containing 11,462 yield ratios between 800 and 1820. While 'European' in its scope, the material is highly biased toward the published historical sources available to the author, while chronological biases are often reinforced by geographical ones: before 1300, most yields come from England, and after 1500 most come from Central and Eastern Europe.

An instructive way of visualizing the lack of geographical representativeness inherent within these historical datasets is to focus on the maps reproduced in recent articles on historical plague outbreaks. Figure 1 shows crucial gaps in spatial coverage in one of the most widely used digital datasets (Büntgen et al., 2012) by comparing it to an appendix of plague references from historical sources in one particular area (the Low Countries) in the period 1349–1500.

4.2 | Questionable representativeness of variables

Understanding the characteristics of the historical records behind the datasets is also essential to ensure the correct use of the specific variables that they purport to represent. In the absence of spatially and temporally extensive *Yersinia pestis* DNA evidence that definitively points to plague, for example, studies in Table 2 often use documentary evidence based on direct anecdotal references by contemporaries as to what they thought was plague. Yet even today the diagnosis of plague through signs and symptoms is difficult for trained medical professionals. In medieval and early modern Europe, terms such as "peste" and "pestilentia" were often indiscriminate references to all sorts of afflictions (Theilmann & Cate, 2007). Historians are rightly cautious about retrospective diagnoses based on written evidence alone (Green, 2015). For the already-mentioned Chinese studies (Xu et al., 2014; Xu et al., 2011), all the supporting documents tell us is that "plague data were collected from multiple sources of historical records," which says nothing about the methods and criteria used by the original compilers of the data, nor the context behind the original diagnoses. A more careful approach has been taken by another study linking epidemic
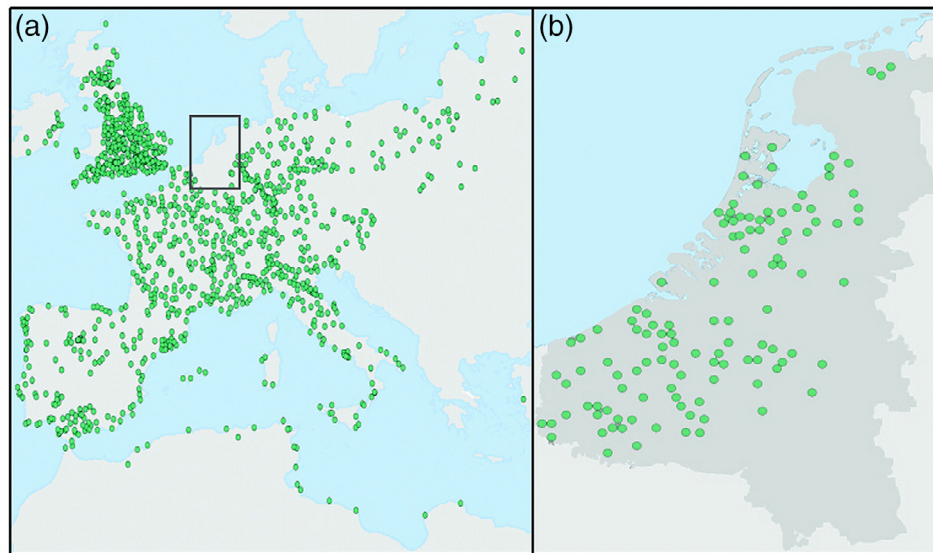
FIGURE 1   Illustration of geographical gaps in digitized Biraben plague dataset. Part (a) shows localities in Europe and North Africa reporting plague outbreaks in the period 1347–1760 according to the digitized version of the Biraben dataset (image courtesy of Yue & Lee, 2018; based on digitization by Büntgen et al., 2012). The gaps in spatial coverage are immediately visible when taking into account data for the Low Countries, indicated in the inset. When contrasted with an appendix of locations reporting plague outbreaks in the Low Countries just for the period 1349–1500 (part b) (Roosen & Curtis, 2019), the extent of the spatial gap for this region becomes apparent—and this appendix is far from exhaustive

outbreaks in Ming and Qing China to climate—by simply conceding that it is impossible to accurately identify precise cause of death from documentary evidence alone (Brook, 2017; Tian et al., 2017).

Issues over the representativeness of variables are also found in research on the long-term link between climate and agricultural productivity. For example, the van Bath series (1963) uses yields per seed as an indicator of productivity, which were highly dependent on seeding ratios (declining with more seed per surface) and the amount of cropland under cultivation (increasing cultivation leads to declining yields), as well as other factors such as labor input and fertilization. Today, agricultural historians favor yields per unit area of land instead of yields per seed (Campbell, 2007; Thoen & Soens, 2015). Similarly, while the start of the grape harvest might indicate weather conditions during the growing season of the grapes, each data series needs careful source criticism and contextualization. Rather than neutral reflections of the grape "ripeness" and hence the growing conditions, grape harvest dates are usually the result of a collective decision-making process, which might evolve over time with evolutions in taste, institutional configurations or production methods (Daux et al., 2012). While such criticism is usually included in the original publication of the dataset, it is often omitted in later studies, simply presuming "that management practices have changed relatively little over time" (Cook & Wolkovich, 2016, p. 720). Finally, some of the observed patterns can be a direct result of the registration process behind the data. Chronological gazetteers on scientific and technological "innovations" used by De Dreu and van Dijk (2018), for instance, can be highly dependent on the way societies appreciated 'innovation' and "improvement" as positive contributions to society. In Europe, such appreciation grew fundamentally from the seventeenth century onwards (Mokyr, 2016). Both the evolving enthusiasm for inventions and inventors and the publication of key-treatises might produce surges in the number of recorded inventions, which needs to be falsified before investigating its potential link with climate-induced agricultural stress.

5 | USE OF HISTORICAL DATASETS

5.1 | Falsely uniform data

In the absence of source criticism, quantitative methods can lead to an artificial "uniformity" in the coverage and quality of historical data across space and time, as seen in the majority of articles in our bibliometric analysis (Table 2). For example, although we now have a large body of temperature and precipitation reconstructions covering the last millennium, we find little recognition that similar changes in temperature can have completely different meaning and impact between different

historical societies depending on their environmental and social context (Brönnimann & Wintzer, 2019). First, different crops in different social-environmental contexts do not respond in the same manner to temperature variability (de Vries, 1980). But second, even in neighboring areas, similar increases in food prices or intensity of harvest failures did not always produce equivalent effects—certainly with regard to excess mortality, displacement, and social unrest (Curtis & Dijkman, 2019). Points such as these are problematic when we consider that, in each of the three subdomains, multiple linkages are often packaged into single statistical tests or composite series for large areas (Pei et al., 2013; Pei et al., 2014; Pei et al., 2015). This can create an illusion that aggregated data are either a complete or sufficiently random sample over a defined spatial area, but instead they are a highly selective sample dependent on the legacy of recording.

5.2 | Problems with the spatial scale

Similarly, there exists a lack of spatial contextualization when selecting and categorizing variables (Brönnimann & Wintzer, 2019). Lee and Zhang (2015), for example, consider historical temperature-conflict links in the geographical groupings of eastern and southern Africa, and western and central Africa, finding a negative correlation between conflict incidence in these regions and a Southern Hemisphere temperature reconstruction. However, the authors omit the crucial point that it is precipitation, not temperature, which is the key climatic variable for food production across much of sub-Saharan Africa (Hannaford & Nash, 2016; McCann, 1999). Moreover, as long-term temperature variability influences precipitation variability in the tropics and subtropics, precipitation is partially embedded in records of temperature, albeit in a nonlinear manner, in southern and eastern Africa (Nash et al., 2016). Grouping these two regions together thus makes little sense from a dynamical-climatological perspective. In the absence of well-defined causal models identifying the specific meteorological variables and critical periods of the farming year in certain geographical contexts, any statistical analysis may offer spurious results.

5.3 | Problems with the temporal scale

An equally fundamental issue relates to the temporal scale at which historical datasets are used in relation to climate variability. Hsiang et al. (2013, p. 9) argue that "climatic anomalies of all temporal durations, from the anomalous hour to the anomalous millennium have been implicated in some form of human conflict," yet it is often unclear precisely why a certain temporal resolution is chosen for statistical tests, and how this connects to the underlying theory (Degroot, 2018). Examples are studies that use running totals of collective conflicts, rather than annual or seasonal conflict outbreaks, as their dependent variable (Tol & Wagner, 2010; Zhang, Brecke, et al., 2007; Zhang, Zhang, et al., 2007), and find that colder average climatic conditions in the Northern Hemisphere are correlated with increased numbers of these running conflicts. Colder decades did not simply translate into repeated seasonal harvest failures, so why were colder average conditions more likely to precipitate conflict than, for example, the shorter-term incidence of back-to-back years of late-spring frosts outside of a colder decade? What specific role did these colder average conditions play in escalating or sustaining conflict, and where is the mechanism? Such questions of temporal scale are largely left unaddressed, but considering the aggregate number of conflicts provides only the most limited insight into the nature of climate-conflict relationships.

5.4 | Problems with the weight of data points

Many of the historical plague databases cited above also suffer from a restrictive uniformity where datapoints are weighted equally. Not only are the original sources given equitable credibility and value, regardless of whether they are eyewitness accounts, or whether they were contemporaneous or noncontemporaneous, but they are often read out of context. These issues must be considered if we are to reconstruct symptoms and epidemiology from the written sources. Only the study on climate-disease interactions in Ming-Qing China has tried to factor in severity into the original database (Tian et al., 2017), but this is based on a subjective ranking system of descriptions from observers where the term "countless people died" is deemed to indicate a much more severe epidemic than the term "many people died," for example. Terms such as "no people left in the villages" are assumed to refer to plague mortality, but can also reflect displacement. Similar problems of assessing mortality from face-value descriptions of contemporaries have been found in recent work on the Black Death in Europe (Gómez & Verdú, 2017), when it is clear they are subject to inexactness and sometimes deliberate exaggeration.

These historical databases of descriptive mentions of plague are not the same as showing differences in plague characteristics across the same type of source material. This is important because the climate-plague connection cannot merely be

demonstrated in spread—it also has to take into account severity, seasonality, longevity, and selectivity, and recent research has shown that these features could differ considerably, not only between outbreaks, but also between localities facing the same outbreak (Alfani, 2013; Alfani & Murphy, 2017; Curtis, 2016). At the moment, most climate-plague research is conducted at the subcontinental scale, but while the global climate may "drive" plagues, it is local environmental and societal conditions that dictate actual epidemiologic outcomes and variability of plague characteristics (Brook, 2017).

Overall then, the risk is that these historical datasets command an undeserved sense of reliability and authority when employed without proper regard for the outlined pitfalls. Our bibliometric analysis shows that such problems beset the majority of relevant publications and particularly the more influential ones. The results presented in these articles are not always passively accepted: quantitative approaches to climate-conflict research covering late-twentieth-century timeframes, for instance, have received robust criticism on their methodology, as highlighted earlier. However, the specific problems we emphasize are generally overlooked because of a lack of engagement with the science of history, and notwithstanding notable exceptions (Degroot, 2018), a lack of visible engagement with these studies by historians. Furthermore, the affected datasets do not merely suffer from some small random error or unresolvable minor issues, but systematic biases from which spurious results can emerge.

6 | INTERPRETATION OF HISTORICAL DATASETS

6.1 | Inductive reasoning

A third set of pitfalls lies in the interpretation of quantitative analyses derived from historical datasets. Many quantitative analyses adopt an inductive approach by simply searching for causes to fit observed patterns with minimal theoretical grounding or justification. The simple correlation of two variables is readily interpreted as evidence for a causal relationship (Haldon, Elton, et al., 2018). As argued above, this might be problematic because of the questionable nature of the variables which are employed as proxies for basic historical indices. It is often the variable that is easiest to quantify that is elevated to the level of dominant predictor, while other factors are ignored or marginalized. Such insights tend to give a "one-eyed" view of history, offering overstated conclusions, which are artificially precise and reductionist (Hulme, 2011).

This problem can be seen in those works linking drops in temperature or extreme rainfall with increased conflict, which tend to assume that pressure on food resources is a "logical" intermediary mechanism (Tol & Wagner, 2010; Zhang, Brecke, et al., 2007), without the realization that, in the pre-industrial period at least, extreme weather on its own was not always enough to tip harvest failures over into famine symptoms—that is, not inevitably producing excess mortality or migration (de Vries, 1980; Slavin, 2016). In certain parts of the world the links were clear between weather, production, and famine (Alfani & Ó Gráda, 2018), but in other places these links were not strong at all or waned during or after certain periods in time (Curtis & Dijkman, 2019). This becomes further complicated given that conflict itself has often been cited as a cause of food crises (Alfani & Ó Gráda, 2017), so the chain of causation then becomes completely reversed. These kinds of issues are best highlighted by comparing two recent conceptual models developed to interpret and explain the links between climate and economy (Butzer, 2012; Pei et al., 2014) (Figure 2). Clear oversimplifications in the 2014 model—while likely driven by the need to simplify in order to quantify—could easily undo more comprehensive linkages established in the 2012 model. For the sake of quantification, simpler causal models can be useful to test relationships between two variables, but even if the result is statistically significant, the model will tell us very little about the underlying mechanisms producing such a relationship.

6.2 | Outdated or absent historiography

Inductive thinking in the construction of models or identification of potential causal links is likely because many of the dependent variables are frequently identified by way of historical scholarship that is absent or not up-to-date. Less than half of the identified articles in our bibliometric analysis utilized up-to-date historical scholarship, but more than half—the considerably more influential segment—either included no historiography or historiography that was not up-to-date. While high-impact and generalist journals of the type privileged by our enquiry tend to impose tighter restrictions on the number of references and words than publication outlets favored by historians, it remains problematic that articles on climate shocks and technological innovation (De Dreu & van Dijk, 2018) or climatic change and wine harvests (Cook & Wolkovich, 2016) are published without any reference to recent historical literature on changing cultural preferences and social practices surrounding technological 'innovation' or wine production. Furthermore, in their selection of historical narratives invoked to explain either drivers of societal change or the outcome of climatic disruptions, authors tend to privilege clear political or military events such as
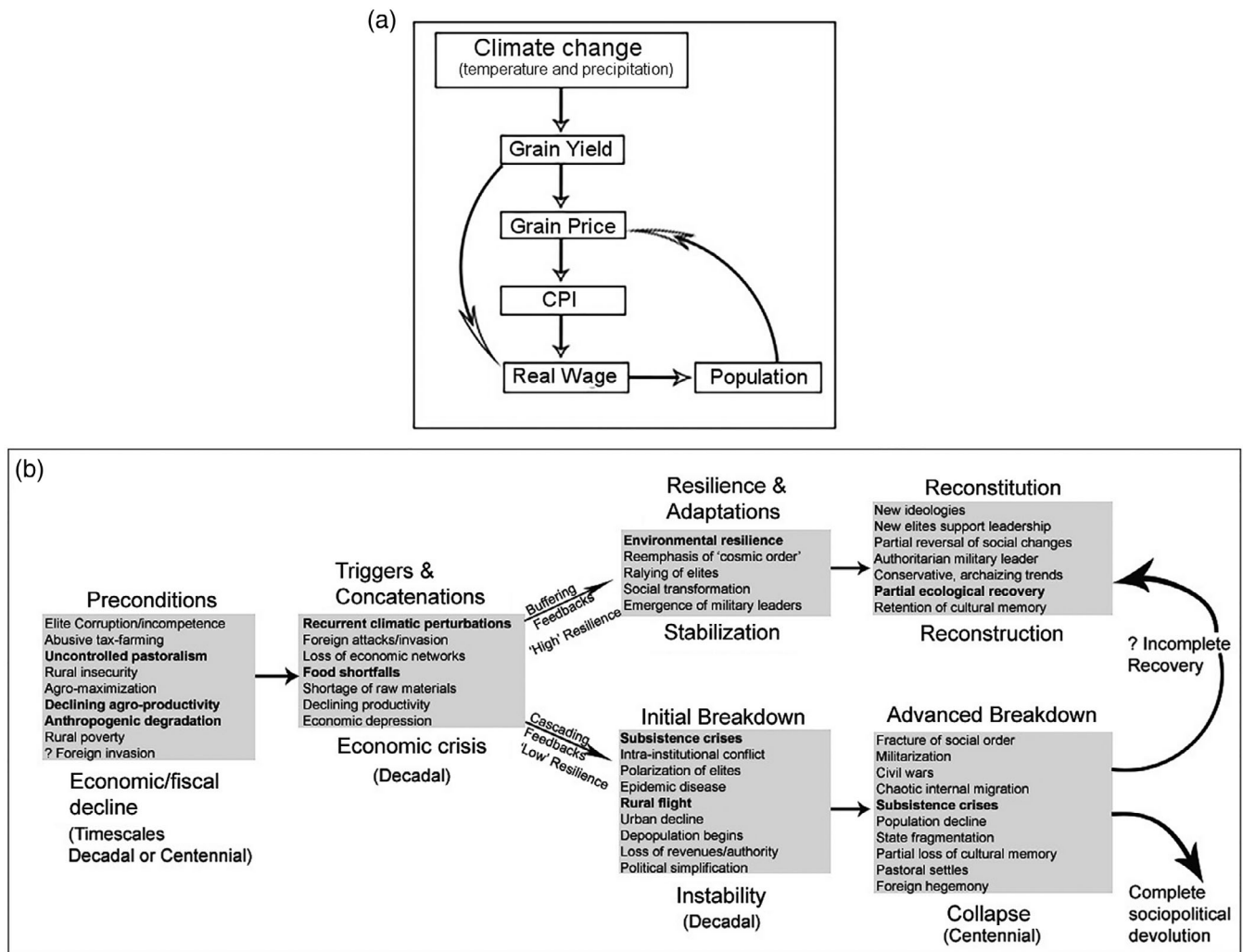
FIGURE 2   Comparison of two recent conceptual frameworks. Panel (a) shows the conceptual model of climate change and macro-economic cycles in pre-industrial Europe as used in Pei et al. (2014). The arrows indicate that "change in x is associated with change in y." This framework focuses on unilinear and direct effects and does not consider the complex social and institutional contexts of societies affected by climate change. A more nuanced overview of climatic (and other) factors influencing historical collapse can be found in Butzer (2012) (panel b). The text in bold is elaborated by the subscripts below each box. This conceptual model considers a range of variables and processes of stress and interaction and also reflects on a multitude of possible outcomes

dynastic changes or the disintegration of empires, while incremental, complex, and systemic changes, which often predominate historical interpretations of the same events, are put aside. The fall of the Roman Empire is such a clear event in popular imagination which has been linked to climatic transitions (Büntgen et al., 2011; Drake, 2017). However, while recent historiography rightly stresses the importance of climate when explaining the full complexity of socio-environmental transitions unfolding in the Roman West between the fourth and seventh centuries CE (Harper, 2017), it becomes problematic when the analysis isolates one causal variable (the climate), while downplaying others, just as previous generations of historians have been refuted for isolating political, religious or economic drivers of the same event. Furthermore, these "fall and collapse" narratives have been criticized for focusing on change and disruption, while ignoring the many signs of continuity (Haldon, Mordechai, et al., 2018; Wickham, 2010). Overall then, when recent scientific literature employs historical data, more sophisticated mapping, statistical, and digitization techniques often mask simplistic events-based explanations at the heart of the work.

7 | INTERVENTIONS AND GUIDELINES

The preceding pitfalls and problems may induce scholars to abort attempts to use large-scale historical datasets altogether. Some have argued, for instance, that human interaction with weather and climate involves such a high degree of human

agency that quantitative analysis is unsuited to climate-conflict research (Selby, 2014). We do not share this dismissive view, however, since historical datasets and quantitative analyses represent one of the main possibilities to obtain a systematic understanding of the long-term relationship between climatic change and society. Instead, we advocate three sets of interventions:

A. Enhance multidisciplinary data collection and construction of datasets

Proper selection, use, contextualization and interpretation of historical data requires the input of historians—and more precisely those historians well acquainted with the specific spatial and temporal context under investigation—within multidisciplinary scientific teams (Haldon, Elton, et al., 2018; Kwok, 2017; Holm & Winiwarter, 2017). One apparent feature of our bibliometric search is that a minority of articles include historians within the authorship teams, yet it was these articles where source critique and up-to-date historiography tended to be present (e.g., Brook, 2017; Haldon, Elton, et al., 2018; Manning et al., 2017). This is not necessarily a lack of inclusivity per se, but may owe much to disciplinary traditions, where history still largely favors lone scholarship and the monograph as the primary form of publication. For example, we have criticized the reliance of the climate-plague literature on Biraben (1975/6) and gazetteers, but partially this reliance stems from a lack of initiative on the part of historians to produce better databases, or even work in teams. Incentives are therefore needed to change outcomes. This could include open-access platforms for historical data to be published and assessed, rather than simply made available as appendixes. In the geosciences, for example, the *Geoscience Data Journal* has recently appeared to this effect, but such fora are lacking for historical scholarship. A related aspect to this is making sure more scientific journals with articles using historical data involve historians as part of the peer-review team (King & Green, 2018).

Including historians within research and peer-review teams and providing new outlets for the publication of datasets is important because it will in turn incentivize historians to start compiling more relevant historical indicators of interest. For example, given the recent trend toward original manuscript digitization, it would be possible for a large project to collect epidemiologic data for almost the entirety of seventeenth-century Western Europe using just one standardized source (burial records)—important given this coincides with some of the coolest phases of the Little Ice Age (Degroot, 2019). Such ventures can be supported by "citizen science" data rescue schemes, which involve the transcription of scanned images from historical sources by teams of volunteers through dedicated websites (Sangster, Jones, & Macdonald, 2018). In recent years these projects have contributed valuable sources of data to climatology, such as the millions of weather observations contained within historical ships' logbooks (the "Old Weather" project and "Weather Detective" (Freeman et al., 2017), though their potential still remains largely unrealized in history. Of course, in order to support the generation of new data on historical climate-society interactions, digital research infrastructure needs to facilitate general access to the primary manuscripts, transcriptions, and the analyses that follow (Allan et al., 2016).

B. Increase specificity and transparency about uncertainty or potential biases in the data

Constructing historical datasets that consist only of average values extracted from the original observations is one side of the problem. Instead, all unique observations that appear in the original sources should be included in historical datasets, thus allowing for further analysis and for quantification of margins of error when using an average value. Moreover, estimates of the underlying uncertainties should be made available along with the historical datasets. All of this can be aided by developing clear rules for good practice in building and using historical big datasets. These guidelines and quality control procedures are routinely followed in the construction of historical climatic datasets (Freeman et al., 2017); datasets on historical human activity should strive for similar levels of rigor.

Further, the use of historical data requires a more precise clarification of the appropriate temporal and geographical scales. In some cases, this necessitates narrowing down the chronological frame of analysis, and ensuring that the geographical frame of analysis is aligned to the social and political context of the societies under consideration—noting that this may be the region (as most often in pre-industrial Europe), or may be of a larger-scale in vast agrarian states. These contextual factors are equally as important as, and directly related to, careful selection of the sources. This requires us to build historical indicators of human activity using a more consistent and standardized source—mortality information from a series of burial records, for example, should take precedence over a database of plague mentions from diverse fragmented sources where we are unclear what each datapoint represents. In turn, these indices can be analyzed against the growing body of high-resolution climate reconstructions available at regional and subregional geographical scales (Pribyl, Cornes, & Pfister, 2012 for East Anglia in the thirteenth-fifteenth centuries; Camenisch et al., 2016 for the fifteenth-century Low Countries).

C. Reveal the causal mechanisms linking climate and society

Finally, it is now time to move beyond the statement that climate mattered in history. Most authors will agree that it did in one way or another. Rather than asking *whether* climate mattered, we should ask *how* humans and climate interacted: we urgently need more insight into the complex causal mechanisms linking climate and society (Haldon, 2016), and correlation alone is not enough to reach such an insight (Contreras, 2017). Through comparative research, history has the potential to show different pathways and trajectories between societies experiencing similar climatic conditions, and there is an urgent need for such research, as stressed by consecutive IPCC reports (Holm & Winiwarter, 2017). In order to reveal these causal mechanisms behind the climate-society nexus, we need to work on a wider range of spatial and temporal scales: on the one hand focusing in on more localized environments, for more limited periods in time, with societal contexts offering fewer variables, and yet also receptive to the notion that climate could have impacts that were either delayed, lingered or occurred far from the actual site of the weather anomalies. All too often, scholarship either privileges the "long-distance" connections, linking for instance drought on the Central Asian plains to political upheaval in Europe or China, or the very localized interactions between weather conditions and agriculture in a specific region, but seldom integrate both perspectives.

Accordingly, we point to the limitations of inductive approaches that simply "mine" data to determine whether certain variables are related to one another. Regressing the most easily quantifiable variables against one another to see if they are correlated at a level of statistical significance may yield interesting insights—and in recent years, statistical methods focusing on nonlinear relationships and impacts for historical societies have been increasingly refined—but still this should at least be complemented by other approaches and methods. Aside from the abovementioned issues concerning the construction and robustness of historical datasets, this approach may conflate statistical association with evidence of causality—the result being correlations with predictive power but little explanatory power. Instead, deductive approaches that start from a set of premises grounded in historiographical and social context offer a more promising way forward for quantitative climate-society research.

8 | CONCLUSION

Historical datasets are being increasingly used to investigate climate-society interactions over long timeframes and yet, as this article has demonstrated, progress is hindered by insufficient critical engagement with datasets' contents and their potential use. These pitfalls should not discourage future researchers from using these data in their work, but rather highlight the need to employ current historical knowledge and debate. In particular, a more critical multidisciplinary approach to the construction of historical datasets, and increased specificity and transparency about uncertainty or biases—including a more precise and considered clarification of appropriate temporal and spatial scales—should improve the validity and robustness of interpretations on the long-term relationship between climate and society.

Overall then, it becomes clear that historical societies are just as complex as current societies, and that simplistic causal relationships between climatic variables and conflict, disease or economic development are untenable. Only by taking into account the specificities of the environmental, economic, political and cultural context can causal relationships be established, which explain why similar environmental or climatic fluctuations often produced very different outcomes—even in regions very close to one another. Climatic change may prove to be the driver behind significant developments in human history, but only the integration of this information with contextually constrained historical datasets on human activity can inform us on the intensity, longevity and direction of the outcomes.

Annelies Tukker (supporting); Supervision, Bas van Bavel; Funding Acquisition, Bas van Bavel, Daniel Curtis, Tim Soens. Data and materials availability: All data is available in the main text or can be provided upon request.

CONFLICT OF INTEREST

The authors have declared no conflicts of interest for this article.

ENDNOTES

[1] Table S1 in the appendix contains the keywords, and section S2 of the appendix contains all the stemming words used. Also, the exact commands used in Scopus are available to facilitate the reproduction of our search results.

[2] We have used the raw keyword-based approach for its systematic qualities, even though the listed journals may feature a few additional articles that fit our scope, but are not picked up by this search method.

[3] In case any doubts remain, we scored Y (avoided the pitfall), in order to guard against any negative bias on our behalf.

[4] The FWCI shows how an article's citation count compares to similar articles in the same field. A score of 1 means the article is cited as expected, greater than 1 means better than expected, and less than 1 is worse than expected.

[5] There are 41 total entries, but eight studies are categorized in two fields. Therefore, the total number of unique studies is 33, of which 23 employ quantitative analysis.

ORCID

Bas J. P. van Bavel https://orcid.org/0000-0001-8686-0279
Daniel R. Curtis https://orcid.org/0000-0001-8355-4153
Matthew J. Hannaford https://orcid.org/0000-0002-9982-9868
Michail Moatsos https://orcid.org/0000-0002-3132-0651
Joris Roosen https://orcid.org/0000-0003-1056-0418
Tim Soens https://orcid.org/0000-0003-1040-5266

REFERENCES

Adams, C., Ide, T., Barnett, J., & Detges, A. (2018). Sampling bias in climate-conflict research. *Nature Climate Change*, *8*(3), 200–203. https://doi.org/10.1038/s41558-018-0068-2

Adamson, G. C., Hannaford, M. J., & Rohland, E. J. (2018). Re-thinking the present: The role of a historical focus in climate change adaptation research. *Global Environmental Change*, *48*, 195–205. https://doi.org/10.1016/j.gloenvcha.2017.12.003

Alfani, G. (2013). Plague in seventeenth-century Europe and the decline of Italy: An epidemiological hypothesis. *European Review of Economic History*, *17*(4), 408–430. https://doi.org/10.1093/ereh/het013

Alfani, G., & Murphy, T. (2017). Plague and lethal epidemics in the pre-industrial world. *Journal of Economic History*, *77*(1), 314–343. https://doi.org/10.1017/S0022050717000092

Alfani, G., & Ó Gráda, C. (2017). Famines in Europe: An overview. In G. Alfani & C. Ó. Gráda (Eds.), *Famine in European history* (pp. 1–24). Cambridge, England: Cambridge University Press. https://doi.org/10.1017/9781316841235.001

Alfani, G., & Ó Gráda, C. (2018). The timing and causes of famines in Europe. *Nature Sustainability*, *1*, 283–288. https://doi.org/10.1038/s41893-018-0078-0

Allan, R., Endfield, G., Damodaran, V., Adamson, G., Hannaford, M., Carroll, F., … Bliuc, A.-M. (2016). Toward integrated historical climate research: The example of atmospheric circulation reconstructions over the earth. *WIREs Climate Change*, *7*(2), 164–174. https://doi.org/10.1002/wcc.379

Allen, R. C. (2001). The great divergence in European wages and prices from the middle ages to the first world war. *Explorations in Economic History*, *38*(4), 411–447. https://doi.org/10.1006/exeh.2001.0775

Battisti, D. S., & Naylor, R. L. (2009). Historical warnings of future food insecurity with unprecedented seasonal heat. *Science*, *323*(5911), 240–244. https://doi.org/10.1126/science.1164363

Ben-Ari, T., Gershunov, A., Gage, K. L., Snäll, T., Ettestad, P., Kausrud, K. L., … Stenseth, N. C. (2008). Human plague in the USA: The importance of regional and local climate. *The Royal Society: Biology Letters*, *4*(6), 737–740. https://doi.org/10.1098/rsbl.2008.0363

Ben-Ari, T., Neerinckx, S., Agier, L., Cazelles, B., Xu, L., Zhang, Z., … Stenseth, N. C. (2012). Identification of Chinese plague foci from long-term epidemiological data. *Proceedings of the National Academy of Sciences*, *109*(21), 8196–8201. https://doi.org/10.1073/pnas.1110585109

Beveridge, W. (1922). Wheat prices and rainfall in Western Europe. *Journal of the Royal Statistical Society*, *85*(3), 412–478.

Biraben, J. N. (1975/6). *Les hommes et la peste en France et dans les pays Européens et Méditerranéens* (Vol. 2). Paris, France: Mouton.

Brádzil, R., Pfister, C., Wanner, H., Von Storch, H., & Luterbacher, J. (2005). Historical climatology in Europe – The state of the art. *Climate Change*, *70*(3), 363–430. https://doi.org/10.1007/s10584-005-5924-1

Brecke, P. (1999). *Violent conflicts 1400 A.D. to the present in different regions of the world*. Centre for Global Economic History. Retrieved from http://www.cgeh.nl/sites/default/files/Conflict%20Catalog%2018%20vars.xls

Brecke, P. (2012). *Notes regarding the conflict catalog*. Centre for Global Economic History. Retrieved from www.cgeh.nl/sites/default/files/Notes%20about%20Conflict%20Catalog.pdf

Brönnimann, S., & Wintzer, J. (2019). Climate data empathy. *WIREs Climate Change*, *10*(2), e559. https://doi.org/10.1002/wcc.559

Brook, T. (2017). Differential effects of global and local climate data in assessing environmental drivers of epidemic outbreaks. *Proceedings of the National Academy of Sciences*, *114*(49), 12845–12847. https://doi.org/10.1073/pnas.1717723114

Buhaug, H. (2010). Climate not to blame for African civil wars. *Proceedings of the National Academy of Sciences*, *107*(38), 16477–16782. https://doi.org/10.1073/pnas.1005739107

Buhaug, H., Nordkvelle, J., Bernauer, T., Böhmelt, T., Brzoska, M., Busby, J. W., … von Uexkull, N. (2014). One effect to rule them all? A comment on climate and conflict. *Climatic Change*, *127*(3-4), 391–397. https://doi.org/10.1007/s10584-014-1266-1

Büntgen, U., Tegel, W., Nicolussi, K., McCormick, M., Frank, D., Trouet, V., … Esper, D. (2011). 2500 years of European climate variability and human susceptibility. *Science*, *331*(6017), 578–582. https://doi.org/10.1126/science.1197175

Büntgen, U., Ginzler, C., Esper, J., Tegel, W., & McMichael, J. (2012). Digitizing historical plague. *Clinical Infectious Diseases*, *55*(11), 1586–1588. https://doi.org/10.1093/cid/cis723

Büntgen, U., Kyncl, T., Ginzler, C., Jacks, D.S., Esper, J., Tegel, W., … Kyncl, J. (2013). Filling the Eastern European gap in millennium-long temperature reconstructions. *Proceedings of the National Academy of Sciences*, *110*(5), 1773–1778. https://doi.org/10.1073/pnas.1211485110

Burke, M. B., Miguel, E., Satyanath, S., Dykema, J. A., & Lobell, D. B. (2009). Warming increases the risk of civil war in Africa. *Proceedings of the National Academy of Sciences*, *106*(49), 20670–20674. https://doi.org/10.1073/pnas.0907998106

Butzer, K. W. (2012). Collapse, environment, and society. *Proceedings of the National Academy of Sciences*, *109*(10), 3632–36392. https://doi.org/10.1073/pnas.1114845109

Camenisch, C., Keller, K. M., Salvisberg, M., Amann, B., Bauch, M., Blumer, B., … Wetter, O. (2016). The 1430s: A cold period of extraordinary internal climate variability during the early Spörer minimum with social and economic impacts in north-western and Central Europe. *Climate of the Past*, *12*(11), 2107–2126. https://doi.org/10.5194/cp-12-2107-2016

Campbell, B. (2007). *Three centuries of English crop yields, 1211-1491*. Medieval Crop Yields Database. Retrieved from www.cropyields.ac.uk/

Carleton, T. A., & Hsiang, S. M. (2016). Social and economic impacts of climate. *Science*, *353*(6304), aad9837. https://doi.org/10.1126/science.aad9837

Contreras, D. A. (2017). Correlation is not enough: Building better arguments in the archaeology of human-environment interactions. In D. A. Contreras (Ed.), *The archaeology of human-environment interactions: Strategies for investigating anthropogenic landscapes, dynamic environments, and climate change in the human past* (pp. 3–22). Oxon, England: Routledge.

Cook, B., & Wolkovich, E. (2016). Climate change decouples drought from early wine grape harvests in France. *Nature Climate Change*, *6*(7), 715–720. https://doi.org/10.1038/nclimate2960

Curtis, D. R. (2016). Was plague an exclusively urban phenomenon? Plague mortality in the seventeenth-century Low Countries. *Journal of Interdisciplinary History*, *47*(2), 139–170. https://doi.org/10.1162/JINH_a_00975

Curtis, D. R., & Dijkman, J. (2019). The escape from famine in the Northern Netherlands: A reconsideration using the 1690s harvest failures and a broader Northwest European perspective. *The Seventeenth Century*, *34*(2), 229–258. https://doi.org/10.1080/0268117X.2017.1410494

Daux, V., de Cortazar-Atauri, I. G., Yiou, P., Chuine, I., Garnier, E., Ladurie, E. L. R., … Tardaguila, J. (2012). An open-access database of grape harvest dates for climate research: Data description and quality assessment. *Climate of the Past*, *8*(5), 1403–1418. https://doi.org/10.5194/cp-8-1403-2012

De Dreu, C. K. W., & van Dijk, M. A. (2018). Climatic shocks associate with innovation in science and technology. *PLoS One*, *13*(1), e0190122. https://doi.org/10.1371/journal.pone.0190122

Degroot, D. (2018). Climate change and conflict. In S. White, C. Pfister, & F. Mauelshagen (Eds.), *The Palgrave handbook of climate history* (pp. 367–385). London, England: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-43020-5_29

Degroot, D. (2019). Climate change and society in the 15th to 18th centuries. *WIREs Climate Change*, *9*(3), e518. https://doi.org/10.1002/wcc.518

de Vries, J. (1980). Measuring the impact of climate on history: The search for appropriate methodologies. *Journal of Interdisciplinary History*, *10*(4), 599–630. https://doi.org/10.2307/203061

Drake, D. L. (2017). Changes in North Atlantic oscillation drove population migrations and the collapse of the Western Roman empire. *Scientific Reports*, *7*, 1227. https://doi.org/10.1038/s41598-017-01289-z

Emile-Geay, J., McKay, N. P., Kaufman, D. S., vonGunten, L., Wang, J., Anchukaitis, K. J., … Zinke, J. (2017). A global multiproxy database for temperature reconstructions of the common era. *Scientific Data*, *4*, aad17088. https://doi.org/10.1038/sdata.2017.88

Endfield, G. H. (2012). The resilience and adaptive capacity of social-environmental systems in colonial Mexico. *Proceedings of the National Academy of Sciences*, *109*(10), 3676–3681. https://doi.org/10.1073/pnas.1114831109

Federico, G. (2005). *Feeding the world. An economic history of agriculture, 1800–2000*. Princeton & Oxford, England: Princeton University Press.

Freeman, E., Woodruff, S. D., Worley, S. J., Lubker, S. J., Kent, E. C., Angel, W. E., … Smith, S. R. (2017). ICOADS release 3.0: A major update to the historical marine climate record. *International Journal of Climatology*, *37*(5), 2211–2232. https://doi.org/10.1002/joc.4775

Gartzke, E. (2012). Could climate change precipitate peace? *Journal of Peace Research*, *49*(1), 177–192. https://doi.org/10.1177/0022343311427342

Gleditsch, N. P. (2012). Whither the weather? Climate change and conflict. *Journal of Peace Research*, *49*(1), 3–9. https://doi.org/10.1177/0022343311431288

Gómez, J. M., & Verdú, M. (2017). Network theory may explain the vulnerability of medieval human settlements to the Black Death pandemic. *Scientific Reports*, *7*, 43467. https://doi.org/10.1038/srep43467

Green, M. (2015). Editor's introduction. In M. Green (Ed.), *Pandemic disease in the medieval world: Rethinking the Black Death* (pp. 9–26). Leeds, England: Arc Medieval Press.

Haldon, J. (2016). Cooling and societal change. *Nature Geoscience*, *9*(3), 191–192. https://doi.org/10.1038/ngeo2659

Haldon, J., Mordechai, L., Newfield, T. P., Chase, A. F., Izdebski, A., Guzowski., P., … Roberts, N. (2018). History meets palaeoscience: Consilience and collaboration in studying past societal responses to environmental change. *Proceedings of the National Academy of Sciences*, *115*(13), 3210–3218. https://doi.org/10.1073/pnas.1716912115

Haldon, J., Elton, H., Huebner, S. R., Izdebski, A., Mordechai, L., & Newfield T. P. (2018). Plagues, climate change, and the end of an empire: A response to Kyle Harper's The Fate of Rome (3): Disease, agency, and collapse. *History Compass*, *16*(12), e12507. https://doi.org/10.1111/hic3.12507

Hannaford, M. J., & Nash, D. J. (2016). Climate, history, society over the last millennium in Southeast Africa. *WIREs Climate Change*, *7*(3), 370–392. https://doi.org/10.1002/wcc.389

Harper, K. (2017). *The fate of Rome: Climate, disease, and the end of an empire*. Princeton, NJ: Princeton University Press.

Helama, S., Helama, S., Arppe, L., Uusitalo, J., Holopainen, J., Mäkelä, H. M., Mäkinen, H., … Oinonen, M. (2018). Volcanic dust veils from sixth century tree-ring isotopes linked to reduced irradiance, primary production and human health. *Scientific Reports*, *8*, 1339. https://doi.org/10.1038/s41598-018-19760-w

Holm, P., & Winiwarter, V. (2017). Climate change studies and the human sciences. *Global and Planetary Change*, *156*, 115–122. https://doi.org/10.1016/j.gloplacha.2017.05.006

Hsiang, S. M., Burke, E., & Miguel, E. (2013). Quantifying the influence of climate on human conflict. *Science*, *341*(6151), 1235367. https://doi.org/10.1126/science.1235367

Hulme, M. (2011). Reducing the future to climate: A story of climate determinism and reductionism. *Osiris*, *26*, 245–266. https://doi.org/10.1086/661274

Kaniewski, D., Marriner, N., Ilan, D., Morhange, C., Thareani, Y., & Van Campo E. (2017). Climate change and water management in the biblical city of Dan. *Science Advances*, *3*(11), e1700954. https://doi.org/10.1126/sciadv.1700954

King, H., & Green, M. (2018). On the misuses of medical history. *The Lancet*, *391*(10128), 1354–1355. https://doi.org/10.1016/S0140-6736(18)30490-2

Klomp, J., & Bulte, E. (2013). Climate change, weather shocks, and violent conflict: A critical look at the evidence. *Agricultural Economics*, *44*(s1), 63–78. https://doi.org/10.1111/agec.12051

Kohn, G. C. (1999). *Dictionary of wars* (Revised ed.). New York, NY: Facts On File.

Kreppel, K. S., Caminade, C., Telfer, S., Rajerison, M., Rahalison, L., Morse, A., …. Baylis, M. (2014). A non-stationary relationship between global climate phenomena and human plague incidence in Madagascar. *PLoS Neglected Tropical Diseases*, *8*(10), e3155. https://doi.org/10.1371/journal.pntd.0003155

Kukal, M. S., & Irmak, S. (2018). U.S. agro-climate in 20th century: Growing degree days, first and last frost, growing season length, and impacts on crop yields. *Scientific Reports*, *8*, 6977. https://doi.org/10.1038/s41598-018-25212-2

Kwok, R. (2017). Historical data: Hidden in the past. *Nature*, *549*, 419–421. https://doi.org/10.1038/nj7672-419

Lee, H. F., & Zhang, D. D. (2015). Quantitative analysis of climate change and human crises in history. In M. P. Kwan, D. Richardson, D. Wang, & C. Zhou (Eds.), *Space-time integration in geography and GIScience* (pp. 235–267). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-94-017-9205-9_14

Lee, H. F., Fei, J., Chan, C. Y., Pei, Q., Jia, X., & Yue, R. P. (2017). Climate change and epidemics in Chinese history: A multi-scalar analysis. *Social Science & Medicine*, *174*, 53–63. https://doi.org/10.1016/j.socscimed.2016.12.020

Lewnard, J. A., & Townsend, J. P. (2016). Climatic and evolutionary drivers of phase shifts in the plague epidemics of colonial India. *Proceedings of the National Academy of Sciences*, *113*(51), 14601–14608. https://doi.org/10.1073/pnas.1604985113

Luard, E. (1986). *War in international society: A study in international sociology*. London, England: Tauris.

Manning, J. G., Ludlow, F., Stine, A. R., Boos, W. R., Sigl, M., & Marlon, J. R. (2017). Volcanic suppression of Nile summer flooding triggers revolt and constrains interstate conflict in ancient Egypt. *Nature Communications*, *8*(1), 900. https://doi.org/10.1038/s41467-017-00957-y

Mauelshagen, F. (2018). Migration and climate in world history. In S. White, C. Pfister, & F. Mauelshagen (Eds.), *The Palgrave handbook of climate history* (pp. 413–444). London, England: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-43020-5_31

McCann, J. C. (1999). Climate and causation in African history. *International Journal of African Historical Studies*, *32*(2/3), 261–279. https://doi.org/10.2307/220342

McEvedy, C., & Jones, R. (1978). *Atlas of world population history*. London, England: Allen Lane.

McMichael, M. (2012). Insights from past millennia into climatic impacts on human health and survival. *Proceedings of the National Academy of Sciences*, *109*(13), 4730–4737. https://doi.org/10.1073/pnas.1120177109

Mokyr, J. (2016). *A culture of growth: The origins of the modern economy*. Princeton, NJ: Princeton University Press.

Nash, D. J., De Cort, G., Chase, B. M., Verschuren, D., Nicholson, S. E., Shanahan, T. M., … Grab, S. W. (2016). African hydroclimatic variability during the last 2000 years. *Quaternary Science Reviews*, *154*, 1–22. https://doi.org/10.1016/j.quascirev.2016.10.012

Nelson, M. C., Ingram, S. E., Dugmore, A. J., Streeter, R., Peeples, M. A., McGovern, T. H., … Spielmann, K. A. (2016). Climate challenges, vulnerabilities, and food security. *Proceedings of the National Academy of Sciences*, *113*(2), 298–303. https://doi.org/10.1073/pnas.1506494113

Olmstead, A. L., & Rhode, P. W. (2011). Adapting north American wheat production to climatic changes. *Proceedings of the National Academy of Sciences*, *108*(2), 480–485. https://doi.org/10.1073/pnas.1008279108

Paprotny, D., Sebastian, A., Morales-Nápoles, O., & Jonkman, S. N. (2018). Trends in flood losses in Europe over the past 150 years. *Nature Communications*, *29;9*(1), 1985. https://doi.org/10.1038/s41467-018-04253-1

Pei, Q., Zhang, D., Li, G., & Lee, H. (2013). Short- and long-term impacts of climate variations on the agrarian economy of pre-industrial Europe. *Climate Research*, *56*(2), 169–180. https://doi.org/10.3354/cr01145

Pei, Q., Zhang, D., Lee, H., & Li, G. (2014). Climate change and macro-economic cycles in pre-industrial Europe. *PLoS One*, *9*(2), e88155. https://doi.org/10.1371/journal.pone.0088155

Pei, Q., Zhang, D., Li, G., & Lee, H. (2015). Climate change and the macroeconomic structure in pre-industrial Europe: New evidence from wavelet analysis. *PLoS One*, *10*(6), e0126480. https://doi.org/10.1371/journal.pone.0126480

Pribyl, K., Cornes, R., & Pfister, C. (2012). Reconstructing medieval April-July mean temperatures in East Anglia. *Climate Change*, *113*(2), 393–412. https://doi.org/10.1007/s10584-011-0327-y

Roosen, J., & Curtis, D. R. (2019). The 'light touch' of the Black Death in the southern Netherlands: An urban trick? *The Economic History Review*, *72*(1), 32–56. https://doi.org/10.1111/ehr.12667

Sangster, H., Jones, C., & Macdonald, N. (2018). The co-evolution of historical source materials in the geophysical, hydrological and meteorological sciences: Learning from the past and moving forward. *Progress in Physical Geography*, *42*(1), 61–82. https://doi.org/10.1177/0309133317744738

Schmid, B. V., Büntgen, U., Easterday, W. R., Ginzler, C., Walløe, L., Bramanti, B., … Stenseth, N. C. (2015). Climate-driven introduction of the Black Death and successive plague reintroductions into Europe. *Proceedings of the National Academy of Sciences*, *112*(10), 3020–3025. https://doi.org/10.1073/pnas.1412887112

Selby, J. (2014). Positivist climate conflict research: A critique. *Geopolitics*, *19*(4), 829–856. https://doi.org/10.1080/14650045.2014.964865

Shennan, S., Downey, S. S., Timpson, A., Edinborough, K., Colledge, S., Kerig, T., … Thomas, M. G. (2013). Regional population collapse followed initial productivity booms in mid-Holocene Europe. *Nature Communications*, *4*, 2486. https://doi.org/10.1038/ncomms3486

Simonton, D. K. (1980). Techno-scientific activity and war: A yearly time-series analysis 1500-1903 AD. *Scientometrics*, *2*(4), 251–255. https://doi.org/10.1007/BF02016346

Slavin, P. (2016). Climate and famines: A historical reassessment. *WIREs Climate Change*, *7*(3), 433–447. https://doi.org/10.1002/wcc.395

Slicher van Bath, B. (1963). *Yield ratios, 810–1820*. Afdeling Agrarische Geschiedenis, Landbouwhogeschool: Wageningen, Netherlands.

Stenseth, N. C., Samia, N. I., Viljugrein, H., Kausrud, K. L., Begon, M., Davis, S., … Klassovskiy, N. L. (2006). Plague dynamics are driven by climate variation. *Proceedings of the National Academy of Sciences*, *103*(35), 13110–13115. https://doi.org/10.1073/pnas.0602447103

Stenseth, N. C., Atshabar, B. B., Begon, M., Belmain, S. R., Bertherat, E., Carniel, E., … Rahalison, L. (2008). Plague: Past, present, and future. *PLoS Medicine*, *5*(1), 9–13. https://doi.org/10.1371/journal.pmed.0050003

Streeter, R., Dugmore, A. J., & Vésteinsson, O. (2012). Plague and landscape resilience in premodern Iceland. *Proceedings of the National Academy of Sciences*, *109*(10), 3664–3669. https://doi.org/10.1073/pnas.1113937109

Tan, L., Cai, Y., An, Z., Cheng, H., Shen, C. C., Breitenbach, S. F., … Du, Y. (2015). A Chinese cave links climate change, social impacts, and human adaptation over the last 500 years. *Scientific Reports*, *5*, 12284. https://doi.org/10.1038/srep12284

Theilmann, J., & Cate, F. (2007). A plague of plagues: The problem of plague diagnosis in medieval England. *Journal of Interdisciplinary History*, *37*(3), 371–393. https://doi.org/10.1162/jinh.2007.37.3.371

Thoen, E., & Soens, T. (2015). Contextualizing 1500 years of agricultural productivity and land use in the North Sea area. Regionally divergent paths towards the world top. In E. Thoen & T. Soens (Eds.), *Struggling with the environment: Land use and productivity* (pp. 455–499). Turnhout, Belgium: Brepols. https://doi.org/10.1484/M.RES-EB.5.108045

Tian, H., Yan, C., Xu, L., Büntgen, U., Stenseth, N. C., & Zhang, Z. (2017). Scale-dependent climatic drivers of human epidemics in ancient China. *Proceedings of the National Academy of Sciences*, *114*(39), 12970–12975. https://doi.org/10.1073/pnas.1706470114

Tol, R. S. J., & Wagner, S. (2010). Climate change and violent conflict in Europe over the last millennium. *Climatic Change*, *99*(1-2), 65–79. https://doi.org/10.1007/s10584-009-9659-2

Webb, J. L. A. (2018). Climate, ecology, and infectious human disease. In S. White, C. Pfister, & F. Mauelshagen (Eds.), *The Palgrave handbook of climate history* (pp. 355–365). London, England: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-43020-5_28

Welford, M. R., & Bossak, B. H. (2009). Validation of inverse seasonal peak mortality in medieval plagues, including the Black Death, in comparison to modern Yersinia pestis-variant diseases. *PLoS One*, *4*(12), e8401. https://doi.org/10.1371/journal.pone.0008401

White, S. (2018). Climate, weather, agriculture, and food. In S. White, C. Pfister, & F. Mauelshagen (Eds.), *The Palgrave handbook of climate history* (pp. 331–353). London, England: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-43020-5_27

Wickham, C. (2010). *The inheritance of Rome: Illuminating the dark ages, 400–1000*. London, England: Penguin.

Wig, T. (2016). Peace from the past: Pre-colonial political institutions and civil wars in Africa. *Journal of Peace Research*, *53*(4), 509–524. https://doi.org/10.1177%2F0022343316640595

World Health Organization (WHO). (2009). *Operational guidelines on plague surveillance, diagnosis, prevention and control*. New Delhi, India: WHO Regional Office for South-East Asia.

Wright, Q. (1942). *A study of war* (Vol. 2). Chicago, IL: University of Chicago Press.

Xu, L., Liu, Q., Stige, L. C., Ari, T. B., Fang, X., Chan, K. S., … Zhang, Z. (2011). Nonlinear effect of climate on plague during the third pandemic in China. *Proceedings of the National Academy of Sciences*, *108*(25), 10214–10219. https://doi.org/10.1073/pnas.1019486108

Xu, L., Stige, L.C., Kausrud, K.L., Ben Ari, T., Wang, S., Fang, X., … Zhang, Z.. (2014). Wet climate and transportation routes accelerate spread of human plague. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1780), 1–9. https://dx.doi.org/10.1098%2Frspb.2013.3159

Yin, J., Su, Y., & Fang, X. (2016). Climate change and social vicissitudes in China over the past two millennia. *Quaternary Research*, *86*(2), 133–143. https://doi.org/10.1016/j.yqres.2016.07.003

Yue, R. P. H., Lee, H. F., & Wu, C. Y. H. (2016). Navigable rivers facilitated the spread and recurrence of plague in pre-industrial Europe. *Scientific Reports*, *6*, 34867. https://doi.org/10.1038/srep34867

Yue, R. P. H., Lee, H. F., & Wu, C. Y. H. (2017). Trade routes and plague transmission in pre-industrial Europe. *Scientific Reports*, *7*, 12973. https://doi.org/10.1038/s41598-017-13481-2

Yue, R. P. H., & Lee, H. F. (2018). Pre-industrial plague transmission is mediated by the synergistic effect of temperature and aridity index. *BMC Infectious Diseases*, *18*(1), 134. https://doi.org/10.1186/s12879-018-3045-5

Zhang, D. D., Brecke, P., Lee, H. F., He, Y. Q., & Zhang, J. (2007). Global climate change, war, and population decline in recent human history. *Proceedings of the National Academy of Sciences*, *104*(49), 19214–19219. https://doi.org/10.1073/pnas.0703073104

Zhang, D. D., Zhang, J., Lee, H. F., & He, Y. (2007). Climate change and war frequency in eastern China over the last millennium. *Human Ecology*, *35*(4), 403–414. https://doi.org/10.1007/s10745-007-9115-8

Zhang, D. D., Lee, H. F., Wang, C., Li, B., Pei, Q., Zhang, J., … An, Y. (2011). The causality analysis of climate change and large-scale human crisis. *Proceedings of the National Academy of Sciences*, *108*(42), 17296–17301. https://doi.org/10.1073/pnas.1104268108

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.