

In my opinion, this manuscript constitutes a valuable and potentially important contribution to European historical climatology, especially due to the considerable work of recovering, critiquing, and systematizing documentary sources carried out by the authors. The construction of a database with 542 entries and 525 episodes of strong wind, supplemented by 14 years of complete daily observations, provides exceptional information on a phenomenon scarcely studied in 17th-century Poland and allows for the analysis of its seasonality, regional distribution, intensity, and effects on buildings, forests, and society.

Reply. Thank you for your good opinion about our work.

However, its main weakness lies in the fact that the availability, resolution, and reliability of the sources vary substantially throughout the century, meaning that the reconstructed frequencies may reflect both meteorological variability and changes in documentary coverage. This problem does not invalidate the study, but it does require a clearer distinction between the general inventory of the century, the periods with daily information, the case studies, and the records of varying quality.

Reply. We fully agree with the Reviewer's opinion. This is a common weakness of elaborations based on the documentary evidence. Nevertheless, as the reviewer also emphasises, this type of proxy data is crucial for understanding the climate of past epochs before the advent of instrumental observations. Regarding the strong winds analysed in the article, identifying high-velocity winds associated with damage is far more reliable than identifying winds below 17 ms^{-1} , which do not cause such consequences and are frequently overlooked by chroniclers. Such lower-speed winds were typically recorded only in daily weather observations. Therefore, as we emphasise in the article, this constitutes the primary advantage of this type of data and represents a novel finding not previously published. It allows for an estimation of the biases inherent in the irregular information provided by traditional documentary evidence.

In the database attached to the manuscript (<https://repod.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/ESPWII>), the reader can easily find a clearer distinction between strong wind events identified from irregular documentary evidence and those identified from regular daily weather observations. The latter comes from only two sources and two areas: Gdańsk (Buthner, 1693) and northeastern Poland (Chrapowicki, 1978, 1988). We would like to clarify that we did not evaluate the quality of the weather records themselves, but rather the quality of the sources from which they originate. While weather records contain varying amounts of information about the meteorological element of interest, assessing the quality of an individual record is difficult.

Furthermore, it should be explained more reproducibly how historical expressions such as "gale," "storm," or "squall" are transformed into the five intensity categories and to what extent these can be compared with modern instrumental observations.

Reply. Thank you for this comment. Reviewer 1 made a similar comment, and we addressed it in detail in point 2; therefore, we do not repeat that response here.

The novelty compared to the work of Przybylak et al. (2025) should also be better defined: it lies not so much in a new methodology as in the chronological expansion of the database, the

incorporation of daily series, and the analysis of new historical evidence. In short, I consider it a solid manuscript, original in its sources and worthy of publication, but it needs a significant methodological revision, an explicit assessment of documentary bias, and more cautious conclusions, referring to "documented episodes during the 17th century" and not to the homogeneous climatic behavior of the century as a whole.

Reply. Thank you for these suggestions. We agree with your assessment of the novelty of our paper, and we emphasised this more clearly in the revised version of the manuscript.

The main weaknesses that I find in this manuscript are the following:

A. Event frequency is confounded with source availability. This is the most important weakness. The statement that most reconstructed events occurred in the second half of the century may reflect some combination of actual meteorological variability, more surviving sources, higher temporal resolution, the emergence of daily observations, geographical differences between observers and preferential documentation of damaging events. Unless documentary coverage is quantified, a graph of event counts may partly be a graph of source density. The manuscript should avoid treating raw counts as direct climatic frequencies without an explicit correction, sensitivity analysis or strong qualification.

Reply. Thank you for these comments. All researchers working with historical data to reconstruct meteorological variables are aware of these limitations, and we note them in our work as well. Nevertheless, we propose adding the following passage to the *Discussion and conclusion* section that emphasises this point even more strongly.

It should be noted that the presented analysis, in particular of the absolute number of strong wind occurrences and their consequences, requires cautious interpretation due to the strong correlation between source availability and recorded instances of strong winds, particularly when it is based on information derived from irregular documentary evidence. It is also worth emphasising that the biases are significantly lower for very strong winds that caused damage to buildings, trees, orchards, and the like; one reason is that chroniclers sometimes recorded only damaging events, omitting strong winds that caused no destruction. Significantly more reliable information about the occurrence of strong winds is available from continuous daily documentary evidence. But this kind of source also has an important weakness: the information about strong winds is limited to a single location or a small area. Therefore, they do not allow for an analysis of the phenomenon's spatial distribution across the entire country. For this reason, combining both types of sources is the only reasonable methodological approach. For more details on the strengths and weaknesses of using documentary evidence in historical climatology, please see Brázdil et al. (2005).

B. The different analytical datasets are not sufficiently separated. The manuscript appears to combine or alternate among different sub-databases. Each subset answers a different question. Presently, the boundary between them is not always obvious.

Reply. Thank you for this comment. A note has been added to the manuscript stating that all available documentary evidence was used to analyse strong wind events across Poland. Meanwhile, for specific locations or areas (Gdańsk, northeastern part of Poland), the daily data are clearly separated and labelled as case studies. Please also see our reply to your previous comments, which outlines the strengths and weaknesses of the information about strong winds available in the daily data and where to find it in the database.

C. The classification system is not reproducible enough. A reader needs a decision protocol showing how a documentary description becomes one of the five intensity categories.

For example, classification could depend on the historical wind term, movement or destruction of objects, damage to buildings, uprooting of trees, effects on ships or water bodies, duration, spatial extent or an explicit contemporary comparison. If these indicators disagree, the manuscript should state which takes precedence. Author can see a remarkable example in the CLIWOC dictionary (García-Herrera, R., Prieto, L., Gallego, D., Hernández, E., Gimeno, L., Können, G.P., Koek, F.B., Wheeler, D.A., Wilkinson, C., Prieto, M.R., Báez, C. and Woodruff, S. (2003): CLIWOC multilingual meteorological dictionary, EU contract EVK2-CT-2000-00090; García-Herrera, R., Können, G.P., Wheeler, D.A. et al. CLIWOC: A Climatological Database for the World's Oceans 1750–1854. *Climatic Change* 73, 1–12 (2005). <https://doi.org/10.1007/s10584-005-6952-6>).

Reply. Thank you for this comment. A response to this point has already been provided by us to the first reviewer under item 2. Additionally, in the methodology section of the article, we explained the procedure used to classify specific strong wind events into intensity categories or to categorise them based on the extent and type of damage (see text below):

“The classification of strong winds into the above categories was originally made by the team member entering the record of strong wind occurrences into the database. In the following step, all propositions were discussed and ultimately accepted by all team members. The same procedure was also used to estimate the degree of damage caused by strong winds.”

D. Risk of circularity between intensity and damage. If damage is used to assign wind intensity and the paper later analyses damage by intensity category, the two variables are not independent. The authors need to make clear whether intensity classification was derived independently of the damage variable.

Reply. Thank you for this comment. We believe the reviewer misinterpreted the method we employed. We assessed the intensity of strong winds based on the damage caused—exactly as is done when using the Beaufort scale—rather than the other way around. **The article does not present the extent and nature of the damage categorized by intensity categories.**

E. Insufficient treatment of uncertainty. The three-level source-quality score is useful, but it is not the same as uncertainty in event date, event location, event duration, intensity, event identity, translation, or meteorological interpretation. A high-quality contemporary source may still use ambiguous terminology. Conversely, a later source may preserve a precise description. Separate confidence fields would improve the analysis.

Reply. Thank you for this comment. There appears to be a misunderstanding regarding the nature of the sources here as well. In research relying on documentary evidence, the most reliable accounts of the weather are those recorded by eyewitnesses or by individuals who had direct contact with such witnesses. Later sources and compilations merely repeat the original material; they are often abridged, translated, or misinterpreted, rendering them significantly less reliable. However, instances do occur where the original, earliest source has been lost, and the content is available only through later compilations—where an older compilation might

actually contain less information than a more recent one (as is the case, for example, with surviving 18th- and 19th-century copies of Jan Chrapowicki's diary). In such cases, it is justifiable to use a source that is technically less reliable but contains information closer to the original. We adopted this approach, incorporating all relevant information—sometimes drawn from multiple sources—into the indexing process.

F. Modern comparison may suggest greater comparability than exists. Historical observers and modern stations do not share the same instruments, exposure, sampling frequency, terminology, and so on. Thus, the modern comparison is defensible as a contextual or calibration exercise, but not necessarily as a direct comparison of climatological frequencies.

Reply. Thank you for this comment. We absolutely agree with your opinion. Therefore, in our study, we explicitly state and emphasise that the primary basis for comparison is the annual pattern and spatial distribution of strong winds across Poland. Only by using daily data is it possible to capture virtually all extreme wind events, including those in the category that do not cause damage (fresh and strong breeze).

G. Novelty is not articulated clearly. Because the methodology deliberately follows the earlier paper, the novelty must be presented as new evidence, improved temporal resolution, new historical sources and new substantive findings. Without that framing, the manuscript can look like the previous analysis shifted forward by one century.

Reply. Thank you for this comment. We already replied earlier to this suggestion; see the beginning of your review.

H. The Discussion is broad but not sufficiently analytical. The Discussion would benefit from organising itself around three or four testable questions, rather than reviewing numerous possible explanations.

Reply. Thank you for this comment. However, it is not entirely clear to us what exactly the reviewer expects—specifically, which questions need to be answered. Discussing the issue at hand is challenging due to the scarcity of available studies reconstructing strong winds in Europe—not to mention the availability of comprehensive databases like the one accompanying this manuscript. To the best of our knowledge, we have included the majority of currently available research findings, which vary in detail. In the discussion, we addressed, among other things, the uncertainty in the results and the benefits of using daily weather data; finally, we compared our findings with other results available for Europe.

I. Terminology and grammar occasionally overstate the scope. Expressions such as “strong winds in Poland in the seventeenth century” are acceptable as a topic label, but conclusions should preferably use formulations such as: (i) “documented strong-wind events during the seventeenth century”; (ii) “the available seventeenth-century evidence”; (iii) “the analysed documentary record”; (iv) or, where appropriate, “selected periods of the later seventeenth century.” This would avoid implying complete and homogeneous century-long observation.

Reply. Thank you for this comment. We have introduced some of the proposed terminology where appropriate.

My main proposals to the authors are:

1. The single most effective improvement: add a source-to-result audit framework. I would recommend adding both a workflow figure and an accompanying dataset table.

Reply. Thank you for these suggestions; however, the mechanism for indexing and classifying strong winds by intensity and the extent of damage is briefly described in the text, so there is no need to add a figure to an already lengthy article. Regarding the data table, we opted instead to use colour-coding in the database to distinguish daily data from other entries.

2. Reframe the title cautiously. A possible title would be: "Documented strong winds in seventeenth-century Poland and their consequences: evidence from historical sources". "Documented" immediately signals that the study analyses the surviving record, not a homogeneous instrumental climatology.

Reply. Thank you for this suggestion. We decided to change the title a little to:

Strong winds in seventeenth-century Poland and their consequences based on documentary evidence.

In historical climatology, *documentary evidence* is a well-established and widely used term; therefore, we prefer to retain it in the title of our paper.

On the other hand, your proposition is more concise and modern, but the word "documented" can be somewhat ambiguous. It could imply either "documented instances of strong winds" or "strong winds described in documents." The final phrase, "evidence from historical sources," partially repeats the meaning conveyed by "documented."

3. Add a documentary-coverage figure. For each year, show: (i) number of sources available; (ii) number of days covered by daily records; (iii) number of wind-related notes; (iv) number of accepted events; and (v) proportion of records in each quality class. This could reveal whether peaks in reconstructed events coincide with peaks in source availability.

Reply. Thank you for this suggestion. However, in our view, performing such detailed calculations to document the link between the number of strong wind events and the number of available sources is unnecessary, as we—along with other researchers specialising in historical climatology—have repeatedly demonstrated this in numerous studies.

4. Moderate causal interpretations. The Maunder Minimum and Little Ice Age provide important climatic context, but the manuscript should not imply that low solar activity directly explains the reconstructed wind patterns unless an atmospheric mechanism is tested. Temporal coincidence is not sufficient. Note that the Maunder minimum is not properly cited (J. Eddy, 1976 "The Maunder Minimum" *Science* 192, Issue 4245, pp. 1189-1202, DOI: 10.1126/science.192.4245.118; I. G. Usoskin et al., 2015, "The Maunder Minimum (1645–1715) was indeed a grand minimum: A reassessment of multiple datasets" *A&A*, 581 (2015) A95, DOI: <https://doi.org/10.1051/0004-6361/201526652>).

Reply. Thank you for these suggestions. However, we do not see in our manuscript a statement such as the one you write, namely that "low solar activity directly explains the reconstructed wind patterns". On the other hand, we added two of your literature propositions to the references.

5. Reorganise the paper. A clearer structure would be: (i) Introduction, (ii) Study área, (iii) Documentary sources, (iv) Data construction and source criticism, (v) Wind classification and analytical methods, (vi) Results, (vii) Discussion and (viii) Conclusions.

Reply. Thank you for this suggestion. In our opinion, however, the proposed structure is overly detailed; our current format is more typical of scientific journals. Therefore, and also considering that the other reviewers did not suggest such changes, we would prefer to retain the article's current structure, as it better aligns with the subject matter and the scope of the content.

Dear Reviewer,

To conclude our response, we would like to thank you for finding the time to read our paper. Thank you very much for your constructive comments and suggestions. We hope we have adequately addressed at least most of them.