

Review of An open-database of Grape Harvest dates for climate research: data description and quality assessment

by

V. Daux, I. Garcia de Cortazar-Atauri, P. Yiou, I. Chuine, E. Garnier, E. Le Roy Ladurie, O. Mestre, and J. Tardaguila

For Climates of the Past

General Comments:

This paper introduces an important update to, and compilation into an easily accessible database, a set of historical climate proxy data, in the form of annual grape harvest dates over the past five centuries in western Europe (mainly France). As such, it is an important scientific contribution and should be published. It is generally clear and well-written. The authors are to be highly commended for building a database intended to be open and easily accessible.

I have some concerns that the aim of the authors and the potential uses of this dataset are not completely clear from the text of this paper. The grape harvest dates are presented as being of interest to the climatological community, and it seems to be implicitly assumed that the harvest dates are a proxy for temperature. It seems likely, then, that users will consider the database as a temperature proxy. However, the authors state explicitly “We strongly recommend the contextualization of the series before interpreting them in terms of climate change”. This statement needs to be expanded upon and clarified: should users undertake research in local archives before using the data? How can users interpret the data in terms of temperature, if this is the intended purpose of the dataset? If this is not the intended purpose, how do the authors suggest the data to be used? If these points can be established, it would make interpretation of the paper, and the data, clearer for readers and users.

The authors also state that: “While our aim is neither to discuss the reliability of GHD as climate proxies, nor to interpret the GHD-RCS in terms of climate dynamics, we assess the quality of the series by pairwise correlations. The former point has been already debated in two papers of the same research team (Garcia de Cortazar-Atauri et al., 2010; Garnier et al., 2011), and some aspects of the latter are addressed in a companion paper (Yiou et al., 2011).”

Given the potential importance of this paper as the main reference and descriptor of a valuable database, there needs to be more discussion of the climate-GHD relationship in *this* paper to elucidate for potential users the strengths and weaknesses of the proxy indicator. As well as more discussion on how the GHD series should be interpreted, the likely errors and uncertainties inherent in using these data should also be addressed. This might be done by adding in a section summarizing some of the main points from previous papers written by the authors and cited in the text. For example, a brief discussion of the maximum estimate of 30% of the harvest dates being possibly affected by non-climatic issues as described in Garnier et al, 2010; the confidence intervals and applications of the different climate-GHD models by García de Cortázar-Atauri et al, 2010, and the uncertainty values obtained by Yiou et al, 2011 (all papers cited in the article), could be used to provide examples of the methods users might use to related GHD series to climate, and of the magnitude and type of uncertainties potential users would need to consider in using this database.

A part of the database could also include an on-going quality assessment, as various series are

used and evaluated in different studies.

The Figures, especially the Figures presenting time series, need to be improved, with proper legends, descriptors in the captions, distinguishing sub-titles and axes labels provided (see separate document with technical corrections).

Specific comments:

Abstract:

1. Lines 12-13: Vague wording in sentence “strong correlations exist between most of them”. I think the authors are referring to high correlations between the pair-wise comparisons of all regional series?

2. Line 25: The median value of all the regional series? Why is this described as a “general synthetic” series, is it not calculated from the observations?

Main text:

3. Page 3825, Line 25: Does the “its” for the standard deviation refer to the harvest date or the delay between veraison and harvest date?

4. Page 3828, line 12: See general comments. A short sentence describing the possible variation in dates due to the various changes in law affecting the GHD would be extremely helpful for potential users in assessing the suitability of this dataset for climatic studies and the potential for inhomogeneities. For example, is there an associated +/- error estimate in degrees C per of day possible change in harvest date? If this is discussed in previous papers, a reference and the citation of a few error estimates would be sufficient.

5. Page 3829, line 19: How do these changes affect or improve the estimations?

6. Page 3832, lines 22-23: What does aggregated mean in this sentence? Surely the dates are specific to the year in which the grapes are harvested: how is the time-sensitive data aggregated into other time periods? This sentence is not clear.

7. Page 3833, line 23: See general comments for a discussion of quality issues.

8. Page 3837, lines 18-21: If there are no temperature data which overlap with these series, how are they verified as a climate proxy? Could the authors discuss why they include these series, and how they expect them to be interpreted? Are biological models or other means of verifying the GHD and climate relationships used? This is quite confusing; perhaps a sentence in the introduction explaining how the authors expect these data to be used would be helpful. See general comments.

10. Page 3838, line 4: Give the values or ranges of values for the correlations in the text.

11. Page 3839, line 6: Explain in more detail the procedure by which the “general” z-score series was obtained, both in the text and in the figure caption.

12. Figure 6: it looks like there is a step discontinuity c1980, or a plateau discontinuity c1950-1980. Are the possible historical/anthropogenic reasons for this sudden change?

Technical corrections:

Figures:

1. Figure 2, Figure 4 and Figure 6: Provide a legend for each series. The y-axes labels are missing; the labels, which presumably refer the series shown, are illegible. The series can be also distinguished by labeling the panels a),b),c), etc, and describing the series in the Figure caption. The full name of the series should be written out in the Figure captions.
2. Figure 6: The series are not adequately described, in fact panel b) is not described at all, nor is the purple line in panel a). Or else panel A is not described, in either case, the caption does not describe the series. The figure needs to be explained more clearly, both in the caption and the text.

Supplementary files:

Descriptors of the data the sheets contain should be given near the first columns (i.e. “Correlation values”) of each sheet, or even better, as the sheet name.

Grammatical corrections:

1. 1. Abstract: Line 2 and elsewhere: “non-translated” not really necessary. Page 3825 : line 4 “be useful for” rather than “allow”
2. Main text: page 3826, line 1 and elsewhere: In general, there is confusion in the paper between the singular and the plural. If the GHD is to be considered as a plural noun, then the first line of the manuscript should refer to grape harvest dates (plural). Otherwise, reference should be made to GHD series or GHD indicators or a similar term. The same thing applies for references to the plants: if referring to “grape” or “vine” in the singular, an article should be attached to the noun (*a* grape or *the* vine). Otherwise the noun should be in the plural, with the verb also conjugated in the plural. See also page 3826, line 9, line 15, line 19, line 21, and throughout the text. Nouns should generally have an article, for example line 19: “*the* grapes”.
3. Page 3827, line 4: “Enlightened” may not be the correct term. “Careful” or “Considered” decision, perhaps?
4. Page 3827, line 18 “new diseases, *such as...*”
5. Page 3834, line 10: Not sure “clues” is the right word. Perhaps “results” or “investigations”.

6. Page 3835, line 12: “Pairwise correlations *between all series and the Burgundy...*”; line 14: “With *the Southern...*”; line 20, “Burgundian” should be capitalized; line 22 “most *of the time...*”
7. Page 3835, lines 12-27: Inconsistent verb tense. Present tense is fine for this discussion.
8. Page 3836, line 14: “series *may* be faulty...”; line 25: “importantly” should perhaps be “significantly”.
9. Page 3837, lines 18-21: It might help the clarity and flow of these sentences to include the full names of the series, rather than the abbreviation.
10. Page 3838, line 23: “decrease” or “decline” instead of “degrade”; line 26: “are *lower (or smaller)* than...”.
11. Page 3839, line 9: “extraction of”, not “extracting”.