

Interactive comment on “An open-database of Grape Harvest dates for climate research: data description and quality assessment” by V. Daux et al.

V. Daux et al.

valerie.daux@lsce.ipsl.fr

Received and published: 13 July 2012

Answer to reviewers

We are grateful to reviewer 1 and 2 for their comments, corrections and suggestions. We agreed on most of them. We have largely rewritten the introduction and the first part of the results-discussion section to explicit the aim of this paper and address the ‘uncertainty’ issue in more detail. Explanations of the change are given below.

Review 1

- Title: is "open-database" catchy and precise enough? In case of changes, use the

C2810

same term also in the first sentence of the abstract. It was changed in “open-access database”. The same term is now used in the first sentence

- Abstract, L5 ff.: include the temporal coverage of the data and give a mean series length and a mean ratio of observations/series length. The requested precisions are provided (underlined in the text below): “As of June 2011, this GHD dataset comprises 380 series mainly from France (93% of the data) as well as series from Switzerland, Italy, Spain and Luxemburg. The series have variable length (from 1 to 479 data, mean length of 45 data) and contain gaps of variable sizes (mean ratio of observations/series length of 0.74). The longest and most complete ones are from Burgundy, Switzerland, Southern Rhône valley, Jura and Ile-de-France. The most ancient GHD of the dataset is in 1354 in Burgundy.”

- L18: “quality of the GHD series”. For what? Elaborate here. “Quality” may not be appropriate indeed. We changed it for reliability and explained what we mean in a few words. “The GHD regional composite series (GHD-RCS) were calculated and compared pairwise to assess their reliability assuming that series close to one another must be highly correlated.”

- P3825 Delete the last paragraph of the abstract. If not include the webaddress here. We included the webaddress.

- P3826 L9: omit decimals for lon/lat descriptions. L14: start a new paragraph Accepted

- P3827 I’m not sure how the spelling and naming of laws are correct. Please check with a native speaker. If you mean a French native speaker, all the authors are French.

- P3828 L15: check sentence structure. Very complicated now. The sentence was simplified: “The database is composed of 380 individual GHD series. 93% correspond to French vineyards, the 7% left from neighbouring countries (Germany, Switzerland, Spain and Italy).”

- P3829 L7: do you plan to make accessible the original documents in the internet?

C2811

I hope so. Please explain here.

A lot of documents were used to complete the database, and we do not plan to make them accessible in the internet. But the authors will provide copies of the original documents, or information for helping to find them (in case the documents could not be found from the information already provided in the database). Specific request will have to be sent to the first author.

- L8: check spelling of reference and brackets. Done

- P3831 Consider www.ncdc.noaa.gov/paleo/phenology.html for making the data accessible and note here. Thank you for this suggestion. We will indeed propose the regional series to the world center for paleoclimatology. The data will be available there as soon as the paper is accepted for publication. The reference to the NOAA site has been added to the abstract and in the section: 2.3. Database structure;

- P3832 L13/14: add more examples of "special conditions" Some information was added in the paragraph (see below) and in the supplement S1

"Then, for each region, a mean composite series (hereafter grape harvest dates regional composite series, GHD-RCS; Tables S1 and S2 in the Supplement; Fig. 2) was calculated as the median of the individual adjusted series of the region. Only the GHD series of early varieties were taken into consideration. Some series were excluded from the calculation of the composite series of their region because they come from vineyards at special locations (for instance, higher elevation than the other datasets of the region), they contain too few data, they do not correspond to early varieties or because they are poorly correlated to the other series of the same area, which could introduce some noise in the GHD-RCS (Garcia de Cortazar-Atauri et al., 2010).

- L22/23: "aggregated"??? I have no clue how you treated the data here. This sentence was relevant in a previous version of the text but was left by mistake here, where it is indeed meaningless. It was removed in the new version.

C2812

- P3834 L5: highlight the 8 selected series in Table 1. Bold face, star symbol... Done

- Table 1: Consider adding one line at the bottom of the table with average values. Done

- Fig 2: Consider landscape format and larger printing area. Include a horizontal reference to link time the series with the corresponding y-lab. Horizontal lines were included. The format of the submitted version was landscape. . .

- Fig 4: Include significance level as a line for easier reading. Done. The figure caption was modified accordingly. A few words were added in the text in 3.1., in order to underline the period where the correlation is not significant.

- Fig 5 Use same graphics as in Fig 4. Dots are hard to read. Done

Review 2:

Abstract: - 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 pairwise comparisons of all regional series? 'Strong' is an appropriate term to describe a correlation with a high correlation coefficient. In order to be more precise, we have rephrased it as: "Most of the pairwise correlations are significant (p -value <0.001) and strong (mean pairwise correlation coefficient of 0.58)"

- 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? Yes, indeed it is. By synthetic we mean 'which is a synthesis of all the data'. We understand that this term is confusing and removed it.

Main text: - Page 3825, Line 25: Does the "its" for the standard deviation refer to the harvest date or the delay between veraison and harvest date? Grammatically, the 'its' must refer to the subject of the precedent sentence and therefore there there should not be ambiguities. However, we have repeated the word to avoid any ambiguity.

C2813

- 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. We have rewritten the first part of the results and discussion section in order to deal with the "uncertainty and quality issues" better than we did in the previous version. In particular, we have proposed some error estimation based on previous studies.

- Page 3829, line 19: How do these changes affect or improve the estimations? The data were verified and completed. We have modified the text to be more precise: "Some original documents (Chobaut's collection in the multimedia library of Avignon) used in Le Roy Ladurie (1967) were also revisited; some gaps were filled in the Northern and Southern Rhône valley series and three new series of the Burgundy region were found: Chaignay, Marsanay-la-Côte, Marsannais-le-Bois which are respectively 18, 9 and 17-year long. We also obtained information about the changes through time in the varieties grown in the southern and northern Rhône valley".

- 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. This sentence was relevant in a previous version of the text but was left by mistake here, where it is indeed meaningless. It was removed in the new version.

- Page 3833, line 23: See general comments for a discussion of quality issues. See point 4 please.

- 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

C2814

include these series, and how they expect them to be interpreted? For series with no overlapping temperature data, process-based phenological models, which do not need to be calibrated for each location but parametrized for each variety, may be used if the phenology of the variety is known.

- Are biological models or other means of verifying the GHD and climate relationships used? Yes, we have added a few words in the introduction about phenological models (section 1.3., see below for more information on the changes made in the introduction).

- 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. We have largely rewritten the introduction. In particular sub-sections were added in order to present the rationale behind using GHD for temperature reconstruction. The introduction is now presented as follows: Grapevine and temperature; 1.2. Scheduling of the grape harvest and 1.3. Grape harvest date and temperature. The last paragraph of the introduction now explicitly expresses what is the database intended for: "Considering the strong link they have with temperature, GHD are valuable proxies for this meteorological parameter. They can be used for reconstructing temperature variations through time using linear regression or process-based phenological models. The present paper describes a GHD database that we will make available on Internet. This database includes datasets mentioned above as well as many unpublished series. It gathers 380 GHD series of variable length from all the French vineyards and from some Spanish, Swiss, Italian and German ones. Over the description of the structure of the database, we propose a grouping of series into 27 composite series according to geological, geomorphological, historical and viticultural criteria in order to produce series representative of a region. The quality of these regional series is assessed by correlations of the GHD-series two-by-two and with instrumental temperature series."

- Page 3838, line 4: Give the values or ranges of values for the correlations in the text. We have included this precision: "...the strongest correlations obtained with the north-western GHD-RCS (the five lowest PCC, all close to -0.7, were obtained for Vendée

C2815

- Poitou - Charente, Low Loire Valley, High Loire Valley, Champagne 1 and Ile-de-France).”

- 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. The procedure is very classical. It is now explicitly written in the text, but we do not think it is necessary to recall it again in the figure caption. “To identify extreme years, we first transformed GHD-RCS in z-scores series as follows: $z\text{-score}_i = (\text{GHD}_i - \text{GHD}_{\text{mean}}) / \sigma$ where GHD_i is a date in a GHD-RCS, GHD_{mean} the mean date of this series and σ its standard deviation. We then calculated a “general” z-score series as the median of the z-scores of the GHD-RCS (Fig. 6). . .”

- 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? We have performed the same z-score calculation with the meteorological data used in the paper (see figure 3 for the distribution of the meteorological stations) as the one performed for the GHD-RCS series. The results were reported on figure 6. The two curves are very similar. In particular, they show the same plateau discontinuity in the 1950-1980 interval. Modelling has shown that this plateau could be explained by the combined effect of volcanic forcing (from 1950; diminishes temperature), stationary solar forcing, increasing sulfate effect (from 1940's; diminishes temperature) and increasing greenhouse gas effect (slight increase from ca. 1940 to ca. 1960, increasing sharply since then; increases temperature). The greenhouse effect compensates volcanic and sulfate effects since ca 1980 (e.g. Stott et al., 2000; Meehl et al., 2004). We have added the figure and the following lines: "Calculation of z-scores was also performed with the instrumental TmaxAMJJAS series. The medians of the z-scores of the temperature series and of the GHD-RCS series present many similarities (Fig. 6B). In particular, they show the same plateau discontinuity in the 1950-1980 interval and the same steep increase since ca. 1980. The 5 hottest years in the instrumental record (which extends only to 2006) are 1896, 1945, 1947, 1979 and 2003. All these years

C2816

are above the average or extreme according to the GHD record. The general agreement between the general z-score of the GHD-RCS and the instrumental temperature records confirms the usefulness of GHD as climate proxies."

Meehl, Gerald A., Warren M. Washington, Caspar M. Ammann, Julie M. Arblaster, T. M. L. Wigley, Claudia Tebaldi, 2004: Combinations of Natural and Anthropogenic Forcings in Twentieth-Century Climate. *J. Climate*, 17, 3721–3727. doi: [http://dx.doi.org/10.1175/1520-0442\(2004\)017<3721:CONAAF>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2004)017<3721:CONAAF>2.0.CO;2) Stott PA, et al. (2000) External control of 20th century temperature by natural and anthropogenic forcings. *Science* 290:2133–2137.

Technical corrections:

- Figure 2, Figure 4 and Figure 6: Provide a legend for each series. The y-axis 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. We agree that the labels are not very readable but our figure was supposed to appear as a landscape and not to be so small. . . . The figure was changed according to the recommendations.

- 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. The z-scores of Tmax AMJJAS were added and the different panels of the figure were described more precisely in the caption. Description of figure 6 was also extended in the text in 3.3.

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. It was modified.

C2817

Grammatical corrections: - Abstract: Line 2 and elsewhere: "non-translated" not really necessary. OK Page 3825 : line 4 "be useful for" rather than "allow". This sentence was removed - 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". We corrected throughout the text. EN COURS - Page 3827, line 4: "Enlightened" may not be the correct term. "Careful" or "Considered" decision, perhaps? We accept "considered" - Page 3827, line 18 "new diseases, such as..." Added - Page 3834, line 10: Not sure "clues" is the right word. Perhaps "results" or "investigations". "Investigation" was adopted - 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..." accepted - Page 3835, lines 12-27: Inconsistent verb tense. Present tense is fine for this discussion. It was checked - Page 3836, line 14: "series may be faulty..." ; line 25: "importantly" should perhaps be "significantly". Accepted - 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. Accepted - Page 3838, line 23: "decrease" or "decline" instead of "degrade"; line 26: "are lower (or smaller) than...". c - Page 3839, line 9: "extraction of", not "extracting. Accepted

Please also note the supplement to this comment:

<http://www.clim-past-discuss.net/7/C2810/2012/cpd-7-C2810-2012-supplement.zip>

C2818

Interactive comment on Clim. Past Discuss., 7, 3823, 2011.

C2819

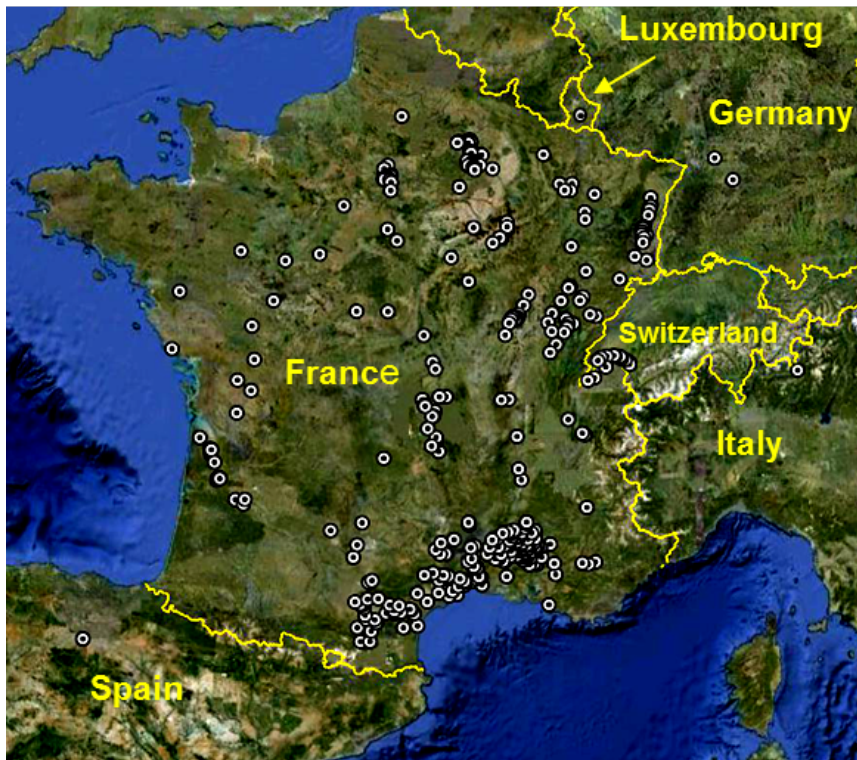


Fig. 1. Figure 1A

C2820

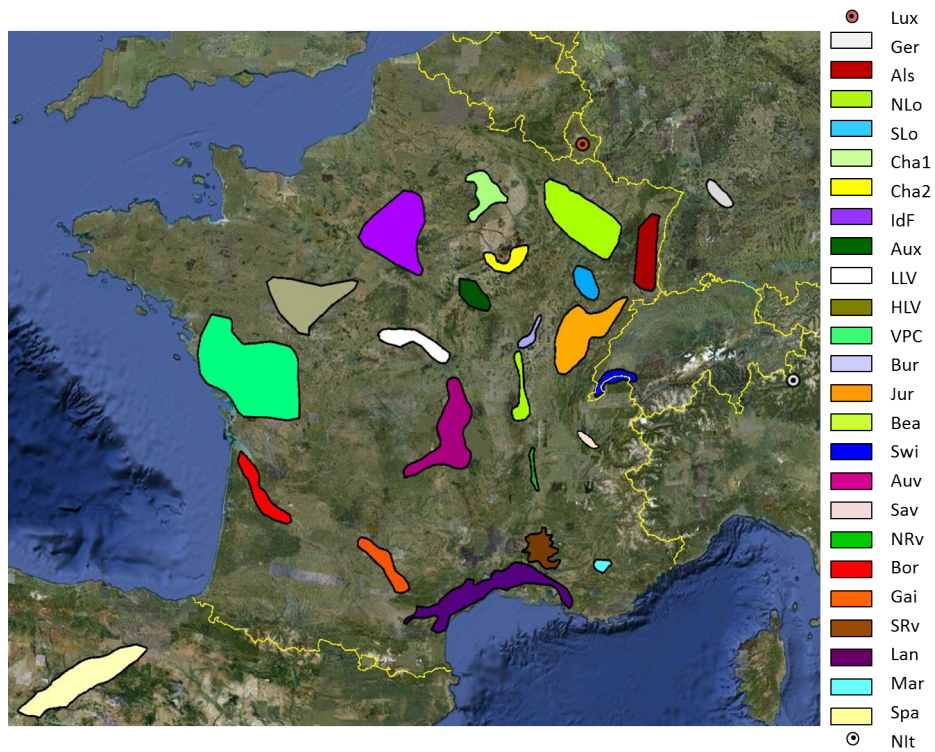


Fig. 2. Figure 1B

C2821

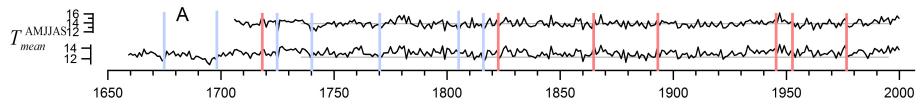


Fig. 3. Figure 2A

C2822

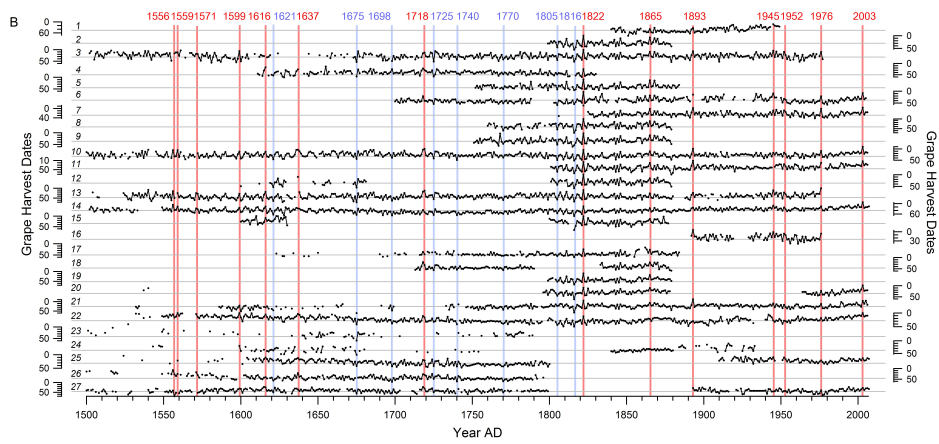


Fig. 4. Figure 2B

C2823

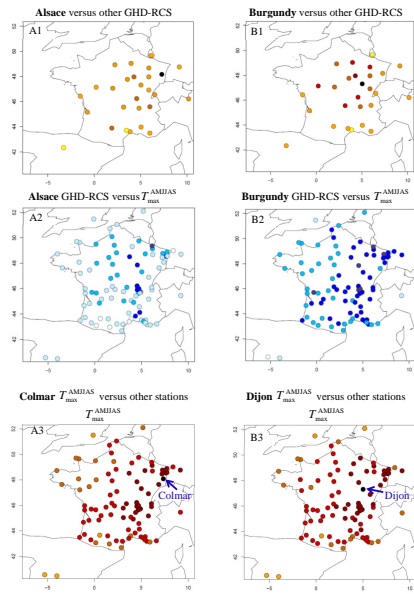


Fig. 5. Figure 3-1

C2824

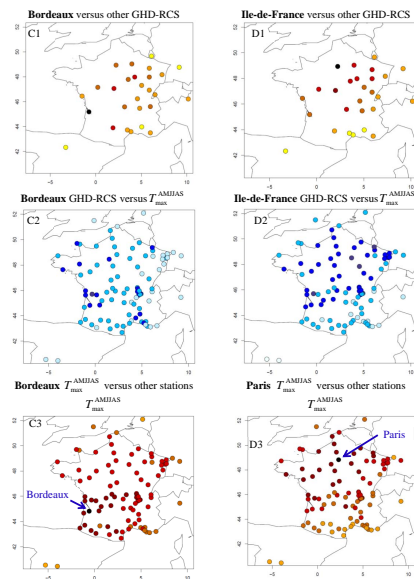


Fig. 6. Figure 3-2

C2825

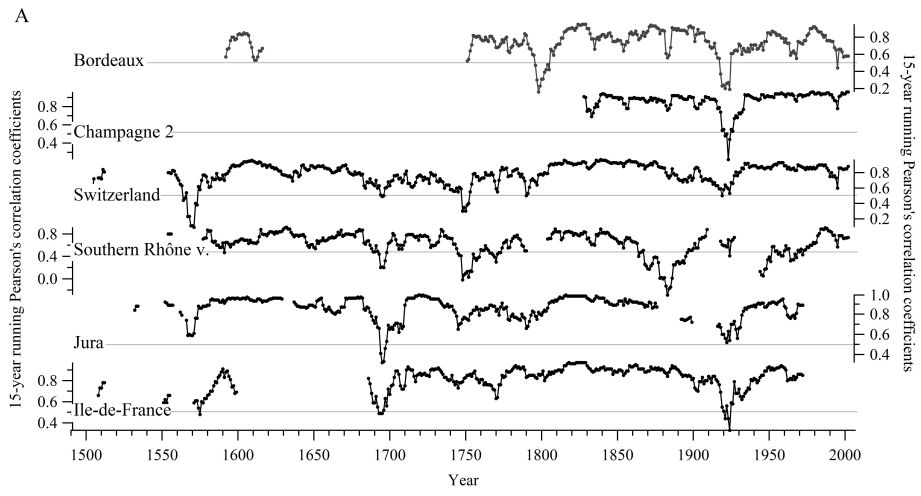


Fig. 7. Figure 4A

C2826

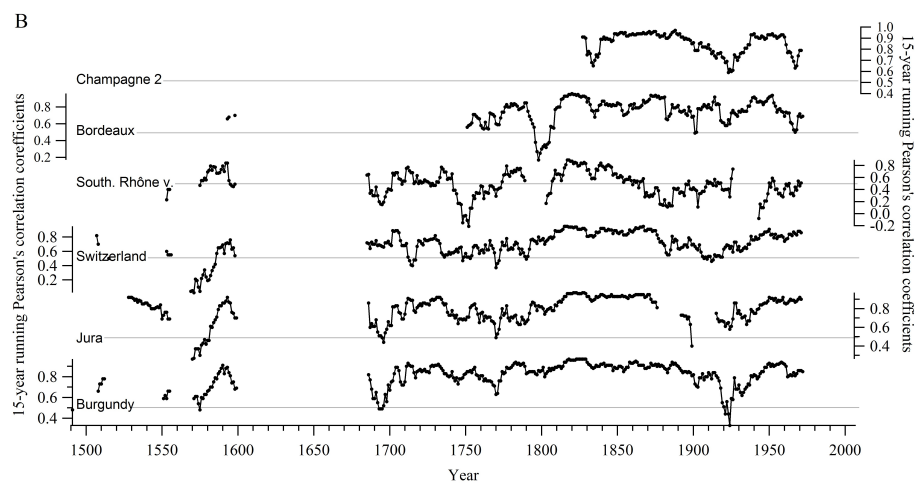


Fig. 8. Figure 4B

C2827

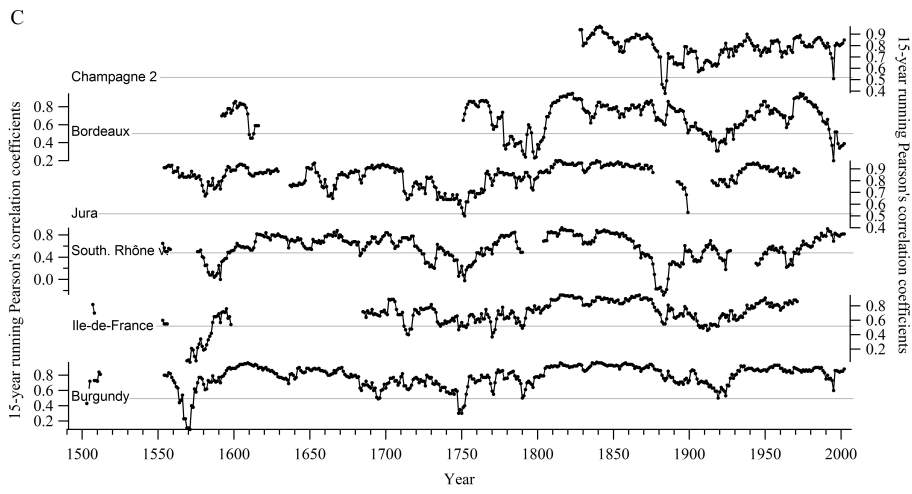


Fig. 9. Figure 4C

C2828

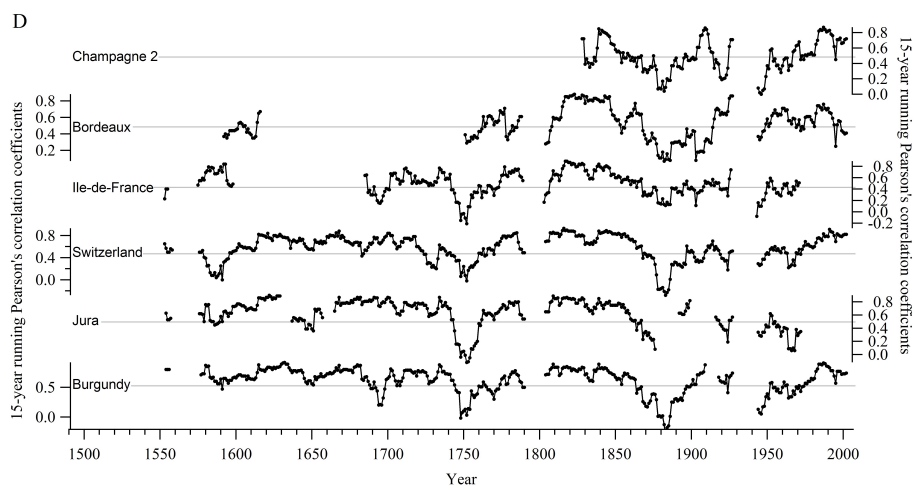


Fig. 10. Figure4D

C2829

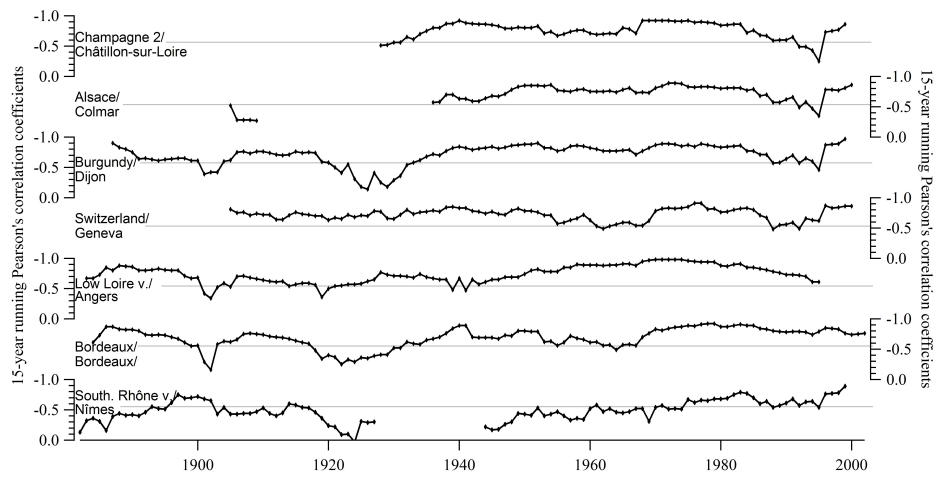


Fig. 11. Figure 5

C2830

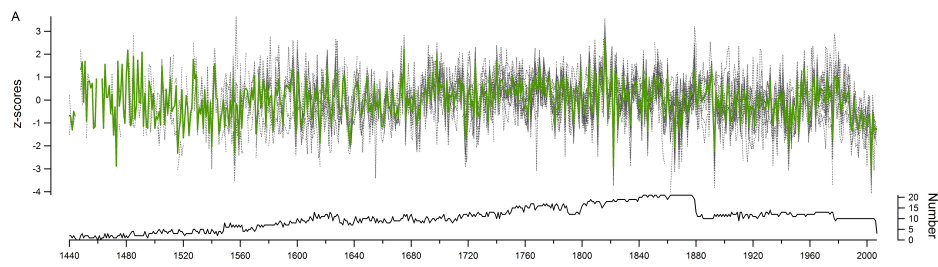


Fig. 12. Figure 6A

C2831

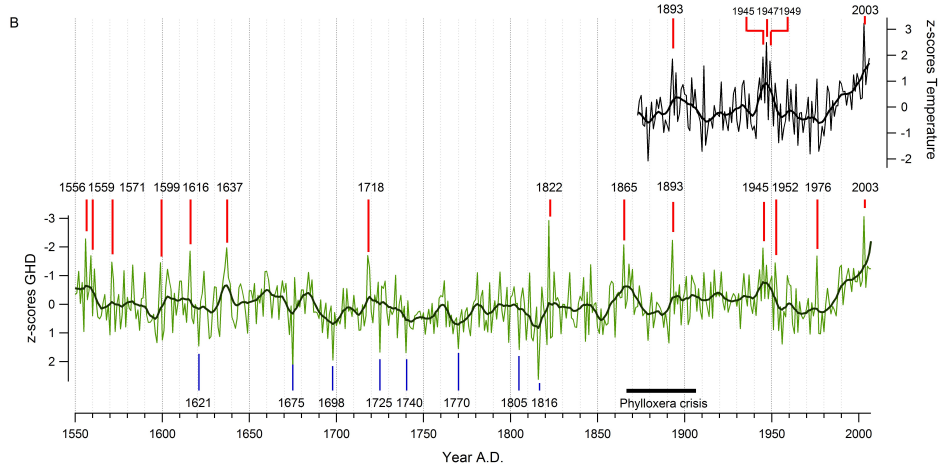


Fig. 13. Figure 6B

C2832