Appendix I - Handling missing GVP dates.

Many eruption dates in the GVP dataset are incomplete: only (part of) the start/end date is given. We discarded the data for which no start year was known. To be able to plot the data, the following convention is adopted for dates after 1679:

- If an end month is known, in which case an end year is known, then:
  - If the end year is not the start year, the start month is the end month.
  - If the end year is the start year and the end month is later than April, then the start month is 3 months earlier than the end month, otherwise the start month is January.
- If both the start month and the end month are missing, then the start month is set to June. After this step, all start months are populated.
  - If the end year is missing, and the start month is earlier than October, then the end year is set to be the start year, otherwise the end year is set to be one year later.
- By now, all start months and years, and all end years are populated.
  - If the end year is different from the start year, the end month is the start month.
  - If the end year is the same as the start year, either the start month is earlier than October, in which case the end month is 3 months later, or it is December. Then all months and years are populated.
- We are left to set the missing days:
  - Missing start days are set to 1.
  - Missing end days are set to 28, this is to avoid the problem of having to distinguish whether the end month is February (of which year), or a month with 30 versus 31 days.
- If the start and end day of the eruption are the same, we artificially added one day just for making the eruption data more visible.

For dates before 1679, the following convention was adopted:

- Where the end year is missing, the same year as start year is used. Months/days are not used in this category, the data are not precise enough. However, for plotting purposes, an additional 2 years were added to the length of the period when it falls after BC and before 1679. For BC eruptions, 20 years are added. This is because the time span is long, to make the data more visible (the correct data are available on hover above the data points).

The above conventions are used to be able to visualize eruption periods, the exact content of GVP dates is preserved.

Three case examples of Etna volcano eruptions in 1820, 1874 and 1444 were selected to illustrate the implementation of this technique, as shown in Figure S1, Figure S2 and Figure S3 of this Appendix.

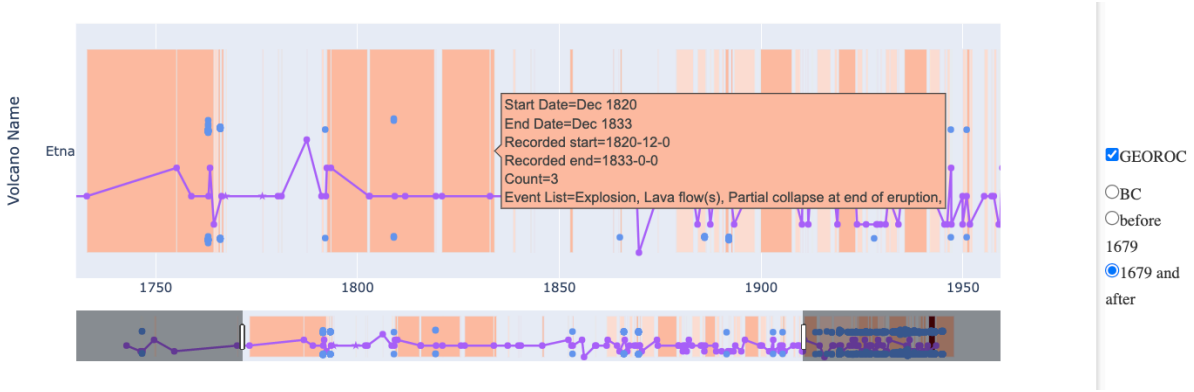

Figure S1. A case example for Etna volcano, an eruption record from GVP with start-date on December 1820, with no start day (Recorded start) and end date in 1833, with no end day and no end month (Recorded end). For the purpose of plotting, the assigned start date and end date became 1 December 1820 (Start Date) and 28 December 1833 (End Date) respectively. Each red bar represents the duration of an eruption, with detailed eruption dates and a list of corresponding events of the eruption on hover.

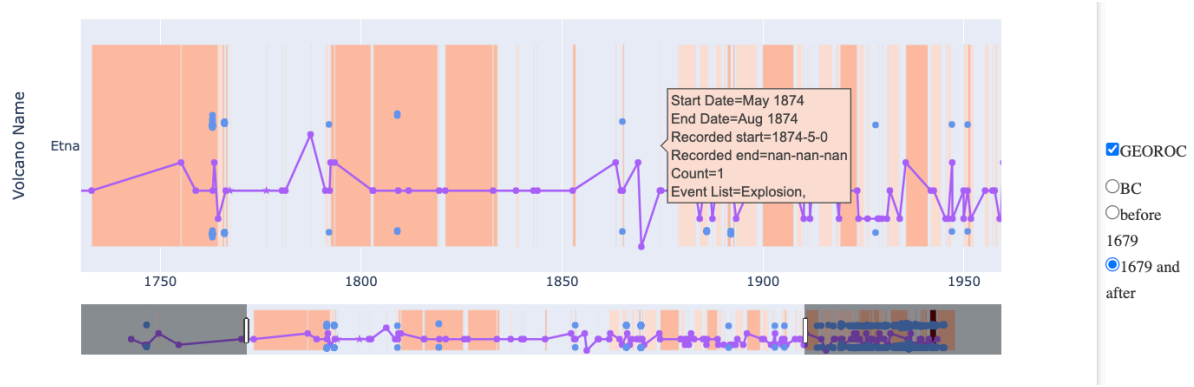

Figure S2. An eruption record for Etna volcano with GVP start-date on May 1874 with no start day (Recorded start) and unknown end date (Recorded end). For the purpose of plotting, the assigned start date and end date became 1 May 1874 (Start Date) and 28 August 1874 (End Date) respectively.

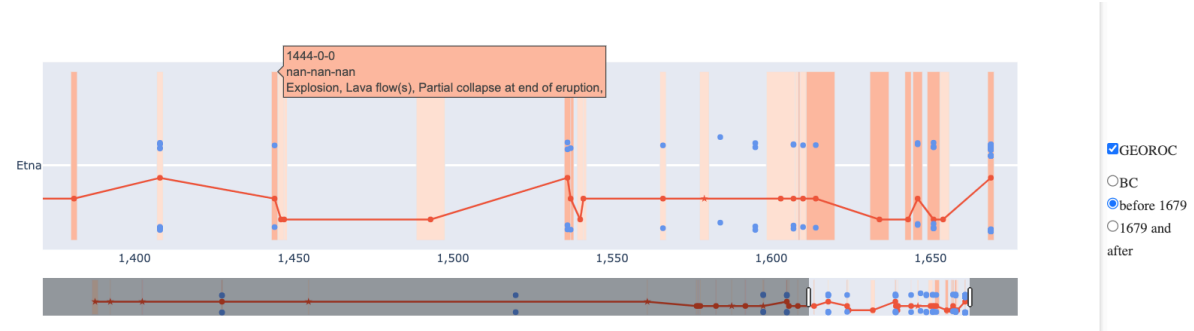

Figure S3. Another case example for Etna volcano, an eruption record before 1679 with known start date of 1444 and no end date. The end date becomes 1444.