Assessment made by CL to check for representativity and saturation of data collection on the categories:
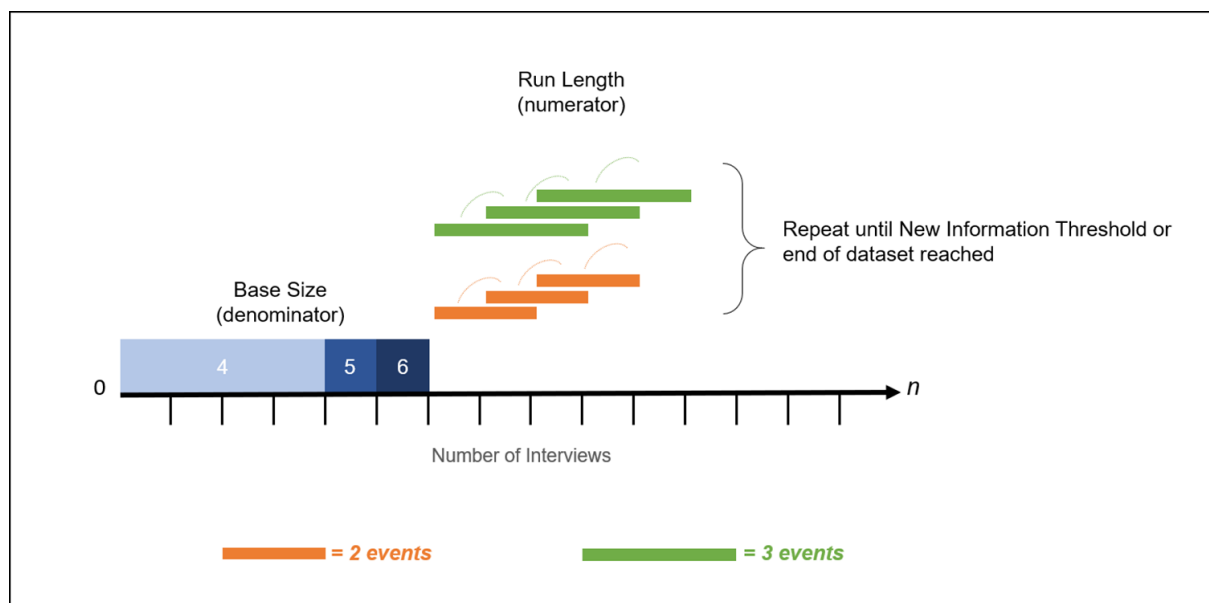1. Journal-name
2. Year (first published online)
3. Type-of-study
4. Period/Date*
5. Geographic-location*

*These were used to widen saturation test as with only 3 categories it looked a bit poor, although since Journal-name is what drives diversity the results are basically the same.

Representativity - assessment made on the basis of the distribution of answers as automatically given by the form (Summary of responses).

Saturation - calculated according to Guest, Namey and Chen (2020) with the following parameters:
- Base size: 6
- Run length: 3
- New Information thresholds: ≤5% and 0%

European dataset

Saturation
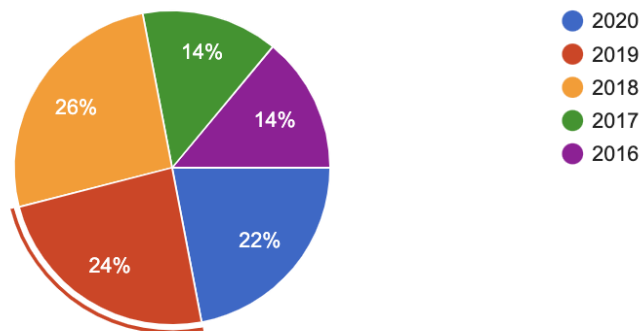Categories 1 to 3:
- 0% reached at paper 38+3

Categories 1 to 5:
- ≤5% reached at paper 13+3
- 0% reached at paper 38+3

Representativity:

- *Journal name*: we have 28 single entries over a potential list of 61 (46%). Five journals have more than one entry: Quaternary International (7), Vegetation History and Archaeobotany (7), Archaeological and Anthropological Science (6), Journal of Archaeological Science Reports (5) and Environmental Archaeology (2). All other journals are represented by 1 entry but we have representation of generalists and specialist journals, a good representation of disciplines (archaeology, geology, crop science, geochemistry, etc) as well as a couple of instances of journals centred on specific geographic locations (Bulletin of the Geological Society of Finland, Sprawozdania Archeologiczne).

- *Year*: slight unbalance towards more recent year but all in all quite good representation of the 5 years selected for study.

Year (first published online)

50 responses

- *Type of study*: archaeology is definitely predominant (60%), followed by paleoecology (18%), methodology (16% - combining all the different methodological sub-categories), geochemistry (8%) and plant physiology (2%) [total over 100% as some papers were put in 2 categories].
  This might be regarded as the more skewed data, and one that might require assessment of further paper. However, in order to really evaluate these results we would need to know what is the overall percentage of archaeological papers over all phytolith papers published (i.e., whether archaeological papers are in general c. 60% of all phytoliths papers published). As the search criteria did not include the type of study, in my opinion we can regard this category as representative of the general picture of phytolith publications.

South American dataset

<u>Saturation</u>
Categories 1 to 3:
- 0% reached at paper 17+3

Categories 1 to 5:
- ≤5% reached at paper 13+3
- 0% reached at paper 17+3

<u>Representativity</u>:
- *Journal name*: we have 28 single entries over a potential of 55 (51%). Six journals have more than one entry: Quaternary International (9), Journal of South American Earth Sciences (6), Flora (4), Review of Paleobotany and Palynology (4), PPP (3), and Vegetation History and Archaeobotany (2). All other journals have one entry each. As for Europe we have good representativity of subjects and local journals.

- *Year*: very similar to Europe, slight unbalance towards more recent years but all in all good representation of all the five years.

Year (first published online)

50 responses
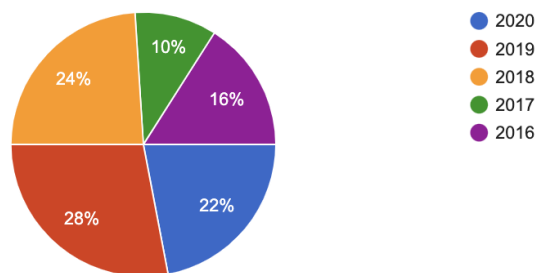
- *Type of study*: here paleoecology is predominant (49%), followed by methodological - modern reference (33%), archaeology (20%), ecology (4%) and all other subjects at 2%. Same as for Europe, I think that the type of search criteria we used makes this as balanced as it gets.

In general, between Europe and South America, we have at least 1 paper for all the categories included in "Type of Study", except for "Methodological- radiocarbon".

Guest G, Namey E, Chen M (2020) A simple method to assess and report thematic saturation in qualitative research. PLoS ONE 15(5): e0232076.
https://doi.org/10.1371/journal.pone.0232076