

This document details metadata and workflow for the ‘Database of foodstuffs and food systems in Africa, 1497-1840. Version 1’.

Abstract

This dataset catalogues observations of foodstuffs and food systems in eastern Africa over a period of three centuries (1497-1840 CE). It compiles 2,695 reports of the presence/absence of foodstuffs, domesticates, and non-food crops across 278 sites. These observations were supplemented by information on foodstuff quantity in >650 instances, with >250 unique detailed extracts on methods of production and relative dietary, social-cultural and economic importance. Importantly, this combination of the breadth and depth of ‘big qualitative data’ within the historical record enables data on the materiality of foodstuffs to be investigated alongside social-cultural perspectives (e.g. gender, belief systems, the effects of colonial ways of knowing on indigenous knowledges), and the possible drivers of transitions and transformations in African food systems.

Database assemblage

Data collection was underpinned by extensive consultation of published and unpublished historical documentary records written by colonial officials, chroniclers, missionaries, and traveller-diarists, who regularly recorded observations on the myriad crops, domestic animals, and agricultural systems in the societies they encountered, colonised and administered. Up to 13 variables were logged for each foodstuff observation (Table 1) following the workflow in Fig. 1. Nine of these were required fields, including presence/absence of foodstuff, locational information, author, and start/end year, for which 2,695 observations were compiled. The four non-mandatory fields comprised less commonly recorded but invaluable open-ended descriptions of the quantity, relative importance and method(s) of production, and a general “Notes” field. The full source bibliography of the written records from which the data were extracted is provided in a separate document. The database is available as an Excel file to enable wide readability and straightforward transferability into GIS applications via .csv format.

All evidence within the documents was considered contextually. The background of the author, the amount of time spent in the region and the purpose and genre of writing was examined in relation to their potential familiarity with the crops and agricultural systems they described. Importantly, given the varying nomenclature for crops within and between languages as well as over time, almost every source was consulted in its original language and recorded verbatim to ensure accuracy and transparency. One key example concerns the Portuguese term *milho*, which was erroneously translated to maize in some English-language translations of Portuguese documents,¹ but in previous centuries (including the timespan of the database) the term was used to describe grains of African origin, highly likely sorghum.^{5,2} The terms *meixoeira* and *nachenim/naxenim*, which frequently appear alongside *milho* and are also typically described as African grains, represent pearl millet and finger millet, respectively.⁶⁷ However, as non-Portuguese writers seldom distinguished between the two

millet, an unspecified millet category was also used. Added caution is needed when interpreting the geography of millet relative to other grains, as not every writer may have known or used the relevant terminology, while occasionally *milho* may have been used by less familiar writers as a catchall term to represent African grains. Maize likely fell under various other terms, including *milho grosso* and Indian corn.¹⁵ While many observers identified sorghum and the two millets, other African foodstuffs, such as the Bambara groundnut (*Vigna subterranean*), were usually referred to imprecisely as a type of bean rather than the more specific term *jugo*. This led to underrepresentation in the historical record, which must not be confused with absence of cultivation.

Table 1. Database variables and descriptions.

Variable	Required?	Description
Place Name	Yes	Name of the location described
Foodstuff	Yes	Name of the foodstuff/domesticated, identified through direct description and linguistic evidence
Author	Yes	Author of the document or publication
Start Year	Yes	Earliest year that the observation refers to; if unknown then date of publication is listed and stated in “Notes” column
End Year	Yes	Latest year that the observation refers to; if unknown then date of publication is listed and stated in “Notes” column
Lat and Lon	Yes	Coordinates (decimal degrees) of the place referred to in the text to four decimals, identified through known historical sites, textual descriptions and/or historical cartographic evidence
Site Type	Yes	Point: descriptions are precise and easily matched to known historical sites; Midpoint: a midpoint of the strip of land between the two known points described; Area: an approximate centre point of an area or territory described
Location Accuracy	Yes	A value indicative of the precision of the written report and the location identification process. 3: exact locational information obtained through sites still in existence, precise descriptions, or georeferenced maps; 2: historical sites identified via an element of inference, such as reported distance (in leagues or miles) from a known site, or non-georeferenced maps; 1: site classifications with a higher degree of inference from descriptions of areas or tracts of land, or unnamed rivers identified via narrative descriptions
Location Notes	Yes	Information on how the location of the site, midpoint, or area was derived, including (where relevant) the sources or maps used to identify the location
Quantity	No	Information on the quantity of the foodstuff. In many cases this is verbatim information from the document; however, numbers, weights and measures are available in some cases

Relative Importance	No	Relative importance of the foodstuff, including to different social groups; usually verbatim qualitative information
Production Methods	No	Descriptions of cultivation or production methods, harvest cycle, and labour inputs (e.g. gender, class)
Notes	No	Any other relevant information, for example foodstuff name in original language, species, prices, locational information, or author notes on database entry

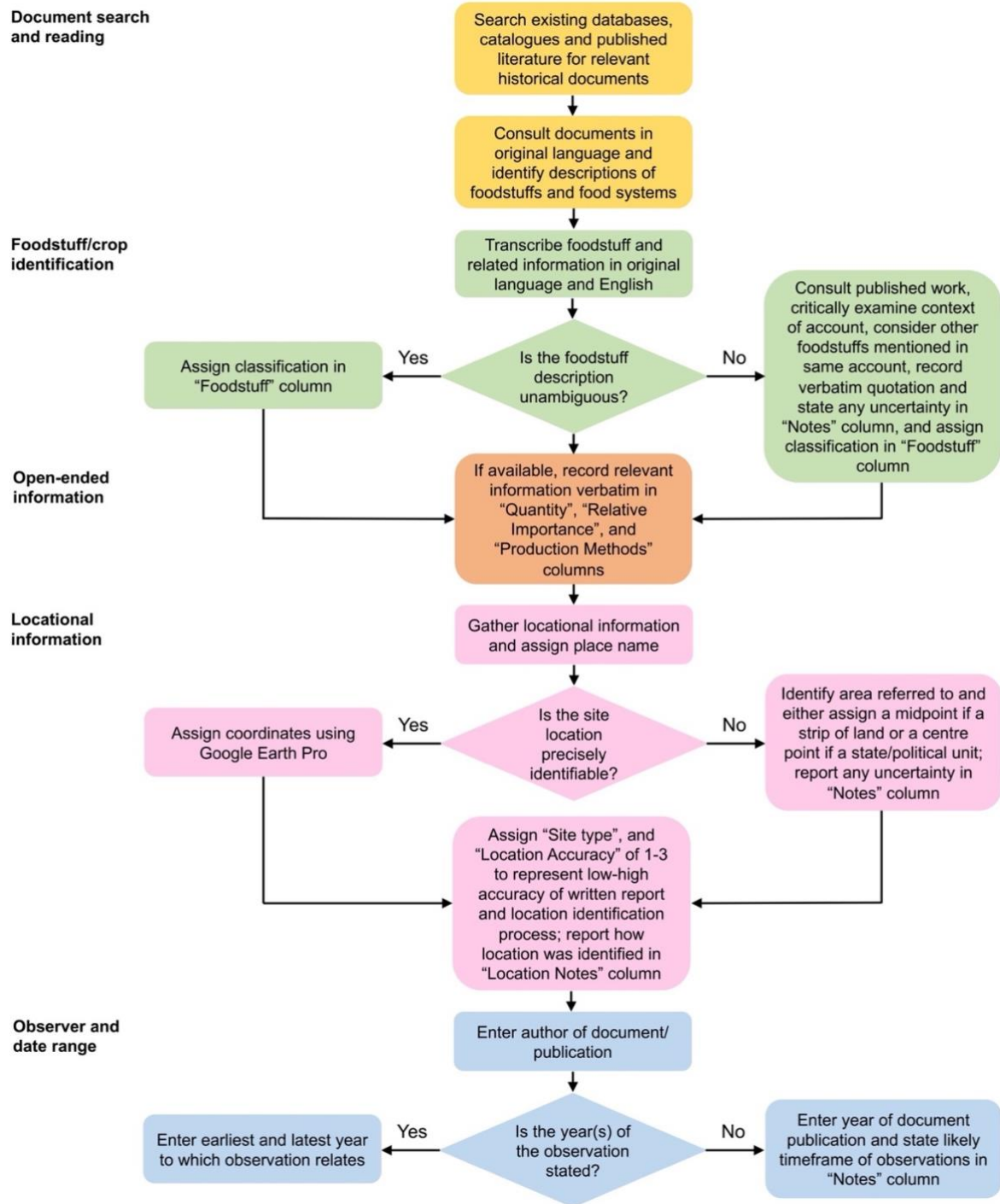


Fig. 7. Database workflow.

Locational data within documentary material were matched to known historical sites using historical or published cartographic evidence and then assigned latitude and longitude coordinates via Google Earth Pro. Many writers gave precise descriptions that could be matched to specific sites (e.g. settlements) or small areas (e.g. individual colonial landholdings) to a high degree of accuracy. However, a smaller but not insignificant number of descriptions simply referred to a strip of land, usually between two rivers or settlements, or an entire territory, usually an African polity. To avoid discarding valuable data for these two categories, which typically concerned areas with a lesser colonial presence, a midpoint of the tract of land or approximate centre point of the territory was assigned (as shown in Fig. 1) and specified in the “Site Type” variable. This permitted identification of a total of 278 sites, with 218 points, 39 midpoints, and 21 areas. All site types are useful for analysing regional foodstuff distribution, though non-point sites are less suitable for finer-grained analyses such as the relationship between soil types and crop distributions.

A further measure adopted to ensure the transparency of site classifications was the “Location Accuracy” variable, adapted from Bird and colleagues.³ This categorical value was generated based on the precision of the written report *and* the location identification process. A value of 3 represents more precise locational information obtained either through identification of sites still in existence, precise descriptions, or georeferenced maps. 2 refers to historical sites identified via an element of inference, such as reported distance (in leagues, miles, yards, or days’ march) from a known site or non-georeferenced maps. 1 relates to site classifications with a higher degree of inference from descriptions of areas or tracts of land, or unnamed rivers identified via narrative descriptions. 122 sites had a location accuracy of 3, 119 with 2, and 37 with 1. In all cases, the basis on which locations were identified is explained in the “Location Notes” column.

A key advantage of written data is its temporal precision. In the majority of historical accounts, a start and end year was readily available. However, a small number of important accounts (for example, Antonio Gomes’s 1648 *Viagem ao Império de Manomotapa*) only contained the year of publication rather than observation. In almost all of these cases the date of publication and probable time window of the observations were found to have been a maximum of c. 10-20 years apart, thus the date of publication was used as the timeframe and a note detailing the likely time window of the observations was added in the “Notes” field.

Of the non-required variables, “Quantity” captured verbatim descriptions of foodstuff quantity. These data are a mixture of qualitative descriptions (e.g. “abundant” or “rare”) and numerical data (e.g. numbers of cattle), which enable comparison between sites with common reporting styles. More detailed information was usually provided for the “Relative Importance” and “Production Methods” variables, although these were fewer in number overall. These data nevertheless provide valuable descriptions of economic, social-cultural and distributional aspects of food systems, including production for consumption versus production for exchange, commercialisation, dietary favourability, exchange networks, and types of cultivation including rainfed, flood recession and wetland agriculture. Other descriptions of production methods also permit studies of early colonial discourses on indigenous ecological knowledges. The “Notes” field captured other contextual information,

including verbatim quotations, the name of the foodstuff in its original language, further detail on the foodstuff, and clarificatory detail on the time period of the description.

¹ Silva Rego, A. & Baxter, T.W. *Documents on the Portuguese in Mozambique and Central Africa, 1497-1840*. 9 Vols. (Salisbury, National Archives of Zimbabwe, 1962-1989).

² Hair, P.E.H. Milho, Meixoeira and Other Foodstuffs of the Sofala Garrison, 1505-1525. *Cahiers D'études Africaines* 17, 353–63 (1977).

³ Bird, D. *et al.* p3k14c, a synthetic global database of archaeological radiocarbon dates. *Sci Data* **9**, 27 (2022)