Bayesian inference of prehistoric population dynamics from multiple proxies: a case study from the North of the Swiss Alps

Martin Hinz1,2,✉, Joe Roe1, Julian Laabs3, Caroline Heitz3, and Jan Kolář4,5

28 Mai, 2022

Robust estimates of population are essential to the study of human–environment relations and socio-ecological dynamics in the past. Population size and density can directly inform reconstructions of prehistoric group size, social organisation, economic constraints, exchange, and political and social institutions. In this pilot study, we present an approach that we believe can be usefully transferred to other regions, as well as refined and extended to greatly advance our understanding of prehistoric demography. Here, we present a Bayesian hierarchical model that uses Poisson regression and state-space representation to produce absolute estimates of past population size and density. Using the area North of the main ridge of the Swiss Alps in prehistoric times (6000–1000 BCE) as a case study, we show that combining multiple proxies (site counts, radiocarbon dates, dendrochronological dates, and landscape openness) produces a more robust reconstruction of population dynamics than any single proxy alone. The model’s estimates of the credibility of its prediction, and the relative weight it affords to individual proxies through time, give further insights into the relative reliability of the evidence currently available for paleodemographic research. Our prediction of population development of the case study area accords well with the current understanding in the wider literature, but provides a more precise and higher-resolution estimate that is less sensitive to spurious fluctuations in the proxy data than existing approaches, especially the popular summed probability distribution of radiocarbon dates. The archaeological record provides several potential proxies of human population dynamics, but individually they are inaccurate, biased, and sparse in their spatial and temporal coverage. Similarly, current methods for estimating past population dynamics are often simplistic: they work on limited spatial scales, tend to rely ona single proxy, and are rarely able to infer population size or density in absolute terms. In contemporary demography, it is becoming increasingly common to use Bayesian statistics to estimate population trends and project them into the future. The Bayesian approach is popular because offers the possibility of combining heterogenous data, and at the same time qualifying the uncertainty and credibility attached to forecasts. These same characteristics make it well-suited to applications to archaeological data in paleodemographic studies.

1 Institute of Archaeological Sciences, University of Bern  
2 Oeschger Centre for Climate Change Research, University of Bern  
3 CRC 1266 - Scales of Transformation, University of Kiel  
4 Department of Vegetation Ecology, Institute of Botany of the Czech Academy of Sciences  
5 Institute of Archaeology and Museology, Faculty of Arts, Masaryk University

✉ Correspondence: [Martin Hinz <[martin.hinz@iaw.unibe.ch](mailto:martin.hinz@iaw.unibe.ch)>](mailto:martin.hinz@iaw.unibe.ch)

Keywords: Prehistoric demography; Bayesian modelling; Multi-proxy; Settlement dynamics

Highlights: - Bayesian modelling can integrate multiple, heterogeneous population proxies from the archaeological record - Our initial model produces more robust, high-resolution estimates of past population dynamics than previous, single-proxy approaches - We provide absolute estimates of population size and density on the area north of the Swiss Alpes in prehistoric times (6000–1000 BCE)

# 1 Introduction

## 1.1 Prehistoric demography

Prehistorians have long recognised demography as a fundamental force in human cultural evolution (Childe, 1936). Despite decades of interest in the demographic dynamics of prehistoric societies, concrete estimates of population size and density before written records remain elusive. Though the archaeological record provides multiple possible demographic proxies (Müller and Diachenko, 2019), a lack of access to this data and methodological tools for turning it into quantitative estimates has left the conclusions drawn from it vague and superficial (Hassan, 1981). As a result, ‘expert estimates’ transferred from ethnographic parallels have often taken the place of direct inference from archaeological evidence Turchin et al. (2015).

Prehistoric demography has experienced a resurgence in recent years (Riede, 2009 and others in same issue; Shennan, 2000), partly explained by a renewed interest in human–environment relations and human impact, necessarily requiring an assessment of population size. Kintigh et al. (2014) list human influence, dominance, population size, and population growth amongst their ‘grand challenges’ for archaeology in the 21st century.

In the last decade, the ‘dates as data’ technique (Rick, 1987), using the frequency of radiocarbon dates as a proxy for population dynamics, was taken up and significantly developed (e.g. Shennan et al., 2013), and has since been widely applied to archaeological contexts worldwide (Crema, 2022). This approach has contributed greatly to our understanding of prehistoric demography, but is not without its critics Carleton and Groucutt (2021). While the methodology continues to evolve and address these critiques (Crema, 2022), it remains subject to fundamental problems common to all approaches relying on a single proxy Schmidt et al. (2021). We believe that these problems cannot be overcome by methodological refinements in this area alone. Instead, a Bayesian approach offers a robust, quantitative methodology for inferring prehistoric population dynamics from multiple proxies, including summed radiocarbon dates.

# 2 Background

## 2.1 Existing approaches to population estimation

Proxies currently used for the estimation of population size in prehistory (following Müller and Diachenko, 2019) can roughly be divided into three groups: ethnographic analogies; deductive estimates from ecological/economic factors; and the interpolation of frequencies of archaeological features (e.g. settlements, structures, individual finds). Three basic problems are common to all these approaches:

1. **Reliance on single proxy**: Most investigations are using only one source of evidence. Although multi-proxy approaches exist, there the individual proxies only serve to support each other or the main estimator, without explicitly combining them to reduce the inaccuracy and biases inherent to single proxies.
2. **Uncertainty in measurements**: All archaeological evidence is inherently uncertain which is carried through to derived measurements. However, in most studies, single curves are presented as estimates, and the potential error associated is almost never specified.
3. **Lack of a transfer function**: By ‘transfer function’, we mean something that allows for the proxy data to be interpreted in terms of actual population size or density. These could be absolute, i.e. a numerical estimate of population, or relative, i.e. a means of scaling changes in the proxy value to changes in population. Lack of suitable frameworks and ‘calibration’ data means that this is rarely presented alongside proxy estimates. In the best cases, there is a qualitative assessment of the informative value of the proxy, not sufficiently accounting for the complex nature of archaeological data.

Furthermore, the types of archaeological data commonly used as population proxies share a number of problematic characteristics, being:

* **Limited**: We have only incomplete data, and they are usually not very informative.
* **Unevenly distributed**: For example, although there is a good data on settlement frequencies for some regions, these regions are very unevenly distributed over time and space.
* **Noisy**: Frequently individual proxies are strongly influenced by factors unrelated to population, for example taphonomic conditions or depositional biases.
* **Unreliable**: Research strategies, research history and varying levels of resources available to researchers strongly affect the nature of compiled datasets. Systematic distortions are the rule rather than the exception.
* **Heterogeneous**: All potential proxies have different spatio-temporal scales, granularity, information value, scales, and data formats.
* **Indirect**: We will never have direct data on prehistoric population; only proxy data that is thought to be a reasonable substitute. The transfer functions linking the proxy data with the desired quantity (population) are unknown.
* **Contradictory**: When considering several proxies, differences in transfer functions, data quality and noisiness inevitably lead to different results.

We argue that many, if not all, of these problems can be ameliorated through a) the explicit, quantitative integration of multiple proxies; and b) the use of a Bayesian approach to take account of and estimate uncertainty.

## 2.2 Hierarchical Bayesian demographic models

Many of the problems with archaeological population proxies are shared with contemporary demography. In response, demographers have increasingly turned to Bayesian methods to estimate and forecast contemporary population dynamics. For example, Bryant and Zhang (2018) consider Bayesian data modelling a solution to exactly the kind of problems that affect archaeological data. Bayesian approaches are well suited for limited, unreliable and noisy data. Various data sources, even contradictory data, can be brought into a common framework and used to support one another. These methods also provide a quantitative estimate of the likelihood and uncertainty of the model’s resulting predictions (or in our case retro-dictions). Bayesian approaches are also capable of accounting for spatially and temporally incomplete data: where this data is missing, the uncertainty increases, but this does not prevent general modelling and estimation. Finally, hierarchically-structured model suites, with sub-models for each individual proxy, can be used to estimate transfer functions between them and the value to be modelled, thanks to the interaction of a large number of evidence.

This modeling technique can thus be used to join different lines of evidence horizontally and vertically and combine their results into a overall estimate, including an assessment of their reliability: contradicting data lead to a lower overall reliability, while a mutual support to smaller confidence intervals. If there is no systematic bias that affects all data sources to the same extent, this results in the most reliable estimate possible through the most heterogeneous set of data sources.

The well-established application of a Bayesian approach in radiocarbon calibration, where it is used to model radiometric uncertainty based on stratigraphic information, represents a very similar use case in archaeology. More recently, archaeologists have also used Bayesian modelling techniques for testing hypotheses relating to demographic models based on 14C data (e.g. Crema and Shoda, 2021). This approach differs from the one presented here in that, in these analyses, deductive models are generated and their plausibility is tested on the basis of 14C data only. This is a clear step forward to a model-based, scientific analysis. However, the use of only one proxy, exclusively for testing hypotheses developed independently, creates problems comparable to those of the inductive approaches used so far: lacking a combination with other indicators, one is limited to the problems and conditions of sum calibration. Furthermore, this approach loses significant potential information that would be gained by a direct evaluation of time series.

We attempt to make Bayesian hierarchical techniques usable for archaeological reconstructions. We want to show, in a reproducible and practical form using a case study, how Bayesian methods can make a decisive contribution to a better assessment of population development, crucial for the reconstruction of the human past, even in for periods for which we only have very patchy, noisy and unreliable data.

## 2.3 Assumptions of a Bayesian approach

Bayesian statistics rely on the premise that there is always some prior assumption, even if very rough, about the probability of an event. This assumption is adjusted by observing data, by checking how credible these priors are (likelihood, see also Bryant and Zhang, 2018, p. 66). This is Bayesian updating (cf. also Kruschke, 2015, especially 15–25), resulting in the posterior probability distribution, which represents not a point prediction. Small amounts of data lead to a broad distribution not strongly localised and restricted. Thus, we simultaneously obtained a result and an estimate the credibility interval, given the data.

This Bayesian learning is iterative and sequential, so that the result of one Bayesian inference can form the prior of another (Kruschke, 2015, p. 17). This allows different information to be combined (Bryant and Zhang, 2018, pp. 219–224), as it has long been exploited by archaeology in using stratigraphic information to make radiometric dating more accurate (Ramsey, 1995).

This also makes a hierarchical formulation of problem domains possible. Parameters that are necessary for an estimation, such as the relationship of population density to the deforestation signal in pollen data, need not be specified explicitly, but can be given by probability distributions and then estimated in the model itself (Bryant and Zhang, 2018, p. 186). The more data available, the more degrees of freedom can be estimated with a reasonable width of credibility intervals (Kruschke, 2015, p. 112). For the estimation of these parameters, submodels have to be created describing the relationship of the data to the characteristics of the parameter (Kruschke, 2015, pp. 221–222).

# 3 Materials: Population proxy data

Our case study area north of the Swiss Alps (Figure 3.1) covers about one third of Switzerland’s territory and comprises the partly flat, but largely hilly area between the Jura Mountains and the Alps. It is favourable for settlement and agriculture, especially the Swiss Plateau between Lake Zurich and Lake Geneva, by far the most densely populated region of the Switzerland today. This serves as core region in our case study, because here archaeological data is most abundant and accessible. Depending on the region, it has a very diverse natural landscape: shaped by glaciers during the ice ages, the many lakes and bogs provide excellent preservation conditions for the numerous Neolithic and Bronze Age lakeside settlements, and a rich source for vegetation reconstructions by means of pollen analyses. Thanks to the very active and efficient archaeological research and heritage management there is an abundance of archaeological information, including known sites as well as dendrochronological and 14C data.

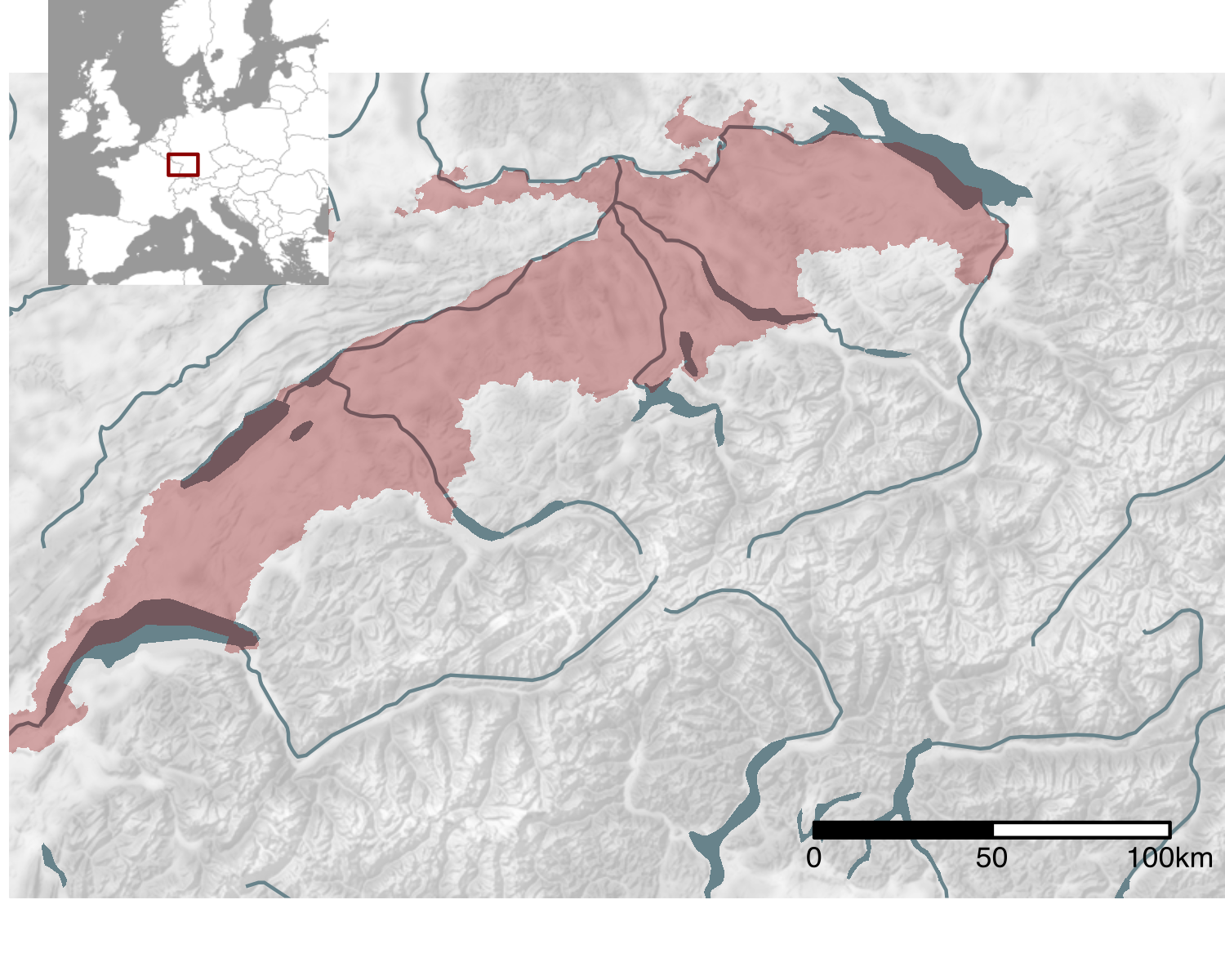


Figure 3.1: Location and extent of the Swiss Plateau as biogeographical region (based on swisstopo) including additional low altitude areas in the north of Switzerland (regions along the High Rhine between Schaffhausen and Basel).

Our case study targets the time window between 6000–1000 BCE. Its lower limit was chosen to avoid the so-called ‘Hallstatt plateau’ on the Northern Hemisphere radiocarbon calibration curve, causing difficulties for the 14C proxy. The upper limit coincides with post-glacial changes in pollen spectra, before which the openness indicator is highly unlikely to reflect human influence.

A large number of different proxies can be integrated into an overall model, provided that these observations a) can be understood as dependent on the population density in the past, and b) a model-like description of this dependence can be created. Table 3.1 provides a non-exhaustive list. For our case study, we used: a landscape openness indicator; an aoristic sum of typological dated sites; a sum calibration; and frequency data for dendro-dated lakeshore settlements in the Three Lakes region (western Swiss Plateau).

(#tab:tableproxies) A incomplete list of possible observation that can be linked to population developments in the past. Proxies used in this study are highlighted.

| Proxies |
| --- |
| Expert estimates |
| Ethnographic Analogies |
| Carrying Capacity |
| Economic modelling |
| Extrapolation of buried individuals |
| Burial anthropology |
| Settlement data, number of houses |
| Settlement data, settlement size |
| **Aoristic analysis** |
| **Dendro dates** |
| Amount of archaeological objects |
| **Radiocarbon sum calibration** |
| Estimates based on specific object types |
| **Human impact from pollen or colluvial data** |
| aDNA based estimates |
| … |

## 3.1 Dendro-dated lakeshore settlements

From the Neolithic onwards, known settlement areas in Switzerland concentrate along its rivers and lakes (Christian Lüthi, 2009). Thus, our working region offers excellent data for demographic estimation, but poses very specific problems for such an undertaking. We have high-resolution information on the temporal sequence of individual lakeside settlements by means of dendro data. In these cases, 14C data are not as abundant simply because they are inferior to dendro dating.

The dataset we use for the number of dendro-dated wetland settlements in the Three Lakes region was collected by Julian Laabs for his PhD thesis (Laabs, 2019). The time series used here runs from 3900 to 800 BCE, and contains the number of chronologically registered fell phases at individual settlements.

## 3.2 Summed radiocarbon

The dataset for the 14C sum calibration primarily consists of data from the XRONOS database (<https://xronos.ch>), supplemented by dates from the unpublished PhD thesis of Julian Laabs (Laabs, 2019) and the data collection of Martínez-Grau et al. (2021). It contains a total of 1135 single 14C data from 246 sites (see Figure 3.2). The dates in the dataset range in 14C years from 10730 to 235 uncal BP. This time window extends beyond the study horizon in order to minimise boundary effects.

We binned the data at site levels to obtain a temporally dispersed count and thus an expected value of contemporaneous 14C dated sites. For the creation of the sum calibration, the corresponding functions of the R package rcarbon (Crema and Bevan, 2021) were used with their default settings.

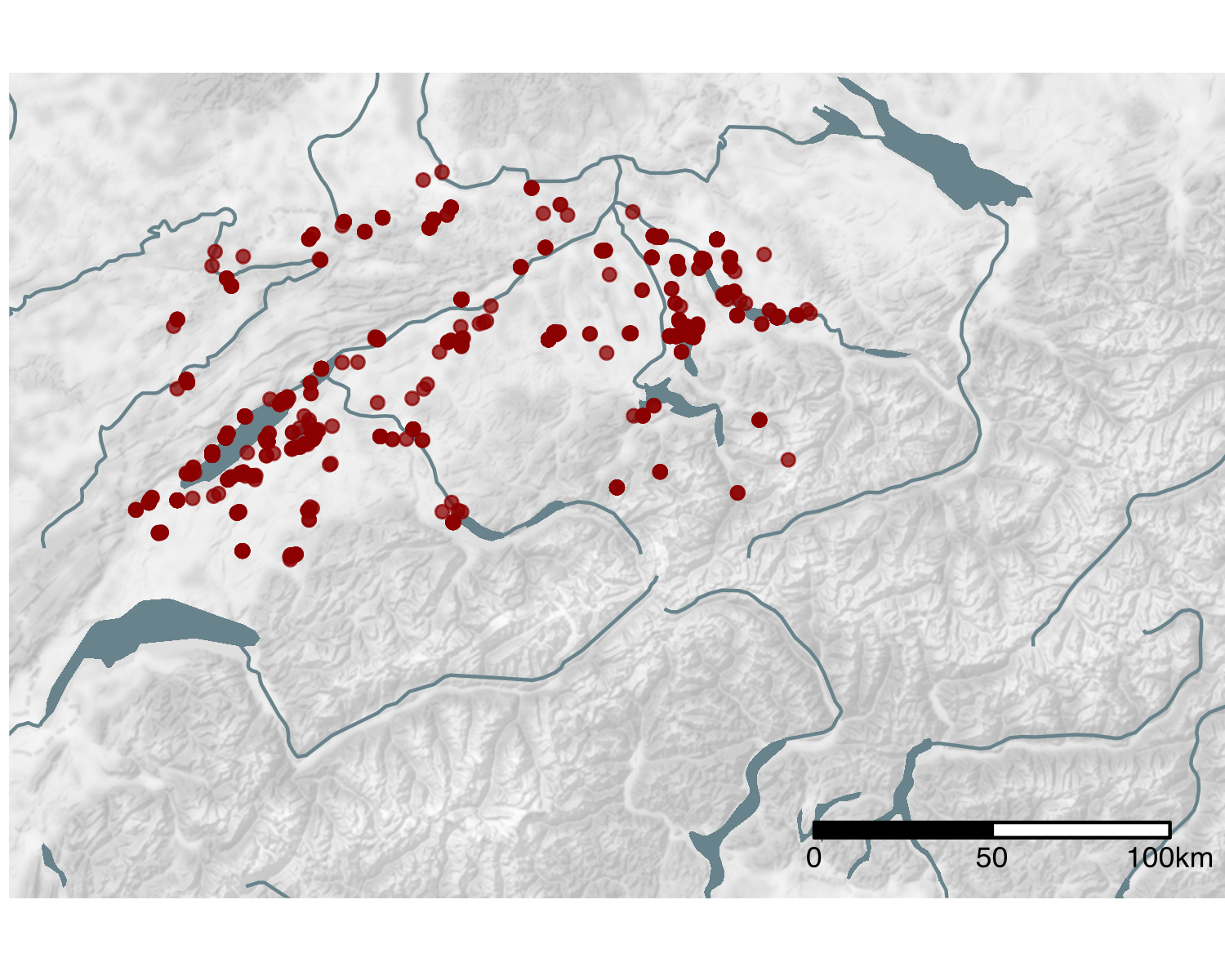


Figure 3.2: The location of the 14C dated sites in the dataset.

## 3.3 Aoristic sum

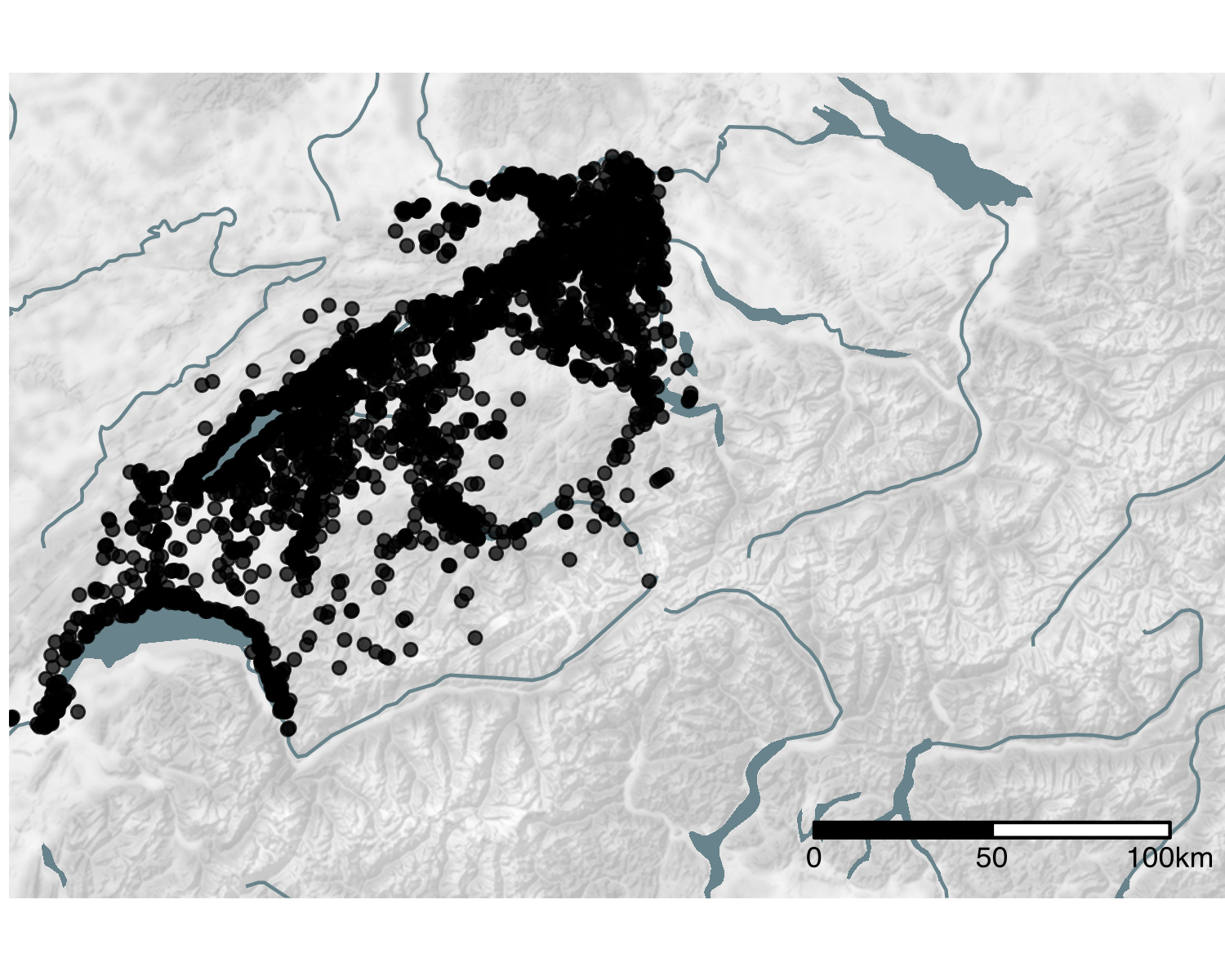


Figure 3.3: The location of the sites from the find reports of the cantonal archaeology (heritage management). The locations are ‘fuzzed’ by ~1km.

We include relative dating information obtained from the Swiss cantonal heritage management (Figure 3.3). These are primarily derived from scattered surface finds, often with a low dating accuracy (only in the range of archaeological periods), incorporate into our model as a typologically dated aorist time series. It is not bound to the problems of radiocarbon dating and thus methodological issues of sum calibration can be avoided. Data from 4321 sites were included in the aoristic sum.

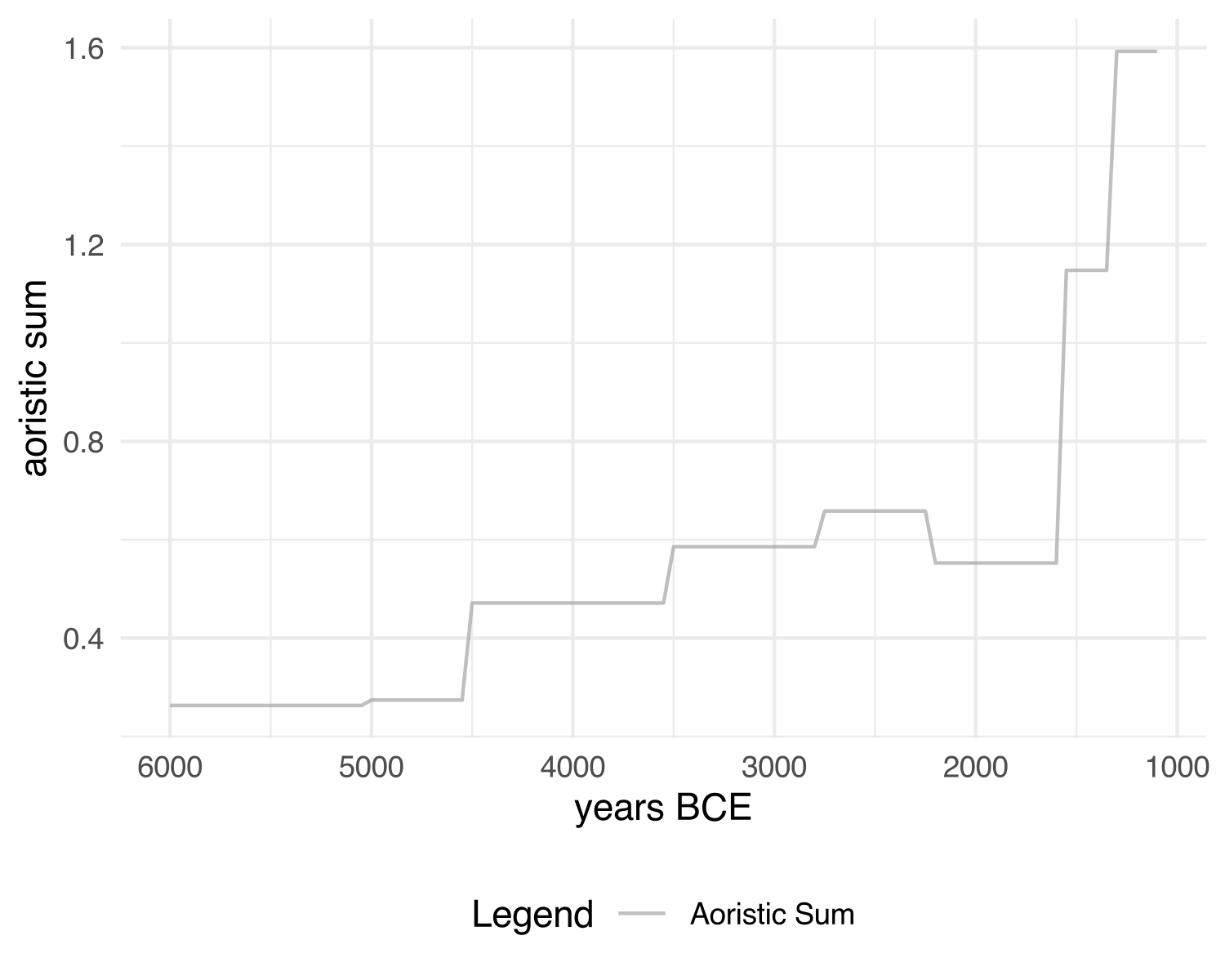


Figure 3.4: Aoristic sum of the archaeological sites used in this analysis.

## 3.4 Landscape openness

The natural conditions of the lakes of the Swiss Plateau enable not only highly precise dating of archaeological sites, but also a very dense network of pollen analysis. We make use of this by generating a supra-regional openness indicator for the vegetation from the pollen data (Figure 3.5). This proxy has the specific advantage that it is not dependent on archaeological preservation conditions, making it particularly valuable for compensating systematic distortions that result from archaeological taphonomy and temporally specific settlement patterns.

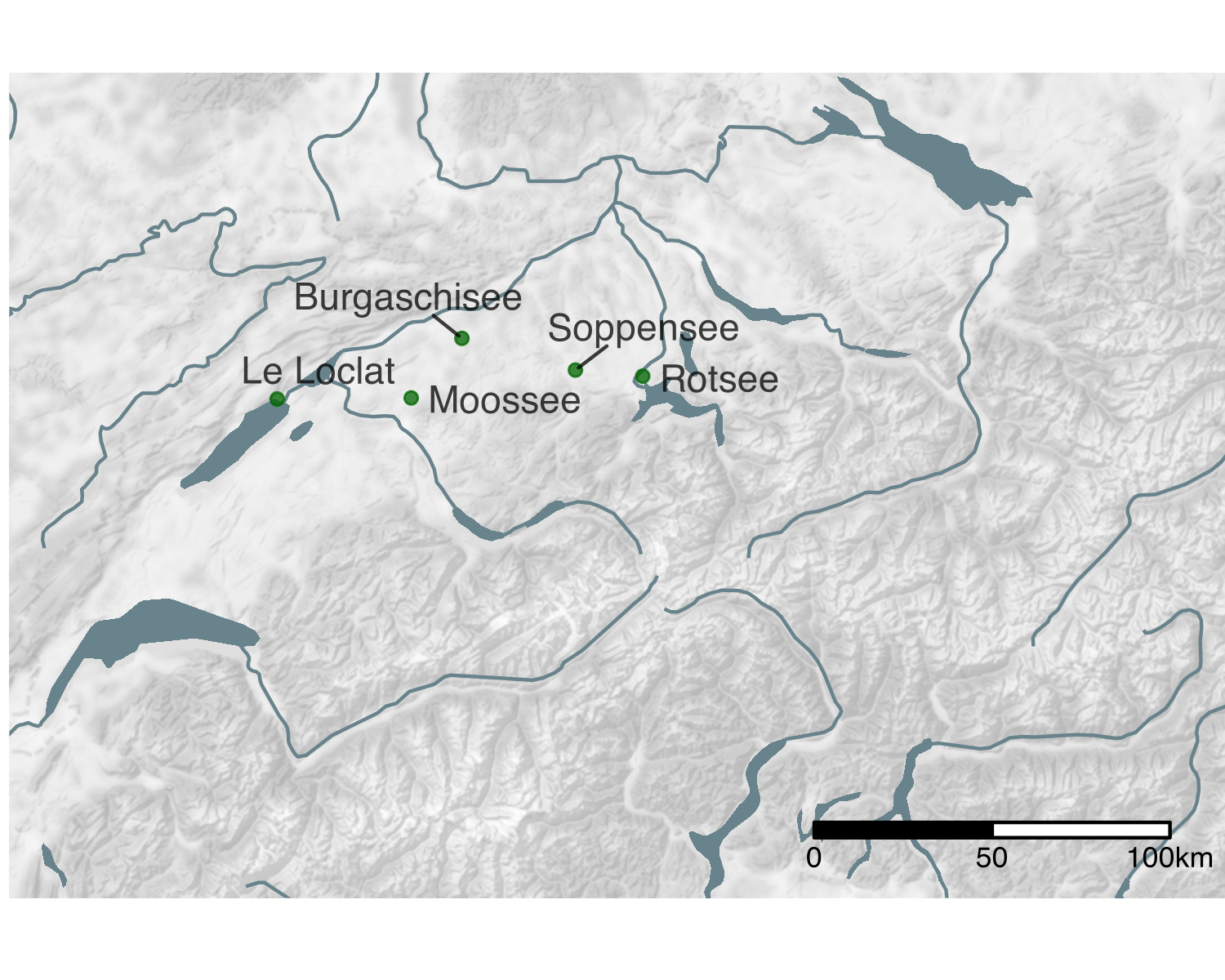


Figure 3.5: Location of the pollen profiles used for the openness indicator.

We assume that the higher the population density in an area, the greater the human influence on the natural environment Lechterbeck et al. (2014). Evidence of deforestation can therefore provide indications of population dynamics. The full procedure for deriving this proxy from several different pollen diagrams is detailed in a previous publication (Heitz et al., 2021). Here, we use five pollen diagrams from sites mainly in the hinterland of the large Alpine lakes.

![Figure 3.6: Value on the first dimension of the PCA against dating of the samples for the individual pollen profiles and their combined average value as the openness indicator.](data:application/pdf;base64,)

Figure 3.6: Value on the first dimension of the PCA against dating of the samples for the individual pollen profiles and their combined average value as the openness indicator.

# 4 Methods: Bayesian model

The sum calibration, openness and the dendro-dated settlement data was smoothed by a moving average with a 50 years window, corresponding to the unified sampling interval for all proxies. The aoristic sum was not smoothed, because it already has a very coarse temporal resolution. In the construction of our ‘observational model’, we considered all these proxies as informative of the number of settlements located in the north of the Swiss Alps. Population development is simulated in a ‘process model’ using a Poisson process.

## 4.1 Process model

A special class of Bayesian hierarchical models are so-called State Space Models, specifically designed for time series following two principles. First, a hidden or latent process is assumed, representing the state of the variable of interest through the entire time series. Every state of variable x in the future, as well as in the past, is bound by a Markov process to the state of variable at time . Second, it is assumed that certain observations, represented in variable y, are dependent on the state of variable at time . This implies that a relationship between the individual states of variable is generated over time via the hidden variable , which itself is not observable.

This structure makes these models particularly suitable for demographic reconstruction using archaeological and other data. The population density itself is not directly measurable. All we have at our disposal are observations derived by unknown transfer functions.

The overall model for the estimation of demographic developments is broken down into several hierarchically interconnected individual elements. The a process model represents the demographic development itself, without already being explicitly parameterised with data. Here we assume that the latent variable ‘number of sites’ is strongly autocorrelated across different time periods. The number of sites in 3000 BCE is strongly conditioned by the number of sites in 3050 BCE, and so on. The population at time results from the population at time times a parameter , which represents the population change at this time.

The univariate discrete Poisson distribution is particularly suitable for modelling frequencies, numbers of events that occur independently of each other at a constant mean rate in a fixed time interval or spatial area. It is determined by a real parameter >0, describing the expected value and the variance. Thus, the relationship shown above can be rearranged as follows:

If we now have information about the change in population development (the proxies), this can enter into the model via a change in in form of a regression: for all proxy values — represented as a vector of independent variables , with as an n-dimensional Euclidean space defined by the n variables — the model takes the form:

Using the logarithm as a link function ensures that , which must always be positive, can also be described by variables (proxies) that may also be negative. serves as slope factor, as in a normal linear regression. Here, it functions as a scaling factor for the individual proxies. is to be understood as an intercept, representing a baseline when there were no change due to the variables. This is the desired behaviour: is equal to the value of the population in the previous time period, plus or minus the changes resulting from the variables.

Since and are essentially in the same range (e.g. if , the expected value for would also be 1), must also be log-transformed for the congruence of both values. Population size as well as population change are time-dependent. At each individual point in time, these variables will take on different values. But we can assume that the population change will not exceed certain limits (), though it is not possible to specify this at this point.

The gamma distribution centres the probability in a range , adding 1 makes this range . This prevents the number of sites from explosively increasing between two time periods, which would lead to problems for the convergence of the model. The estimation of this parameter for the entire model, as well as the estimation of the respective population change per time section, results from the modelling and the interaction with the data.

## 4.2 Observational model

The implementation of the observational model in this paper is a Poisson regression, where the proxies are used to inform the change in the number of settlements between time steps. The individual proxies were z-normalised, absolute differences between time steps were then computed. If the value of the proxy increases, this results in a positive difference from the previous time step, and vice versa.

The sum of the resulting differences between the time steps, together with the settlement number of the previous step as the expected value, then forms as the expected value for the settlement number of the current time step.

Here, is a scaling factor that represents the influence of the respective proxy. It is a confidence value of the model for the respective proxy, so that the sum of all results in 1.

A Dirichlet distribution, a multivariate generalization of the beta distribution, is commonly used for this purpose in a hierarchical Bayesian model. Its density function gives the probabilities of different exclusive events. It has a parameter vector , for which we have chosen a weakly informative log-normal prior. The priors for the log-normal distribution in turn come from a weakly informative exponential distribution for the mean and a log-nomal distribution with of 1 and of 0.1:

As an intuition, we consider the sum of the proxies as determinant for the number of settlements, whereby the share of each individual proxy is variable and is estimated within the model. This share is recorded within the model as the parameter p.

The error value is represented by the Poisson process in the process model. In this implementation, the model finds the best possible combination between the individual proxies to describe a settlement dynamic. The number of sites is converted into population density using (certainly debatable) parameters defined by us, but which are only scaling factors for the intermediate value of number of settlements. We assume that each site represents a number of people that is poisson distributed around the value 50, a compromise, as both Mesolithic and Neolithic and Bronze Age settlement communities need to be represented. An evidence-based estimate data series of the temporal development of settlement sizes could enhance this specification. From the number of sites and the mean number of individuals a population density can be calculated using the case study area (12649 km²), making the models estimate comparable with estimates from other sources or the literature.

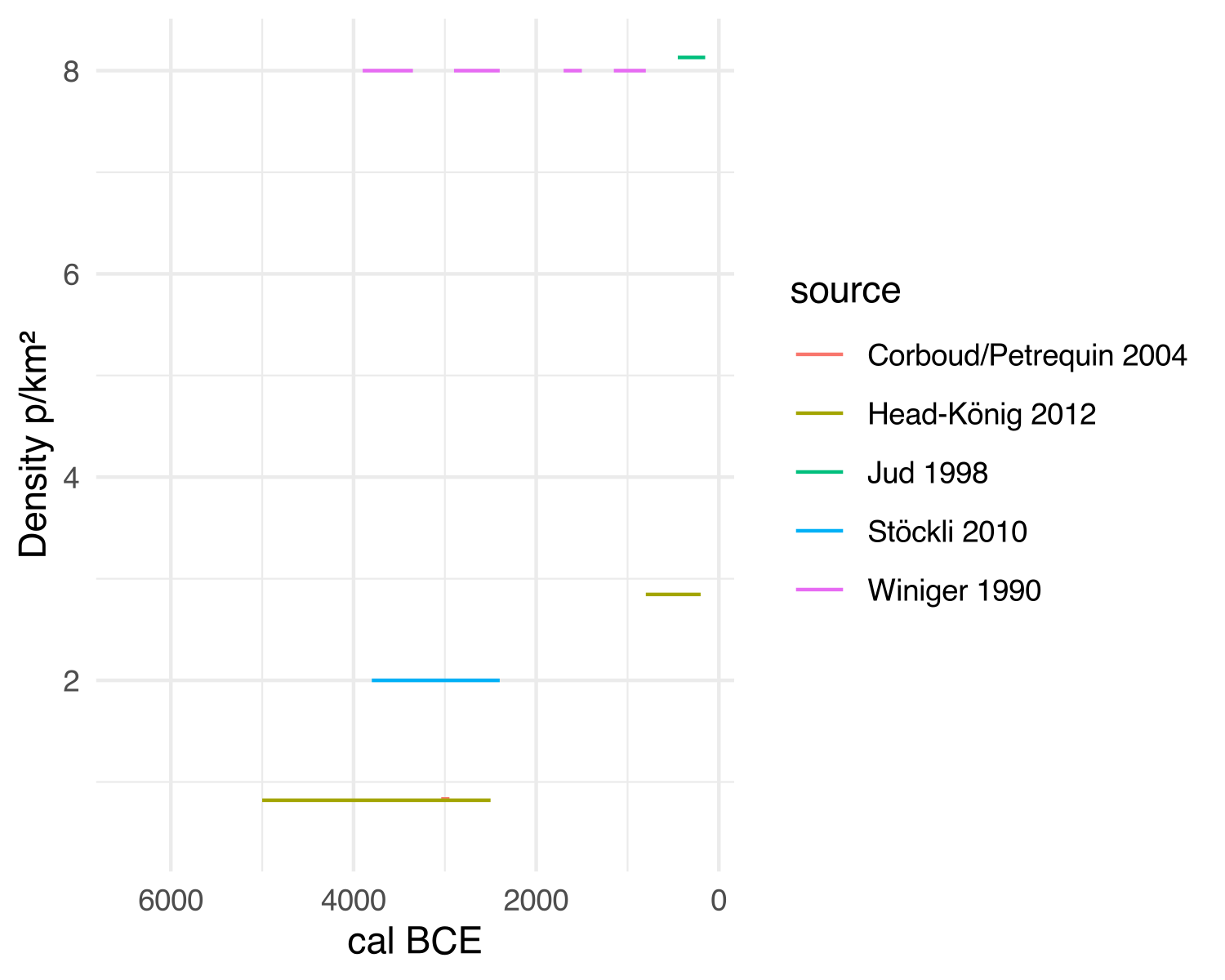


Figure 4.1: Expert estimations for the population density on the Swiss Plateau from different authors.

## 4.3 Model fitting

The model was fitted using the R package *nimble* (version 0.11.1, R version 4.1.3), using 4 parallel chains. Achieving and ensuring convergence and sufficient effective samples (10000) for a reliable assessment of the highest posterior density interval was carried out in steps.

1. the model was initialised for each chain and run for 100000 iterations (with a thinning of 10). On a reasonably capable computer (Linux, Intel(R) Xeon(R) CPU E3-1240 v5 @ 3.50GHz, 4 cores, 8 threads), this takes approximately a minute.
2. the run was extended until convergence could be determined using Gelman and Rubin’s convergence diagnostic, the criterion being that a potential scale reduction factor of less than 1.1 was achieved for all monitored variables. Convergence occurred after about thirty seconds.
3. Due to the high correlation of the parameters and thus a low sampling efficiency, the collection of at least 10,000 effective samples for all parameters took about five hours.

A starting value of 5 p/km² for a population density of the Late Bronze Age (1000 BCE) was taken from the literature, which may represent a general average value for all prehistoric population estimates (Nikulka, 2016, p. 258). For the model, this was set as the mean of a normal distribution with a standard deviation of 0.5, which should give enough leeway for deviations resulting from the data. Nevertheless, especially the late part of the reconstruction is of course clearly influenced by this predefined value.

For the traceplots and the prior-posterior overlap as well as the density functions of the posterior samples of the individual parameters, please refer to the supplementary material.

(#tab:priorsandparameters) Priors and fixed parameters used in the model.

| **Priors** | **Value** | **Plot/Comment** |
| --- | --- | --- |
| MeanSiteSize | dpois(50) | MeanSiteSize |
| max\_growth\_rate | dgamma(shape = 5, scale=0.05) + 1 | max_growth_rate |
| mu\_alpha | dlnorm(1,sdlog=0.1) | mu_alpha |
| a\_alpha | dexp(1) | a_alpha |
| alpha | dlnorm(mu\_alpha[j],sdlog=a\_alpha[j]) | alpha |
| p | ddirch(alpha[1:4]) | p |
| **Parameters** |  |  |
| nEnd | 5 |  |
| AreaSwissPlateau | 12649 km² |  |
| **Initial Values** |  |  |
| lambda |  | exponential increase of the factor 10 |
| PopDens | nEnd (=5) |  |
| nSites | 50 |  |

# 5 Results

![Figure 5.1: The estimate of the population density resulting from the model and the four proxies, which are also plotted (scaled) for comparison.](data:application/pdf;base64,)

Figure 5.1: The estimate of the population density resulting from the model and the four proxies, which are also plotted (scaled) for comparison.

The population density estimated by the model (Figure 5.1) ranges between 0.2 p/km² for the beginning (6000 BCE) and 4.8 p/km² for the end of the estimate (1000 BCE), reaching a maximum of 6.5 p/km² for around 1250 BCE, remaining within the values considered plausible by the expert estimates. Clear peaks are around 1250 BCE, as well as around 2750 BCE, which corresponds to the beginning of the influence of the Corded Ware ceramic style (Hafner, 2004).

![Figure 5.2: The variability of the estimate of the population density over time, with the estimate itself for reference.](data:application/pdf;base64,)

Figure 5.2: The variability of the estimate of the population density over time, with the estimate itself for reference.

The temporal distribution of the variability in the estimate (Figure 5.2) allows us assess at which time steps the uncertainty is greater due to e.g. contradictions in the proxies. The coefficient of variation is 0.13 for the beginning and 0.1 for the end of the estimate, the greatest variability is reached around 2150 BCE with 0.47. This is not astonishing as there are fewer archaeological contexts recorded connected to the earlier phase of the Early Bronze Age c. 2200-1800 BCE, this picture changes from c. 1800 BCE onwards (David-Elbiali, 2000; Hafner, 1995). The beginning and end of the time series are relatively clearly determined, resulting from the *a priori* setting of final population density, but also from the uniformity of the proxies during these periods. Overall, the variability is relatively stable over the entire estimation and averages overall at 33% of the respective mean.

![Figure 5.3: Distribution of the influence ratio of the different proxies on the final estimation of the number of sites.](data:application/pdf;base64,)

Figure 5.3: Distribution of the influence ratio of the different proxies on the final estimation of the number of sites.

The parameter reflects the relative weight given to the individual proxies. Its posterior distribution (Figure 5.3) shows that the model weights the openness indicator the highest, averaging slightly above 60%, followed by the sum calibration, with an average of about 20%. The aoristic sum is slightly above 10%, whereas the importance of the dendro-dated settlements is below 10%. The reason for the latter is certainly that there are no lakeshore settlements over large areas of the time window, and therefore the overall confidence in the proxy is low. The aoristic sum is flat over large sections, making it difficult to relate to the other estimators. The sum calibration shows very strong short-term fluctuations, at least partly due to the calibration curve, which also make this proxy seem ill-suited to reliably represent a continuous population trend. Its fluctuations have an impact on the overall estimate, albeit to a lesser extent.

# 6 Discussion

## 6.1 Comparison of individual proxies

Comparing the model’s overall estimate with the individual proxies enables insights into the quality of these records. The sum calibration, currently the most frequently used proxy for (relative) population changes in prehistory, has its large fluctuations dampened when considered alongside other proxies, especially the first fluctuation shortly after 4000 BCE. The expected increase in archaeological remains with the onset of Neolithisation is still clearly visible, but the overall curve is much flatter than the sum calibration itself. The period between 3950 and 3700 BCE, contemporaneous with the first major settlement of the Three Lakes regions’ lakeshores, coincides with a noticeable plateau in the calibration curve, producing an overestimation of the 14C density. A second maximum, after 3000 BCE, is supported by the other proxies, and is consequently much more reflected in the overall estimate, coinciding with a smaller and shorter plateau. The rise towards the Middle and Late Bronze Age is also supported by the other proxies, without a significant pattern in the calibration curve. We may conclude that the model is successful in using information from other proxies to sift ‘real’ fluctuations in the summed radiocarbon record from artefacts of the calibration curve.

On average, the model weights the sum calibration at about 20%, significantly less than the 60% afforded to the openness indicator. After an initial increase, which is easily explained by spread of agriculture, the openness indicator tends to fluctuate less and thus has a dampening effect on the overall estimate. In general, this trend in the sum calibration is well reflected in land openness, while changes within the Neolithic and Bronze Age are more gradual.

The aoristic sum remains flat over long spans of time. It is not until the Middle and Late Bronze Age that we see a significant rise, which is also apparent in the model’s overall estimate. It remains to be seen to what extent modelling of the taphonomic loss (Surovell et al., 2009) could be integrated in this approach.

The number of simultaneously existing lakeshore settlements is a temporally and spatially limited estimator, but extremely reliable. Its limitations are reflected in the low overall confidence of the model, since its value is zero over long stretches. However, where it has information potential, such as around and shortly after 3800 BCE, 3200 BCE, 1600 BCE or especially around 2750 BCE, its fluctuations have a noticeable influence on the overall estimate. This highlights another potential of our approach: where a proxy has little structure and thus little significance, or where its trends cannot be linked to other indicators, it consequently has little influence. For periods in which it can provide information, however, this will also feed into the overall model, despite a low overall confidence in the estimator.

## 6.2 Prehistoric population dynamics north of the Swiss Alps

In order to review the reconstruction against the background of established archaeological knowledge, it is useful to overlay conventionally-defined archaeological phase boundaries (Hafner, 2005) on the results of our model (Figure 6.1).

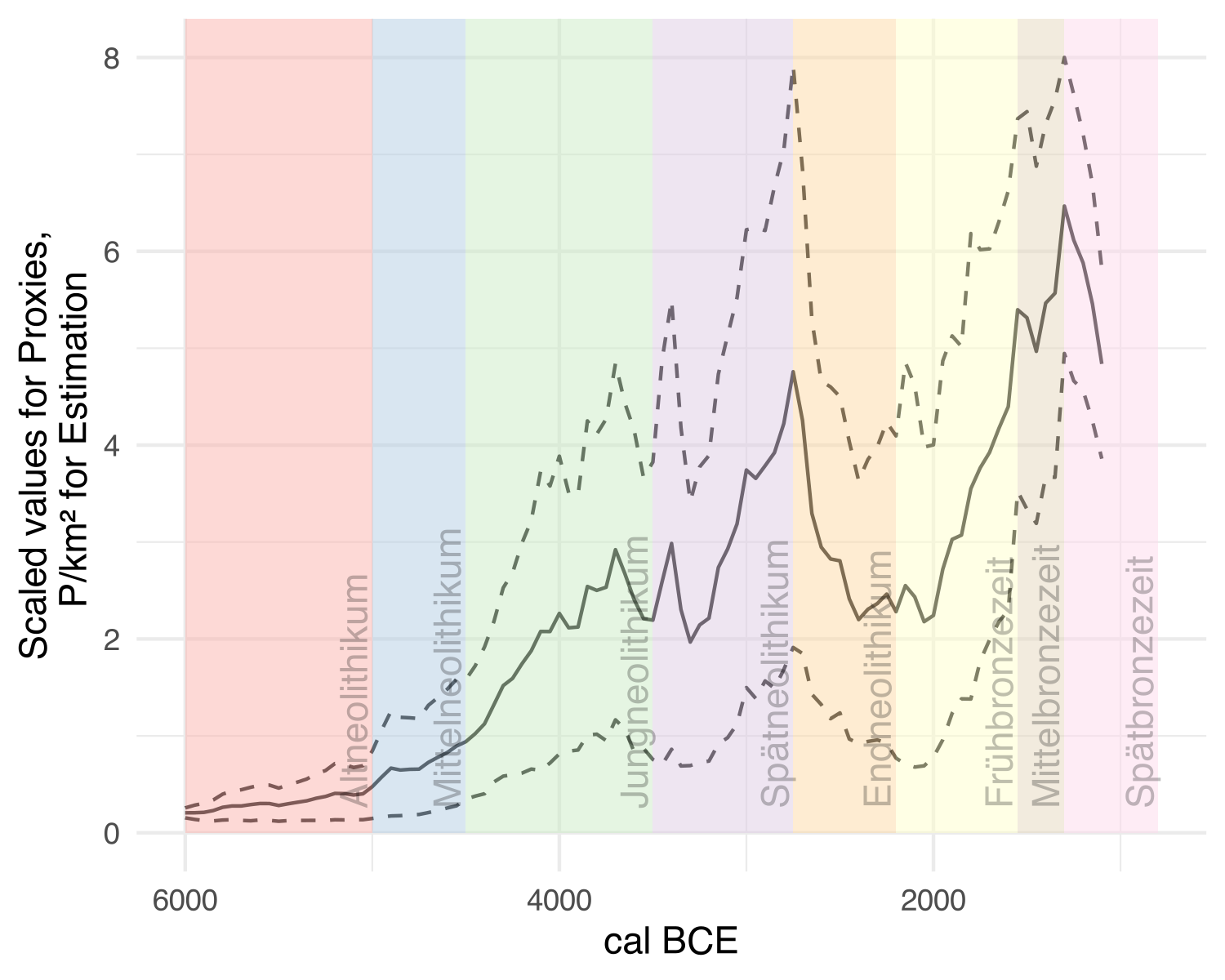


Figure 6.1: The estimate of the population density in relation to the chronology for the case study area north of the Swiss Alps.

The Early and Middle Neolithic are hardly documented with known sites in Switzerland. We must assume a low level of settlement, probably mainly by mobile groups. Isolated Neolithic sites of the LBK and later groups are known in the periphery of Switzerland, but they play a subordinate role (Ebersbach et al., 2012). The evidence of the Neolithic is dense from the so-called Upper Neolithic onwards, connected with the typochronological pottery phases of Egolzwil (late 5th M BCE) and Cortaillod respectively Pfyn (first half of the 4th M BCE). The first lake shore settlements north of the Alps date to this time too. Here we see a clear increase in the estimated population in the model. In the transition to the Late Neolithic we know from the lakeshore settlements the so-called Horgen Gap (Hafner, 2005). This is also visible as a slight decrease in the model. In another study (Heitz et al., 2021) we could show that this is in fact probably not a decline in population. Rather communities relocated their settlements to the hinterland of the large lakes in times of stronger lake level rises due to climatic changes. In the Late Neolithic, associated with the Horgen pottery, we then see a clear increase in the settlement intensity, which peaks and breaks off at the transition to the Final Neolithic (Hafner, 2004). In the second half of the Early Bronze Age, during which lakeshores were resettled to a smaller extend, there is again a clear increase in population size according to the model, continuing until the Late Bronze Age. The general trends fit very well with the previous reconstructions of population development for Switzerland (see eg. Lechterbeck et al., 2014), while offering higher precision and higher resolution.

# 7 Conclusions

The key advance in the model we presented is the ability to estimate, in absolute terms, past population sizes and the uncertainty accompanying our present knowledge. These estimates can be a basis for further studies where relative measures of population development are not helpful, such as long-term land use studies. Modelling of large-scale socio-ecological systems based on archaeological data does not have to rely deductive, asynchronous population models (e.g. carrying capacity or ethnographic analogues).

We have also demonstrated that, with Bayesian hierarchical modelling, it is possible to achieve a true multi-proxy analysis – as opposed to a juxtaposition of different indicators. This opens up the possibility of quantitatively linking different records and assessing their credibility. We are also able to specifying a confidence interval for the overall estimate. The result is a firmer basis for reconstructing population dynamics and settlement patterns in prehistory.

Nevertheless, we consider this model as only the first step towards a more sophisticated Bayesian approach. We have trusted the individual proxies in aggregate, without individualised measurement error. Our estimates are based on a limited number of sources, almost all of which are subject to taphonomic biases in the archaeological record. Consequently, we can only transform the model’s prediction into an absolute estimate of population density with predefined parameters: settlement size and the initial value of the reconstruction. Overcoming this limitation would represent a major refinement of our approach.

Incorporating additional proxies independent of the immediate, time-dependent conditions of the archaeological record could be one way to achieve this. These could be data on settlement sizes, parameters for economic-ecological carrying capacity, demographic data from burial groups or archaeogenetic data on population sizes. This data is available to varying degrees in different regions. On the Swiss Plateau, for example, we have too little data on human remains over large spans of prehistory compared to the abundance of wetland settlements.

To apply the model to different regions, the proxies we use here would have to be adapted to fit local conditions and research histories. By means of large-scale modelling, however, it would be possible to supplement gaps in the data in one region with data from another by regionalisation and a partial transfer of information (partial pooling). Such an extension would be the next logical step in the improvement of the model, to which end we hope to be able to contribute a further study in the near future.

# 8 Acknowledgements

Data collection were conducted as part of the project ‘Beyond lake settlements’ in the doctoral thesis of Julian Laabs, funded by the SNF (project number 152862, PI Albert Hafner) and as part of the XRONOS project, also funded by the SNSF (project number 198153, PI Martin Hinz). The development of the openness index took place within the framework of the project Time and Temporality in Archaeology (project number 194326, PI Caroline Heitz), inspired by the cooperation within the project QuantHum (project 169371, PI Marco Conedera), both also funded by the SNSF. Jan Kolář was supported by a long-term research development project (RV67985939) and by a grant from the Czech Science Foundation (19-20970Y). We also thank the Institute of Archaeological Sciences of the University of Bern for its support and faith in the outcome of our modelling project. Finally, we thank (already) the unknown reviewers for their helpful comments, which will certainly improve this manuscript significantly.

# 9 References

Attenbrow, V., Hiscock, P., 2015. Dates and demography: Are radiometric dates a robust proxy for long-term prehistoric demographic change? Archaeology in Oceania 50, 30–36. <https://doi.org/10.1002/arco.5052>

Bryant, J., Zhang, J.L., 2018. Bayesian Demographic Estimation and Forecasting. Chapman; Hall/CRC. <https://doi.org/10.1201/9780429452987>

Carleton, W.C., Groucutt, H.S., 2021. Sum things are not what they seem: Problems with point-wise interpretations and quantitative analyses of proxies based on aggregated radiocarbon dates. The Holocene 31, 630–643. <https://doi.org/10.1177/0959683620981700>

Childe, V.G. (Ed.), 1936. Man Makes Himself. Watts & Co., London.

Christian Lüthi, 2009. [Mittelland (region)](https://hls-dhs-dss.ch/de/articles/008566/2009-08-25/).

Crema, E.R., 2022. Statistical Inference of Prehistoric Demography from Frequency Distributions of Radiocarbon Dates: A Review and a Guide for the Perplexed. Journal of Archaeological Method and Theory. <https://doi.org/10.1007/s10816-022-09559-5>

Crema, E.R., Bevan, A., 2021. Inference from large sets of radiocarbon dates: software and methods. Radiocarbon 63, 23–39. <https://doi.org/10.1017/RDC.2020.95>

Crema, E.R., Shoda, S., 2021. A Bayesian approach for fitting and comparing demographic growth models of radiocarbon dates: A case study on the Jomon-Yayoi transition in Kyushu (Japan). PloS One 16, e0251695. <https://doi.org/10.1371/journal.pone.0251695>

David-Elbiali, M., 2000. La Suisse occidentale au IIe millénaire av. J.-C. : Chronologie, culture, intégration européenne. Cahiers d’archéologie romande, Lausanne.

Ebersbach, R., Kühn, M., Stopp, B., Schibler, J., 2012. Die Nutzung neuer Lebensräume in der Schweiz und angrenzenden Gebieten im 5. Jtsd. V. Chr. : Siedlungs- und wirtschaftsarchäologische Aspekte. Jahrbuch Archäologie Schweiz 95, 7–34. <https://doi.org/10.5169/SEALS-392483>

French, J.C., Riris, P., Fernandéz-López de Pablo, J., Lozano, S., Silva, F., 2021. A manifesto for palaeodemography in the twenty-first century. Philosophical Transactions of the Royal Society B: Biological Sciences 376, 20190707. <https://doi.org/10.1098/rstb.2019.0707>

Hafner, A., 2005. Neolithikum: Raum/Zeit-Ordnung und neue Denkmodelle. Archäologie im Kanton Bern 6B, 431–498.

Hafner, A., 2004. Vom Spät- zum Endneolithikum. Wandel und Kontinuität um 2700 v. Chr. In Mitteleuropa. - Vom Endneolithikum zur Frühbronzezeit: Wandel und Kontinuität zwischen 2400 und 1500 v.Chr, in: Beier, H.J., Einicke, R. (Eds.), Varia Neolithica. 3. Beiträge Zur Ur- Und Frühgeschichte Mitteleuropas. Beier&Beran, Weissbach, pp. 213–249.

Hafner, A., 1995. Die Frühe Bronzezeit in der Westschweiz. Funde und Befunde aus Siedlungen, Gräbern und Horten der entwickelten Frühbronzezeit. Staatlicher Lehrmittelverlag Bern 1995, Bern.

Hassan, F.A., 1981. Demographic archaeology, Studies in archaeology. Academic Press, New York.

Heitz, C., Hinz, M., Laabs, J., Hafner, A., 2021. Mobility as resilience capacity in northern Alpine Neolithic settlement communities. <https://doi.org/10.17863/CAM.79042>

Kintigh, K.W., Altschul, J.H., Beaudry, M.C., Drennan, R.D., Kinzig, A.P., Kohler, T.A., Limp, W.F., Maschner, H.D.G., Michener, W.K., Pauketat, T.R., Peregrine, P., Sabloff, J.A., Wilkinson, T.J., Wright, H.T., Zeder, M.A., 2014. Grand challenges for archaeology. Proceedings of the National Academy of Sciences 111, 879–880. <https://doi.org/10.1073/pnas.1324000111>

Kruschke, J.K., 2015. Doing Bayesian Data Analysis. Elsevier. <https://doi.org/10.1016/B978-0-12-405888-0.09999-2>

Laabs, J., 2019. Populations- und landnutzungsmodellierung der neolithischen und bronzezeitlichen westschweiz (PhD thesis). Bern.

Lechterbeck, J., Edinborough, K., Kerig, T., Fyfe, R., Roberts, N., Shennan, S., 2014. Is Neolithic land use correlated with demography? An evaluation of pollen-derived land cover and radiocarbon-inferred demographic change from Central Europe. The Holocene 24, 1297–1307. <https://doi.org/10.1177/0959683614540952>

Martínez-Grau, H., Morell-Rovira, B., Antolín, F., 2021. Radiocarbon Dates Associated to Neolithic Contexts (Ca. 5900 2000 Cal BC) from the Northwestern Mediterranean Arch to the High Rhine Area. Journal of Open Archaeology Data 9, 1. <https://doi.org/10.5334/joad.72>

Morris, I., 2013. The measure of civilization: How social development decides the fate of nations. Princeton University Press, Princeton.

Müller, J., Diachenko, A., 2019. Tracing long-term demographic changes: The issue of spatial scales. PLOS ONE 14, e0208739. <https://doi.org/10.1371/journal.pone.0208739>

Nikulka, F., 2016. [Archäologische Demographie. Methoden, Daten und Bevölkerung der europäischen Bronze- und Eisenzeiten](https://www.sidestone.com/books/archaeologische-demographie). Sidestone Press, Leiden.

Ramsey, C.B., 1995. [Radiocarbon calibration and analysis of stratigraphy; the OxCal program.](https://journals.uair.arizona.edu/index.php/radiocarbon/article/download/1690/1694) Radiocarbon 37, 425430.

Rick, J.W., 1987. Dates as Data: An Examination of the Peruvian Preceramic Radiocarbon Record. American Antiquity 52, 55–73. <https://doi.org/10.2307/281060>

Riede, F., 2009. Climate and demography in early prehistory: Using calibrated (14)C dates as population proxies. Human Biology 81, 309–337. <https://doi.org/10.3378/027.081.0311>

Schmidt, I., Hilpert, J., Kretschmer, I., Peters, R., Broich, M., Schiesberg, S., Vogels, O., Wendt, K.P., Zimmermann, A., Maier, A., 2021. Approaching prehistoric demography: Proxies, scales and scope of the Cologne Protocol in European contexts. Philosophical Transactions of the Royal Society B: Biological Sciences 376, 20190714. <https://doi.org/10.1098/rstb.2019.0714>

Shennan, S., 2000. Population, Culture History, and the Dynamics of Culture Change. Current Anthropology 41, 811–835.

Shennan, S., Downey, S.S., Timpson, A., Edinborough, K., Colledge, S., Kerig, T., Manning, K., Thomas, M.G., 2013. Regional population collapse followed initial agriculture booms in mid-Holocene Europe. Nature Communications 4, 2486. <https://doi.org/10.1038/ncomms3486>

Surovell, T.A., Byrd Finley, J., Smith, G.M., Brantingham, P.J., Kelly, R., 2009. Correcting temporal frequency distributions for taphonomic bias. Journal of Archaeological Science 36, 1715–1724. <https://doi.org/10.1016/j.jas.2009.03.029>

Turchin, P., Brennan, R., Currie, T., Feeney, K., Francois, P., Hoyer, D., Manning, J., Marciniak, A., Mullins, D., Palmisano, A., Peregrine, P., Turner, E.A.L., Whitehouse, H., 2015. Seshat: The Global History Databank. Cliodynamics 6. <https://doi.org/10.21237/C7clio6127917>

Zimmermann, A., 2004. Landschaftsarchäologie II. Überlegungen zu prinzipien einer landschaftsarchäologie.

### 9.0.1 Colophon

This report was generated on 2022-05-28 15:16:21 using the following computational environment and dependencies:

#> ─ Session info ───────────────────────────────────────────────────────────────  
#> setting value  
#> version R version 4.2.0 (2022-04-22)  
#> os Manjaro Linux  
#> system x86\_64, linux-gnu  
#> ui X11  
#> language (EN)  
#> collate C  
#> ctype de\_DE.UTF-8  
#> tz Europe/Zurich  
#> date 2022-05-28  
#> pandoc 2.17.1.1 @ /usr/bin/ (via rmarkdown)  
#>   
#> ─ Packages ───────────────────────────────────────────────────────────────────  
#> package \* version date (UTC) lib source  
#> assertthat 0.2.1 2019-03-21 [1] CRAN (R 4.2.0)  
#> bitops 1.0-7 2021-04-24 [1] CRAN (R 4.2.0)  
#> bookdown 0.26 2022-04-15 [1] CRAN (R 4.2.0)  
#> brio 1.1.3 2021-11-30 [1] CRAN (R 4.2.0)  
#> cachem 1.0.6 2021-08-19 [1] CRAN (R 4.2.0)  
#> callr 3.7.0 2021-04-20 [1] CRAN (R 4.2.0)  
#> class 7.3-20 2022-01-16 [2] CRAN (R 4.2.0)  
#> classInt 0.4-3 2020-04-07 [1] CRAN (R 4.2.0)  
#> cli 3.3.0 2022-04-25 [1] CRAN (R 4.2.0)  
#> colorspace 2.0-3 2022-02-21 [1] CRAN (R 4.2.0)  
#> cowplot \* 1.1.1 2020-12-30 [1] CRAN (R 4.2.0)  
#> crayon 1.5.1 2022-03-26 [1] CRAN (R 4.2.0)  
#> curl 4.3.2 2021-06-23 [1] CRAN (R 4.2.0)  
#> DBI 1.1.2 2021-12-20 [1] CRAN (R 4.2.0)  
#> desc 1.4.1 2022-03-06 [1] CRAN (R 4.2.0)  
#> devtools 2.4.3 2021-11-30 [1] CRAN (R 4.2.0)  
#> digest 0.6.29 2021-12-01 [1] CRAN (R 4.2.0)  
#> dplyr 1.0.9 2022-04-28 [1] CRAN (R 4.2.0)  
#> e1071 1.7-9 2021-09-16 [1] CRAN (R 4.2.0)  
#> ellipsis 0.3.2 2021-04-29 [1] CRAN (R 4.2.0)  
#> evaluate 0.15 2022-02-18 [1] CRAN (R 4.2.0)  
#> fansi 1.0.3 2022-03-24 [1] CRAN (R 4.2.0)  
#> farver 2.1.0 2021-02-28 [1] CRAN (R 4.2.0)  
#> fastmap 1.1.0 2021-01-25 [1] CRAN (R 4.2.0)  
#> foreign 0.8-82 2022-01-16 [2] CRAN (R 4.2.0)  
#> fs 1.5.2 2021-12-08 [1] CRAN (R 4.2.0)  
#> generics 0.1.2 2022-01-31 [1] CRAN (R 4.2.0)  
#> ggmap \* 3.0.0 2019-02-05 [1] CRAN (R 4.2.0)  
#> ggplot2 \* 3.3.6 2022-05-03 [1] CRAN (R 4.2.0)  
#> ggrepel \* 0.9.1 2021-01-15 [1] CRAN (R 4.2.0)  
#> ggsn \* 0.5.0 2019-02-18 [1] CRAN (R 4.2.0)  
#> glue 1.6.2 2022-02-24 [1] CRAN (R 4.2.0)  
#> gtable 0.3.0 2019-03-25 [1] CRAN (R 4.2.0)  
#> here \* 1.0.1 2020-12-13 [1] CRAN (R 4.2.0)  
#> highr 0.9 2021-04-16 [1] CRAN (R 4.2.0)  
#> htmltools 0.5.2 2021-08-25 [1] CRAN (R 4.2.0)  
#> httr 1.4.3 2022-05-04 [1] CRAN (R 4.2.0)  
#> jpeg 0.1-9 2021-07-24 [1] CRAN (R 4.2.0)  
#> KernSmooth 2.23-20 2021-05-03 [2] CRAN (R 4.2.0)  
#> knitr 1.39 2022-04-26 [1] CRAN (R 4.2.0)  
#> labeling 0.4.2 2020-10-20 [1] CRAN (R 4.2.0)  
#> lattice 0.20-45 2021-09-22 [2] CRAN (R 4.2.0)  
#> lifecycle 1.0.1 2021-09-24 [1] CRAN (R 4.2.0)  
#> magrittr 2.0.3 2022-03-30 [1] CRAN (R 4.2.0)  
#> maptools 1.1-4 2022-04-17 [1] CRAN (R 4.2.0)  
#> memoise 2.0.1 2021-11-26 [1] CRAN (R 4.2.0)  
#> munsell 0.5.0 2018-06-12 [1] CRAN (R 4.2.0)  
#> pillar 1.7.0 2022-02-01 [1] CRAN (R 4.2.0)  
#> pkgbuild 1.3.1 2021-12-20 [1] CRAN (R 4.2.0)  
#> pkgconfig 2.0.3 2019-09-22 [1] CRAN (R 4.2.0)  
#> pkgload 1.2.4 2021-11-30 [1] CRAN (R 4.2.0)  
#> plyr 1.8.7 2022-03-24 [1] CRAN (R 4.2.0)  
#> png 0.1-7 2013-12-03 [1] CRAN (R 4.2.0)  
#> prettyunits 1.1.1 2020-01-24 [1] CRAN (R 4.2.0)  
#> processx 3.5.3 2022-03-25 [1] CRAN (R 4.2.0)  
#> proxy 0.4-26 2021-06-07 [1] CRAN (R 4.2.0)  
#> ps 1.7.0 2022-04-23 [1] CRAN (R 4.2.0)  
#> purrr 0.3.4 2020-04-17 [1] CRAN (R 4.2.0)  
#> R6 2.5.1 2021-08-19 [1] CRAN (R 4.2.0)  
#> RColorBrewer 1.1-3 2022-04-03 [1] CRAN (R 4.2.0)  
#> Rcpp 1.0.8.3 2022-03-17 [1] CRAN (R 4.2.0)  
#> remotes 2.4.2 2021-11-30 [1] CRAN (R 4.2.0)  
#> rgdal 1.5-32 2022-05-09 [1] CRAN (R 4.2.0)  
#> RgoogleMaps 1.4.5.3 2020-02-12 [1] CRAN (R 4.2.0)  
#> rjson 0.2.21 2022-01-09 [1] CRAN (R 4.2.0)  
#> rlang 1.0.2 2022-03-04 [1] CRAN (R 4.2.0)  
#> rmarkdown 2.14 2022-04-25 [1] CRAN (R 4.2.0)  
#> rnaturalearth \* 0.1.0 2017-03-21 [1] CRAN (R 4.2.0)  
#> rprojroot 2.0.3 2022-04-02 [1] CRAN (R 4.2.0)  
#> rstudioapi 0.13 2020-11-12 [1] CRAN (R 4.2.0)  
#> s2 1.0.7 2021-09-28 [1] CRAN (R 4.2.0)  
#> scales 1.2.0 2022-04-13 [1] CRAN (R 4.2.0)  
#> sessioninfo 1.2.2 2021-12-06 [1] CRAN (R 4.2.0)  
#> sf \* 1.0-7 2022-03-07 [1] CRAN (R 4.2.0)  
#> sp \* 1.4-7 2022-04-20 [1] CRAN (R 4.2.0)  
#> stringi 1.7.6 2021-11-29 [1] CRAN (R 4.2.0)  
#> stringr 1.4.0 2019-02-10 [1] CRAN (R 4.2.0)  
#> testthat 3.1.4 2022-04-26 [1] CRAN (R 4.2.0)  
#> tibble 3.1.7 2022-05-03 [1] CRAN (R 4.2.0)  
#> tidyr 1.2.0 2022-02-01 [1] CRAN (R 4.2.0)  
#> tidyselect 1.1.2 2022-02-21 [1] CRAN (R 4.2.0)  
#> units 0.8-0 2022-02-05 [1] CRAN (R 4.2.0)  
#> usethis 2.1.5 2021-12-09 [1] CRAN (R 4.2.0)  
#> utf8 1.2.2 2021-07-24 [1] CRAN (R 4.2.0)  
#> vctrs 0.4.1 2022-04-13 [1] CRAN (R 4.2.0)  
#> withr 2.5.0 2022-03-03 [1] CRAN (R 4.2.0)  
#> wk 0.6.0 2022-01-03 [1] CRAN (R 4.2.0)  
#> xfun 0.31 2022-05-10 [1] CRAN (R 4.2.0)  
#> yaml 2.3.5 2022-02-21 [1] CRAN (R 4.2.0)  
#>   
#> [1] /home/martin/R/x86\_64-pc-linux-gnu-library/4.2  
#> [2] /usr/lib/R/library  
#>   
#> ──────────────────────────────────────────────────────────────────────────────