Event Date Model : A Robust Bayesian Tool for Chronology Building

Philippe Lanos, Anne Philippe

EVENT DATE MODEL: A ROBUST BAYESIAN TOOL FOR CHRONOLOGY BUILDING

LANOS PHILIPPE AND PHILIPPE ANNE

Abstract. We propose a robust event date model aiming to estimate the date of a target event by the combination of individual dates obtained from archaeological artifacts assumed to be contemporaneous. These dates are affected by errors of different types: laboratory and calibration curve errors, irreducible errors related to contaminations, taphonomic disturbances, etc, hence the possible presence of outliers. This modeling, based on a hierarchical Bayesian statistical approach, provides a very simple way to automatically penalize outlying data without having to remove them from the dataset. Prior information on the individual irreducible errors is introduced using a uniform shrinkage density with minimal assumptions about Bayesian parameters. We show that the event date model is more robust than models implemented in BCal or OxCal, although it generally yields less precise credibility intervals. The model is extended in the case of stratigraphic sequences which involve several events with temporal order constraints (relative dating), or with duration, hiatus constraints. Calculations are based on MCMC numerical techniques and can be performed using the ChronoModel software which is freeware, open source and cross-platform. Features of the software are presented in Vibet et al. (2016). We finally compare our prior on event dates implemented in ChronoModel with the prior in BCal and OxCal which involves supplementary parameters defined as boundaries to phases or sequences.

Target event date model ; robust combination of dates ; prior information on dates ; hierarchical Bayesian statistics ; individual errors ; outlier penalization ; MCMC computation ; Chronomodel software

1. Introduction

Bayesian chronological modeling appears as an important issue in archeology and palaeo-environmental sciences. This methodology has been developed since the 1990s (Bayliss, 2009, 2015) and is now the method of choice especially for the interpretation of radiocarbon dates. Most applications are undertaken using the flexible software packages, BCal (Buck et al., 1999), Datelab (Nicholls and Jones, 2002) and OxCal (Bronk Ramsey, 1995, 1998, 2001, 2008, 2009a,b; Bronk Ramsey et al., 2001, 2010; Bronk Ramsey and Lee, 2013).

All of these models provide an estimation of a chronology of dated events (DE), the estimated dates correspond to the dates of events that are actually dated by any chronometric technique. However calibrated radiocarbon dates and date estimates from other chronometric dating methods such as thermoluminescence, archaeomagnetism, dendrochronology, etc. can be combined with prior archeological

Date: November 21, 2017Corresponding author: Professor, Laboratoire de mathématiques Jean Leray, Université Nantes, UFR Sciences et techniques 2 rue de la Houssiniere 44322 Nantes France. E-mail anne.philippe@univ-nantes.fr .

information of various kinds to produce a combined chronology that should be more reliable than its individual components.

We propose a new chronological Bayesian model based on the concept of target event. This is related to the concepts of the dated event and the target event proposed by Dean (1978). The target event (TE) is the event to which the date is to be applied by the chronometrician. Usually, the target events are not directly related to the dated events. According to Dean (1978), the dated event is contemporaneous with the target event when there is convergence (the two dates are coeval: DE = TE) and when the date is relevant to the target event. "Relevance" refers to the degree to which the date is applicable to the TE : it must be demonstrated or argued on the basis of archaeological or other evidence. Because it is generally not easy or possible to assure the relevance of the dated event to the target event date of interest, it is recommended to get many dates of dated event , and if possible from different dating techniques.

These dates can be outliers without having a means to determine if they are dating anomalies. In other words, we do not have any convincing archaeological arguments for rejecting them before modeling. This motivates the development of a robust statistical model for combining theses dates in such a way that it is very little sensitive to outliers (see Lanos and Philippe, 2017). .

The target event model is a statistical model introduced in (Lanos and Philippe, 2017) for estimating the date of an event called "target event" . This model allows to combine in a robust way the dates of artifacts, which are assumed to be contemporary to this target event. To validate the robustness of this model to outliers, we provide a comparison of our model with the t-type outlier model implemented in Oxcal application. Numerical experiments also illustrate the sensibility to the outliers.

The target event model is then integrated to a most global model for constructing chronologies of target events. Prior information is brought upon the dates of the target events. We show how the prior archeological information based on relative dating between the target events (in a stratigraphic sequence for instance) or based on duration , hiatus or *Terminus post quem* (TPQ) or *terminus ante quem* (TAQ), assessment can be included in the model. The simulations and application illustrates the improvement brought in term of robustness. The proposed model is implemented in ChronoModel software, whose features are described in Vibet et al. (2016). This model can be compared with the standard chronological models when the target event is associated to only one dated event. However the interest of our approach is to get a robust estimation of the date of target event even if the dated events embeded dated events are outliers. We compare in the simulation part this approach with an outlier model based using discrete mixture distributions.

Many issues in archaeology raise the problem of phasing, that is how to characterize the beginning, the end and the duration of a given period. This question can be viewed as a post processing of the chronological model as for instance in Philippe and Vibet (2017a); Guérin et al. (2017). It is also possible to include additional parameters to characterise the phases. This requires the construction of a prior distribution on the dates of dated events, which belong to the phase. This point is discussed in this paper, we analyse the choice of model implemented in Oxcal application. We get an explicit form for the prior distribution of the dates, and we show that this probability distribution behaves in the same way as the dates included in

a target event in the sense that we observe a concentration of the dates belonging to the phase. Our result also shows the difficulty to construct appropriate prior on a sequence of dates.

The paper is organised as follows : In Section 2 we recall the construction of the target event model and we provide a theoretical comparison with the t-outlier model. In Section 3 we propose a new Bayesian model to estimate dates of target event. In Section 4, we analyse the Bayesian modelling of phases implemented in Oxcal application, and we compare this model with the target event model. In Sections 5 - 6 we apply our models to simulated and real datasets.

2. A robust combination of dates : the target event date model

We propose a statistical Bayesian approach for combining dates, which is based on a robust statistical model. This combination of dates is aiming to date a called "target" event defined on the basis of archaeological/historical arguments.

2.1. Description of the Bayesian model.

We propose to use a hierarchical Bayesian model to estimate the date θ of a target event Et. It combines dates t_i $(i = 1, ..., n)$ of dated events Ed in such a way that it is robust to outliers. This model is based on very few assumptions and does not need to tune hyper-parameters. It is described as follows.

The event dates t_i are estimated from n independent measurements (observations, also called determinations in Buck's terminology) M_i yielded by the different chronometric techniques. Each measurement M_i, obtained from a specific dating technique, can be related to an individual date t_i through a calibration curve g_i and its error σ_{g_i} (see Section 2 in Lanos and Philippe, 2017). Here this curve is supposedly known with some known uncertainty.

At this step, the random effect model can be written as follows:

$$\begin{aligned} M_i &= \mu_i + s_i\epsilon_i, \qquad \forall\, i = 1, ..., n \\ \mu_i &= g_i(t_i) + \sigma_{g_i}(t_i)\rho_i \end{aligned} \tag{1}$$

where $(\epsilon_1, ...\epsilon_n, \rho_1, ..., \rho_n)$ are independent and identically Gaussian distributed random variables with zero mean and variance 1, and where :

- $s_i\epsilon_i$ represents the experimental error provided by the laboratory.
- $\sigma_{g_i}(t_i)\rho_i$ represents the error provided by the calibration curve.

We assume that the experimental error provided by the laboratory are independent. Dependence structure could be added between the $(s_i\epsilon_i)_{i=1,...,n}$ in order to take into account the systematic error associated to the dating techniques of each laboratory (see Combs and Philippe (2017) in the particular case of the optically stimulated luminescence dating). However the construction of such a model requires additional information depending on each dating method and each laboratory, which is not available in most of the application of chronological models.

Remark 1. *If we assume that all the measurements can be calibrated with a common calibration curve (i.e. $g_i = g$ for all $i = 1, ..., n$), for example when the same object is analyzed by different laboratories, the measurements can be combined according to the R-combine model (Bronk Ramsey, 2009b) before being incorporated in the event date model. This model can be viewed as a degenerated version of the target event model by taking $\sigma_i = 0$ for all i. On the contrary, if several calibration*

curves are involved, the R-combine model is no longer valid and chronometric dates are directly incorporated in the event date model.

The target event date θ is estimated from the dated events t_i. The main assumption we make is the contemporaneity (relevance and convergence, following Dean (1978)), of the dates t_i with the event date θ. However, because of error sources of unknown origin, it can exist an "over-dispersion" or dating anomalies of the dates with respect to θ. These errors can come from the way of ensuring that the samples studied can realistically provide results for the events that we wish to characterize, the care in sampling in the field, the care in sample handling, preparation and measure in the laboratory, or other non-controllable random factors that can appear during the process (Christen, 1994). This over-dispersion is also similar to the irreducible error described in (Niu et al., 2013) in the framework of radiocarbon calibration curve building.

Consequently, we model the over-dispersion by an individual error σ_i according to the following random effect model:

$$t_i = \theta + \sigma_i \lambda_i \tag{2}$$

where $(\lambda_1, ..., \lambda_n)$ are independent and identically Gaussian distributed random variables with zero mean and variance 1. The individual error σ_i measures the degree of disagreement which can exist between a date (Ed) and its target event date (Et). It will be a posteriori small when the dates (Ed) are consistent with the target date θ and high when the date t_i is far from the target date θ. Consequently, the posterior distribution of the individual error σ_i will give some information about the "outlying" state of a date with respect to the target date.

Finally, the joint distribution of the probabilistic model can be written according to a Bayesian hierarchical structure:

$$p(M_1, ..., M_n, \mu_1, ..., \mu_n, t_1, ..., t_n, \sigma_1^2, ..., \sigma_n^2, \theta) = p(\theta) \prod_{i=1}^{n} p(M_i|\mu_i) p(\mu_i|t_i) p(t_i|\sigma_i^2, \theta) p(\sigma_i^2), \tag{3}$$

where the conditional distributions that appear in the decomposition are given by:

$$\begin{aligned} M_i|\mu_i &\sim \mathcal{N}(\mu_i, s_i^2), \\ \mu_i|t_i &\sim \mathcal{N}(g_i(t_i), \sigma_{g_i}^2(t_i)), \\ t_i|\sigma_i^2, \theta &\sim \mathcal{N}(\theta, \sigma_i^2), \\ \sigma_i^2 &\sim Shrink(s_0^2). \end{aligned} \tag{4}$$

The parameter of interest θ is assumed to be uniformly distributed on an interval $T = [T_a, T_b]$

$$\theta \sim \text{Uniform}(T). \tag{5}$$

This interval, called "study period", is fixed by the user based on historical or archeological evidences. This appears as an important a priori temporal information. Note that we do not assume that the same information is imparted to the dates t_i. Consequently their support is the set of real numbers $\mathbb{R}$.

The uniform shrinkage distribution for the variance σ_i^2, denoted $Shrink(s_0^2)$, admits as density

$$p(\sigma_i^2) = \frac{s_0^2}{(s_0^2+\sigma_i^2)^2} 1_{[0,\infty]}(\sigma_i^2), \tag{6}$$

where $1_A(x)$ is the indicator function ($= 1$ if $x \in A, = 0$ if $x \notin A$) and where the parameter s_0^2 must be fixed.

The motivation for this choice of prior is described in detail in (Lanos and Philippe, 2017). Parameter s_0^2 quantifies the magnitude of error on the measurements. It is estimated according to the following process:

(1) An individual calibration step is done for each measurement $M_i, i = 1, ..., n$. It consists of the simple model

$$M_i|t_i \sim \mathcal{N}(g_i(t_i), s_i^2 + \sigma_{g_i}^2(t_i)),$$
$$t_i \sim \text{Uniform}(T).$$

using the same notation as (4) and (5). For each i,

(a) Sample from the posterior distribution of t_i given M_i

$$p(t_i|M_i) \propto \frac{1}{S_i} \exp\left(\frac{-1}{2S_i^2}(M_i - g_i(t_i))^2\right) 1_T(t_i), \tag{7}$$

where $S_i^2 = s_i^2 + \sigma_{g_i}^2(t_i)$

(b) Approximate the posterior variance $\text{var}(t_i|M_i)$ by its Monte Carlo approximation denoted w_i^2.

(2) Take as shrinkage parameter s_0^2:

$$\frac{1}{s_0^2} = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{w_i^2}.$$

Remark 2. *The calibration curve g_i in radiocarbon or in archaeomagnetic dating is always defined on a bounded support $[T_m, T_M]$. Therefore an extension of g_i on $\mathbb{R}$ is required to well defined the conditional distribution of t_i in (4).*

We suggest to extend the calibration curve by an arbitrary constant value with a very large variance with respect to the known reference curve: for instance

$$g_i(t) = (g_i(T_m) + g_i(T_M))/2$$

and

$$\sigma_{g_i}^2 = 10^6 \left(\sup_{t\in[T_m;T_M]}(g_i(t)) - \inf_{t\in[T_m;T_M]}(g_i(t))\right)^2$$

In the case of TL/OSL or Gauss measurements, there is no need for an extension because $g_i(t)$ is defined whatever t in $\mathbb{R}$.

This statistical approach does not model the outliers, i.e. we do not estimate the posterior probability that a date is an outlier. Outlier modeling (See section 2.2 for such an approach) can provide more accurate results, but it often requires two (maybe more) estimations of the model: the outliers are identified after a first estimation and thus discarded from the dataset. Then the final model is estimated again from the new dataset (however, the question remains to know when to stop, because new outliers can appear during the subsequent runs). The event model which is based on the choice of robustness, avoids this two-step procedure.

2.2. **Comparison with alternative outlier event models.** Several outlier models have been implemented since the 90's, in the framework of radiocarbon dating. So we can distinguish between mainly three outlier modelings:

(1) Outlier model with respect to the measurement parameter (Christen, 1994; Buck et al., 2003).
(2) Outlier model with respect to the laboratory variance parameter (Christen and Pérez, 2009).
(3) Outlier model with respect to the time parameter. This corresponds to the t-type outlier model (Bronk Ramsey, 2009b)

The t-type outlier model can be compared to the event model, by considering that the date parameter t directly coincides with our target event date parameter θ. It means that the hierarchical level between t and θ does no longer exist in the t-type outlier model. Hence the observed measurement will be noted M_j instead of M_i, the true unknown measurement noted μ_j instead of μ_i and a date noted t_j instead of t_i in order to avoid any confusion with index i used for chronometric dates (dated events) in the target event date model.

The model with random effect becomes :

$$\begin{aligned} M_j &= \mu_j + s_j\epsilon_j, \\ \mu_j &= g_j(t_j + \delta_j\phi_j 10^u) + \sigma_{g_j}(t_j + \delta_j\phi_j 10^u)\rho_j \end{aligned} \tag{8}$$

where

- $(\epsilon_1, ...\epsilon_r, \rho_1, ..., \rho_r)$ are independent and identically Gaussian distributed random variables with zero mean and variance 1.
- $s_j\epsilon_j$ represents the experimental error provided by the laboratory and $\sigma_{g_j}(t_j + \delta_j\phi_j 10^u)\rho_j$ the calibration error.
- the prior on ϕ_j is the Bernoulli distribution with parameter p_j. A priori, ϕ_j takes the value 1 if the measurement requires a shift and 0 otherwise. In practice p_j must be chosen and the recommended values are 0.1 in Christen (1994); Buck et al. (2003) or 0.05 in Bronk Ramsey (2009b).
- δ_j corresponds to the shift on the measurement M_j if it is detected as an outlier.
- u is a scale parameter to offset δ_j. Parameter u can be fixed (for instance 0) or prior distributed as Uniform([0,4])

When parameter u is fixed, the joint distribution of the probabilistic model for r dates t_j $(j = 1, ..., r)$ can be written in the form:

$$p(M, \mu, t, \delta, \phi) = \prod_{j=1}^{r} p(M_j|\mu_j)p(\mu_j|t_j, \delta_j, \phi_j)p(t_j)p(\delta_j)p(\phi_j), \tag{9}$$

where the conditional distributions that appear in the decomposition are given by:

$$\begin{aligned} M_j|\mu_j &\sim \mathcal{N}(\mu_j, s_j^2) \\ \mu_j|t_j &\sim \mathcal{N}(g_j(t_j + \delta_j\phi_j), \sigma_{g_j}^2(t_j + \delta_j\phi_j)) \\ t_j &\sim \text{Uniform}([T_a, T_b]) \\ \delta_j &\sim \mathcal{N}(0, \sigma_\delta^2) \quad \text{or} \quad \sim \mathcal{T}(\nu) : \text{ Student's t- distribution with } \nu \text{ degrees of freedom} \\ \phi_j &\sim \text{Bernoulli}(p_j) \end{aligned}$$

This modeling can be compared to the event date model if we consider several dates nested in the Oxcal function "Combine" with t-type outlier model. The dates t_j $(j = 1, ..., r)$ then becomes a common date t. Consequently, (9) is transformed into:

$$p(M, \mu, t, \delta, \phi) \propto p(t) \prod_{j=1}^{r} p(M_j|\mu_j) p(\mu_j|t, \delta_j, \phi_j) p(\delta_j) p(\phi_j), \tag{10}$$

If we consider linear calibration curves with constant errors, that is, by setting:

- $g_j(t) = t$
- $\sigma_{g_j}(t) = \sigma_g$,

and knowing that

- $\delta_j \sim \mathcal{N}(0, \sigma_\delta^2)$
- $\phi_j \sim \text{Bernoulli}(p)$,

it is possible to analytically integrate (10) with respect to δ_j , ϕ_j, and μ_j. The posterior probability density of t is then given by:

$$p(t|M) \propto p(t) \prod_{j=1}^{r} p(M_j|t), \tag{11}$$

where the conditional distribution of M_j given t is a finite mixture distribution defined by

(12)

$$p(M_j|t) = p \frac{1}{\sqrt{2\pi}\sqrt{s_j^2 + \sigma_g^2 + \sigma_\delta^2}} e^{\frac{-(M_j - t)^2}{2(s_j^2 + \sigma_g^2 + \sigma_\delta^2)}} + (1 - p) \frac{1}{\sqrt{s_j^2 + \sigma_g^2}\sqrt{2\pi}} e^{\frac{-(M_j - t)^2}{-2(s_j^2 + \sigma_g^2)}}.$$

Note that the outlier model operates well as a model averaging thanks to the structure of mixture distribution in (12).

On the other hand, posterior density of event date $\theta = t$ deduced from (3) can be compared to density of time t in (11) after integration with respect to $\mu_i = \mu_j$ and $\sigma_i^2 = \sigma_j^2$. The posterior probability density of θ is given by:

$$p(\theta|M) \propto p(\theta) \prod_{j=1}^{r} p(M_j|\theta), \tag{13}$$

where

$$p(M_j|\theta) = \int_0^\infty \frac{1}{\sqrt{2\pi(s_j^2 + \sigma_g^2 + \sigma_j^2)}} e^{\frac{1}{-2(s_j^2 + \sigma_g^2 + \sigma_j^2)}(M_j - \theta)^2} \frac{s_{0j}^2}{(s_{0j}^2 + \sigma_j^2)^2} \, \mathrm{d}\sigma_j^2 \tag{14}$$

A graphical representation of the densities defined in (12) and (14) are shown in Figure 1 with the following parameter values: $s_j = 30$, $\sigma_g = 10$, $\sigma_\delta = 10^2$ and $\theta = t = 1000$. The density (12) is plotted for three different values $p = 0.01, 0.05, 0.10$. We can observe that shrinkage modeling in (14) leads to a more diffuse density making it possible to better take into account the possible presence of outliers.

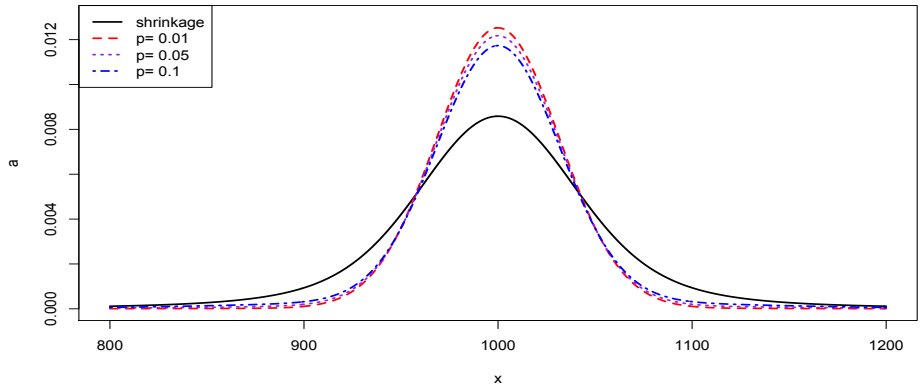

FIGURE 1. Graphical representation of the densities defined in (12) and (14) with the following parameter values: $s_j = 30$, $\sigma_g = 10$, $\sigma_\delta = 10^2$ and $\theta = t = 1000$

3. Prior information on groups of target event dates

We now consider a group of dates θ_j $(j = 1, ..., r)$ of target events which belong either to a "stratigraphic phase defined as a group of ordered contexts, or to a "chronological phase defined as a set of contexts build on the basis of archaeological, architectural, geological, environmental,...criteria. In practice, a "context" is defined by the nature of the stratification at a site, and the excavation approach used by the archaeologist. Together, these two determine the smallest units of space and time -the context- that can be identified in the stratigraphic record at an archaeological site. Because the term "phase" is also used to describe Bayesian chronological models, Dye and Buck (2015) use the term "stratigraphic phase to refer to a group of ordered contexts, and the term "chronological phase to refer to a time period in a chronological model.

The concept of phase is currently implemented in chronological modeling by using a specific parametrization which will be discussed in details in Section 4. Hereafter, we extend the target event date model in a simple way to the case of a group of dates related to each other by order or duration relationships. Hence we propose to estimate the beginning, the end and the duration of a phase directly from the group of target dates, without adding any supplementary parametrization.

3.1. Characterization of a group of target event dates.

We consider a collection of r event dates denoted θ_j with $j = 1, ..., r$. For each event date θ_j, we suppose that n_j measurements denoted $\mathbf{M}_j = (M_{j1}, \ldots, M_{jn_j})$ charactering the event are observed. We assume the independence of the measurements conditionnally to the event dates $(\theta_1, ..., \theta_r)$

$$p(\mathbf{M}_1, \ldots, \mathbf{M}_r, \theta_1, ..., \theta_r) = p(\theta_1, ..., \theta_r) \prod_{j=1}^{r} p(\mathbf{M}_j|\theta_j)$$

where $p(\mathbf{M}_j|\theta_j)$ is the conditionnal distribution of $\mathbf{M}_j$ given θ_j in the hierarchical model defined in (3). The Bayesian model is summarized by Figure 2.

When there is no supplementary information, we assumed that the dates θ_j are independent and uniformly distributed on the time interval (study period) $T = [T_a, T_b]$. We remind that this interval is fixed by the user based on historical or archeological evidences. Consequently, the prior on vector θ is such that:

$$p(\theta_1, ..., \theta_r) = \frac{1}{(T_b - T_a)^r} \prod_{j=1}^{r} 1_T(\theta_j) \tag{15}$$

The overall Bayesian model is described in Figure 2 using a directed acyclic graph (DAG). Such a graph describes the dependencies in the joint distribution of the probabilistic model. Each random variable of the model (that is an observation or a parameter) appears as a node in the graph. Any node is conditionally independent of its non-descendants given its parents. The circles correspond to all the random variables of the model. With the color of the circles, we distinguish between observations (red), parameters (blue) and exogenous variables (green).

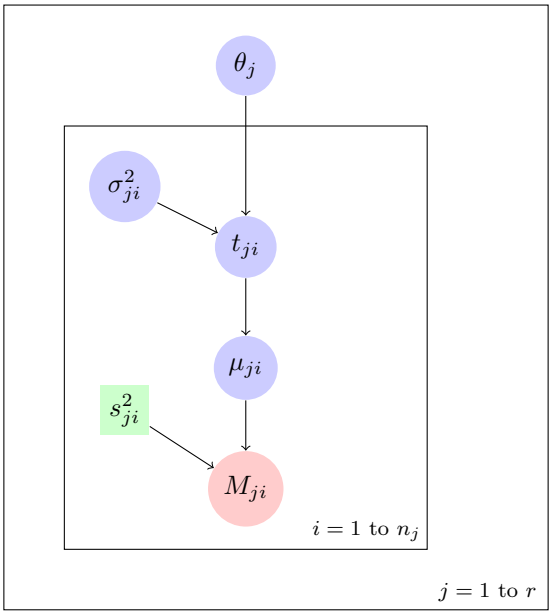

FIGURE 2. DAG for the hierarchical model of a group of event dates.

Almost always we do not know how a group of event dates can a priori be distributed in a phase. It means that, in our approach, a phase does not respond to a statistical model. From a chronological point of view, all the information is carried by the target event dates themselves. Consequently, we estimate the beginning and the end of a phase as a measurable application of the parameters θ_j $(j = 1, ..., r)$. Thus the beginning of a phase is estimated by the minimum $\theta_{(1)}$ of the r event dates included in the phase

$$\theta_{(1)} = \min(\theta_j, \ \ j = 1, \ldots, r). \tag{16}$$

In the same way, the end of a phase is estimated by the maximum $\theta_{(r)}$ of the r event dates included in the phase:

$$\theta_{(r)} = \max(\theta_j, \ j = 1, \ldots, r). \tag{17}$$

By the plug-in principle, we estimate the duration of the phase by

$$\tau = \theta_{(r)} - \theta_{(1)} \tag{18}$$

Considering two phases $P_k = \{\theta_{k,1}, ..., \theta_{k,r_k}\}$, $k = 1, 2$, the hiatus between P_1 and P_2 is the time gap between the end of P_1 and the beginning P_2. The hiatus is estimated by :

$$\gamma = \max(\theta_{2,(1)} - \theta_{1,(r_1)}\ , 0) \tag{19}$$

where $\theta_{2,(1)} = \min(\theta_{2,1}, ..., \theta_{2,r_2})$ and where $\theta_{1,(r_1)} = \max(\theta_{1,1}, ..., \theta_{1,r_1})$. Note that the estimate γ takes the value 0 if $\theta_{1,(r_1)} \leq \theta_{2,(1)}$, that corresponds to the absence of hiatus. The conditional distributions of the parameters $\theta_{(1)}, \theta_{(r)}, \gamma$ and τ given the observations can be easily derived from the joint posterior distributions of the event dates.

Remark 3. *Let us remind that these parameters are estimated knowing the data, that is to say from the target event dates which are available in the phase of interest. In particular, a good estimation of the beginning (resp. the end) of a phase requires that the archaeologist has sampled artifacts belonging to target events which are very near to this beginning (resp. this end).*

These estimates are valid whatever the prior on the event dates are. Precision on the estimation of these parameters can be gained if it is possible to add supplementary information on the target event dates within the study period T. For instance some temporal orders induce some restrictions on the support of the distribution.

Different prior on dates θ_j can be defined according to the following circumstances:

- Relative dating based on stratigraphy as defined in Harris (1989) and Desachy (2005, 2008), can imply antero-posteriority relationships between target dates θ_j. This can also imply antero-posteriority relationships between groups of dates θ_j, these groups defining different phases.
- We can have some prior information about the maximal duration of a group of target events.
- We can also have some prior information about the minimal temporal hiatus between two groups of target events.

It has been demonstrated that these types of supplementary prior information can significantly improve chronometric dates (see initial works of Naylor and Smith (1988); Buck et al. (1991, 1992, 1994, 1996) and Christen (1994)). In the next three subsections, we discuss each of these prior information in the framework of the target event date model. Our modeling approach is very soft thanks to the fact that all the relationships operate directly onto the target event dates θ_j. As a consequence, in a global modeling project, different phasing systems (multiphasing) can be defined on the basis of different criteria and so a phasing system can intersect an other phasing system: for example a ceramic phasing can intersect a lithic phasing in the sense that some target events can belong to two or several phases. Such a constraints network, when available, can significantly contribute to improve the precision of the estimates.

3.2. **Prior information on temporal order.** Target events or phases (groups) of target events can have to check order relationships. This order can be defined in different ways: by the stratigraphic relationship (physical relationship observed in

the field) or by stylistic, technical, architectural,... criteria which may be a priori known. Thus, the constraint of succession is equivalent to a hiatus of unknown amplitude put between event dates or groups (phases) of event dates.

If we consider a stratigraphic sequence composed of target events, the prior on vector θ becomes:

$$p(\theta_1, ..., \theta_r) \propto \frac{1}{(T_b - T_a)^r} 1_C(\theta_1, ..., \theta_r) \tag{20}$$

with $C = S \cap T^r$ and

- $T^r = [T_a, T_b]^r$ the support which defines the study period,
- $S =$ the group of r-uplets event dates θ_j which respect total or partial order relationships.

We can also consider two groups of target events, $P_k = \{\theta_{k,1}, ..., \theta_{k,r_k}\}$, $k = 1, 2$, containing r_k event dates such that all the dates of P_1 are before all the dates of P_2. The following equations should be verified by all the events included in both phases.

$$\forall j \in \{1, ..., r_1\}, \forall l \in \{1, ..., r_2\}, \theta_{1,j} < \theta_{2,l}$$

or

$$\max(\theta_{1,j},\, j = 1, \ldots, r_1) < \min(\theta_{2,l},\, l = 1, \ldots, r_2)$$

or

$$\min(\theta_{2,l},\, l = 1, \ldots, r_2) - \max(\theta_{1,j},\, j = 1, \ldots, r_1) > 0 \tag{21}$$

The r-uplets event dates θ_j will then have to check total or partial stratigraphic constraints and also to satisfy the inequality (21) between event dates of the two phases. We can see that a succession constraint between phases operates in the same mathematical way as for a set of stratigraphic constraints put between all the individual target events.

Remark 4. *Estimating the date of a target event needs to incorporate several dates (Ed), otherwise the event date modeling will not yield a better posterior information. However, it is possible to nest only one date per target event provided that the group of events is constrained by temporal order.*

It is not rare to encounter dating results which contradict the stratigraphic order: one speaks of "stratigraphic inversion". This situation often occurs when some artifact movements are provoked for example by bioturbations or establishment of backfill soils. The target event date model makes it possible to manage such situations thanks to the individual variances σ_i^2 which automatically penalize the dates that are inconsistent with the stratigraphic order. This is illustrated in Section 5.2.

3.3. **Prior information about the duration.** Prior information can be included on the duration of a phase, that is on a group of target events. We can impose a maximal duration τ_0. This means that all the event dates θ_j $(j = 1, ..., r)$ in the phase have to verify the constraint of duration according to the following equation:

$$\max(\theta_j,\, j = 1, \ldots, r) - \min(\theta_j,\, j = 1, \ldots, r) \leq \tau_0 \tag{22}$$

The r-uplets event dates θ_j will have to check total or partial stratigraphic constraints and also to satisfy the inequality 22 between event dates in the phase. This means that any r-uplet θ_j sampled during the MCMC process has a duration of at the most τ_0.

3.4. **Prior information about the amplitude of a hiatus.** Prior information about a hiatus γ between two phases $P_k = \{\theta_{k,1}, ..., \theta_{k,r_k}\}$, $k = 1, 2$, may be available. Hence we can impose that the amplitude of the hiatus is higher than a known value $\gamma_0 > 0$. All the event dates of phase P_1 and phase P_2 have to verify the following constraint :

$$\min(\theta_{2,l},\, l = 1, ..., r_2) - \max(\theta_{1j},\, j = 1, ..., r_1) \geq \gamma_0 \tag{23}$$

The event dates will have to check total or partial stratigraphic constraints and also to satisfy the inequality (23). This means that any r-uplet $\theta_{1,j}$ from phase P_1 and sampled during the MCMC process is separated by a time span of at least γ_0 from the r-uplet $\theta_{2,l}$ of the next phase. Note that it is obviously not possible to impose a hiatus between two phases when a same event belongs to these two phases.

3.5. **Prior information about known event dates: the bounds.** Bounds, such as historical dates, *Terminus post quem* (TPQ) or *terminus ante quem* (TAQ), may also be introduced in order to constrain one or several event dates θ_j. If we consider a set of r events assumed to happen after a special event with true calendar date B such that

$$B < (\theta_1, ..., \theta_r)$$

This condition must be included in the set of constraints that define the support of the prior distribution of the event dates.

The bound can be also defined with an uncertainty, i.e. $B \in [B_a, B_b] \subset [T_a, T_b]$, and so it is included in the set of the parameters. The prior density of the event dates can then be written as:

$$p(\theta_1, ..., \theta_r, B) = p(\theta_1, ..., \theta_r | B) p(B)$$

where

$$p(\theta_1, ..., \theta_r | B) = \frac{1}{(T_b - B)^r} \prod_{j=1}^{r} 1_{[B, T_b]}(\theta_j)$$

and

$$p(B) = \frac{1}{B_b - B_a} 1_{[B_a, B_b]}(B)$$

It is important to note that the introduction of bounds B in a global model composed of groups (phases) of event dates must remain consistent with other constraints of temporal order, of duration or of hiatus described in previous sections.

4. Discussion on date prior probabilities

In this section we discuss the question of the specification of the prior on the event dates θ_j. When there is no supplementary information, it may seem natural to assume that the dates θ_j are independent and uniformly distributed on the time interval (study period) $T = [T_a, T_b]$. However, from an archaeological point of view, it may seem more natural to assert that the span $\Delta = \max(\theta_j) - \min(\theta_j)$ has a uniform prior distribution We can see that this assumption is not checked when starting from the prior density (15). After an appropriate change of variable and

letting $R = T_b - T_a$, we can determine the density of the span according to the number of dates r and regardless their order:

$$p(\Delta) = \frac{r(r-1)}{R^r}(R-\Delta)\Delta^{r-2} \tag{24}$$

A span of 2Δ is favoured over a span of Δ by a factor of about Δ^{r-2} when $\Delta \ll R$ (spreading tendency when r becomes large). This behavior occurs regardless of the order between dates.

Naylor and Smith (1988); Buck et al. (1992); Christen (1994); Buck et al. (1996); Nicholls and Jones (2002) propose a specific phase modeling which aims to avoid this spreading bias. We study the properties of this modeling implemented in BCal and OxCal software. The Naylor-Smith-Buck-Christen (NSBC) prior is defined for a group of event dates which are placed between two additional hyperparameters α and β in the Bayesian hierarchical structure. These hyperparameters α and β represent boundaries where α is the beginning and β the end of the group of events called Phase. The dates θ_j are assumed to be independent conditionally to these two boundaries. Then, the prior density on vector θ becomes:

$$p(\theta_1, \ldots, \theta_r | \alpha, \beta) = \frac{1}{(\beta-\alpha)^r} 1_{[\alpha,\beta]^r}(\theta_1, \ldots, \theta_r) \tag{25}$$

In the absence of supplementary information, a non informative prior density is assigned to α and β :

$$p(\alpha, \beta) = \frac{2}{(T_b - T_a)^2}.1_P(\alpha, \beta) \tag{26}$$

with $P = \{(\alpha, \beta) \mid T_a \leq \alpha \leq \beta \leq T_b\}$. The pairs (α, β) are uniformly distributed on the triangle P.

From these hypotheses, it is possible to calculate the prior joint probability density of the event dates $(\theta_1 \ldots \theta_r)$. This is carried out by integration against α and β between the two limits T_a and T_b. For 2 (ordered or non ordered) event dates θ_1, θ_2, we obtain:

$$\begin{aligned}
p(\theta_1, \theta_2) &= \int p(\theta_1, \theta_2 | \alpha, \beta).p(\alpha, \beta)\, \mathrm{d}\alpha \mathrm{d}\beta \\
&= \int \frac{1}{(\beta-\alpha)^2}.1_{[\alpha,\beta]}(\theta_1).1_{[\alpha,\beta]}(\theta_2) \frac{2}{(T_b-T_a)^2}.1_{[T_a \leq \alpha \leq \beta \leq T_b]}(\alpha, \beta).\, \mathrm{d}\alpha \mathrm{d}\beta \\
&= \frac{2}{(T_b-T_a)^2} \int_{T_a}^{\min(\theta_1,\theta_2)} \int_{\max(\theta_1,\theta_2)}^{T_b} \frac{1}{(\beta-\alpha)^2}\, \mathrm{d}\beta\, \mathrm{d}\alpha \\
&= \frac{2}{(T_b-T_a)^2} \int_{T_a}^{\min(\theta_1,\theta_2)} [\frac{1}{(\max(\theta_1,\theta_2)-\alpha)} - \frac{1}{(T_b-\alpha)}]\, \mathrm{d}\alpha
\end{aligned}$$

Finally, we have:

$$\begin{aligned}
p(\theta_1, \theta_2) = \frac{4}{(T_b-T_a)^2}(&- \ln(\max(\theta_1,\theta_2) - \min(\theta_1,\theta_2)) + \ln(T_b - \min(\theta_1,\theta_2)) \\
&+ \ln(\max(\theta_1,\theta_2) - T_a) - \ln(T_b - T_a))1_{[T_a,T_b]^2}(\theta_1, \theta_2)
\end{aligned} \tag{27}$$

The joint probability density for r (ordered or non ordered) event dates $(\theta_j, j = 1, \ldots, r)$ with $r \geq 3$, is obtained by following the same integration process:

$$p(\theta_1, ..., \theta_r) = \frac{2\ r!}{(r-1)(r-2)(T_b - T_a)^2} \left(\frac{1}{[\theta_{(r)} - \theta_{(1)}]^{r-2}} - \frac{1}{[T_b - \theta_{(1)}]^{r-2}} - \frac{1}{[\theta_{(r)} - T_a]^{r-2}} + \frac{1}{[T_b - T_a]^{r-2}} \right) 1_{[T_a, T_b]^2}(\theta_{(1)}, \theta_{(r)}) \quad (28)$$

where $\theta_{(r)} = \max(\theta_j,\ j = 1, \ldots, r)$ and $\theta_{(1)} = \min(\theta_j,\ j = 1, \ldots, r)$.

Equations (27) and (28) clearly show that the priors (25) and (26) provoke a strong concentration effect of the dates θ_j as illustrated in Figure 3 which is calculated for two event dates with formula (27). The region of high probability of (θ_1, θ_2) is concentrated around the first diagonal $(\theta_1 = \theta_2)$. In conclusion, the NSBC prior clearly favors the fact that dates $(\theta_1, ..., \theta_r)$ become near each other. This property of the phase can be compared with the assumption of contemporaneity imposed on the dates in the event model.

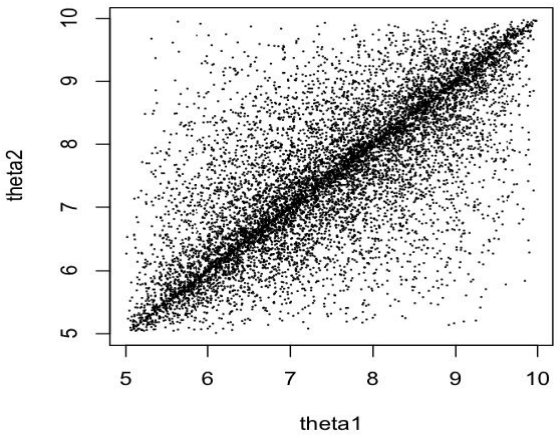

FIGURE 3. "NSBC" phase model: prior joint density on two (non ordered) event dates θ_1 and θ_2 between two boundaries α and β and after integration against α and β: a concentration effect appears around values $\theta_1 = \theta_2$.

Starting from the prior densities (27) and (28), and after an appropriate change of variable, we can also determine the density of the span $\Delta = \max(\theta_j,\ j = 1, ..., r) - \min(\theta_j,\ j = 1, ..., r)$, according to the number of dates r. Letting $R = T_b - T_a$, we obtain :

$$r = 2 : p(\Delta) = \frac{4}{R^2} \left[(R + \Delta)(\ln(R) - ln(\Delta)) + 2(\Delta - R) \right] \quad (29)$$

$$r = 3 : p(\Delta) = \frac{6}{R^2} \left[(R - \Delta)(1 + \frac{\Delta}{R}) + 2\Delta(\ln(\Delta) - ln(R)) \right] \quad (30)$$

$$r \geq 4 : p(\Delta) = \frac{2r}{(r-2)R^2} \left[(R - \Delta)(1 + \frac{\Delta^{r-2}}{R^{r-2}}) - \frac{2}{r-2)}(\Delta - \frac{\Delta^{r-2}}{R^{r-3}}) \right] \quad (31)$$

These distributions are shown in Figure 4. When r is small ($r = 2, 3, 4$), the densities are very high for Δ near zero, hence a high concentration effect. When r tends towards infinity, the distribution of Δ tends to $p(\Delta) = \frac{2}{R^2}(R - \Delta)$. This latter triangular distribution is still maximal when $\Delta = 0$, and it becomes more and more flat as R increases. Nicholls and Jones (2002) proposed to make the prior distribution on the dates θ_j uniform by multiplying the prior in (26) by $\frac{R^2}{2(R-\Delta)}$ where $\min(\theta_j,\ j = 1, ..., r)$ is replaced by α and $\max(\theta_j,\ j = 1, ..., r)$ by β . This option is implemented in OxCal when setting `UniformSpanPrior=true'` in a modeling project (Bronk Ramsey, 2009a).

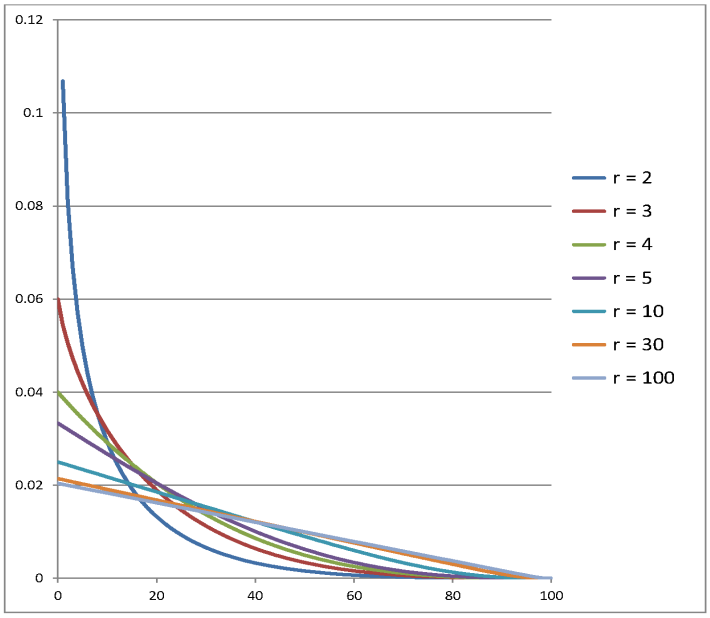

FIGURE 4. Distributions of the span Δ pour different values of the number r of event dates in a NSBC phase model with: $T_a \leq \alpha \leq \theta_1, \ldots, \theta_r \leq \beta \leq T_b$. When r is small ($r = 2, 3, 4$), the densities are very high for Δ near zero, reflecting a high concentration effect.

Instead of looking at the span Δ, we can look at the variance of the event dates θ_j, which is more representative of their scattering. This variance is proportional to the sequence of the Euclidian distance of the event dates to the straight line $\theta_1 = \theta_2 = ... = \theta_r$ in the space of dimension r and therefore it gives a good way to characterize the scattering of the dates. We observe that this distance remains near zero whatever the number r of event dates. Figure 5 shows an evaluation of the dispersion of the dates through the density of the statistic

$$\frac{1}{r}\sum_{i=1}^{r}\left(\theta_i - \frac{1}{r}\sum_{j=1}^{r}\theta_j\right)^2 \tag{32}$$

obtained with $r = 10$. Conditionally to (α, β), the event dates are uniformly sampled between α and β and the distribution of (α, β) is given by

$$p(\alpha) = \frac{2(T_b - \alpha)}{(T_b - T_a)^2}$$
$$p(\beta|\alpha), = \frac{1}{(T_b - \alpha)} 1_{[\alpha, T_b]}(\beta),$$

Note that we simulate a sample from the prior distribution of (α, β) as follows :

1. `Generate` $u \sim$ `Uniform`$(0, 1)$
2. `Take` $\alpha = T_b - R\sqrt{1 - u}$
3. `Generate` $\beta \sim$ `Uniform`(α, T_b).

The same result is obtained whatever the partial or total order of the event dates.

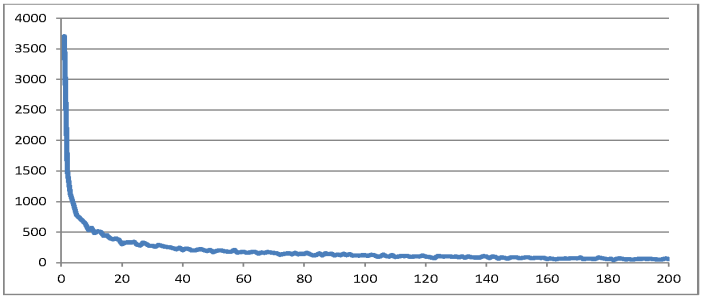

FIGURE 5. An other way to look at the concentration effect in the NSBC phase model: distribution of (32) for $r = 10$ event dates with: $T_a = 0 \leq \alpha \leq \theta_1, \ldots, \theta_{10} \leq \beta \leq 100 = T_b$.

The concentration effect is particularly visible when considering dates with different uncertainties. In this case, the posterior results for α and β are attracted by the most precise dates.

Our result shows that the prior on event dates θ_j given by formulas (27) and (28) does not give an ideal solution for the specification of the prior density on dates θ_j. It allows the correction of the date spreading bias when the uncertainties on the dates are of the same amplitude, but it can generate some undesirable concentration effect when these uncertainties are different from each other. This behavior is illustrated in the following example.

Example 1. *Synthetic data in NSBC phase*

We consider a phase with 5 non ordered Gaussian dates: 3 dates 0 ± 30 *AD, 1 date* -500 ± 200 *AD and 1 date* $+500 \pm 200$ *AD. We can see in Figure 6 [left] that the posterior densities for* α *and* β *calculated with OxCal program are very close to the precise dates* 0 ± 30*, a result predicted by formula* (27) *. It is important to note that, in this case, the agreement index A given by OxCal are low and produce a warning to prevent from an over-interpretation of the result. Figure 3 [right] shows the result obtained with ChronoModel program when the five dates are nested in the event date model. We observe that the NSBC result tends to behave like the event date model in the sense that it favors the contemporaneity of the dates.*

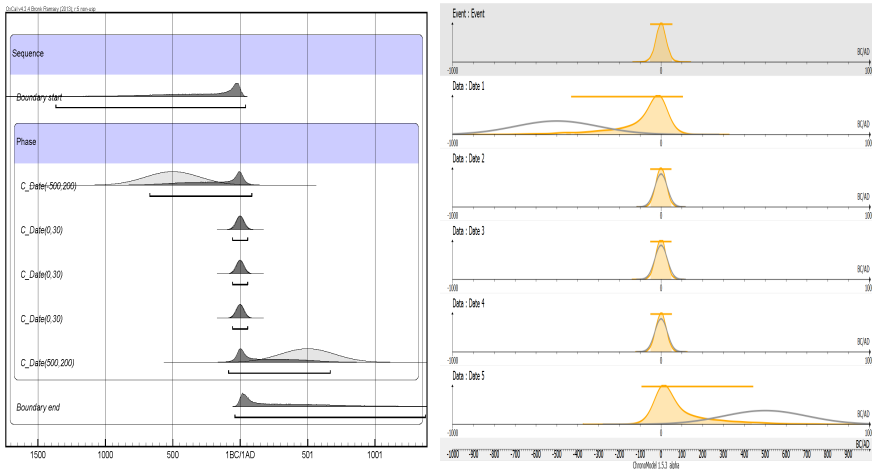

FIGURE 6. Posterior distribution of parameters in NSBC phase model [left] and in event date model [right].

In conclusion, the specification of an appropriate prior on target event dates still remains an open question. We showed that the NSBC phase model favors, in an underlying way, the concentration of the event dates in a phase. We also showed that the uniform prior favors spreading dates, especially when temporal order constraints act.

5. SIMULATIONS

In this section, Monte Carlo experiments are done to illustrate the performance of the proposed model on simulated samples. All the simulations are done using the R package ArchaeoChron (see Philippe and Vibet (2017c)).

5.1. **The target event model.** To illustrate the properties of the event date modeling, especially its robustness, we simulate sample of measurements M_i with outliers using a mixture with two Gaussian components:

$$(1-q)\mathcal{N}(0,1) + q\mathcal{N}(\mu,1) \tag{33}$$

The date of the target event is 0, and q represents the proportion of outliers. We compare the event target model with

- the r-combine model. See Remark 1.
- the t-outlier model described in Section 2.2

We assume that all the measurements have the same calibation curve, this condition is required for the r-combine model. We take $g(x) = x$ and $\sigma_g = 0$. The time range (study period) T is set equal to $[-20\ ,\ 20]$, Figure 7 and 8 represent the boxplot of the Bayes estimates evaluated on 500 independent replications. This confirms that the r-combine model is not robust to outliers. For $q = 0$ (no outlier) the three models gives the same results, we do not observe a significant difference in terms of accuracy. The target event model ant the t-outlier model behave similarly, and provide robust statistical methods. Indeed for $q < 5\%$ the presence of outliers has practically no influence on the target date result in the sense that the boxplots stay centered around the true value

The lengths of the boxplots also indicate that the loss of accuracy if less important for event model than the t-outlier model. To understand the robustness of the target event model, we represent in Figure 8 the boxplot of the Bayes estimate of the individual variance σ_i^2. We compare the values of this parameter when the date t_i of dated event is an outlier or not. Recall that the date t_i is an outlier when it is not contemporaneous of the date θ. Figure 8 shows that the individual standard deviation takes large values for the outliers. This penalizes the contribution of the observation on the estimation of the date, leading to robust procedure. Moreover regardless to the proportion of outliers, the behavior is the same.

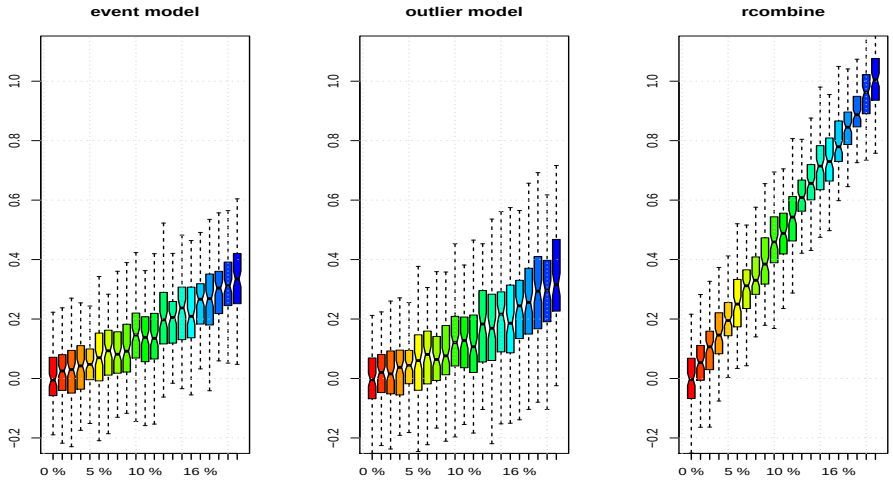

FIGURE 7. Evolution of the Bayes estimate of θ (the date of the target event (5)) as a function of the proportion q of outlier for three Bayesian models. The number of replication is 500. The data are simulated from with $\mu = 5$ and the sample size is 100. The hyperparameter in (12) are fixed $\sigma_\delta = 10$ and $p = 5\%$.

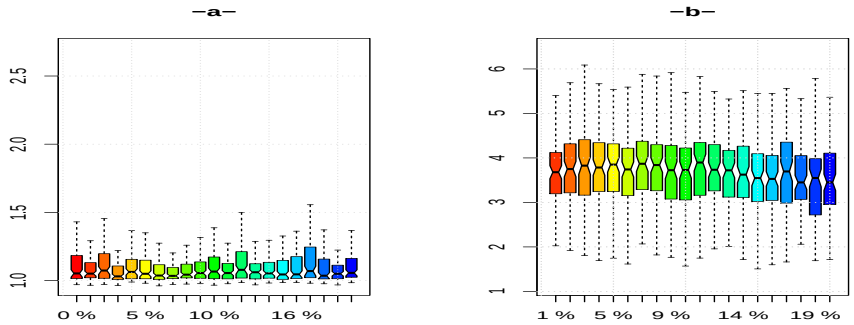

FIGURE 8. Evolution of the Bayes estimator of σ_i in (4). Comparaison between a non- outlier [left] and an outlier [right] as function of the proportion of outlier q. The parameters are the same as in Figure 7

5.2. **Chronological model.** In this part, we illustrate the behavior of our chronological model in presence of stratigraphic inversion. We want to date target events, which satistfy stratigraphy constraints $\theta_1 < \theta_2 < \cdots < \theta_{10}$. To compare our model with the t-outlier model, we construct target event containing only one dated event. We denote $t_1, t_2, \ldots, t_{10}$ the dates of dated event. To evaluate the robustness of event model, we assume that the date t_5 creates a stratigraphic inversion. In other words we take $\theta_i = t_i$ for all $i \neq 5$ and $t_5 > \theta_5$

We simulate sample of 10 measurements such that the true dates satisfy the condition $t_1 < t_2 < \cdots < t_{10}$. Then, we shift the observation associated with the date t_5 to create the stratigraphic inversion in the observations. The prior information on the dates of target events is $\theta_1 < \cdots < \theta_n$ while this constraints is imposed on the dates $t_1 < t_2 < \cdots < t_{10}$ for the t-outlier model. The time range (study period) T is set equal to $[-10\ ,\ 100]$. Figure 9 and 10 compares our model with the t-outlier model.

Figure 9 gives the results for the target event model. The value of t_5 clearly appears as an outlier, as shown by the high values of the standard deviation σ_5 (see Fig 9). The estimations of t_5 are centered around the true value since the stratigraphic constraints is not imposed to the parameters t_i. Due to the stratigraphic constraint, the Bayes estimate of θ_5 is far from the true values of t_5. But this outlier does not disturb the estimation of the rest of the sequence. For all the other dates the stratigraphic constraints on θ_i bring prior information and make more precise the target dates θ_i than the dates t_i.

Figure 10 gives the results of the t-type outlier model is applied to the same datasets for different values of the parameter σ_δ. When $\sigma_\delta = 0$, the outlier are not taken into account in the model. For a good choice of the parameter σ_δ (here 10), the performance of the t-outlier model is quite similar to those of the target event model. But Figure 10 shows that the t-type outlier model is very sensitive to the choice of its hyperparameter. The lack of an adaptative choice for this parameter is the main drawback of this approach.

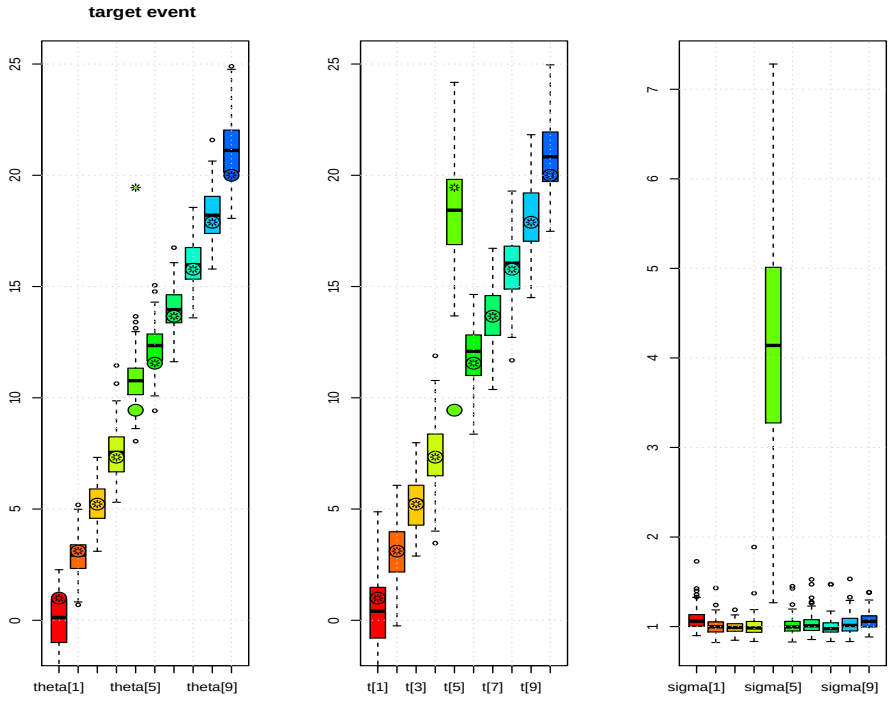

FIGURE 9. Boxplot of the Bayes estimates of parameters $(\theta_i, t_i, \sigma_i)_{i=1,\dots,10}$ in the Target event model (3). The circles (resp. stars) indicate the true values of the dates θ_i. (resp t_i) The number of replication is equal to 500.

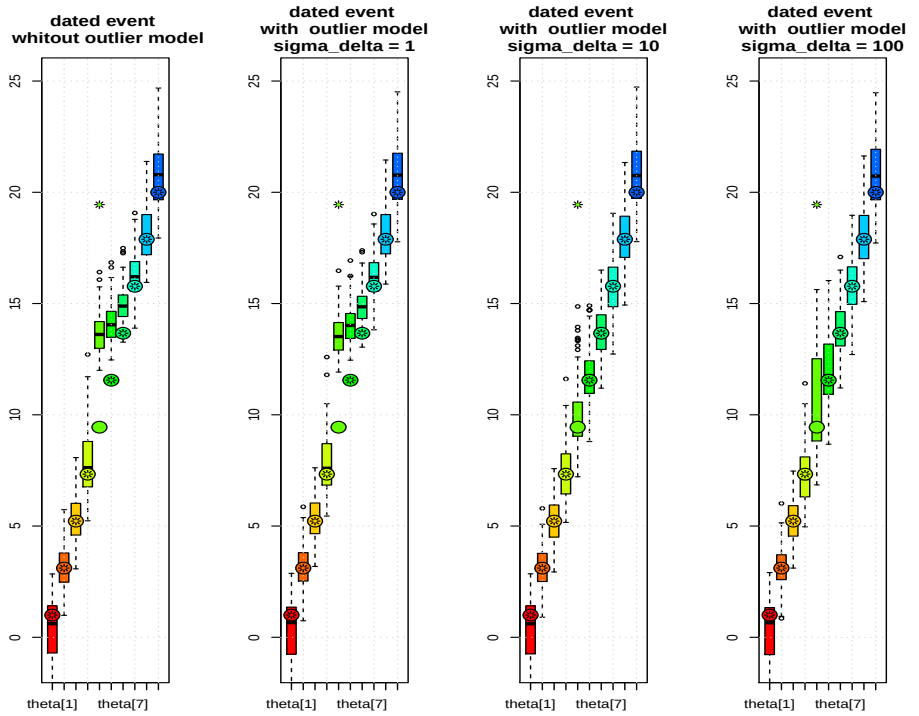

FIGURE 10. Boxplot of the Bayes estimates of t_i for the t-outlier model with different values of the hyperparameter in (12) : σ_δ varies and $p = 5\%$. The simulated sample are the same as in Fig. 9 The circles (resp. stars) indicate the true values of the dates θ_i. (resp t_i)

6. Applications

The model is implemented in the cross-platform application *Chronomodel* which is free and open source software (Lanos et al., 2016). Features of the application are described in details in (Vibet et al., 2016) : graphical user interface for importing data and modeling construction, MCMC options and controls, graphical and numerical results. Different graphical tools are implemented for assessing the convergence of the MCMC: the history plot, the autocorrelation function, the acceptance rate of Metropolis-Hastings algorithms (see Appendix, for details) The user can adjust the length of burn-in, the maximum number of iterations for adaptation and for acquisition, and the thinning rate (Vibet et al., 2016).

6.1. **Lezoux (Auvergne, France): last firing of a potter's kiln.** The aim is to date the last firing of a medieval potter's kiln recovered at the Maison-de-Retraite-Publique site (Mennessier-Jouannet et al., 1995), in Lezoux (Auvergne, France). The target event corresponds to the last firing of the kiln, and this corresponds to the last use of the kiln (i.e. behavioral event, according to Dean (1978)). Some events relevant to this target event are dated by three chronometric techniques: archaeomagnetism (AM), thermoluminescence (TL) and radiocarbon (14C). AM

and TL dates are determined from baked clay and 14C date from charcoals of trees assumed to be felled at the same time as the last firing.

The observations are composed of 3 TL dates (CLER 202a, 202b, 203), 2 AM dates (inclination and declination) and 1 radiocarbon date (Ly-5212) As the kiln belongs to the historical period, the prior time range T is set equal to $[-1000\ ,\ 2000]$. Posterior densities t_i (in color) obtained are greatly shrunken compared to individual calibrated densities (in black line), especially for archaeomagnetic and TL dates. The target date gives a 95% HPD interval equal to [574, 885] AD. Some of the posterior densities for standard deviations σ_i (Fig. 12) are spread. This comes from the multimodality of AM calibrated dates obtained with inclination (Inc) and declination (Dec). More generally, the parameter σ_i takes larger values when the associated date t_i is a possible outlier (see also examples in Lanos and Philippe, 2017).

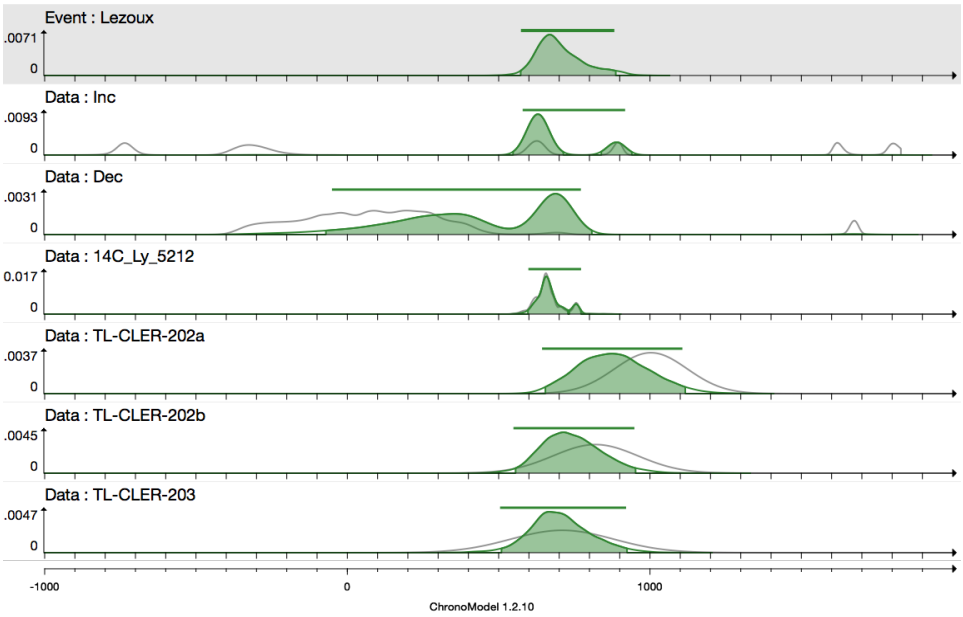

FIGURE 11. Lezoux . [white background] Posterior densities of t_i and individual posterior calibrated densities (black line) obtained for TL dates, for ^{14}C dates and for AM dates. [gray background] Posterior density for event θ. The bar above the density represents the shorter 95% posterior probability interval (credibility interval). The vertical lines, delimiting the colored area under the density curve, indicate the endpoints of the 95% highest posterior density (HPD) region.

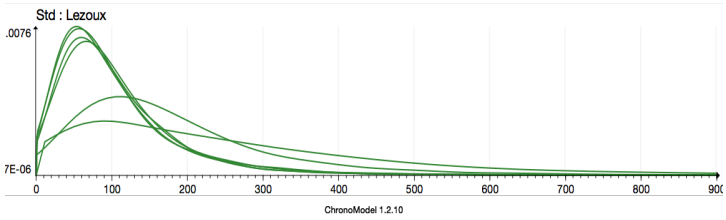

FIGURE 12. Lezoux. Posterior densities obtained for standard deviations σ_i.

,

6.2. **Duration constraint: household cluster from Malpaís Prieto site (Michoacán State, Mexico).** In some favorable archaeological contexts, it is possible to get information about the duration of a phase. It is the case here with an household cluster excavated on the Malpaís Prieto site (Michoacán State, Mexico), in the framework of the archaeological project Uacusecha (Pereira et al., 2016). The archaeological artifacts are typical of a chrono-ceramic phase which is dated to the period 1200 - 1450 AD, so the prior time range (study period) T can be set equal to this interval. Five radiocarbon ages have been obtained from burials (bone samples) and a midden (charcoal samples). Looking at the individual calibrated radiocarbon dates, which appear to be consistent between them (no outliers), the overall date range for the occupation is between 1276 and 1443 AD (at 95% confidence level). Each radiocarbon date corresponds to one target event and these events allow to estimate the beginning and the end of the household cluster phase to which they belong to (Fig. 13, left). This modeling with only one date per target event and without any stratigraphic or duration constraints gives large 95% HPD intervals for the beginning and end estimates:

PHASE	
Begin	$[1209; 1355]$
End	$[1372; 1450]$
Duration	$[48; 211]$

TABLE 1. 95% HPD region for estimating the phase (begin/end/duration)

Thus, this simple modeling does not allow to significantly improve the prior archaeological information. To do this, various archaeological evidences can better constrain the occupation duration: stratigraphic evidences, accumulation processes of the occupation remains, durability of the partly perishable architecture. In this example, these information indicate that the occupation has hardly exceeded one century, the most plausible estimation being between 60 and 90 years. As a trial, we consider that the phase duration cannot exceed $\tau_0 = 60$ years. Applying this constraint during the MCMC sampling according to (22), we obtain the following results (Fig. 13, right) :

PHASE ($\tau_0 = 60$)	
Begin	$[1286; 1397]$
End	$[1335; 1443]$
Duration	$[28, 60]$

TABLE 2. 95% HPD regions for estimating the phase (begin/end/duration)

Posterior densities of the target event dates θ_j in Figure 13 [right] are greatly shrunken compared with the unconstrained modeling. Moreover, this modeling tends to favor the recent part of the study period, by mitigating the bimodal shape of the individual calibrated dates.

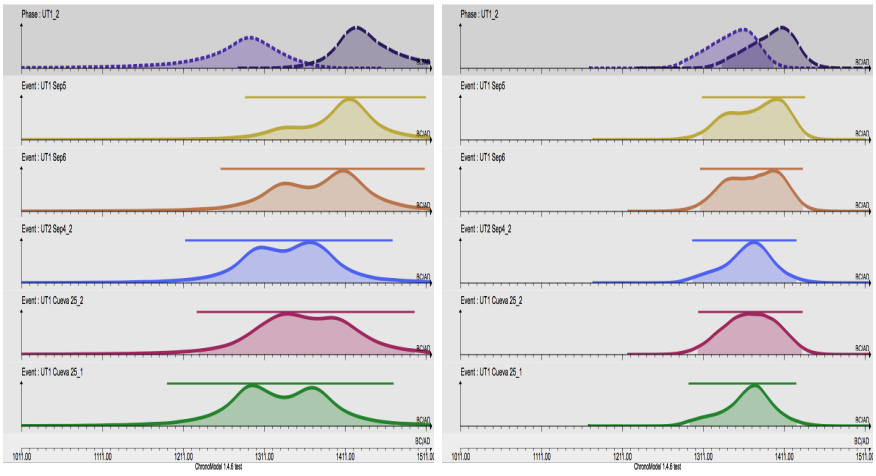

FIGURE 13. Malpaís Prieto, without informative constraint [left] and with constraint of maximal duration $\tau_0 = 60$ [left]. Each picture provides the posterior densities for event dates θ_j [gray background] and the posterior densities for beginning and end of the phase [dark gray background]. The bar above the density represents the shorter 95% posterior probability interval (credibility interval).

7. CONCLUSION

The Bayesian event date model aims to estimate the date θ of a target event from the combination of individual dates t_i coming from relevant dated events. This model has a hierarchical structure which makes it possible to distinguish between target event date θ (the date of interest for the archaeologist) and dates t_i of events (artifacts) dated by chronometric methods, typo-chronology or historical documents. One assumes that these artifacts are all contemporaneous, that is relevant to the date of the target event. The dates can be affected by irreducible errors, hence the possible presence of outliers. To take into account these errors, the discrepancy between dates t_i and target date θ is modeled by an individual variance σ_i^2, which allows the model to be robust to outliers in the sense that individual

variances act as outlier penalization. The posterior distribution of the variance σ_i^2 indicates if an observation is an outlier or not. Thanks to this modeling, it is not necessary to discard outliers because the corresponding high values σ_i^2 will automatically penalize their contributions to the event date estimation. Moreover, this model does not require additional exogenous or hyperparameters. The only parameter involved in prior shrinkage, s_0^2 , comes uniquely from the data analysis via the individual calibration process. So, the approach is adapted to very different datasets. The good robustness properties of the event date model are paid with less precision in the dates. However, this loss of precision is compensated by better reliability of the chronology. The event model constitutes the basic element in our chronological modeling approach. Dating data are nested within target event dates (with or without stratigraphic constraints between them) which in turn may be nested into phases (with or without succession constraints between them). Succession constraint, maximal duration and/or minimal hiatus can be put on the event dates in the phases.

8. Acknowledgments

The authors expressed their gratitude to Grégory Pereira (CNRS, UMR ArchAm, University of Paris 1 Panthéon-Sorbonne) for the exchange of information about the example of "Malpaís Prieto site in Mexico".
We thank Marie Anne Vibet for her very helpful discussions during the preparation of this paper.

References

Bayliss, A. (2009). Rolling out revolution: Using radiocarbon dating in archaeology. *Radiocarbon*, 51:123–147.

Bayliss, A. (2015). Quality in bayesian chronological models in archaeology. *World Archaeology*, 47(4):677–700.

Bronk Ramsey, C. (1995). Radiocarbon calibration and analysis of stratigraphy : the OxCal program. *Radiocarbon*, 37(2):425–430.

Bronk Ramsey, C. (1998). Probability and dating. *Radiocarbon*, 40:461–474.

Bronk Ramsey, C. (2001). Development of the radiocarbon calibration program OxCal. *Radiocarbon*, 43:355–363.

Bronk Ramsey, C. (2008). Deposition models for chronological records. *Quaternary Science Reviews*, 27:42–60.

Bronk Ramsey, C. (2009a). Bayesian analysis of radiocarbon dates. *Radiocarbon*, 51(1):337–360.

Bronk Ramsey, C. (2009b). Dealing with outliers and offsets in radiocarbon dating. *Radiocarbon*, 51(3):1023–1045.

Bronk Ramsey, C., Dee, M., Nakagawa, T., and Staff, R. (2010). Developments in the calibration and modelling of radiocarbon dates. *Radiocarbon*, 52:953–961.

Bronk Ramsey, C. and Lee, S. (2013). Recent and planned developments of the program OxCal. *Radiocarbon*, 55:720–730.

Bronk Ramsey, C., van der Plicht, J., and Weninger, B. (2001). Wiggle matching radiocarbon dates. *Radiocarbon*, 43:381–389.

Buck, C., Christen, J., and James, G. (1999). BCal : an on-line Bayesian radiocarbon calibration tool. *Internet Archaeology*, 7.

Buck, C., Kenworthy, J., Litton, C., and Smith, A. (1991). Combining archaeological and radiocarbon information : a Bayesian approach to calibration. *Antiquity*, 65:808–821.

Buck, C., Litton, C., and Shennan, S. (1994). A case study in combining radiocarbon and archaeological information : the early bronze age of st-veit-klinglberg, land salzburg, austria. *Germania*, 72(2):427–447.

Buck, C., Litton, C., and Smith, A. (1992). Calibration of radiocarbon results pertaining to related archaeological events. *Journal of archaeological Science*, 19:497–512.

Buck, C. E., Higham, T. F. G., and Lowe, D. J. (2003). Bayesian tools for tephrochronology. *The Holocene*, 13(5):639–647.

Buck, C. E., Litton, C. D., and Cavanagh, W. G. (1996). *The Bayesian Approach to Interpreting Archaeological Data*. Chichester, J.Wiley and Son, England.

Christen, J. and Pérez, S. (2009). A new robust statistical model for radiocarbon data. *Radiocarbon*, 51(3):1047–1059.

Christen, J. A. (1994). Summarizing a set of radiocarbon determinations: a robust approach. *Applied Statistics*, 43(3):489–503.

Combs, B. and Philippe, A. (2017). Bayesian analysis of individual and systematic multiplicative errors for estimating ages with stratigraphic constraints in optically stimulated luminescence dating. *Quaternary Geochronology*, 39(Supplement C):24 – 34.

Dean, J. S. (1978). Independent dating in archaeological analysis. *Advances in Archaeological Method and Theory*, 1:223–255.

Desachy, B. (2005). Du temps ordonné au temps quantifié : application d'outils mathématiques au modele d'analyse stratigraphique d'edward harris. *Bulletin de la Société Préhistorique Franaise*, 102(4):729–740.

Desachy, B. (2008). *De la formalisation du traitement des données stratigraphiques en archéologie de terrain*. Thèse de doctorat de l'université de Paris 1, Paris, France.

Dye, T. and Buck, C. (2015). Archaeological sequence diagrams and bayesian chronological models. *Journal of Archaeological Science*, 1:1–19.

Guérin, G., Antoine, P., Schmidt, E., Goval, E., D., H., Jamet, J., Reyss, J.-L., Shao, Q., Philippe, A., Vibet, M.-A., and Bahain, J.-J. (2017). Chronology of the upper pleistocene loess sequence of havrincourt (france) and associated palaeolithic occupations: A bayesian approach from pedostratigraphy, osl, radiocarbon, tl and esr/u-series data. *Quaternary Geochronology*, 42:15 – 30.

Harris, E. (1989). *Principles of archaeological stratigraphy*. Interdisciplinary Statisticss, XIV, 2nd edition. second ed. Academic Press, London.

Lanos, P. and Philippe, A. (2017). Hierarchical Bayesian modeling for combining dates in archaeological context. *Journal de la Socit Franaise de Statistique*, 158:72–88.

Lanos, P., Philippe, A., Lanos, H., and Dufresne, P. (2016). *Chronomodel : Chronological Modelling of Archaeological Data using Bayesian Statistics. (Version 1.5) http://www.chronomodel.fr*.

Mennessier-Jouannet, C., Bucur, I., Evin, J., Lanos, P., and Miallier, D. (1995). Convergence de la typologie de céramiques et de trois méthodes chronométriques pour la datation d'un four de potier à lezoux (puy-de-dôme). *Revue d'Archéométrie*, 19:37–47.

Naylor, J. C. and Smith, A. F. M. (1988). An archaeological inference problem. *Journal of the American Statistical Association*, 83(403):588–595.

Nicholls, G. and Jones, M. (2002). New radiocarbon calibration software. *Radiocarbon*, 44:663–674.

Niu, M., Heaton, T., Blackwell, P., and Buck, C. (2013). The Bayesian approach to radiocarbon calibration curve estimation: the intcal13, marine 13, and shcal13 methodologies. *Radiocarbon*, 55(4):1905–1922.

Pereira, G., Forest, M., Jadot, E., and Darras, V. (2016). Ephemeral cities? the longevity of the postclassic tarascan urban sites of zacapu malpas and its consequences on the migration process. In Arnauld, M., Beekmann, C., and Pereira, G., editors, *Ancient Mesoamerican Cities: Populations on the move*, page in press. University Press of Colorado, Denver, USA.

Philippe, A. and Vibet, M.-A. (2017a). Analysis of Archaeological Phases using the CRAN Package ArchaeoPhases. preprint hal-01347895.

Philippe, A. and Vibet, M.-A. (2017b). *ArchaeoPhases: Post-Processing of the Markov Chain Simulated by 'ChronoModel', 'Oxcal' or 'BCal'*. R package version 1.3.

Philippe, A. and Vibet, M.-A. (2017c). *Bayesian Modeling of Archaeological Chronologies*. R package version 0.1.

Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). Coda: Convergence diagnosis and output analysis for mcmc. *R News*, 6(1):7–11.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Vibet, M.-A., Philippe, A., Lanos, P., and Dufresne, P. (2016). Chronomodel v1.5 user's manual. www.chronomodel.fr.

MCMC computation

.1. **Algorithms.** The posterior distributions of the parameter of interest θ and of other related parameters t_i and σ_i can not be obtained explicitly. It is necessary to implement a computational method to approximate the posterior distributions, their quantiles, the Bayes estimates and the highest posterior density (HPD) regions. We adopt a MCMC (Markov Chain Monte Carlo) algorithm known as the Metropolis-within-Gibbs strategy because the full conditionals cannot be simulated by standard random generators. For each parameter, the full conditional distribution is proportional to (3). Details on the algorithms used are given in (Lanos and Philippe, 2017).

Here we give a more detailed insight on the way to estimate the date t_i which is defined on the set $\mathbb{R}$ due to the random effect model chosen in (2). The full conditional distribution of t_i $(\forall i = 1, ..., n)$ is given by

$$p(t_i|\bullet) \propto \frac{1}{S_i(t_i)} \exp\left\{\frac{-1}{2S_i^2(t_i)}(M_i - g_i(t_i))^2\right\} \exp\left\{\frac{-1}{2\sigma_i^2}(t_i - \theta)^2)\right\} \tag{34}$$

where $S_i^2(t_i) = s_i^2 + \sigma_{g_i}^2(t_i)$. Symbol $\bullet$ represents the observations and all other parameters according to equation (3).

To estimate the posterior density of t_i, we can choose between three MCMC algorithms. In each case, the support of the proposal density includes the support $\mathbb{R}$ of the target posterior density.

- MH-1: A Metropolis-Hastings algorithm where the proposal is the prior distribution of the parameter. This method is recommended when no calibration is needed, namely for TL/OSL, Gaussian measurements or typochronological references.
- MH-2: A Metropolis-Hastings algorithm where the proposal is an adaptive Gaussian random walk. This method is adapted when the density to be approximated is unimodal. The variance of this proposal density is adapted during the process.
- MH-3: Metropolis-Hastings algorithm where the proposal mimics the individual calibration density. This method is adapted for multimodal densities, such as calibrated measurements.

We are frequently confronted with multimodal target distributions as, for instance, in archaeomagnetic dating (see example 6.1). In this case, algorithms MH1 and MH-2 are not well adapted to ensure a good mixing of the Markov chain. Consequently, an alternative is to choose the MH-3 algorithm with a proposal distribution that mimics the individual calibration density defined in (7). In order to ensure the convergence of MCMC algorithm, the support of the proposal must be $\mathbb{R}$. Thus we can consider a mixture having a gaussian component. This component ensures that the whole support is visited. So we take as proposal :

$$\lambda\mathcal{C} + (1-\lambda)\mathcal{N}(0, \frac{1}{4}(T_b - T_a)^2) \tag{35}$$

where λ is a number fixed close to 1, and the distribution $\mathcal{C}$ approximates the individual calibration density. We can choose for instance the empirical measure calculated on the simulated sample from (7) or a mixture of uniform distribution:

$$\sum_{i=1}^{M} \frac{1}{M} \text{ Uniform}([\tilde{t}_i, \tilde{t}_{i+1}])$$

where $(\tilde{t}_i)_{i=1,\dots,M}$ are the ordered values of the simulated sample.

An alternative is to choose the distribution with density

$$\frac{1}{C}\sum_{i=1}^{p} p_i(\tau_i)\mathbb{I}_{[\tau_i,\tau_{i+1}]}, \qquad \text{where } C = \sum_{i=1}^{p} f(\tau_i)(\tau_{i+1} - \tau_i)$$

where $(\tau_i)_i$ is a deterministic grid of T and p_i is the density of individual calibration density (7).

.2. **Assessing Convergence.** Using the packages `ArchaeoPhases` and `coda` implemented in `R` (see Philippe and Vibet (2017b); Plummer et al. (2006) and R Core Team (2017)), we check the convergence of MCMC samples simulated for different values of the parameters λ in (35).

Figure 14 and 15 provide graphical tools for the diagnostic of MCMC sampler. We consider the example of Lezoux and only give the result for the parameter of interest θ and the date t_{dec} associated with the declinaison measurement in AM dating. Note that the behavior is the same for all the parameters of the model.

Whithin the Gibbs sampler, the t_{dec} is simulated using MH-3 algorithm, and so the proposal distribution depends on the choice of parameter λ in 35. For three values of λ, Figure 14 and 15 show a stationary behavior with good mixing properties. The autocorrelations even at lag 1 are small enough, to provide a good approximation of the posterior distribution and its characteristics (mean/ variance

/ quantiles etc) . Morever it is important to note that the results are not sensitive to the value of λ.

The Gelman diagnostic (evaluated from 5 parallel chains) is equal two 1, and so it confirms the convergence of MCMC samplers.

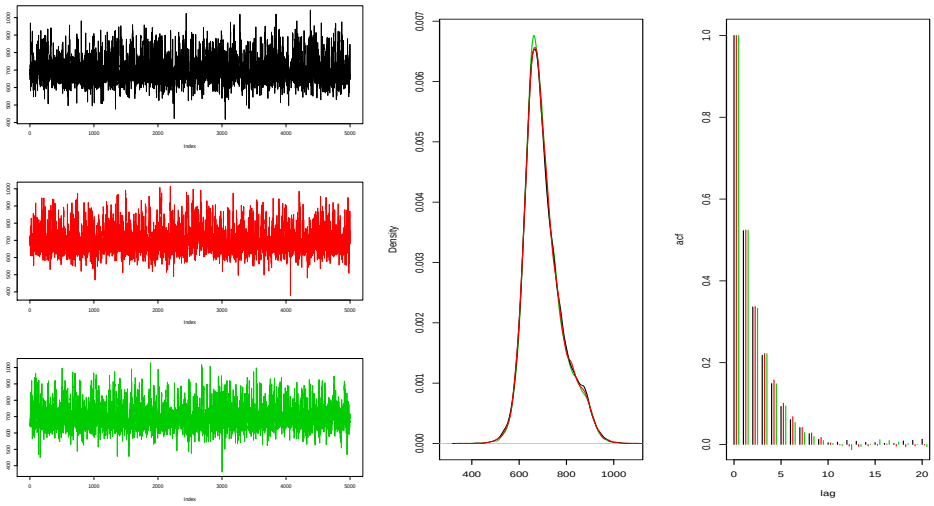

FIGURE 14. Lezoux. Diagnostic for MCMC output from posterior distribution of the date of target event. History plots, autocorrelation and marginal density are represented for different values of $\lambda = .99,\ .9,\ .7$.

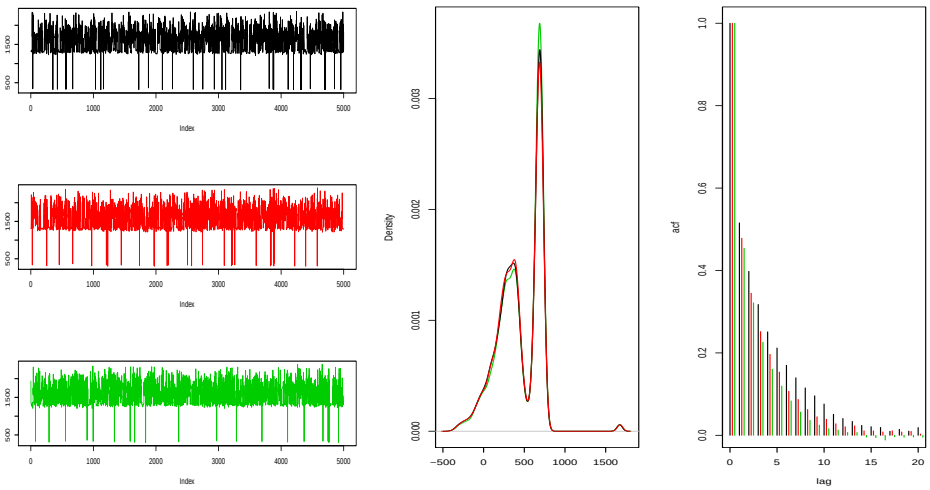

FIGURE 15. Lezoux. Diagnostic for MCMC output from posterior distribution of AM-dec date. History plots, autocorrelation and marginal density are represented for different values of $\lambda = .99, .9, .7$