ChronoModel version 1.5
User's manual

Version : 3
Date : July 19, 2016
Copyright : CNRS, Université de Nantes

Director of project: Philippe Lanos (CNRS)
Co-director : Anne Philippe (Univ. Nantes, ANJA Inria-Rennes)

Authors, Affiliation, Email:
Marie-Anne Vibet, Univ. Nantes, ANJA Inria-Rennes & CHL & Laboratoire Jean Leray , marie-anne.vibet@univ-nantes.fr
Anne Philippe, Univ. Nantes & ANJA Inria-Rennes & Laboratoire Jean Leray, anne.philippe@univ-nantes.fr
Philippe Lanos, CNRS, philippe.lanos@univ-rennes1.fr
Philippe Dufresne, CNRS, philippe.dufresne@u-bordeaux-montaigne.fr

To cite Chronomodel in publications use: Vibet, M.-A., Philippe, A., Lanos, P., and Dufresne, P. (2016). Chronomodel v1.5 user's manual. www.chronomodel.fr.
Ph. Lanos, A. Philippe, H. Lanos, Ph Dufresne (2016). Chronomodel : Chronological Modelling of Archaeological Data using Bayesian Statistics. (Version 1.5). Available from http://www.chronomodel.fr.

A BibTeX entry for LaTeX users is

```
@Manual{chronomodel,
  title =  {ChronoModel V1.5 User's manual},
  author =  {Vibet, M-A and Philippe, A. and Lanos, Ph. and
                Dufresne, Ph },
     url    = {http://www.chronomodel.fr },
  year =  2016
}
```

```
@Manual{chronomodelSoft,
     title  = {Chronomodel  : Chronological Modelling of
              Archaeological Data using Bayesian Statistics (Version 1.5).},
     author = {Ph. Lanos and  A. Philippe and H. Lanos
               and  Ph. Dufresne },
     year   = {2016},
     url    = {http://www.chronomodel.fr },
  }
```

Contents

Chapter 1

Introduction

Chronological modelling with « ChronoModel »

- Events:
  ChronoModel is based on the concept of "Event". An Event is a point in time
  for which we can define a hierarchical Bayesian statistical model.
  An event is determined by its unknown calendar date. In ChronoModel, different types of measurement may be combined in order to estimate the date
  of the event of interest. These measurements may be ^{14}C ages in radiocarbon,
  paleodose measurements in luminescence, inclination, declination or intensity
  of the geomagnetic field in archeaomagnetism, typo-chronological references or
  Gaussian measurements.

- Phases:
  A Phase is a group of Events. It is defined on the basis of archaeological,
  geological, environmental,... criteria we want to locate in time. Unlike "Event"
  model, the Phase does not respond to a statistical model: indeed we do not know
  how Events can be a priori distributed in a phase. However, we can estimate
  the beginning, end or duration of a phase from the Events that are included in
  it. Prior information may be added : the Events of a phase may be constrained
  by the known duration or a hiatus between two phases may be inserted (this
  imposes a temporal order between the two groups of Events).

- Temporal order relationships :
  Events and/or Phases may be linked by temporal order relationships. These
  order relations are defined in different ways: by the stratigraphic relationship
  (physical relationship observed in the field) or by criteria of stylistic, technical,
  architectural *etc.* development which may be a priori known.

Chapter 2

Bayesian modelling

2.1 Event model

2.1.1 Definition of the model

In ChronoModel, an event is determined by its unknown calendar date θ. Assuming that this event can reliably be associated with one or several suitable samples, out of which measurements can be made and calendar dates can be obtained, the Event model combines these calendar dates, assumed to be contemporaneous, in order to estimate the unknown calendar date θ.

The following equation shows the stochastic relationship between θ, the calendar date associated with the event of interest, and $t_1, ...t_n$, the calendar dates obtained from the n measurements, M_i, assumed to be related to the event.

$$t_i = \theta + \sigma_i \, \epsilon_i \tag{2.1}$$

where $\epsilon_i \sim N(0, 1)$, for $i = 1$ to n. ϵ_1, ..., ϵ_n are independent and σ_1, ..., σ_n are the unknown standard deviation parameters.
Such a model means that each calendar date, t_i, assumed to be related to the event of interest and contemporaneous to the other calendar dates $t_{j,\forall j \neq i}$, can be affected by errors σ_i that can come from different sources [1, 2]. Moreover, conditionally on θ and on σ_i^2, t_i is independent of all t_j, for $j \neq i$.

ChronoModel is actually based on a bayesian hierarchical model. Such a model can easily be represented by a directed acyclic graph (DAG) that defines the dependencies between the variables of the model [3]. A DAG is formed by nodes and edges. A node

can either represent an observation (data) or a parameter, that can be stochastic or deterministic. An edge is a directed arc that represents dependencies between two nodes. The edge starts at the parent node and heads to the child node. This relationship is often a stochastic one (single arc) but it may also be a deterministic one (double arc). The DAG can be read as follow, each node of the DAG is, conditionally on all its parent nodes, independent of all other nodes except of its child nodes. The following DAG is a representation of the Event model implemented in ChronoModel.
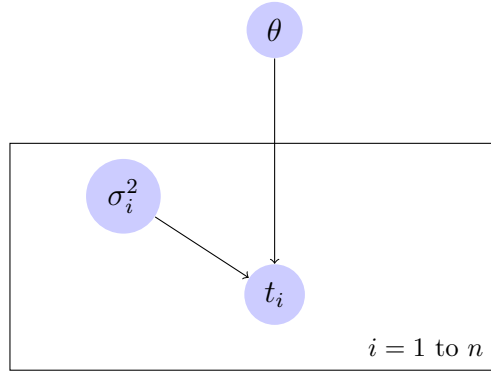
Figure 2.1 – DAG representation of the Event model. Directed edges represent stochastic relationships between two variables, blue circles represent unknown parameters. Rectangular plates are used to show repeated conditionally independent parameters.

Prior information about θ and $\sigma_1, ...\sigma_n$ are defined in section 2.1.2. Then the likelihood function is defined in section 2.1.3. According to the likelihood, three main types of data information may be implemented into ChronoModel: a single measurement with its laboratory error, a combination of multiple radiocarbon measurements or an interval referring to a typo-chronological reference. These different types are explained in section 2.1.3. Then the wiggle matching case, specific to the radiocarbon dating method, is explained in section 2.1.4. Finally, the case where several events are related by a temporal order contraint is explained.

2.1.2   Prior information about θ and $\sigma_1, ..., \sigma_n$

Without any other constraints than the beginning, τ_B, and the end, τ_E, of the study period (τ_B and τ_E are fixed parameters), the unknown calendar date θ is assumed to have a continuous uniform distribution on the study period. Hence the prior density function of θ is:

$$p(\theta) = \frac{1}{\tau_E - \tau_B} 1_{[\tau_B, \tau_E]}(\theta) \tag{2.2}$$

The variances σ_i^2, for $i = 1$ to n, are assumed to have a shrinkage uniform distribution

[4, 5]. Their prior density function is the following one:

$$p(\sigma_i^2) = \frac{s_0^2}{(s_0^2 + \sigma_i^2)^2} \tag{2.3}$$

where
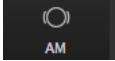
$$\frac{1}{s_0^2} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\hat{s}_i^2} \tag{2.4}$$

with $\hat{s}_i^2 = \text{Var}(t_i|M_i)$ the variance of the posterior distribution of t_i obtained after individual calibration (See section 2.1.3.2) or $\hat{s}_i^2 = \text{Var}(t_i|t_{i,A}, t_{i,B})$ the variance estimated out of a typo-chronological reference (See section 2.1.3.5).

2.1.3 Likelihood

2.1.3.1 Observations

In ChronoModel, these different types of measurement may be combined in order to estimate the date of the event of interest. The measurement may represent :

- a ^{14}C age in radiocarbon

- a paleodose measurement in luminescence

- an inclination, a declination or an intensity of the geomagnetic field in archeao-magnetism

- a typo-chronological reference (for instance, an interval of ceramic dates)

- a Gaussian measurement with known variance

If needed, these measurements, M_i, may be converted by ChronoModel into calendar dates using appropriate calibration curves.

2.1.3.2 Individual calibration

Most measurements have to be calibrated into calendar dates. This is the case of radiocarbon, archeaomanetism, paleodose or even Gaussian ages. ChronoModel may convert these measurements into calendar dates using an appropriate calibration curve. If the observation about t_i comes from only one measurement that needs to be calibrated, then the DAG presented in Figure 2.2 applies.
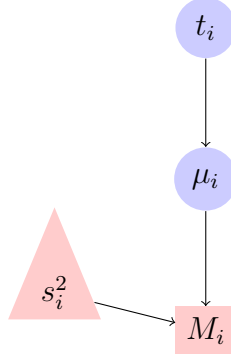
Figure 2.2 – DAG representation of an individual calibration, that is the conversion of one measurement into a calendar date.
Directed edges represent stochastic relationships between two variables, blue circles represent model parameters, pink rectangular nodes represent stochastic observed data, pink triangles represent observed and deterministic data.

Let's use the following notations: M_i is the measurement made by the laboratory and s_i^2, its variance. In ChronoModel, M_i is assumed to follow a normal distribution with mean μ_i, the true age, and with a variance s_i^2, the laboratory error. This may be expressed by the following equation :

$$M_i = \mu_i + s_i\,\epsilon_i^{Lab} \tag{2.5}$$

where $\epsilon_i^{Lab} \sim N(0,1)$.

We assume that μ_i follows a normal distribution with mean $g_i(t_i)$ and variance $\sigma_{g_i}^2(t_i)$, where g_i is the function of calibration associated with the type of measurement of M_i.

$$\mu_i = g_i(t_i) + \sigma_{g_i}(t_i)\,\epsilon_i^{Cal} \tag{2.6}$$

Hence pooling Equations 2.5 and 2.6 together,

$$M_i = g_i(t_i) + s_i\,\epsilon_i^{Lab} + \sigma_{g_i}(t_i)\,\epsilon_i^{Cal} = g_i(t_i) + S_i\,\epsilon_i^{LabCal} \tag{2.7}$$

where $\epsilon_i^{LabCal} \sim N(0,1)$ and $S_i^2 = s_i^2 + \sigma_{g_i}^2(t_i)$.

So, conditionally on t_i, M_i follows a normal distribution with mean $g_i(t_i)$ and variance S_i^2.

2.1.3.3   Individual calibration and reservoir effect

This is a special case of radiocarbon dating. If needed, a reservoir effect may be added to a measurement M_i. Hence, the DAG 2.3 applies.
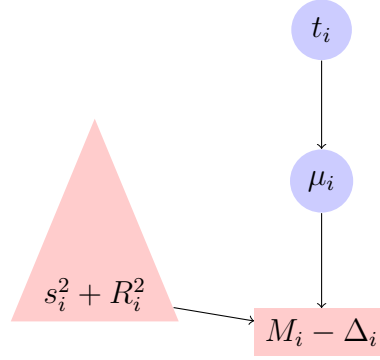
Figure 2.3 – DAG representation of an individual calibration, that is the conversion of one measurement into a calendar date, including a reservoir effect Δ_i and its error R_i.
Directed edges represent stochastic relationships between two variables, blue circles represent model parameters, pink rectangular nodes represent stochastic observed data, pink triangles represent observed and deterministic data.

2.1.3.4 Calibration from multiple measurements

Let's say we have K measurements M_k from a unique sample. For example, a sample may have been sent to K different laboratories. All these measurements refer to the same true radiocarbon age μ. In that case, the Bayesian model first gathers all information about μ before calibrating. Hence, $\forall k = 1, ..., K$,

$$M_k = \mu_i + s_k\, \epsilon_k^{Lab} \tag{2.8}$$

where $\forall k = 1, ..., K$, $\epsilon_k^{Lab} \sim N(0,1)$ and $\epsilon_1^{Lab}, ..., \epsilon_K^{Lab}$ are independent.
Now, as described by Ward and Wilson (1978) [6], a pooled mean, $\overline{M}$, and the variance of the pooled age, $\overline{s}^2$, may be calculated as follow:

$$\overline{M} = (\sum_{k=1}^{K} \frac{M_k}{s_k^2})/(\sum_{k=1}^{K} \frac{1}{s_k^2})$$
$$\overline{s}^2 = \frac{1}{\sum_{k=1}^{K} \frac{1}{s_k^2}}$$

Now, as all M_k refer to the same μ_i, we have

$$\overline{M} = \mu_i + \overline{s}^2 \sum_{k=1}^{K} \epsilon_k^{Lab}/s_k$$
$$\mu_i = g_i(t_i) + \epsilon_i^{Cal}$$

where $\epsilon_i^{Cal} \sim N(0, \sigma_g^2(t_i))$ and g_i is the function of calibration.

Hence,

$$\overline{M} = g_i(t_i) + \overline{s}^2 \sum_{k=1}^{K} \epsilon_k^{Lab}/s_k + \epsilon_i^{Cal} = g_i(t_i) + \epsilon^{LabCalMult} \tag{2.9}$$

where $\epsilon^{LabCalMult} \sim N(0, S^2)$ and $S^2 = \overline{s}^2 + \sigma_{g_i}^2(t_i)$. So, conditionally on t_i, the calibrated measurement has a normal distribution with mean $g_i(t_i)$ and variance S_i^2. Figure 2.4 represents the corresponding DAG.
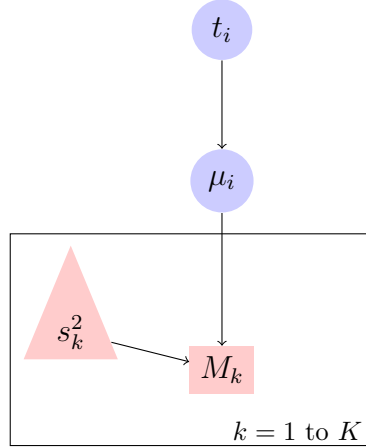
Figure 2.4 – DAG representation of a calibration from multiple measurements. Arrows represent stochastic relationships between two variables, blue circles represent model parameters, pink rectangulars represent stochastic observed data, pink triangles represent deterministic observed data. Rectangular plates are used to show repeated conditionally independent parameters.

Be aware that the pooled mean $\overline{M}$ is not robust to outliers. However, the test proposed by Ward and Wilson [6] is usually not valid as its requires to combine at least 30 measurements.

2.1.3.5   Typo-chronological reference

Let's say that a typo-chronological information is a period defined by two calendar dates $t_{i,A}$ and $t_{i,R}$, with the constraint $t_{i,A} < t_{i,R}$.

The distribution of $(t_{i,A}, t_{i,R})$ conditional on t_i is given by the following equation

$$p(t_{i,A}, t_{i,R}|t_i) = \lambda^2 e^{-\lambda(t_{i,R}-t_{i,A})} 1_{t_{i,A}<t_i<t_{i,R}} \tag{2.10}$$

where λ is a positive constant. Figure 2.5 represents the corresponding DAG.

2.1.4   The wiggle-matching case

This case is specific to radiocarbon dating and dendrochronology. Let's say that we have n radiocarbon datings referring to the unknown calendar date θ shifted by a
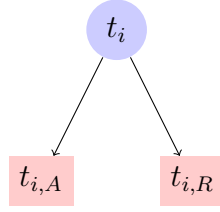
Figure 2.5 – DAG representation of a typo-chronological information. Arrows represent stochastic relationships between two variables, blue circles represent model parameters, pink rectangulars represent stochastic observed data.

known quantity called δ_i. Then, the stochastic relationship between t_i and θ is given by the following equation:
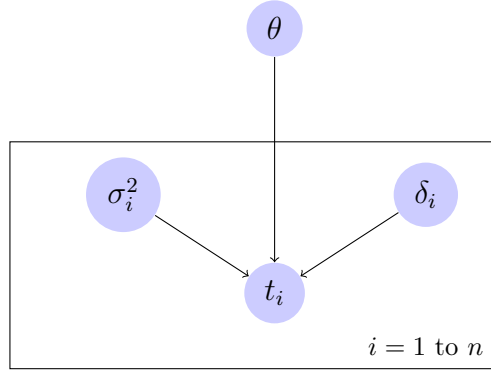
Figure 2.6 – DAG representation of the event model including a wiggle matching. Directed edges represent stochastic relationships between two variables, blue circles represent model unknown parameters. Rectangular plates are used to show repeated conditionally independent parameters.

$$t_i = \theta - \delta_i + \sigma_i \, \epsilon_i^{CM} \tag{2.11}$$

where $\epsilon_i^{CM} \sim N(0,1)$ for i=1 to n and ϵ_1^{CM}, ..., ϵ_n^{CM} are independent.

δ_i may either be a deterministic or a stochastic parameter. Then $t_i + \delta_i$ follows a normal distribution with mean θ and variance σ_i^2.

If δ_i is stochastic, then its prior distribution function may have a uniform distribution on $[d_{1i}, d_{2i}]$.

$$p(\delta_i) = \frac{1}{d_{2i} - d_{1i}} 1_{[d_{2i}, d_{1i}]}(\delta_i)$$

or a Gaussian distribution with mean $\mu_{\delta,i}$ and variance $\sigma_{\delta,i}^2$.

$$p(\delta_i) = \frac{1}{\sqrt{2\pi}\sigma_{\delta,i}} \exp\left(-\frac{1}{2}\left(\frac{\delta_i - \mu_{\delta,i}}{\sigma_{\delta,i}}\right)^2\right)$$

In that case, the associated DAG is presented in Figure 2.6.

2.1.5 Temporal order constraints

Several events may be linked by a temporal order, such as stratigraphic constraints. Let's say that three events happened successively in time. Then their true calendar dates are assumed to verify the following relationship:

$$\theta_1 < \theta_2 < \theta_3$$

Bounds, such as historical dates or so called *Terminus post quem / terminus ante quem* [1], may also be introduced in order to constrain one or several events. Let's say that the three events are assumed to happen after a special event with true calendar date θ^B. Then the following relationship holds :

$$\theta^B < \theta_1 < \theta_2 < \theta_3$$

2.2 Phases

2.2.1 Definition of a phase

A phase is a group of events defined on the basis of objective criteria such as archaeological, geological or environmental criteria. As we do not know how events are distributed in a phase, no model for the phase is computed in ChronoModel. The phase only reflects the group of events included in it.

However, using ChronoModel, we may have information about the beginning, the end and the duration of a phase.

Moreover, we may include prior information about the duration of a phase or about a hiatus between two successive phases.

2.2.2 Beginning, end and duration of a phase

The following information are given for each phase:

The **beginning of a phase**, α, reflects the most ancient event among the r events included in the phase:

$$\alpha = \min(\theta_{j,\,j=1\ldots r})$$

The **end of a phase**, β, reflects the most recent event among the r events included in the phase:

$$\beta = \max(\theta_{j,\,j=1\ldots r})$$

---

[1]Latin for "Limit after / before which"

The **duration**, τ, is the time between the beginning and the end of a phase:

$$\tau = \beta - \alpha$$

The posterior distribution of all these elements may be approximated by MCMC methods (See section 4 for more details) and statistical results such as the mean, the standard deviation and so on, may be estimated.

2.2.3 Prior information about the duration of a phase

The duration of a phase, τ, may be known from prior information. Let's say a phase lasted τ years. In ChronoModel, that means that the time elapsed between the earliest event of the phase and the latest event of the phase is at most of τ years. Hence, if there are r events in a phase, $\theta_{j,\,j=1\ldots r}$, they have to verify the following constraint :

$$\max(\theta_{j,\,j=1\ldots r}) - \min(\theta_{j,\,j=1\ldots r}) \leq \tau$$

2.2.4 Prior information about a succession of phases

Succession or stratigraphic constraints between two phases may be included. These constraints of succession, called hiatus, may be of known duration imposing a temporal order between both groups of events.

2.2.4.1 Temporal succession between phases

Let's say that a phase P, containing r_P events, is constrained to happen before a phase P^*, containing r_{P^*} events. Then, in ChronoModel, this means that all events of phase P are constrained to happen before the events included in phase P^*. The following equations should be verified by all the events included in both phases.

$$\forall i \in [(1, r_P)],\ \forall j \in [(1, r_{P^*})],\ \theta_i^P < \theta_j^{P^*}$$

or

$$\max(\theta_{i,\,i=1\ldots r_P}^P) < \min(\theta_{j,\,j=1\ldots r_{P^*}}^{P^*})$$

or

$$\min(\theta_{j,\,j=1\ldots r_{P^*}}^{P^*}) - \max(\theta_{i,\,i=1\ldots r_P}^P) > 0$$

2.2.4.2   Prior information about a hiatus

Prior information about a hiatus may be included in ChronoModel. Let say we have a hiatus of γ years between two phases : a phase P and a phase P^*. In Chronomodel, that means that the time elapsed between events of phase P are and events of phase P^* is at least of γ years. Consequently, all the events of phase P and phase P^* have to verify the following constraint :

$$\min(\theta_{j,\,j=1...r_{P^*}}^{P^*}) - \max(\theta_{i,\,i=1...r_P}^{P}) \geq \gamma > 0$$

Chapter 3

Use of ChronoModel

3.1   Installation

From the website (www.chronomodel.fr), you may download the software adapted to your computer.

- For MAC. Click on the "MAC Download" button. Then double-click on the package to install the software. Following the wizard window, the software is installed in the Applications directory. Once done, the logo of ChronoModel should be seen in the Applications directory.

- For Windows. Click on the "Windows Download" button. Then double-click on the .EXE to install the software.

- For Linux. The usual way is to install Qt on your computer and compiling yourself the application. Chronomodel is hosted on GitHub.com. You can clone the repository by typing : *clone https://github.com/Chronomodel/chronomodel.git*

3.2   Description of the icons of the main window

The design of ChronoModel is implementing with the following idea : *Information are in the middle of the interface and tools are around.*
This section gives a description of all icons contained in the top of the window.
On the left hand side of the main window, three icons help managing projects and two icons help managing modelling actions.

To create a new project.

To open a pre-existing project. ChronoModel projects are saved using the extension file *.CHR

To save the current project.

To undo the last actions.

To redo the last actions.

In the middle of the window, the icons refer to the different steps of the modelling with ChronoModel.

To design the model.

To open the MCMC settings window.

To run the bayesian modelling.

To show the results of the modelling.

To show summaries of the MCMC methods used for this modelling and its results.

Then on the right hand side of the window:

Help To active or unactive the context-sensitive tips seen in yellow bubbles. By defaults, Help is active.

Manual online To reach the user manual on the web site.

Website To open the ChronoModel website.

3.3 Creating / Opening a project - Defining the study period

After having launched ChronoModel, the following interface appears (See Figure 3.1).
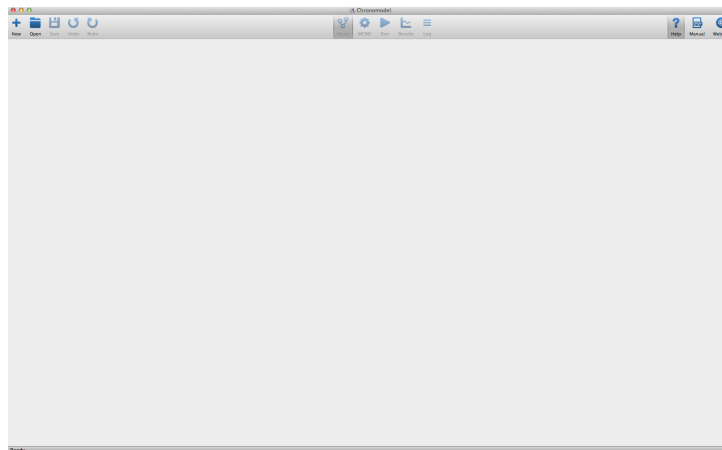
Figure 3.1 – ChronoModel interface

In order to create a new project, click on **New** on the left hand side of the window. This action opens a new window asking you to name this new project and to save it in a chosen directory.

Once done, a dialog box appears and asks to define the study period (see Figure 3.2). This is actually the first thing to do otherwise no other action will be allowed. The
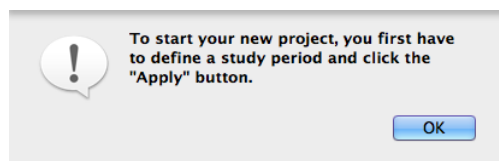
Figure 3.2 – Dialog box to define the study period

study period is the period under which you assume that the unknown calendar dates
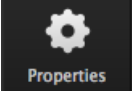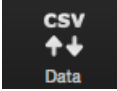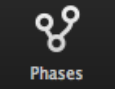
of the events are likely to be. These pieces of information may be modified by filling
the following boxes on the right hand side of the window (See Figure 3.3). The study

Figure 3.3 – Study period definition window

period should be given in the date format BC/AD [1] as the bayesian modelling is
computed in this date format. The "Apply" button will stay red as long as the study
period is undefined and become green afterwards.

For the example, let's use a study period from 0 to 2 000 (BC/AD). Now the interface
looks like Figure 3.4, the left hand side part represents the events' scene and the right
hand side part gives different types of information that will be further detailed in the

following sections. The tab gives information about events and associated

datings. The tab allows to import a .CSV file containing informations about
datings and to export all datings used in the modelling of the current project. The

 tab shows the phases' scene.
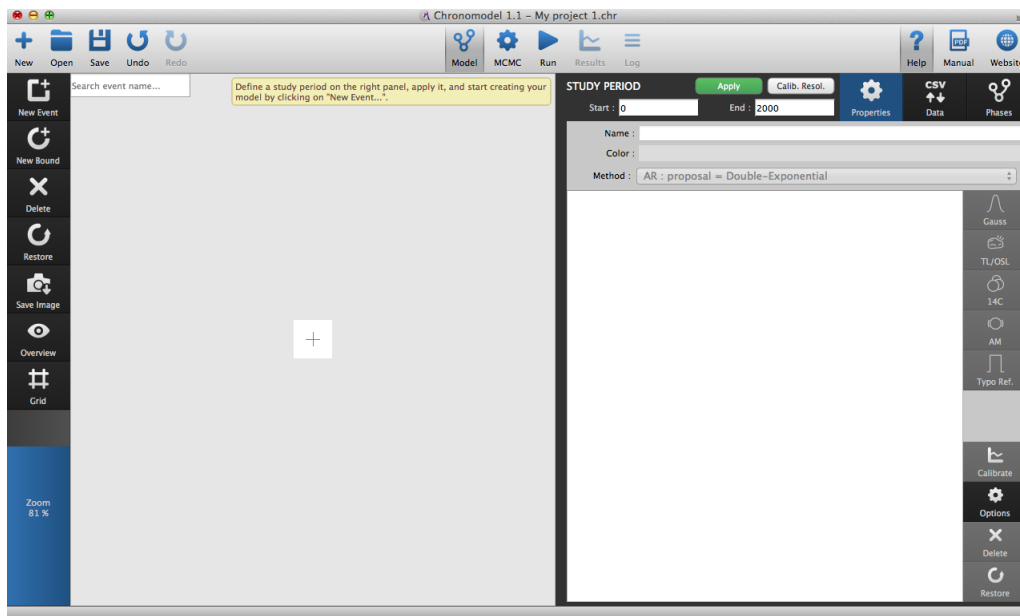
Figure 3.4 – The interface of ChronoModel when starting a new project. On the left handside of the
window, the events' scene, on the rigth handside, data information and phases structure.

---

[1]Before Christ / Anno Domini

To open a pre-existing project, use the **Open** button on the top left of the window and then choose the project. ChronoModel projects are saved with the extension .CHR.

3.4 Creating events, bounds and constraints

3.4.1 Creating a new event

To create a new event, select on the left hand side of the events' scene (see Figure 3.4). A window will be opened asking you to name that new event. For the example, let's call it "My event 1". A default color is given to this new event but you may wish to change it. Click on the color chosen by default and select a new color from the "Colors" window.
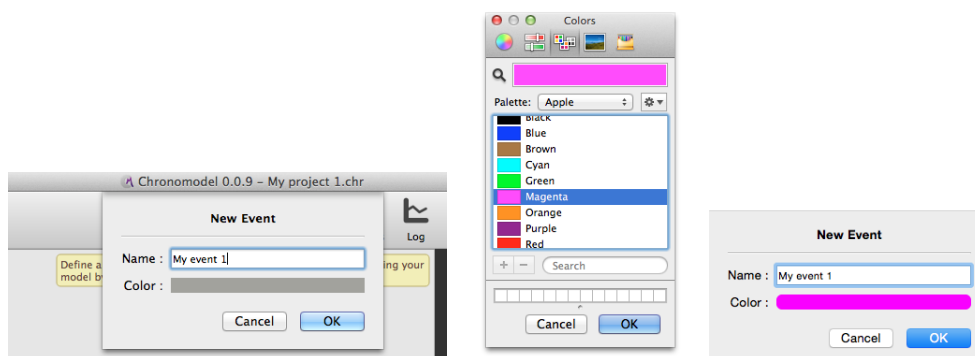
After validation, the event appears in the events' scene (See Figure 3.5). Now, when the event is selected in the event's scene, that is when the event is circled by a red line, the event's properties may be seen on the right hand side of the window, using the tab. From there, the name, the color and the MCMC method associated with this event might be changed (See Section 4 for more details on MCMC methods).
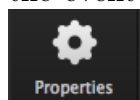
3.4.2 Including measurements

An event may be associated with data information such as measurements or typo-chronological references. There are two ways to insert data with ChronoModel:
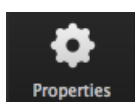
Using the tab
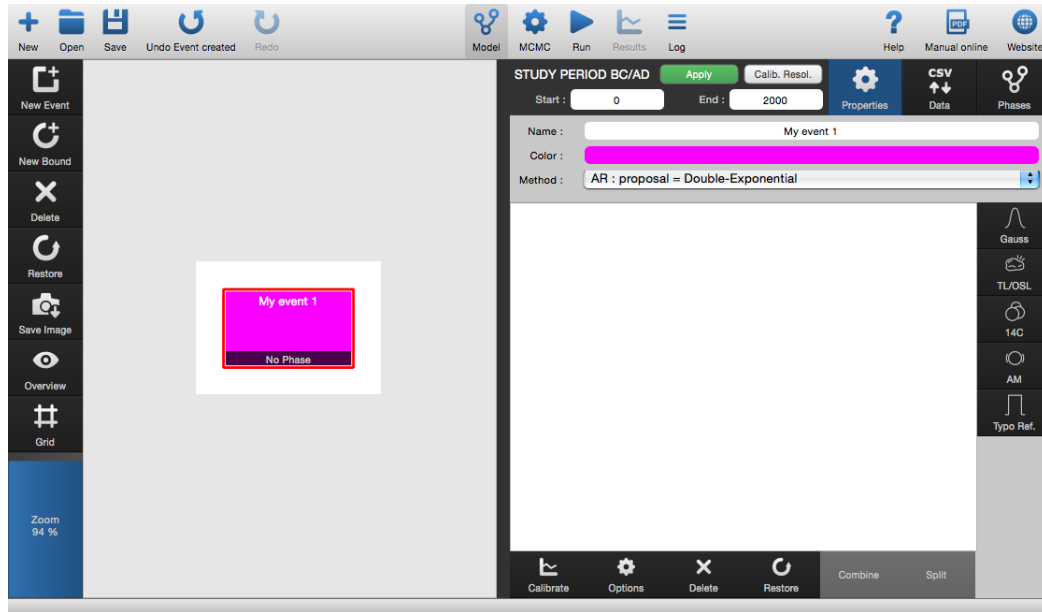
Figure 3.5 – Chronomodel window showing the new event in the events' scene

Select the event in which you want to add a measurement. An event is selected if it is surrounded by a red line. Then select a measurement type by clicking on one measurement icon on the right hand side of the Properties tab and include information measurement by measurement. Details are given for each type of measurement in the followings.
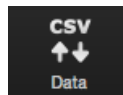
Using the **tab**

Click on "Load CSV file ..." to include information from a CSV file. However, the CSV file has to be organised according to the type of measurements included (more details are given in the followings). Then, you may drag a selected line to the corre-

Figure 3.6 – Inclusion of information dragged from a CSV file

sponding event.

Be aware that by default ChronoModel reads a CSV file using a coma as cell separator and a dot as a decimal separator. In the Application Settings presented in Figure 3.7, these options may be changed.
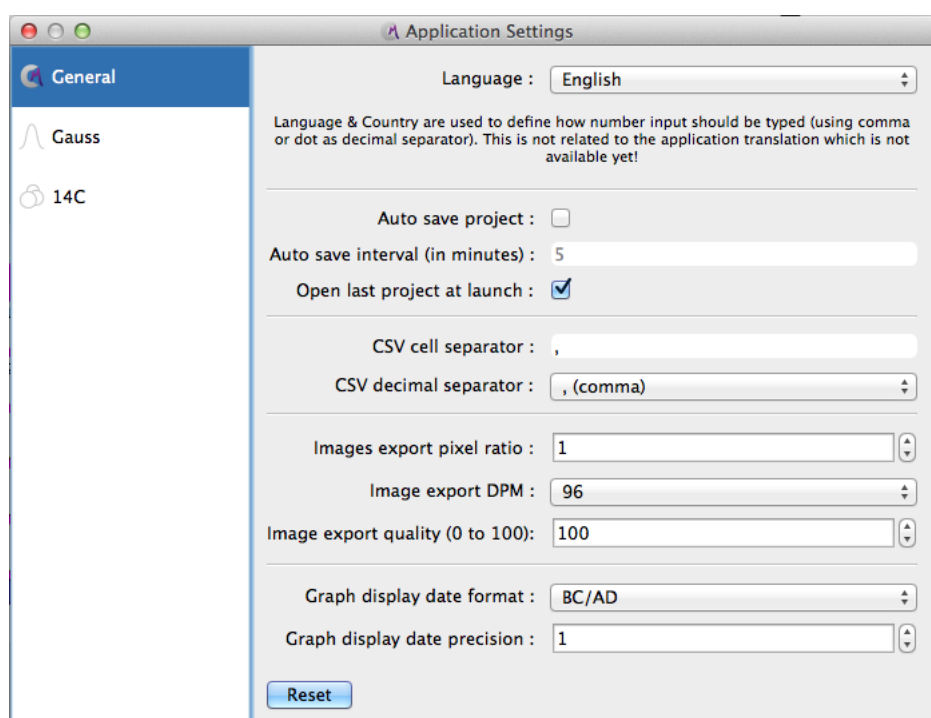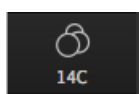
Figure 3.7 – Application settings

3.4.2.1 Radiocarbon dating (14C)
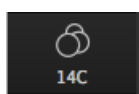
Clicking on , the radiocarbon extension window will be opened (See Figure 3.8). Within this window, you can insert the reference name of the measurement, the age value given by the laboratory ("Age"), its associated standard error ("Error"), the reservoir effect ΔR with its associated standard error $\Delta R\,Error$ (data base: http://radiocarbon.pa.qub.ac.uk/marine/).

Reference curves

In ChronoModel, several calibration curves are already implemented. These curves are listed in Table 3.1. When adding a radiocarbon dating, you may select one of these reference curves using the drop-down menu.

Other curves may be included in ChronoModel from the Application settings. Figure

Figure 3.8 – The radiocarbon measurement extension window with advanced options

3.9 shows the window from which it is possible to add a new reference curve.

Figure 3.9 – Application settings - Reference curves for C14 datations

Name	Reference
Uwsy98	Stuiver *et al.* 1998 [7]
IntCal04	Reimer *et al.* 2004 [8]
Marine04	Hughen *et al.* 2004 [9]
ShCal04	McCormac *et al.* 2004 [10]
IntCal09	Reimer *et al.* 2009 [11]
Marine09	Reimer *et al.* 2009 [11]
IntCal13	Reimer *et al.* 2013 [12]
Marine13	Reimer *et al.* 2013 [12]
ShCal13	McCormac *et al.* 2013 [13]

Table 3.1 – Reference curves for C14 datations implemented in ChronoModel

Advanced options

You may want to include a wiggle matching or you may want to change the MCMC method used to sample out of the posterior distribution of the calibrated date associated with that dating (See 4 section for more details). To do that, click on the "Advanced" menu from the same window (See Figure 3.8).
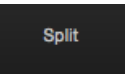
The default MCMC method used for radiocarbon datings is the Metropolos Hastings algorithm using a distribution that mimics the individual calibration density as a proposal. Two other proposals may also be selected, the prior distribution and a Gaussian adaptative random walk (See Section 4 for more details on MCMC methods).

The wiggle matching may be fixed or included within a range of values or even have a Gaussian distribution defined by its mean value and its standard error. By default, the wiggle matching is set to 0.

Once the measurement is validated, its details appear in the Properties tab as shown in Figure 3.10 and might be changed by double-clicking on its name from the list of measurements.

Combining radiocarbon datings

Several radiocarbon measurements referring to a unique sample may be "combined". These radiocarbon measurements should have the same reference curve for the calibration and the same MCMC method. To "combine" several measurements, select the associated event in the events' scene, then select the Properties tab. Now select

all measurements to be combined from the list of measurements and click on at the bottom of the window (See section 2.1.3.4 for more details on the combine

function). Now a unique measurement appears marked by . Click on
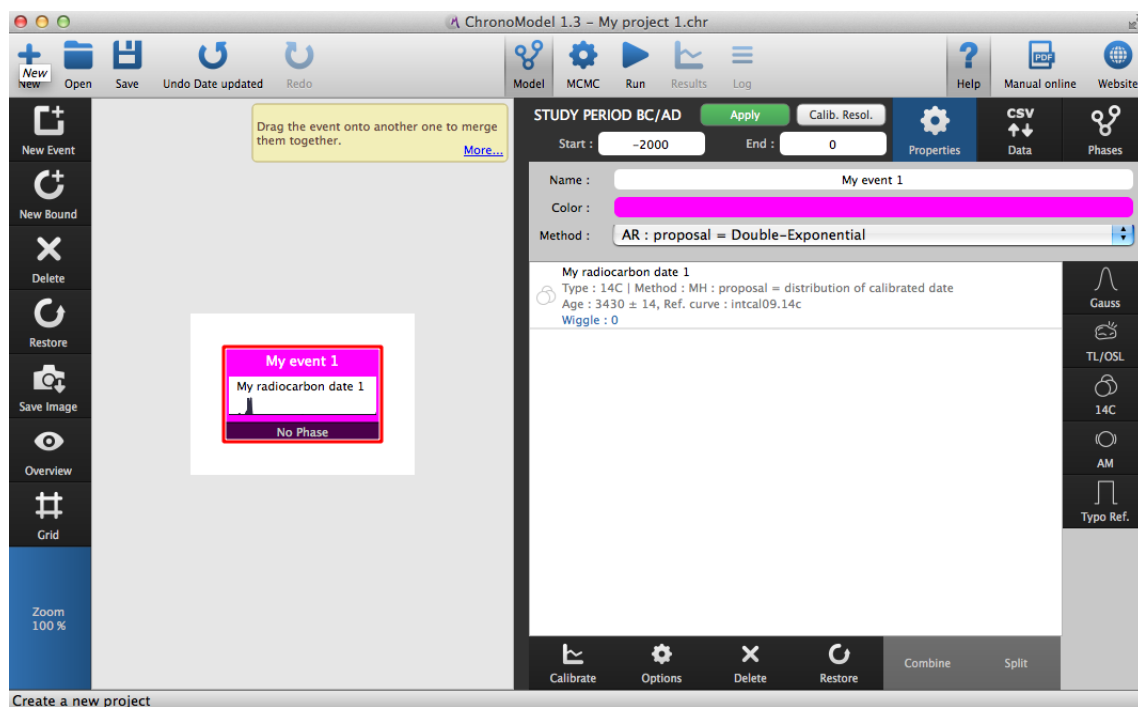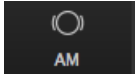
Figure 3.10 – Events' scene and Properties.

to split the measurements.

3.4.2.2 Archeomagnetism dating (AM)

Click on [AM] to open the extension window associated with archeomagnetism datings, as displayed in Figure 3.11. Within this window, you can insert the reference name of the measurement, you can choose the magnetic parameter (inclination, declination or intensity) and insert values, you can also choose the reference curve for the calibration process.
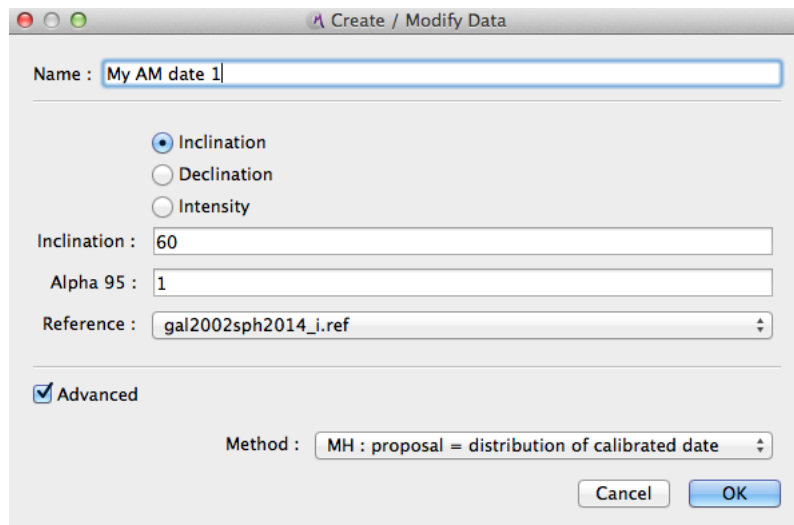
Figure 3.11 – The archeomagnetism extension window with advanced options

- If you choose *Inclination* : enter its value and its *alpha 95* value. The associated error is then calculated by
$$\frac{\alpha_{95}}{2.448}$$

- If you choose *Declination* : enter its value, the value of the *Inclination* associated and its associated *alpha 95* value. The associated error on the declination is calculated by
$$\frac{\alpha_{95}}{\cos(Inclination) * 2.448}$$

- If you choose *Intensity* : enter its value and its *Error*.

Reference curves

Several reference curves are implemented in ChronoModel. They are listed in Table 3.2. Other curves may be included in ChronoModel from Application settings. See Figure 3.9 that shows the windows allowing to add a new reference cur

Name	Reference
Gal2002Sph2014_D	Hervé *et al.* [14, 15]
Gal2002Sph2014_I	Hervé *et al.* [14, 15]
Gws2003uni_F	Hervé *et al.* [14, 15]
Bulgaria_D	Kovacheva *et al.* [16]
Bulgaria_I	Kovacheva *et al.* [16]
Bulgaria_F	Kovacheva *et al.* [16]
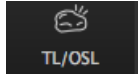
Table 3.2 – Reference curves implemented in ChronoModel

Advanced options

You may want to change the MCMC method used to sample out of the posterior distribution of the calibrated date associated with that dating (See 4 se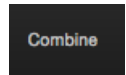ction for more details on MCMC methods). The default MCMC method used for archeomagnetism datings is the Metropolos Hastings algorithm using a distribution that mimics the individual calibration density as a proposal. Two other proposals may also be selected, the prior distribution and a Gaussian adaptative random walk.

3.4.2.3   Luminescence dating (TL/OSL)
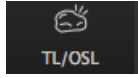
Click on [TL/OSL] to open the TL/OSL extension window as shown in Figure 3.12. A TL/OSL measurement is defined by its name, its Age, its error and the year of reference.

Reference year

The function of calibration associated is then $g(t) = t_0 - t$, where t_0 is the reference year.
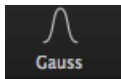
Advanced options

You may want to change the MCMC method used to sample out of the posterior distribution of the calibrated date associated with that dating (See 4 section for more details on MCMC methods). The default MCMC method used for TL/OSL datings is the Metropolos Hastings algorithm using a Gaussian adaptative random walk as a proposal. Two other proposals may also be selected, the prior distribution and a distribution that mimics the individual calibration density.

Figure 3.12 – The paleodose measurement in luminescence extension window with advanced options

3.4.2.4 Gaussian dating (Gauss)
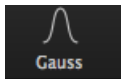
Click on [Gauss] to open the Gauss extension window as presented in Figure 3.13. A Gaussian dating is defined by its name, its mean and its error. A Gaussian date is assumed to be defined in date format BC/AD.

Reference curve

For Gaussian datings, the calibration process follows a quadratic function $g(t) = a * t^2 + b * t + c$.

If $a = c = 0$ and $b = 1$, then $g(t) = t$ that is the calibration function is the identity function (that is no calibration). This is the default choice corresponding to the option "None".

You may also custom your own calibration curve using the option "Build your equation". The parameters a, b, and c may be changed.

Since ChronoModel version 1.3, a third option is implemented. You may load a CSV file containing a different calibration curve. To do that, select the last option "Use custom curve file" and choose a file from the drop-down menu. If your file is not already included in ChronoModel, use the "Application settings" windows to add/delete another file. See Figure 3.9 that shows the windows allowing to add a new reference curve for radiocarbon datings.

Advanced options

You may want to change the MCMC method used to sample out of the posterior distribution of the calibrated date associated with that dating (See 4 section for more details on MCMC methods). The default MCMC method used for Gaussian datings

is the Metropolos Hastings algorithm using a Gaussian adaptative random walk as a proposal. Two other proposals may also be selected, the prior distribution and a distribution that mimics the individual calibration density.
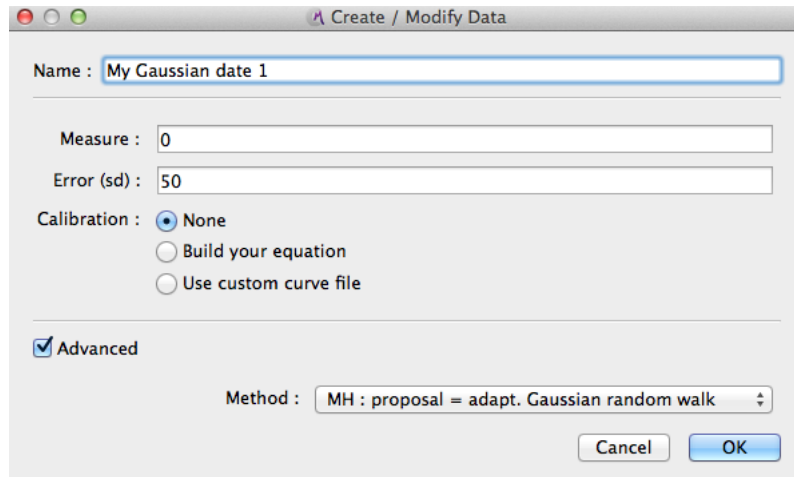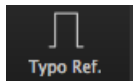
Figure 3.13 – The Gaussian measurement extension window with advanced options

3.4.2.5   Typo-chronological reference (Typo Ref.)
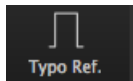
Click on [Typo Ref.] to open the Typo-reference extension window (See Figure 3.14). A Typo-reference is defined by its *Name*, its *Lower date* and its *Upper date*, with the constraint that the *Lower date* is the more ancient date.

Advanced options

You may want to change the MCMC method used to sample out of the posterior distribution of the calibrated date associated with that dating (See 4 section for more details on MCMC methods). The default MCMC method used for typo-chronological reference is the Metropolos Hastings algorithm using a Gaussian adaptative random walk as a proposal. The other proposal that may also be selected is the prior distribution.
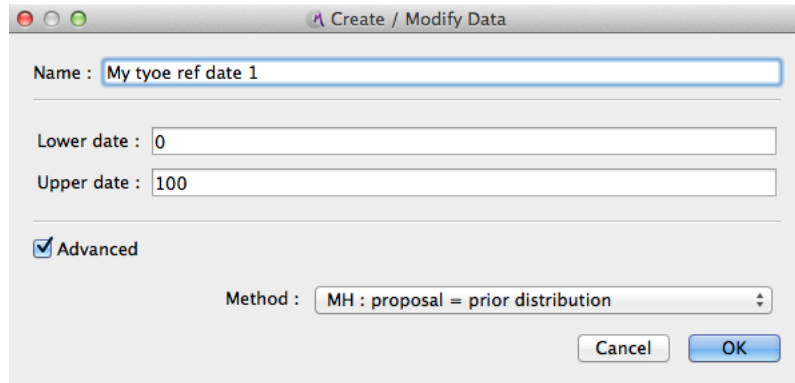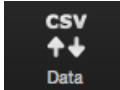
Figure 3.14 – The Typo-reference extension window with advanced options

3.4.2.6 Importing data using a CSV file

Data may also be imported using a CSV file from the tab.
Any CSV file may be imported by clicking on .
However, CSV files have to be organised in a specific way depending on the type of measurements (the type of the dating process). Each of the following information should be written in the same row but in a separate cell.

Gaussian measurements:
Cell 1 : GAUSS, the type of measurement
Cell 2 : Name, the name of the dated sample
Cell 3 : Age, the mean of the Gaussian measurement
Cell 4 : Error (sd), the standard error of the Gaussian measurement
Cell 5 : Ref. curve, one of the following options : "none" (if no calibration is needed) or "equation" (when specifying an equation) or the name of the file containing the reference curve to be used.
Now, if "none" or a CSV file is declared,
Cell 6 : any text will be considered as a comment
Now, if "equation" is declared,
Cell 6 : a, **Cell 7** : b and **Cell 8** : c
where a, b and c refer to the calibration curve. Indeed, in that case the equation of the calibration curve is the following one:

$$g(t) = a * t^2 + b * t + c$$

The content of any following cells will be ignored (so these can be used for comments). See Figure 3.15 for an example.

	A	B	C	D	E	F	G	H	I
1	Gauss	My gaussian data 1	0	50	none				
2	Gauss	My gaussian data 2	1000	50	equation		0	1	0
3	Gauss	My gaussian data 3	1000	50	example_csv_ref.csv				
4									

Figure 3.15 – Example of a CSV file containing Gaussian datings

Thermoluminescence measurements:

Cell 1 : TL/OSL, the type of measurement

Cell 2 : Name, the name of the dated sample

Cell 3 : Age, the value of the measurement

Cell 4 : Error (sd), the standard error of the measurement

Cell 5 : Reference year, the year of the dating of the sample

The content of any following cells will be ignored (so these can be used for comments).
See Figure 3.16 for an example.

	A	B	C	D	E	F
1	TL/OSL	My TL/OSL date 1	520	30	2010	
2	TL/OSL	My TL/OSL date 2	1250	30	2010	
3						

Figure 3.16 – Example of a CSV file containing TL/OSL datings

Radiocarbon measurements:

Cell 1 : 14C, the type of measurement

Cell 2 : Name, the name of the dated sample

Cell 3 : Age, the value of the measurement

Cell 4 : Error (sd), the standard error of the measurement

Cell 5 : Reference curve, the name of the file containing the reference curve associated with the measurement

Cell 6 : Reservoir effect, the value of the reservoir effect if any

Cell 7 : Error of the reservoir effect, the value of the error associated, if any

Cell 8 : Wiggle matching type if any : "fixed", "range" or "gaussian"

Any other word will be ignored by ChronoModel.

Cell 9 : Wiggle matching value, the value if "fixed", the lower value if "range" or the mean value if "gaussian"

Cell 10 : Wiggle matching value, the higher value if "range" or the standard error value if "gaussian"

The content of any following cells will be ignored (so these can be used for comments).
See Figure 3.17 for an example.

	A	B	C	D	E	F	G	H	I	J
1	14C	My radiocarbon date 1	3432.37	13.3934	intcal09.14c	0	0			
2	14C	My radiocarbon date 2	3452	50	intcal04.14c	0	0	gaussian	12	2
3	14C	My radiocarbon date 3	3432.37	13.3934	intcal09.14c	0	0	fixed	15	
4	14C	My radiocarbon date 4	3452	50	intcal04.14c	0	0	range	0	200
5										

Figure 3.17 – Example of a CSV file containing radiocarbon datings

Archeomagnetism measurements:

Cell 1 : AM, the type of measurement

Cell 2 : Name, the name of the dated sample

Cell 3 : Type, "inclination", "declination" or "intensity"

Cell 4 : Inclination value

Cell 5 : Declination value

Cell 6 : Intensity value

Cell 7 : Error (sd) or alpha 95, the degree of alpha 95 associated with the inclination
/ declination value or the standard error associated with the intensity value

Cell 8 : Reference curve, the name of the file containing the reference curve associated with the measurement

The content of any following cells will be ignored (so these can be used for comments).
See Figure 3.18 for an example.

	A	B	C	D	E	F	G	H
1	AM	My AM date 1	inclination	25	0	0	1	gal2002sph2014_i.ref
2	AM	My AM date 2	declination	60	25	0	1	gal2002sph2014_d.ref
3	AM	My AM date 3	intensity	0	0	56	1	gwh2013uni_f.ref
4								

Figure 3.18 – Example of a CSV file containing archeomagnetism datings

Typo-chronological reference:

Cell 1 : Typo ref., the type of measurement

Cell 2 : Name, the name of the dated sample

Cell 3 : Lower date

Cell 4 : Upper date

The content of any following cells will be ignored (so these can be used for comments).
See Figure 3.19 for an example.

Figure 3.19 – Example of a CSV file containing a Typo-chronological reference

Be aware that the cell separator and the decimal separator of CSV files have to be those specified in the Application settings of ChronoModel otherwise the CSV file will not be loaded.

Figure 3.20 displays an example of a CSV file loaded in ChronoModel. This file contains 4 data : 2 radiocarbon measurements, an inclination measurement in archeomagnetism and a Gaussian date. When a data line is selected, the labels of each column is refreshed according to the type of data.

Figure 3.20 – CSV file loaded in ChronoModel - 14C type

Figure 3.21 – CSV file loaded in ChronoModel - AM type

Now, once the CSV file is loaded, data may be included in an already existing event. To do that, select the data line to be included, click on the associated name and then drag and drop it in the event. Multiple selection may also be done. Click on the type of the first dataline to be included, then press and keep pressing "Ctrl" key (or the equivalent for MAC) while selecting the other datalines. Then click on one of the selected line and drag and drop them all in an event.
You may also drag and drop one or several data lines directly into the events' scene. ChronoModel will create an event for each dataline. The name of the events will be the name of the dataline included.

If a line appears in green, that means that the data has been attributed to an event. If the line is colored in green, that means that the data has not been attributed due to errors. In this case, ages and errors should be checked and modified before any use in ChronoModel. .

3.4.2.7   Calibration process

For each measurement, it is possible to see the result from the individual calibration process. First, select the associated event from the events' scene. Then select the measurement from the [Properties] tab and click on [Calibrate] at the bottom of the window. For instance, Figure 3.22 displays the calibration process of a radiocarbon measurement. Several statistical results according to the distribution of the calibrated date are summaries (See section 5.2.4 for more details on statistics). The level of the HPD regions may be changed from this window.

Unclick [Calibrate] to see the events' scene again.
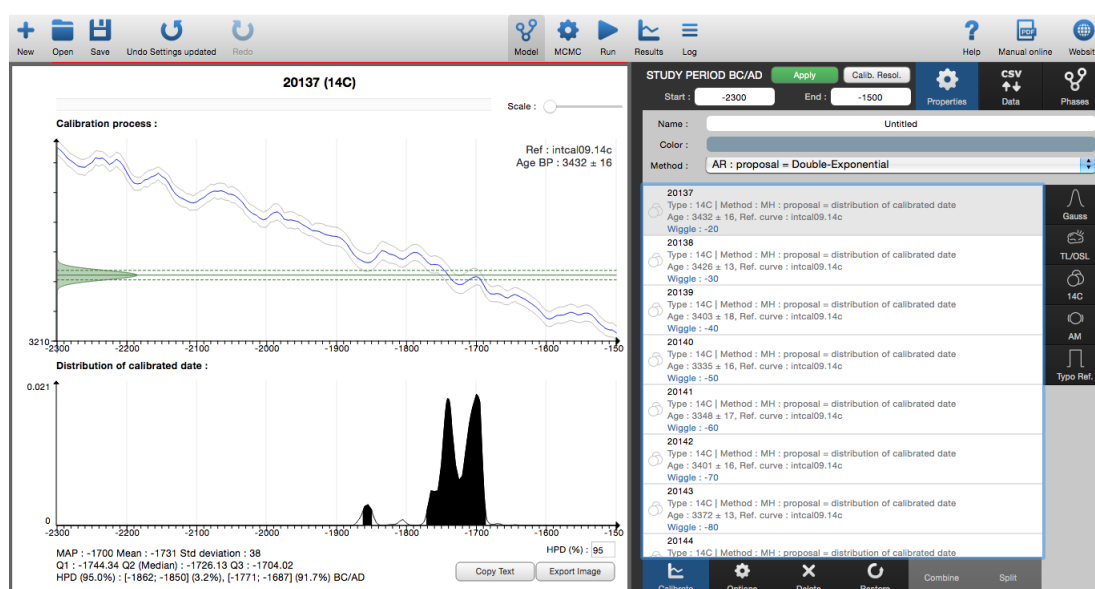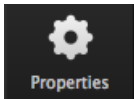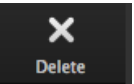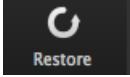
Figure 3.22 – Calibration process of a radiocarbon datation
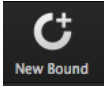
3.4.2.8   Deleting / Restoring a measurement
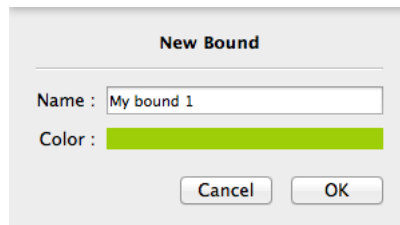
To delete a measurement from an event, first select the event in the event's scene by clicking on its name. The event will be selected only if it is surrounded by a red line. Then using the [Properties] tab, you may select the measurement to be deleted and then click on the [Delete] icon situated on the right hand side of the window to delete it from this event.

A deleted measurement may be restored in a selected event. Click on [Restore] icon at the bottom of the window and choose the measurement to be restored from the list of deleted measurements.
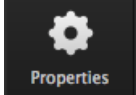
3.4.3   Creating a new bound

To create a new bound in the events' scene, select [New Bound] on the left hand side of the window. A new window will be opened asking you to name this new bound. For example, let's call it "My bound 1".
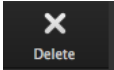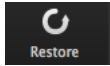
After validation, the bound appears in the events' scene (See Figure 3.23). Now, when the bound is selected in the event's scene, the bound's properties may be seen on the

right hand side of the window, in the [Properties] tab. From there, its name and its color might be changed and values may be added. A bound may either have a fixed value or have a uniform distribution within a range of values. These options may be changed from that window as well.

3.4.4   Deleting / Restoring an event or a bound

An event or a bound may be deleted from the events' scene first by selecting the event and then by clicking on the [Delete] icon placed on the left hand side of the events' scene. An event is deleted with all its measurements.

Any element deleted may be restored by clicking on the [Restore] icon on the left hand side of the events' scene. Then the element to be restored may be picked from the new window presenting all deleted elements (bounds or events). See Figure 3.24 An event containing measurements is restored with all its measurements.
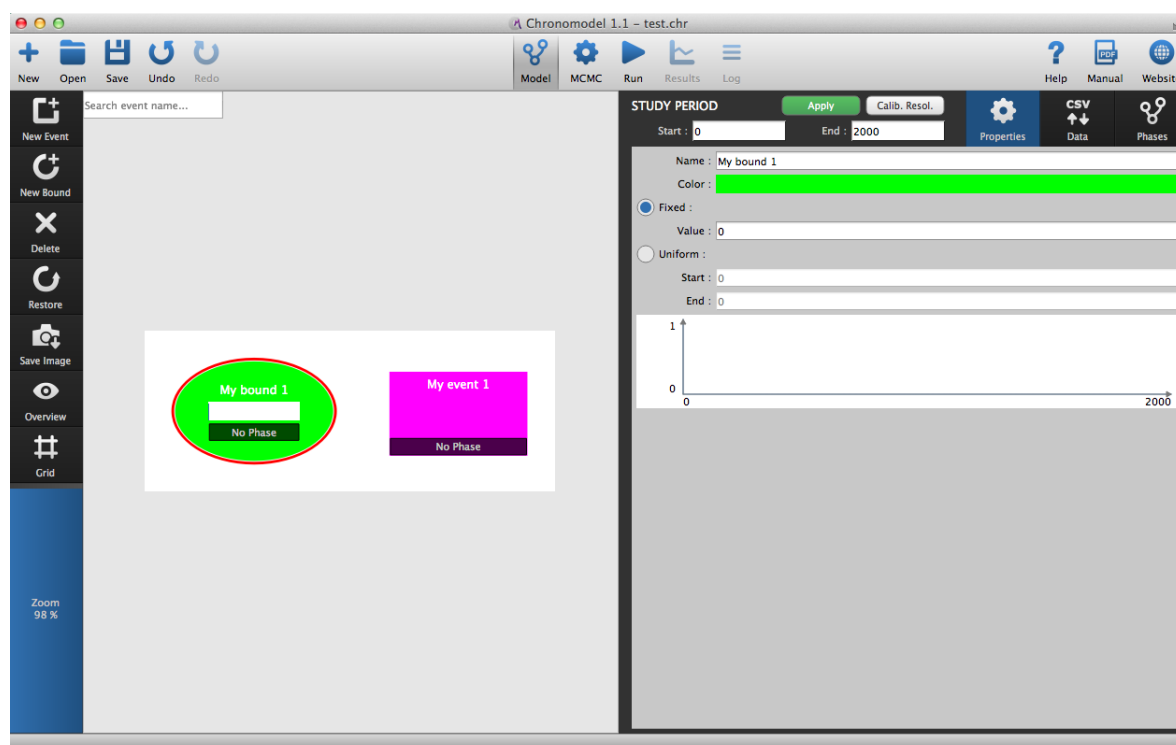
Figure 3.23 – Creation of a new bound in the events' scene
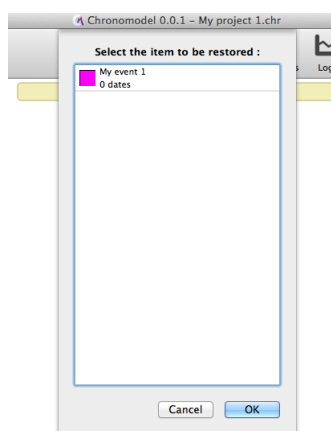
Figure 3.24 – Window presenting all events or bounds that may be restored

3.4.5 Creating / Deleting a constraint

From the events' scene, stratigraphic or succession constraint may be added between any two elements (events and bounds). To create such a constraint, first select the oldest element of the succession. This selected element should now be surrounded by a red line. Now, click and keep clicking on the "Alt" key from your keyboard. A black arrow should now be seen in the events' scene. Now select the youngest element of the succession. To validate the constraint, you need to click on the name of the second element before releasing the "Alt" key. The arrow should be seen in black heading

from the oldest element to the youngest one.
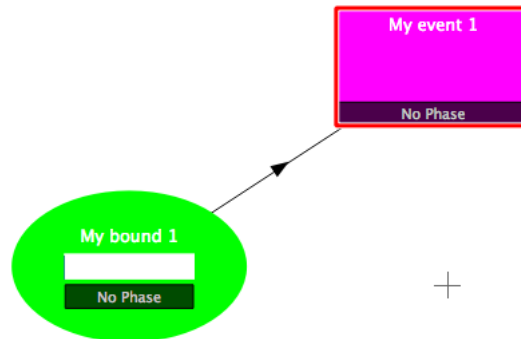
Figure 3.25 – Events' scene showing a constraint between a bound and an event

To delete a constraint, move the mouse over the corresponding black arrow. This arrow should become red and a cross should appear in the middle of the arrow. You may now click on the cross, to begin the deleting process. A confirmation box should appear to validate the deleting action..
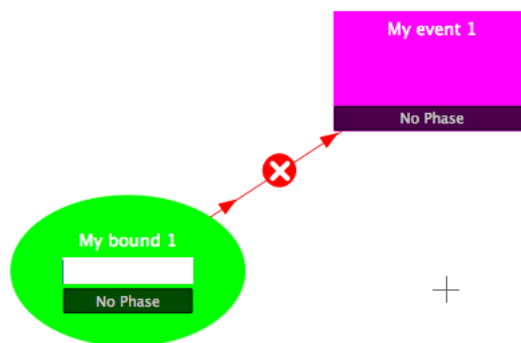
Figure 3.26 – Events' scene showing a constraint ready to be deleted
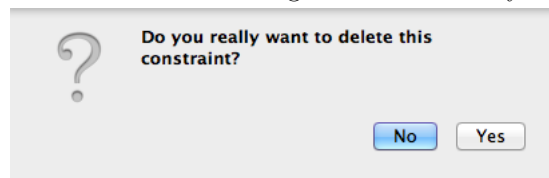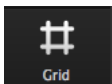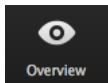
Figure 3.27 – Confirmation box to delete a constraint

3.4.6   Using the grid

By default, the events' scene is white. A grid may be added using the icon situated on the left hand side of the window.

3.4.7 Using the overview

The overview may be useful when many elements are created in the event's scene. To watch the events scene from an overview use the icon [Overview] placed on the left hand side of the window.

3.4.8 Exporting the image of the events' scene

You may also export the image of the events' scene by clicking on the [Save] icon and save it either in PNG format or in Scalable Vector Graphics (SVG) format. In both cases, you will need to name the image and to choose the directory where to save the image (see Figure 3.28). If you choose the PNG format, you will be ask for the image size factor and the number of dots per inch.
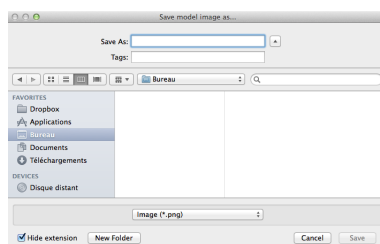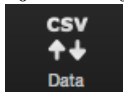
Figure 3.28 – Dialog box used to name the image and to choose the directory where to save the image.

3.4.9 Exporting data from the project

All measurements included in the current project may be exported in a CSV file using [Export all project data as CSV] from the [Data] tab.

3.5   Creating phases and constraints between phases

In ChronoModel, the events' scene and the phases' scene are separated in order to keep the design of the model comprehensible. Indeed an event may belong to several phases. Moreover, the phases' scene may be placed side by side with the events' scene, in order to have a complete look at the design of the model.
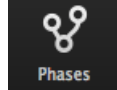
The phases' scene may be seen from the [Phases] tab placed on the right hand side of the window. See Figure 3.29.
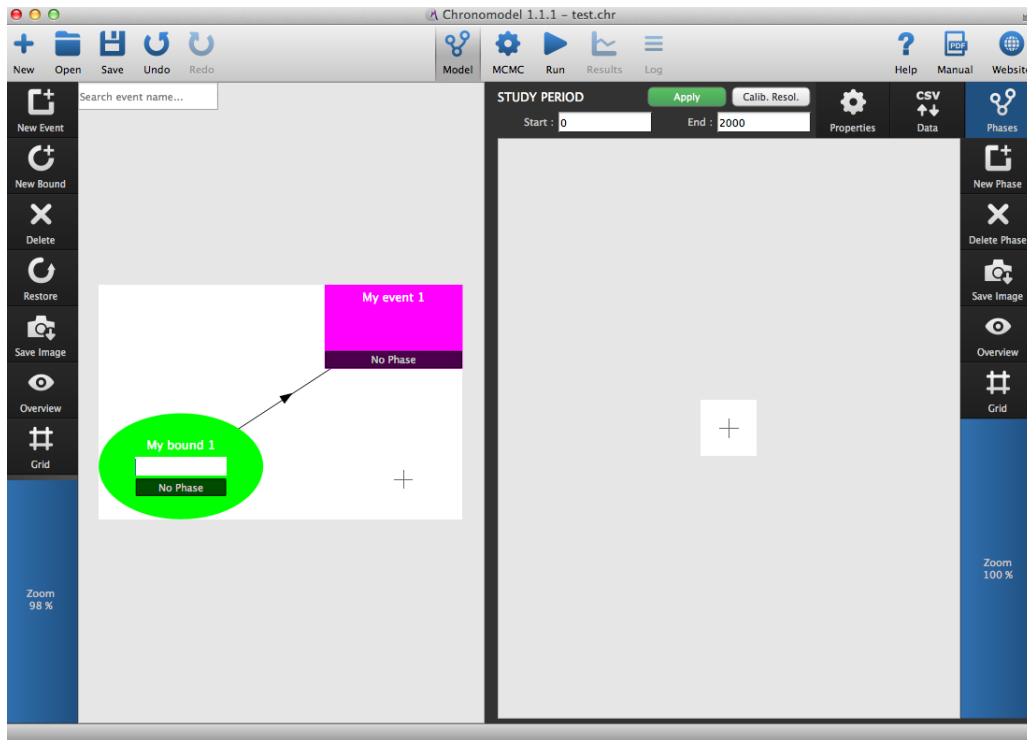
Figure 3.29 – View of the events' scene and the phases' scene

3.5.1   Creating a new phase

To create a phase, click on the [New Phase] icon on the right hand side of the window. This action will open a new dialog box presented in Figure 3.30.

This dialog box asks for the name and the color given to the phase and its maximum duration. By default, the duration of the phase is unknown. If prior information about the maximum duration are available, it may be included in the model. See Figure 3.30.
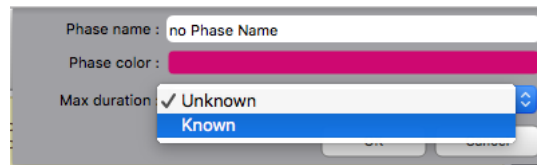
Figure 3.30 – Dialog box to include prior information about the maximum duration of the phase

3.5.2 Modifying / Deleting a phase

To modify a phase, double-click on its name and the dialog box presented in Figure 3.30 will be reopened.

To delete a phase, select the phase to be deleted by clicking on its name and then click on the ⊠ Delete Phase icon on the left hand side of the window.

3.5.3 Including / Removing events or bounds

Events may be included in a phase. You may select one or several events or bounds from the events' scene and then click in the white square on the left of the phase's box. The Figure 3.31 presents the inclusion of a bound in a phase. As a result, the color of the phase appears on the bottom of the event 's (or the bound's) box, and the name and the color of the event (or the bound) appear at the bottom of the phase's box.
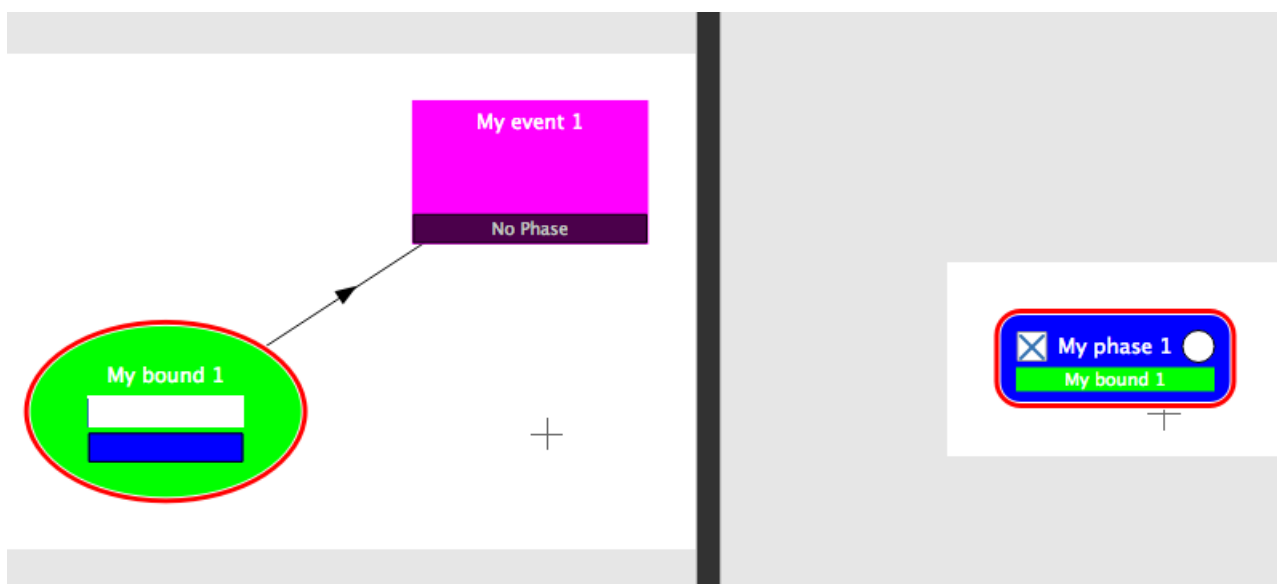
Figure 3.31 – Including a bound in a phase

Several events or bounds may be included in a same phase. An event or a bound may also belong to several phases. Figure 3.32 illustrates these cases. We may see that "My phase 1" contains two elements : "My event 1" and "My bound 1". "My bound 1" belongs to both phases as both colors appear at the bottom of its box, and a cross appears in the white square of each phase's box.
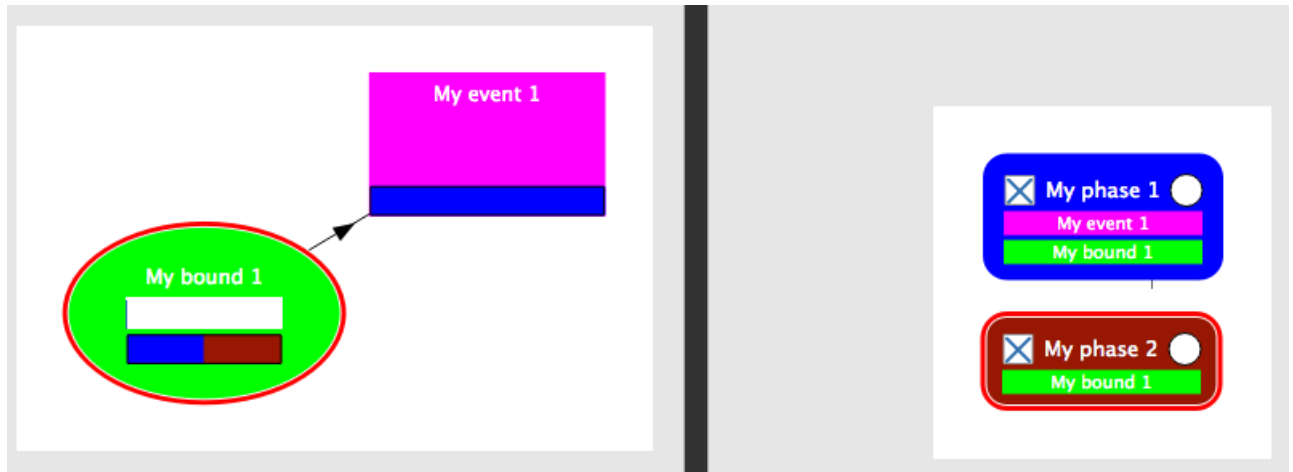
Figure 3.32 – Including several events in a phase and an event in different phases

To remove an element from a phase, first select the element from the events' scene, then click on the cross associated in the phase from which this element should be remove. Once the element is removed from a phase, its name does not appear any more in the phase's box and the phase's color does not appear any longer in the element's box.

3.5.4   Creating / Deleting a constraint between two phases

A succession constraint may relate two phases. To do that, first select the "oldest" phase. Press and keep pressing the "Alt" key from your keyboard. Then move the arrow up to the "youngest" phase and click on its name. Now, you may release the "Alt" key and a black arrow should be seen between both phases heading from the oldest phase to the youngest one. Figure 3.33 displays an example.

In ChronoModel, a hiatus expresses the minimum time elapsed between two phases. By default, the hiatus between these phases is unknown. However, if prior information about the hiatus are available, it might be included in the model.

In order to include prior information about a hiatus, double-click on the constraint (the black arrow) to be modified, and then fill the dialog box shown in Figure 3.34.

To delete a constraint, double-click on it to see the dialog box shown in Figure 3.34
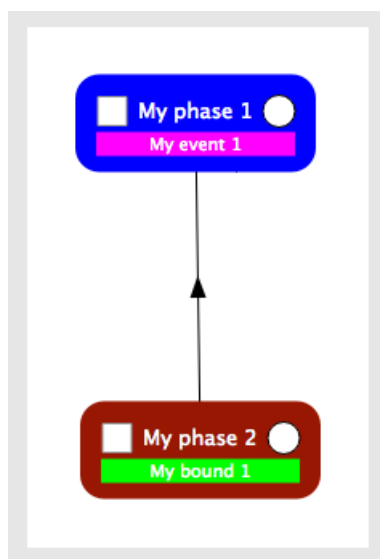
Figure 3.33 – Example of the phases' scene showing a constraint between two phases

and click on "Delete constraint".
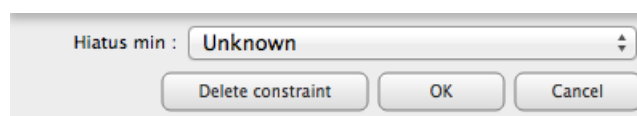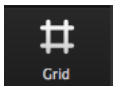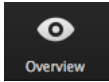
Figure 3.34 – Dialog box to modify a hiatus between two phases

3.5.5 Using the grid

By default, the phases' scene is white. A grid may be added using the icon situated on the right hand side of the window.

3.5.6 Using the overview

The phases' scene may also be seen from an overview using the icon situated on the right hand side of the window.

3.5.7 Exporting the image of the phases' scene

The image of the phases' scene may be saved by clicking on the icon placed on the right hand side of the window. You may save it either in PNG format or in Scalable Vector Graphics (SVG) format. In both cases, you will need to name the image and

to choose the directory where to save the image (see Figure 3.28). If you choose the PNG format, you will be ask for the image size factor and the number of dots per inch.

Chapter 4

Numerical methods

In Bayesian modelling, posterior densities and their marginal densities are of interest. In general, we cannot derive posterior distributions analytically. However, posterior distributions are always known up to proportionnality.

Let's define by θ the parameter of interest and by y the data sample. The posterior distribution $p(\theta|y)$, is known up to proportionnality, by multiplying the prior and likelihood functions as follow:

$$p(\theta|y) \propto p(\theta)p(y|\theta)$$

We can build up a sample of θ from $p(\theta|y)$ even if we only know $p(\theta|y)$ up to proportionality using simulation methods based on Markov chain Monte Carlo (MCMC).

## 4.1 Markov chain Monte Carlo

Monte Carlo methods aim at estimating numerical values using repeated random sampling. By the law of large numbers, integrals described by the expected value of some random variable can be approximated by taking the empirical mean of independent samples of the variable. However, independent samples are usually difficult to compute. That is why Markov chains, more easily computed, are usually implemented.

A Markov chain is a sequence of random variables $\theta^{(1)}$, $\theta^{(2)}$, ..., $\theta^{(I)}$, for which, for any t, the distribution of $\theta^{(t)}$ given all previous θ's depend only on the recent value, $\theta^{(t-1)}$ [17, 18]. Under regularity conditions, when I becomes large, the Markov chain converges to its stationary (or equilibrium) distribution. When a Markov chain has reach its equilibrium (or its stationary state), any sample from the Markov chain is a sample from the stationary distribution.

Now, for any posterior distribution $p(\theta|y)$, it is possible to build a Markov chain $\theta^{(1)}$, $\theta^{(2)}$, ..., $\theta^{(I)}$, whose stationary distribution is the posterior distribution $p(\theta|y)$. Once the equilibrium state is reached, any sample from the Markov chain may be used to estimate the posterior density, the posterior mean ...

As the sampling is done sequentially, with the distribution of the sampled draws depending on the past value drawn, correlation between successive values is high. To reduce this correlation, it is required to thin the Markov chain and use a subsample to estimate the posterior density distribution.

4.2   Building a Markov chain in multiple dimensions: the Gibbs sampler

A convenient algorithm useful in many multidimensional problems is the Gibbs sampler [17, 18].

Let's say we want to approximate the joint posterior probability density function, $p(\theta_1, \theta_2, ..., \theta_d|y)$. The algorithm starts with a sample of initial values $(\theta_1^{(0)}, \theta_2^{(0)}, ..., \theta_d^{(0)})$ randomly selected. The first step of the algorithm is to update the first value by sampling a candidate value of $\theta_1^{(1)}$ knowing $\theta_2^{(0)}, ..., \theta_d^{(0)}$ from the full conditional distribution $p(\theta_1|\theta_2^{(0)}, ..., \theta_d^{(0)})$. The next step is to find a candidate value $\theta_2^{(1)}$ knowing $\theta_1^{(1)}, \theta_3^{(0)}, ..., \theta_d^{(0)}$ using the full conditional distribution $p(\theta_2|\theta_1^{(1)}, \theta_3^{(0)}, ..., \theta_d^{(0)})$. And so on... Then the dst step is to find a candidate value for $\theta_d^{(1)}$ knowing $\theta_1^{(1)}, \theta_2^{(1)}, ..., \theta_{d-1}^{(1)}$. This process is then iteratively repeated.

When the number of iterations, I, becomes large, the distribution of the sample $(\theta_i^{(0)}, ..., \theta_i^{(I)})$ becomes extremely close to the marginal posterior distribution of θ_i. In practical terms, this means that if we can simulate a Markov chain for θ_i until it reaches its equilibrium, then any sample from this Markov chain may be used to estimate information about the marginal posterior distribution.

4.3   Drawing from full conditional posterior distributions

Again, full conditional posterior distributions are not always easy to simulate. Algorithms are required to approximate these posterior distributions. In ChronoModel, two main algorithms are implemented, the **rejection sampling method** (also called

acceptance-rejection) and the **Metropolis Hastings algorithm**. Both algorithms requires a proposal density function that should be easily sampled from in order to generate new candidate values. For the rejection sampling algorithm, it is common to use, if possible, the prior function or the likelihood as a proposal function. For the Metropolis-Hastings algorithm, a common choice is to use a symmetric density, such as the Gaussian density.

Depending on the type of the parameter, event θ, calibrated date t_i, variance of calibrated date σ_i^2 or bound θ^B, different methods are proposed. These methods are described here in turn.

4.3.1 Full conditional posterior distribution of the event θ

Three different methods can be chosen.

- **Rejection sampling with a Gaussian proposal** [19]

- **Rejection sampling with a Double exponential proposal** [19]

- **Metropolis-Hastings algorithm with an adaptative Gaussian random walk** [20]

The two first methods are exact methods. We recommend to use one of them.

4.3.2 Full conditional posterior distribution of calibrated dates t_i

In this case, three different methods can be chosen.

- **Metropolis-Hastings algorithm using a density that mimics the individual calibration density**
  This method is adapted for multimodal densities, such as calibrated measures.

- **Metropolis-Hastings algorithm using the parameter prior distribution**
  This method is recommended when no calibration is needed, namely for TL/OSL, gaussian measurements or typo-chronological references.

- **Metropolis-Hastings algorithm using an adaptative 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.

4.3.3   Full conditional posterior distribution of the variance of a calibrated date σ_i^2

As the full conditional density is unimodal, the Metropolis Hastings algorithm can be implemented here. The proposal density involved is an adaptive Gaussian random walk [20]. The variance of this proposal density is adapted during the process.

4.3.4   Full conditional posterior distribution of a wiggle matching δ_i

If the δ_i is fixed, no sampling is needed.

If the δ_i has a uniform distribution on $[\delta_m,\ \delta_M]$ or if it has a Gaussian distribution, then the rejection sampling is used in order to sample out of the full conditional. In both cases, the proposal is a Gaussian distribution.

4.3.5   Full conditional posterior distribution of a bound θ^B

If the bound is fixed, no sampling is needed.

If the bound has a uniform distribution, the full conditional distribution is also a uniform distribution.

4.4   MCMC settings

The Gibbs sampler described above generates a Markov chain for each parameter. These Markov chains are samples of the marginal posterior distribution of each parameter.

Each Markov chain starts with a given starting value, uses a burn-in phase to forget about this starting value, uses an adaptative phase to adapt all random walk variances. Then, and only then, all values of the chains are used to make estimates of the posterior distribution.

The number of iterations desired for each phase may be modified using MCMC settings. Figure 4.1 shows the dialog box containing the MCMC settings. Each of these settings is describe here in turn.

4.4.1   Burn

The **BURN-in period** is used to start the algorithm. As successive values of a Markov chain depend on the last value, a number of iterations are needed to "forget"
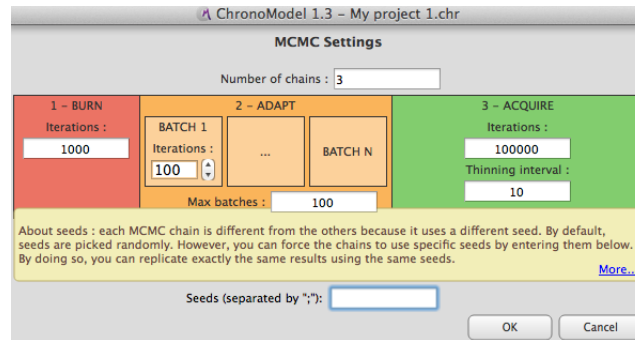
Figure 4.1 – MCMC settings

the initial value. We recommend to use at least 1 000 iterations.

4.4.2 Adapt

During the **ADAPT period**, variances of all adaptative Metropolis-Hastings Gaussian random walks are calibrated so that the corresponding acceptance rates are close to 43%. The adaptation period goes on with another batch until the acceptance rate calculated on the last batch is included between 40% and 46% unless the maximum number of batches is reached.

4.4.3 Acquire

In this period, all Markov chains are assumed to have reached their equilibrium distribution. Of course, this has to be checked and the next section provides useful tools that can help controlling whether the equilibrium is actually reached. If so, Markov chains may be sampled and information about conditional posterior distributions may be extracted.

Sampling from these Markov chains need to be carefully made. Indeed, successive value of a Markov chain are not independent. In order to limit the correlation of the sample, we can choose to thin the sample by only keeping equally spaced values.

4.4.4 Thinning

Successive values of Markov chains are not independent. The idea is to sample a large amount of values and then, once equilibrium reached, to keep every m values. m should be chosen so that remaining values may be considered as independent.

In order to check the correlation between two successive values of the sample kept from the **ACQUIRE period**, autocorrelograms may be seen in the "Autocorrelation

tab" of the Results part. Hence, it is recommended to start with the default thinning interval and to change it if one autocorrelogram is not correct.

4.4.5   Starting values

Each Markov chain starts with a starting value that is randomly selected among all solutions compatible with constraints.
Starting values may be fixed using a seed for each chain.

4.4.6   Number of chains

It is recommended to run several chains in parallel using different starting values when examining a new problem. This may be used to state whether the equilibrium is reached before the end of the **ADAPT period**.

Chapter 5

Results

After having gone through the running process , the results tab appears . Now, before any interpretations, the Markov chains have to be checked in order to known whether they have reach their equilibrium before the acquire period.

5.1 Checking the Markov chains

When Markov chains are generated, two points have to be verified : the convergence of the chains and the absence of correlation between successive values. If the Markov chain has not reach its equilibrium, values extracted from the chains will give inappropriate estimates of the posterior distribution. If high correlation remains between successive values of the chain, then variance of the posterior distribution will be biased. Here are some tools to detect whether a chain has reach its equilibrium and whether successive values are correlated. We also give indications about what can be done in these unfortunate situations.

By default, results shown correspond to the date density functions. To see results corresponding to individual standard deviations (σ_i), select "Individual std deviation"
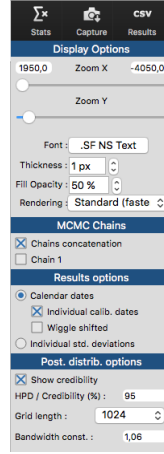
under the "Results options" section   . Now, in order to see information
about all dates included in an event, click on [image] on the right hand side of the
window.

5.1.1   Is the equilibrium reached ? Look at the history plots.

Unfortunately, there is no theoretical way to determine how long a burn-in period
needs to be. The first thing to do is to observe the history plot of the chains (dates
and variances) and to inspect it for signs of convergence. Figure 5.1 displays an
example where the Markov chain explored a first state in the neighbourhood of the
starting value. Then during the **ADAPT period** (orange period), the Markov chain
finally found its equilibirum state that was very different from the starting state.
History plots should have good mixing properties during the **ACQUIRE period**
(green period). They should not show tendencies or constant stages as shown in the
example of Figure 5.2.
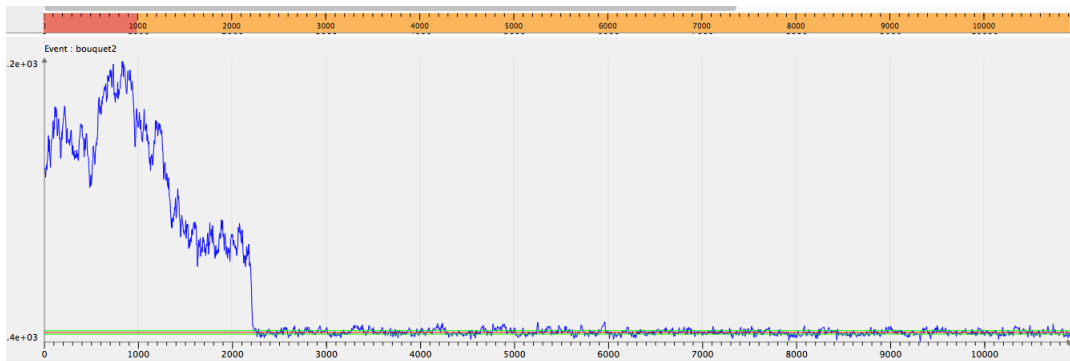
Figure 5.1 – Illustration of the use of the **BURN**-in period and the **ADAPT** period.

Producing parallel Markov chains, all with different starting values, can help de-
ciding whether a chain has reached its equilibrium. When the equilibrium is reached,
the posterior distributions estimated by each chain should be similar.
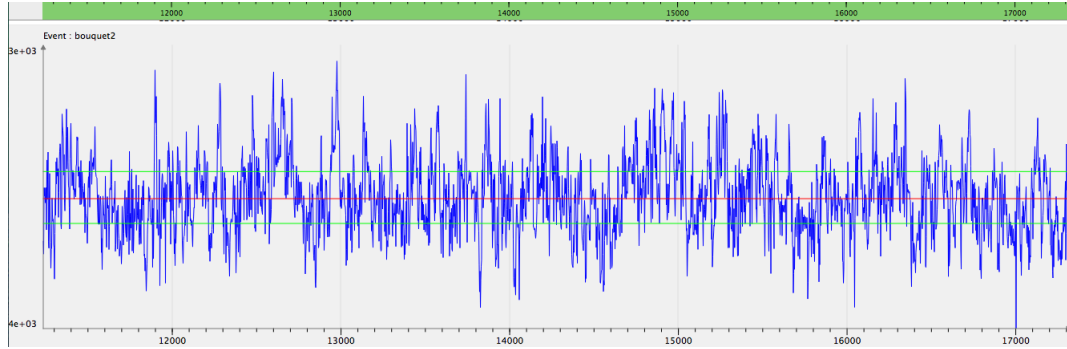
Figure 5.2 – History plot with good mixing properties

What should be done if equilibrium is not reached ? First, the number of iterations of the **BURN period** (red period) should be augmented. A higher number of iterations per batch may also be used.

It might also be of help to change the algorithm used to draw values from full conditional posterior distributions (the MCMC method selected for each parameter).

5.1.2 Correlation between successive values ? Look at the autocorrelation functions.

A Markov chain is a sequence of random variables $\theta^{(1)}$, $\theta^{(2)}$, ..., for which, for any t, the distribution of $\theta^{(t)}$ given all previous θ's depend only on the recent value, $\theta^{(t-1)}$ [17, 18]. Hence, a high correlation between two consecutive values is expected. However, correlation will biased the estimations. That is why thinning the Markov chains is required.

To check whether the chain is correlated, observe the autocorrelation plot. The autocorrelation function should have an exponential decrease. Only the first autocorrelation should be high (at lag 0), the remaining correlations should be negligible. Figure 5.3 displays several autocorrelation functions having a good behaviour.

What should be done if correlation is high ? By default, the thinning interval from the MCMC settings is 10, that is only one value out of 10 is kept from the Markov chains. This thinning interval should be long enough to reduce the correlation between successive values. If not, it should be increased.

5.1.3 Look at the acceptance rates.

The Metropolis Hastings algorithm generates a candidate value from a proposal density. This candidate value is accepted with a probability. An interesting point is the
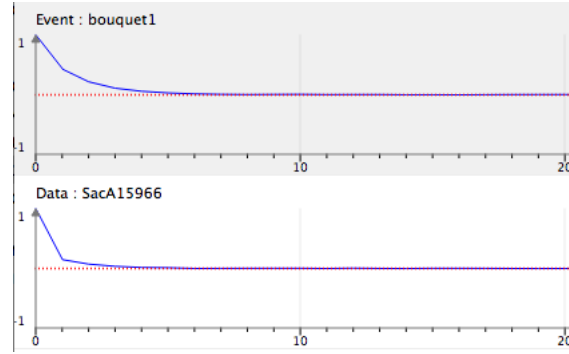
Figure 5.3 – Examples of autocorrelation functions that fall quickly enough under the 95% confidence interval

acceptance rate of these candidate values. This rate gives an idea of the adequation of the MCMC method used. This rate should not be too low or too high. However, there are no theoretical criterion except for adaptative Metropolis Hastings Gaussian random walk.

The theoretical optimal rate for the adaptative Gaussian random walk is 43% [20].

In ChronoModel, this algorithm is used to draw values from the full conditional posterior density of individual variances. But this method may be also choosen for other parameters.

Figure 5.4 displays an example of acceptance rates that are close to 43% (within the interval of 40% and 46%). Acceptance rates of each adaptative Gaussian random walk should be checked.
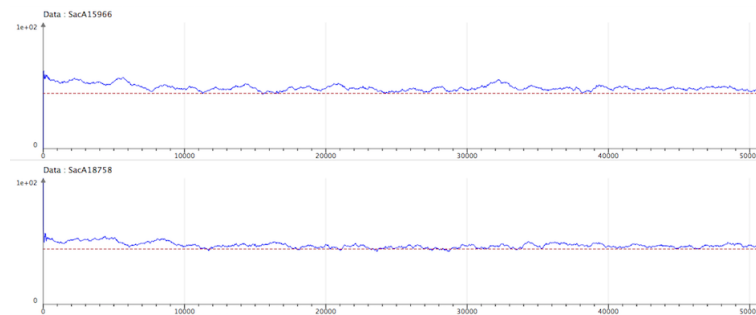
Figure 5.4 – Examples of acceptation rates close to 43%

Be careful, if all batches are used in the adapt period, this might tell that all adaptative Gaussian random walks did not reach the optimal interval. To see that, use the

≡
Log icon on the top of the window. From there in the MCMC part, the number of batches used is stated. Figure 5.5 displays an example where only 18 batches out of 100 were used. That means that all adaptative Gaussian random walks reached the

optimal interval within these 13 batches.
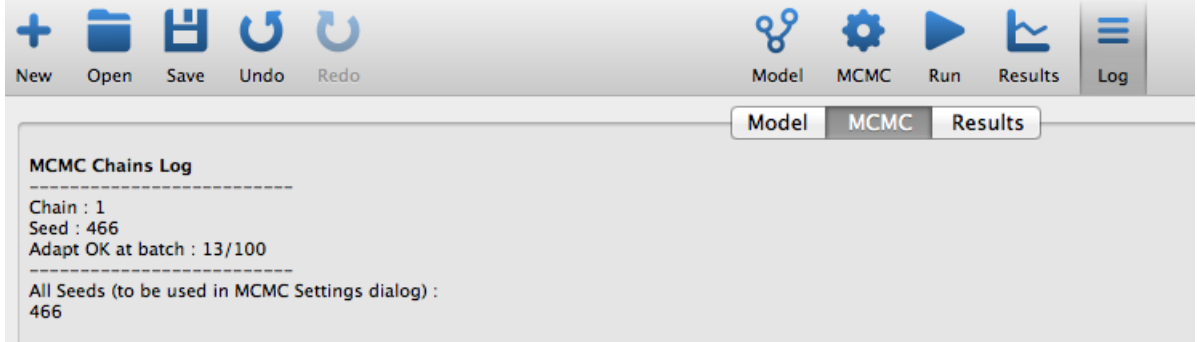
Figure 5.5 – Reading the number of batches used.

What should be done if the acceptance rates are not close to 43% ? From the MCMC settings, ask for a longer number of iterations per batch.

5.2   Marginal posterior densities

If all Markov chains have reached their equilibrium and are not autocorrelated, then the statistical results may be interpreted. These results are estimated using the values of the Markov chains drawn during the Acquire period. By default, results are estimated using values from the concatenation of all chains when several chains were requested.

5.2.1   Estimations

5.2.1.1   Marginal posterior density

When a sample is obtained, let say $X_1, ..., X_n$, the posterior density f is estimated using a Gaussian kernel :

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K(\frac{X_i - x}{h_n})$$

where n is the sample size and K the gaussian density with mean 0 and variance 1. The value of h_n, called the bandwidth, is defined by the Silverman criterion :

$$h_n = \text{bandwidth const.} * s * n^{-1/5}$$

For Gaussian data, the above equation is optimal when the bandwidth const. $= 1.06$ [21]. That is why the default value is 1.06, however it may be changed from the Results options (See Figure 5.6).

This estimation is done on a grid with fixed stepsize. The number of points in this grid is optimal if diatonic. By default, it has 1024 points (10^10). This number of points is called Grid length and may be decreased or increased from the Results options (See Figure 5.6).

Figure 5.6 – Results options

5.2.1.2   Credibility interval (CI)

Once the Markov chain of the parameter of interest is obtained, the credibility interval is constructed using the empirical quantile function. In ChronoModel, a $100(1 - \alpha)\%$ credibility interval is the shortest interval, defined by two empirical quantile function values, out of which $\alpha\%$ of the Markov chain remains.
Figure 5.7 shows an example of a confidence interval.
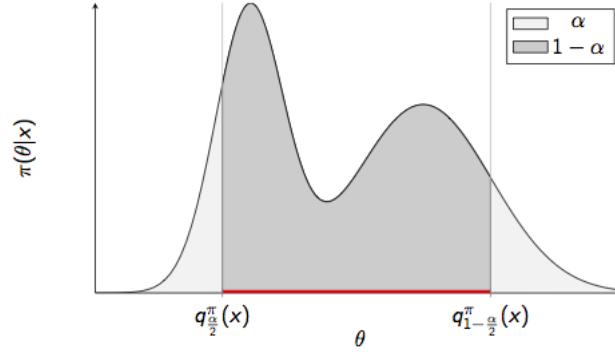
Figure 5.7 – Example of confidence interval (the region in red)

5.2.1.3 Highest posterior density (HPD) region

In ChronoModel, highest posterior density regions are estimated by

$$J = \{x \in Grid | \hat{f}_n(x) \geq k\}$$

The value of k is chosen so that

$$\delta * \sum_{x \in J} \hat{f}_n(x)$$

where δ is the grid stepsize.

However, a better estimation would be the one developped by Samworth and Wand [22].

Figure 5.8 shows an example of a HPD region.
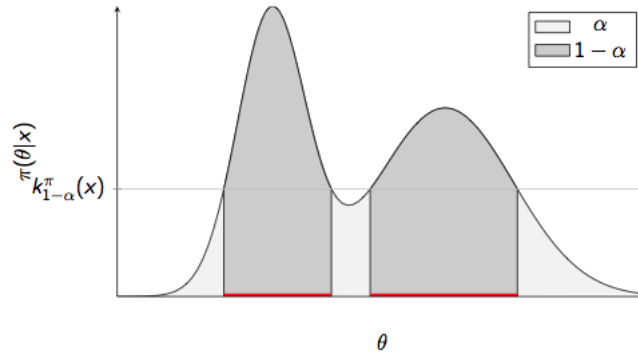
Figure 5.8 – Example of a HPD region (the region in red)

5.2.2   Date format

Results may be seen in different date formats. This might be changed from the "Application Settings" of ChronoModel.

- BC/AD : Before Christ / Anno domini

- Age cal. BP : Ages calibrated before present. Corresponds to 1950 - t in BC/AD

- Date cal. BP : Ages calibrated before present. Corresponds to t in BC/AD - 1950

- Age cal. B2K : Ages calibrated before 2000. Corresponds to 2000 - t in BC/AD

- Date cal. B2K : Ages calibrated before 2000. Corresponds to t in BC/AD - 2000

5.2.3   Graphical results

When the model does not include phases, by default, only marginal posterior distributions of events or bounds are shown. To see marginal posterior distributions of all calendar dates, use the [icon] icon on the left hand side of the window. To see marginal posterior distributions of individual standard deviations, click on "individual std deviations" under "Results options".
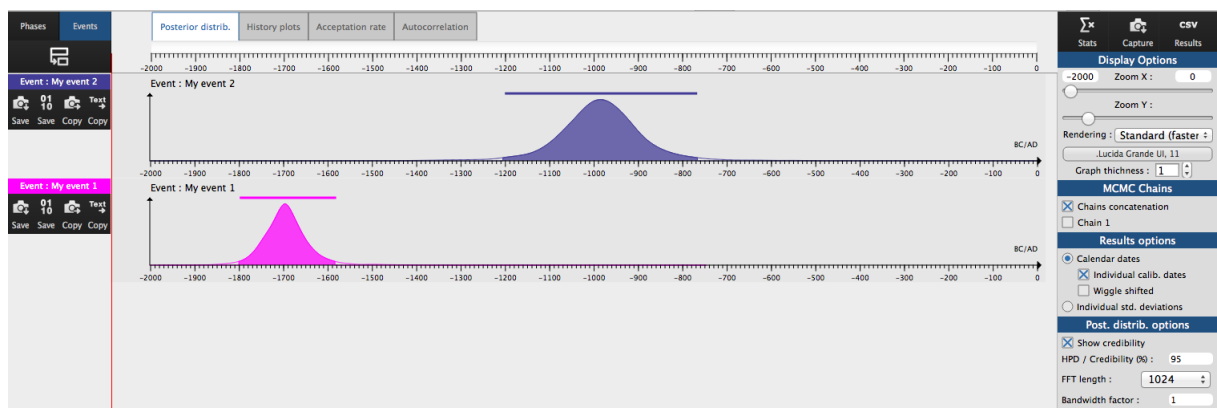
Figure 5.9 – Marginal posterior densities of two events

For any parameter (event, dates and individual standard deviations), the marginal posterior density is associated with its HPD region and its credibility interval as shown in Figure 5.9. HPD regions, regions with the highest probability density, are shown under the curves. Crediblity intervals (CI) are actually the shortest credible intervals

and are shown above the posterior densities by a thicked line. CI may also be hidden, see options under "Post. distrib. options"

By default, HPD regions and credibility intervals are given at 95%. This might be changed under "Post. distrib. options".

By default, results (statistics and posterior density estimations) are given from the concatenation of all chains, but posterior densities may also be seen for each chain using the "MCMC options" on the right hand side of the window.
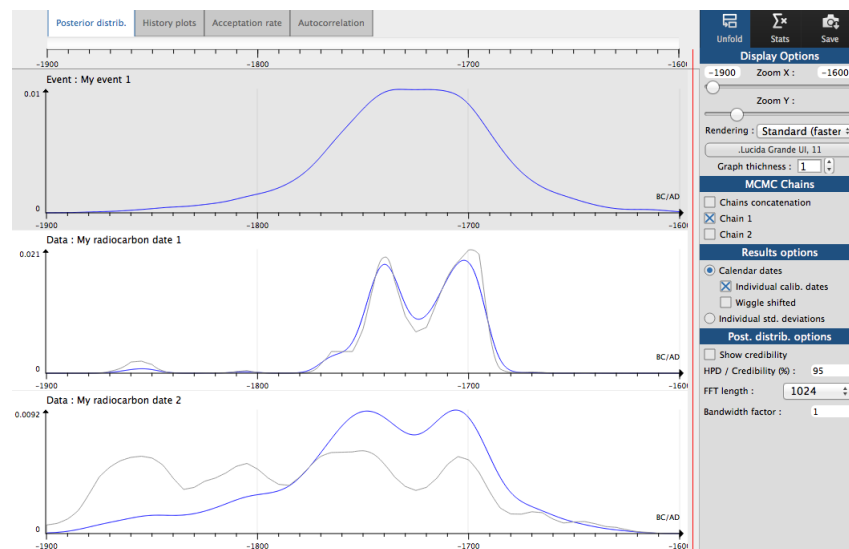
Figure 5.10 – Marginal posterior densities of events and dates

Figure 5.10 displays the posterior densities of event "My event 1" and of both dates included in it. Only one chain is used. On this picture, the grey curves represent densities of the calibration dates, blue curves represent the marginal posterior densities.
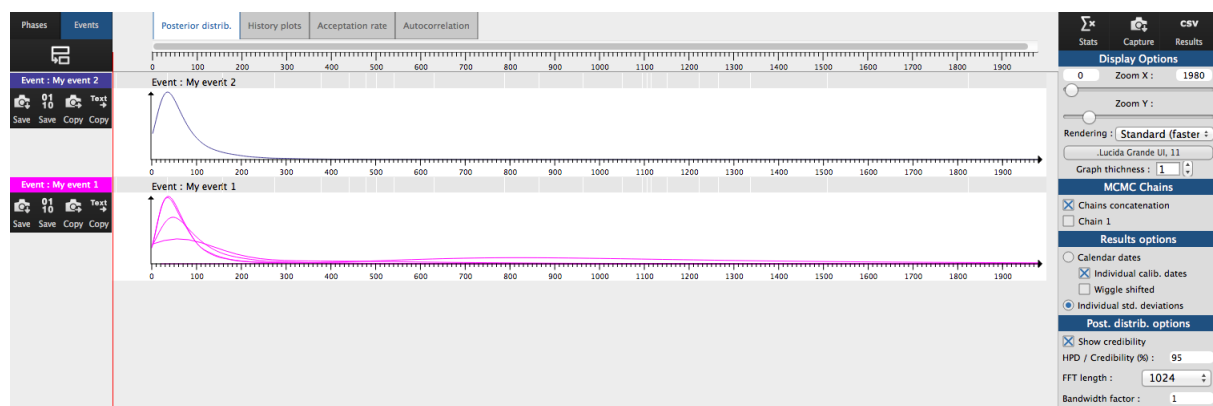
Figure 5.11 – Marginal posterior densities of individual standard deviations

Figure 5.11 shows the posterior densities of individual standard deviations corresponding to the dates. An event is not associated with a standard deviation, however standard deviations from all calendar dates included in the event are overlayed for comparison. Now, to see each individual standard deviation alone, use the

icon.
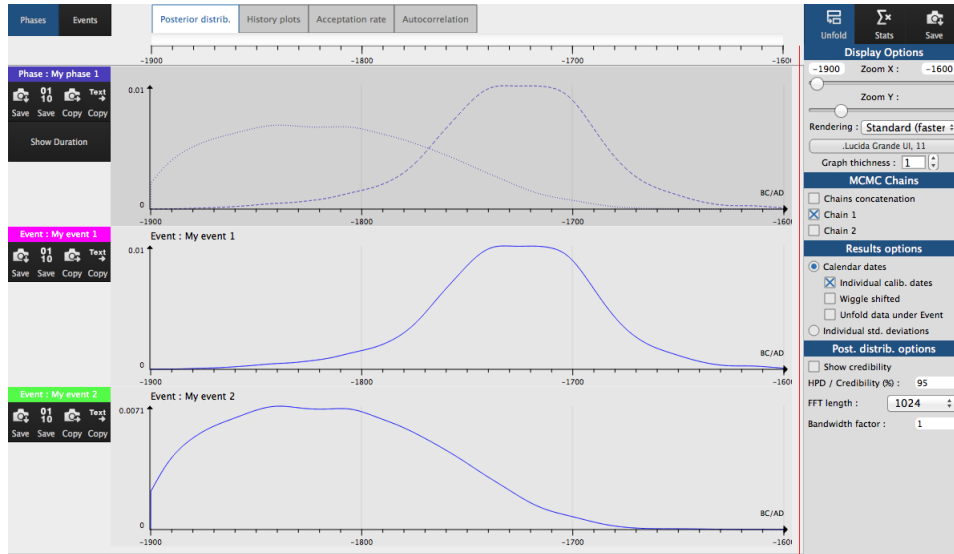
Figure 5.12 – Marginal posterior densities of the beginning and the end of a phase

When a phase is modelled, then the marginal posterior densities of the beginning (dotted line) and the end (dashed line) of the phase are displayed on the same graphic, as presented in Figure 5.12. Posterior densities are marginal densities. These densities should be interpreted parameter by parameter.

To see posterior densities of all the events included in the phase, click on the
icon. To see posterior densities of all the dates included in the events, click on "Unfold data under event", from "Results options".

The density of the duration of the phase may also be seen by clicking on "Show duration" on the left hand side of the window in the part corresponding to the associated phase.

5.2.4   Statistical results

ChronoModel gives also a list of statistical results. For each parameter, the following results are given :
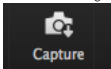
- MAP : Maximum a posteriori, the highest mode of the marginal posterior density. The MAP may be used to give an estimation of the posterior date.

- Mean : the mean of the posterior density function. The MEAN may also be used to give an estimation of the posterior date. The mean is usually preferred to MAP or Median as its statistical properties are better than the others.

- Std deviation : the standard deviation of the chain

- Q1 : the numerical value separating the lower 25% of the data from the higher 75%

- Q2 (MEDIAN): the numerical value separating the lower 50% of the data from the higher 50%. The MEDIAN may also be used to give an estimation of the posterior date.

- Q3 : the numerical value separating the lower 75% of the data from the higher 25%

- Crediblity interval : the shortest credible interval.

- HPD region : the region with the highest posterior density. It might be an interval (if the density is unimodal) or a discontinuous set (if the density is multimodal). When the density is unimodal, HPD region and CI should give similar results. However, due to numerical errors, HPD regions may be less precised than CI. Hence we recommend the use of CI when the density is unimodal.

- Acceptation rate : global acceptation rate

Those statistics may be seen under each posterior density function by clicking on
[Stats] on the right hand side of the window.

5.3 Savings

Each graph may be saved either individually, using the [Save] icon on the left hand side of the window, or globally using the [Capture] icon but from the right hand side of the window. Different type of formats may be choosen : PNG, JPG, SVG or even BMP for global graphs.

Individual graphs may also be copied and pasted using the clipboard thanks to [Copy]

icon on the left hand side of the window.

Statistics related to each parameter may also be saved thanks to [Text Copy] icon on the left hand side of the window.

It is also possible to extract the results from the Markov chains using the [01 10 Save] icon.

Using [CSV Results], on the right hand side of the window, some information may be saved : a summary of the statistical results, and the MCMC sample of the joint distribution in format BC/AD. Only the samples of events and phases parameters (beginning and end) are saved.

Several CSV files are created : a file called "events.csv" containing the MCMC samples of all events, a file called "phases.csv" containing all parameters (beginning and end) of phases if at least one phase is modelled, and a file per phase (if any) containing the MCMC samples of the parameters of the phase and all events included in it.

5.4   Summaries

On the [≡ Log] tab, a list of summaries is displayed.

Under "Model description" is a summary of the model structure along with the type of measurement, age, standard deviation, reference curve for the calibration process, MCMC method, statigraphic constraints, and so on.

Under "MCMC initialization" is the list of initial values used in the MCMC process and for each chain.

Under "Posterior distrib. resuls" are all statistical results according to all posterior distributions.

Chapter 6

Examples

6.1   Radiocarbon datings in Sennefer's tomb (Egypt)

For this illustrative example, we use data published in the article of Anita Quiles [23] regarding Sennefer's tomb. Several bouquets of flowers were found in Sennefer's tomb at Deir el-Medineh. As they were found at the entrance of the tomb, they were assumed to date the same archeological event: one of the three phases of the burial of Sennefer. Samples were extracted from different short-lived plants (leaves, twigs, etc) on each bouquet in order to ensure the consistency of the dates. All samples were radiocarbon dated.

The objective of this study is to date Sennefer's burial using ChronoModel.

6.1.1   Bouquet 1

Let's say we want to estimate the calendar date of the cut of bouquet 1. 6 samples were extracted from this bouquet and radiocarbon dated (Bouquet1.CSV contains all the datings related to Bouquet 1). The event : cut of bouquet 1 is represented by Figure 6.1.

Each radiocarbon measurement is calibrated using IntCal09 curve. No reservoir offset is taken into account. The study period is chosen to start at -2 000 and end at 0 using a step of 1 year.

For the MCMC method, the method used to generate new values at each step of the Gibbs sampler are the following ones (these are the default settings) : for the event Bouquet 1: rejection sampling using a double exponential proposal, for the distributions of the datings : Metropolis-Hastings algorithm using the posterior distribution of calibrated dates.

We start with 1 000 iterations in the Burn-in period, 1 000 iterations in each of the 100 maximum batches in the Adapt period and 10 000 iterations in the Acquire period using thinning intervals of 1. Only one chain is produced.

Figure 6.2 presents the marginal posterior densities of each date parameter (the event and the calendar dates of the calibrated measurements). In this example, 95% intervals (CI and HPD) are represented. We can see that all calendar dates seem to be contemporary. Numerical values, displayed in Figure 6.4, show that the MAP and the mean values were quite close, as well as HPD et CI intervals. The event is dated with at -1370 (mean value) associated with its 95% HPD interval [-1417; -1314]. Figure 6.3 shows the history plots (or the trace of the Markov chains) of each date parameter. During the acquisition period, all chains seem to have good mixing properties. We may assume that all chains have reach their equilibrium before the acquisition period.

Figure 6.1 – Modelling of Bouquet 1 with ChronoModel

Figure 6.5 presents the autocorrelation functions of each date parameter. We can see that all autocorrelation functions decrease exponentially and fall under the 95% confidence interval after a lag of 30 for calibrated dates and after a lag of 50 for the event. This autocorrelation between successive values may be reduced by increasing the thinning interval at 10 for example. In order to keep 10 000 observations in the acquire period, we ask for 100 000 but only 1 out of 10 values were kept for the analysis. The autocorrelation functions obtained decrease exponentially and fall under the 95% confidence interval after a lag of 6 for each parameter (see Figure Figure 6.6). However, with this new MCMC settings, all other results are similar to those already given.

Now, let's look at the individual standard deviations results. The marginal posterior densities of each individual standard deviations, presented Figure 6.7, seem to be of similar behaviour, with a mean about 50 and a standard deviation about 48 (numerical values displayed in Figure 6.8). History plots of these individual standard deviations, presented Figure 6.9, seem to have good mixing properties. Hence, the equilibrium is assumed to be reached. Each acceptance rates, presented Figure 6.10, are close to the optimal rate of 43%. And finally, each autocorrelation function, displayed in Figure 6.10, shows an exponential decrease and all values fall under the 95% interval of signification after a lag of 10.

In conclusion, the modelling of Bouquet 1 seems consistent. All individual standard deviations take values close to 50 compared to -1400 for the event. That is to say, standard deviations are rather small compared to the event's posterior mean. Hence, according to ChronoModel, all datings seem to be contemporary. Now we can
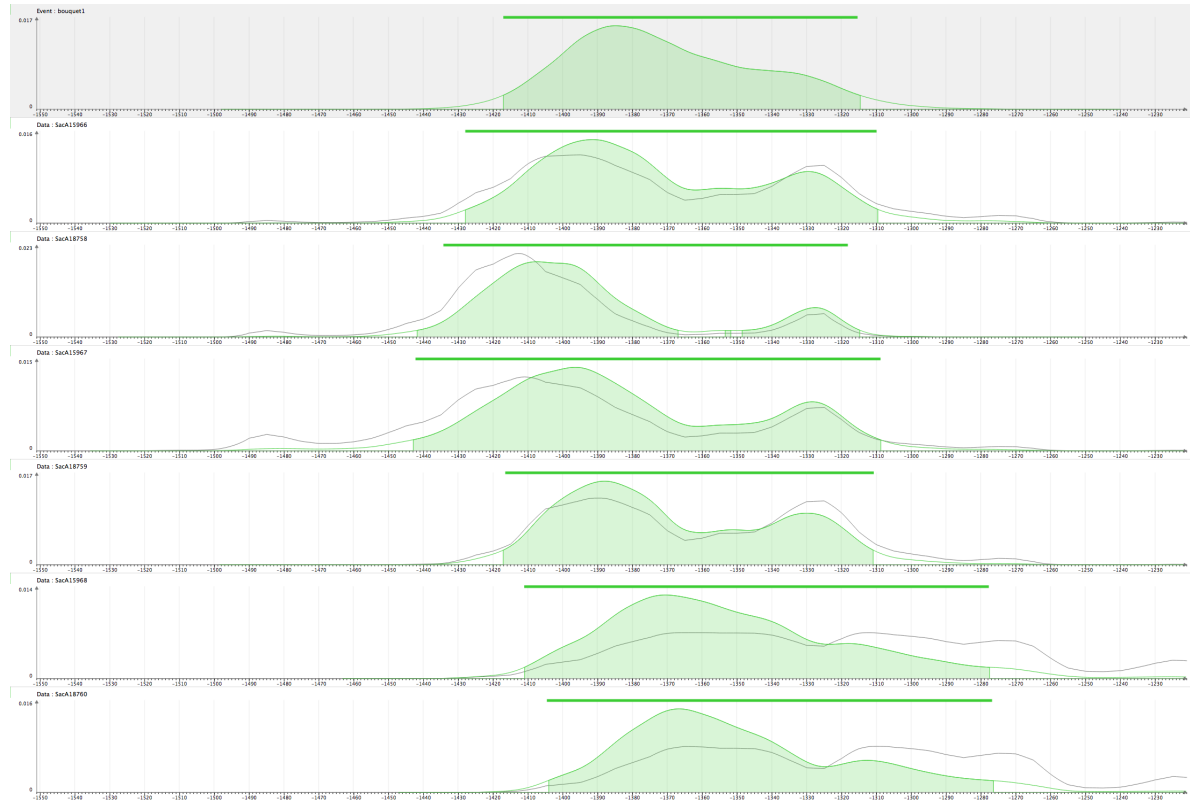
Figure 6.2 – Marginal posterior densities related to the modelling of Bouquet 1 . The dark lines correspond to distribution of calibrated dates, the green lines correspond to posterior density functions. Highest posterior density (HPD) intervals are represented by the green shadow area under the green lines. Credibility intervals are represented by thick lines drawn above the green lines.
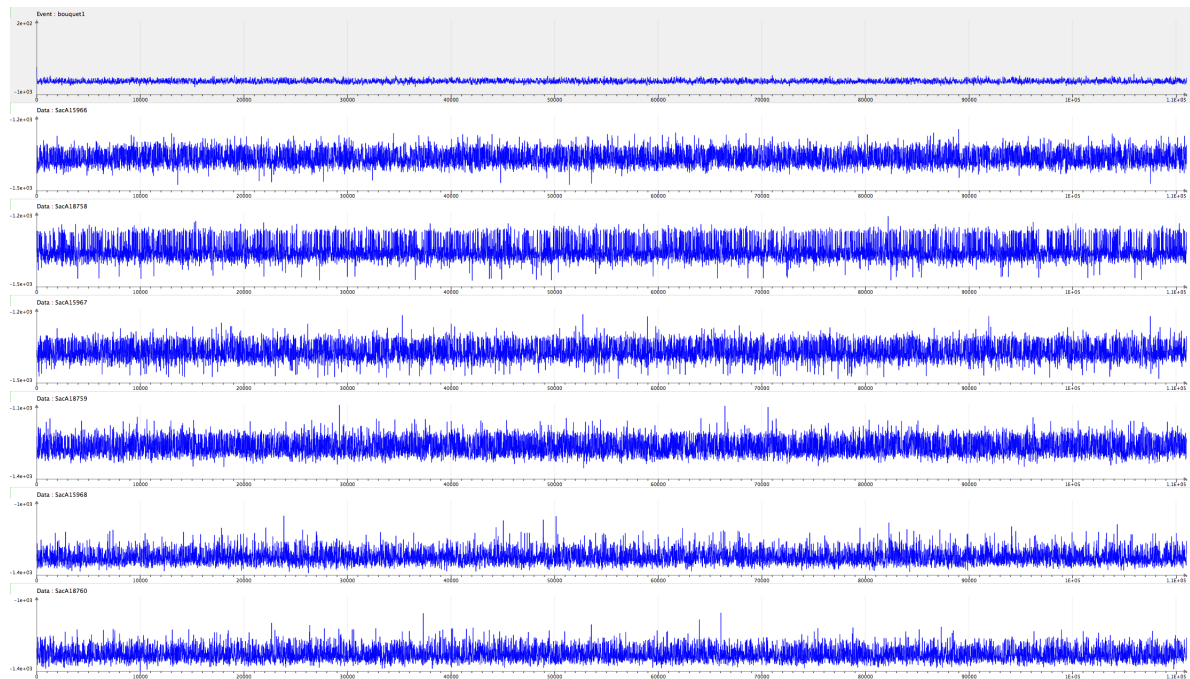
Figure 6.3 – History plots of date parameters related to the modelling of Bouquet 1

Event : bouquet1
Mode : −1385   Mean : −1370   Std deviation : 28
Q1 : −1390.43   Q2 (Median) : −1374.17   Q3 : −1350.88
HPD Intervals (95%) : [−1417, −1314]
Credibility Interval (95.00%) : [−1417, −1315]
Taux d'acceptation global : 100.0% (AR : proposal = Double−Exponential)

Data : SacA15966
Mode : −1392   Mean : −1372   Std deviation : 34
Q1 : −1397.48   Q2 (Median) : −1379.09   Q3 : −1342.56
HPD Intervals (95%) : [−1428, −1309]
Credibility Interval (95.00%) : [−1428, −1310]
Taux d'acceptation global : 62.0% (MH : proposal = distribution of calibrated date)

Data : SacA18758
Mode : −1407   Mean : −1392   Std deviation : 33
Q1 : −1413.62   Q2 (Median) : −1399.97   Q3 : −1382.39
HPD Intervals (95%) : [−1442, −1367] [−1353, −1352] [−1349, −1315]
Credibility Interval (95.00%) : [−1434, −1319]
Taux d'acceptation global : 66.8% (MH : proposal = distribution of calibrated date)

Data : SacA15967
Mode : −1396   Mean : −1382   Std deviation : 38
Q1 : −1408.09   Q2 (Median) : −1389.07   Q3 : −1351.16
HPD Intervals (95%) : [−1443, −1308]
Credibility Interval (95.00%) : [−1442, −1309]
Taux d'acceptation global : 59.7% (MH : proposal = distribution of calibrated date)

Data : SacA18759
Mode : −1388   Mean : −1368   Std deviation : 31
Q1 : −1392.26   Q2 (Median) : −1375.31   Q3 : −1339.86
HPD Intervals (95%) : [−1417, −1311]
Credibility Interval (95.00%) : [−1416, −1311]
Taux d'acceptation global : 63.5% (MH : proposal = distribution of calibrated date)

Data : SacA15968
Mode : −1371   Mean : −1351   Std deviation : 36
Q1 : −1376.55   Q2 (Median) : −1357.24   Q3 : −1330.67
HPD Intervals (95%) : [−1411, −1277]
Credibility Interval (95.00%) : [−1411, −1278]
Taux d'acceptation global : 51.3% (MH : proposal = distribution of calibrated date)

Data : SacA18760
Mode : −1366   Mean : −1348   Std deviation : 35
Q1 : −1372.42   Q2 (Median) : −1355.22   Q3 : −1328.28
HPD Intervals (95%) : [−1404, −1276]
Credibility Interval (95.00%) : [−1404, −1277]
Taux d'acceptation global : 52.9% (MH : proposal = distribution of calibrated date)

Figure 6.4 – Numerical values of each date parameter related to the modelling of Bouquet 1
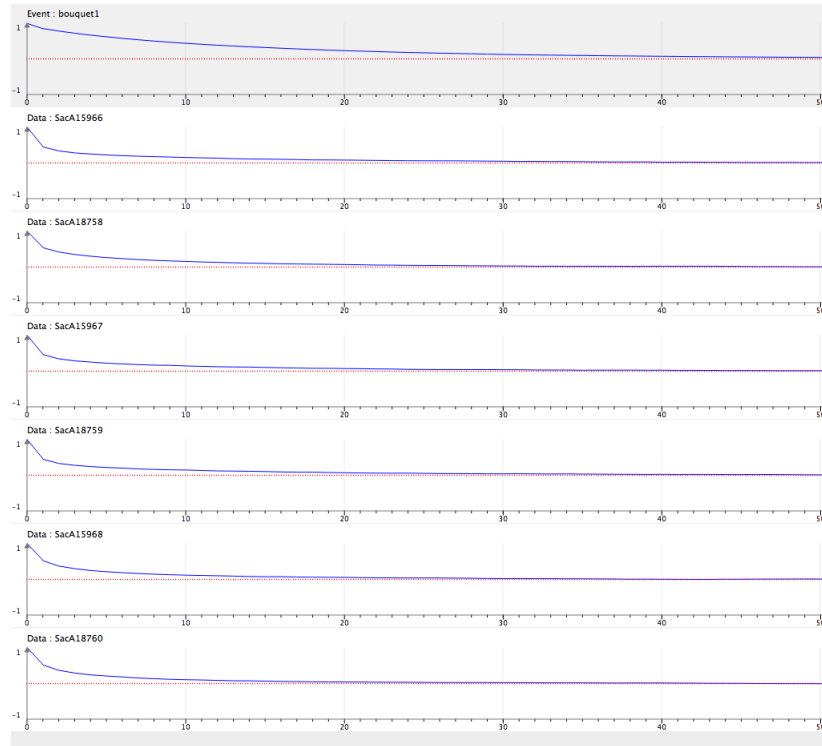
Figure 6.5 – Autocorrelation functions related to the modelling of Bouquet 1 (Thinning interval = 1)
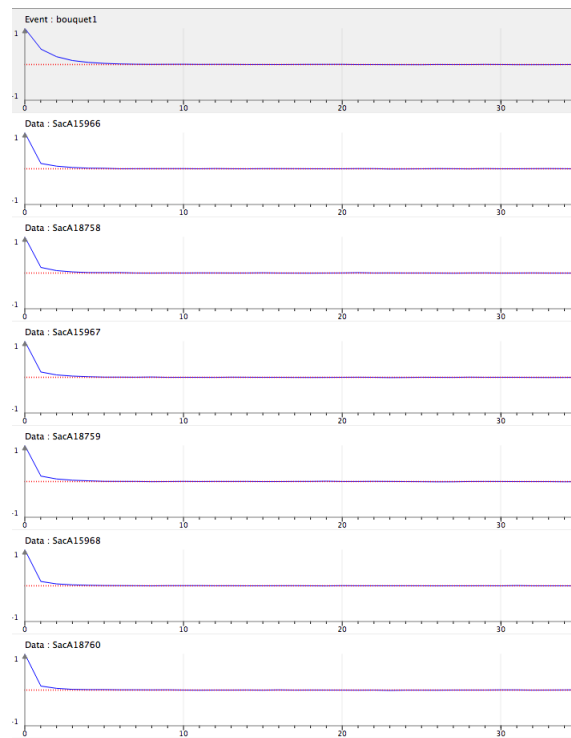
Figure 6.6 – Autocorrelation functions related to the modelling of Bouquet 1 (Thinning interval = 10)

Figure 6.7 – Marginal posterior densities of indivudal standard deviations related to the modelling of Bouquet 1

draw conclusions about the calendar date of the event. The event Bouquet 1 may be dated at -1370 (mean value) with a 95% interval of [-1417; -1314] (HPD interval).

Data : SacA15966
Mode : 27   Mean : 45   Std deviation : 47
Q1 : 20.24   Q2 (Median) : 33.89   Q3 : 54.71
HPD Intervals (95%) : [3, 124]
Credibility Interval (95.00%) : [0, 110]

Data : SacA18758
Mode : 27   Mean : 50   Std deviation : 48
Q1 : 23.58   Q2 (Median) : 38.74   Q3 : 61.27
HPD Intervals (95%) : [1, 127]
Credibility Interval (95.00%) : [0, 125]

Data : SacA15967
Mode : 24   Mean : 48   Std deviation : 48
Q1 : 21.71   Q2 (Median) : 36.32   Q3 : 58.04
HPD Intervals (95%) : [2, 120]
Credibility Interval (95.00%) : [1, 118]

Data : SacA18759
Mode : 22   Mean : 44   Std deviation : 47
Q1 : 19.78   Q2 (Median) : 32.82   Q3 : 53.41
HPD Intervals (95%) : [2, 112]
Credibility Interval (95.00%) : [1, 110]

Data : SacA15968
Mode : 25   Mean : 49   Std deviation : 47
Q1 : 21.85   Q2 (Median) : 36.58   Q3 : 59.31
HPD Intervals (95%) : [1, 128]
Credibility Interval (95.00%) : [1, 126]

Data : SacA18760
Mode : 24   Mean : 49   Std deviation : 47
Q1 : 22.35   Q2 (Median) : 37.41   Q3 : 60.42
HPD Intervals (95%) : [0, 127]
Credibility Interval (95.00%) : [1, 126]

Figure 6.8 – Numerical values of indivudal standard deviations related to the modelling of Bouquet 1
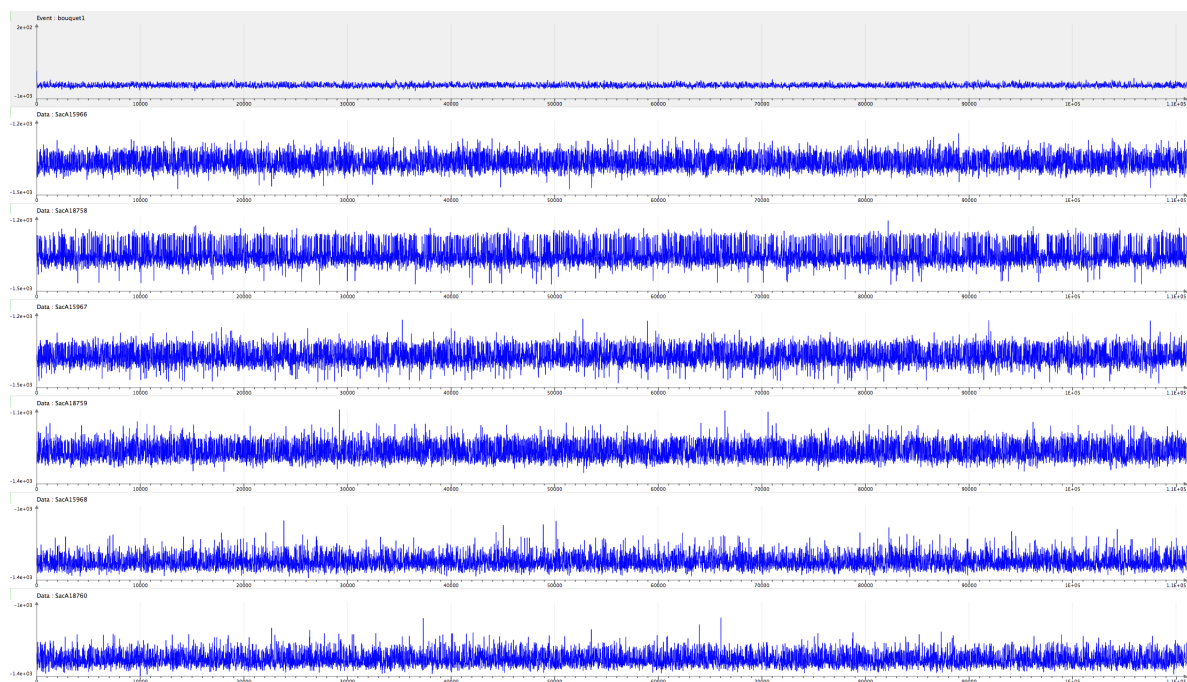
Figure 6.9 – History plots of indivudal standard deviations related to the modelling of Bouquet 1
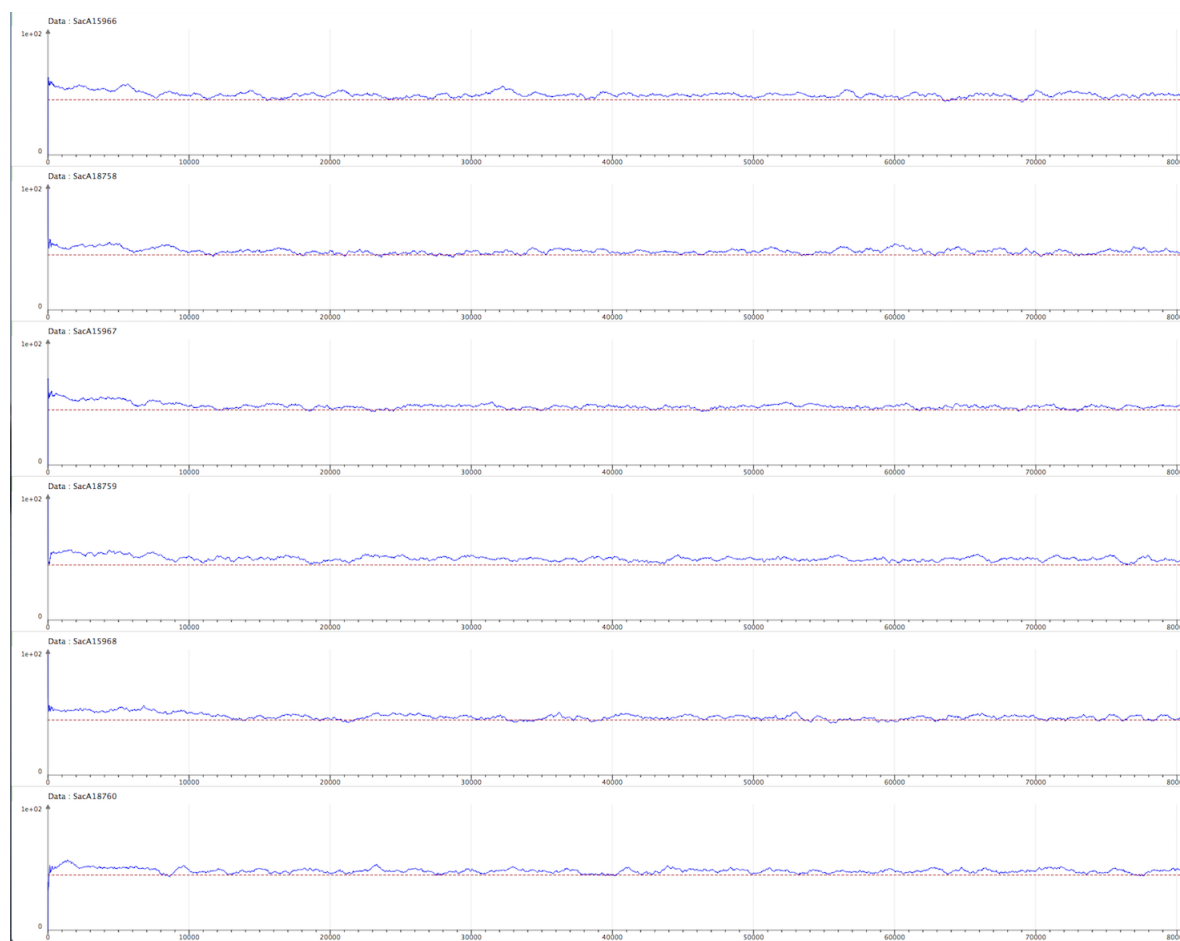
Figure 6.10 – Acceptation rates of indivudal standard deviations related to the modelling of Bouquet 1

Figure 6.11 – Autocorrelation functions of indivudal standard deviations related to the modelling of Bouquet 1

6.1.2 Bouquet 2

Now, let's say we want to estimate the calendar date of bouquet 2. 8 samples were extracted from Bouquet 2 and radiocarbon dated. The modelling of this bouquet by ChronoModel is represented by Figure 6.12.
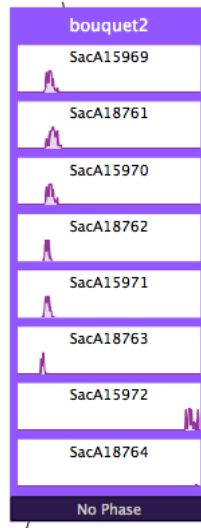
Figure 6.12 – Modelling of Bouquet 2 with ChronoModel

Each radiocarbon measurement is calibrated using IntCal09 curve. No reservoir offset is taken into account. The study period is chosen to start at -2000 and end at 2000 using a step of 1 year. This study period is chosen so that every distribution of the calibrated date is included in this study period.

The default MCMC methods are used. We use 1 000 iterations in the Burn-in period, 1 000 iterations in each of the 100 maximum batches in the Adapt period and 100 000 iterations in the Acquire period using thinning intervals of 10.

The marginal posterior densities, presented in Figure 6.13, are of two sorts. Althought, the first 6 datings seem to be contemporary, the two last ones seem to be some kind of outliers. Indeed their density function takes values about 1600 whereas the other densities take values between -1500 and -1200. All history plots have good mixing properties and autocorrelation functions are correct (results not shown). Looking at individual standard deviations, the marginal posterior densities, displayed Figure 6.15, show three distinct standard deviations. The first 5 samples are associated with a standard deviation density function that takes small values, with mean values about 50. The next sample's individual standard deviation has a mean posterior density at 100. And the 2 last samples are associated with individual standard deviations with a mean higher than 2 000 (See Figure 6.16 for numerical values). Hence these two last datings give a piece of information that has a reduced importance in the construction

of the posterior density function of the event Bouquet 2. All individual standard
deviations have a history plot with good properties, an acceptance rate about 43% or
higher and a correct autocorrelation function (results not shown).

As a conclusion, the first 6 samples seem to be contemporary but the two last ones
seem to be some kind of outliers. Then the event Bouquet 2 may be dated at -1344
(mean value) with a 95% HPD interval [-1405;-1278].

The example shows that the modelling is robust to outliers. Indeed even if two
outliers were included in the analysis, the datation of the event was not affected
by them. ChronoModel do not need any particular manipulation of outliers before
analysing the datings. Indeed there is no need to withdraw any datation or to use a
special treatment to them.

Figure 6.13 – Marginal posterior densities related to the modelling of Bouquet 2

Event : bouquet2
Mode : −1351   Mean : −1344   Std deviation : 33
Q1 : −1367.71   Q2 (Median) : −1346.16   Q3 : −1321.81
HPD Intervals (95%) : [−1405, −1278]
Credibility Interval (95.00%) : [−1404, −1280]
Taux d'acceptation global : 100.0% (AR : proposal = Double−Exponential)

Data : SacA15969
Mode : −1357   Mean : −1329   Std deviation : 42
Q1 : −1360.57   Q2 (Median) : −1336.24   Q3 : −1300.47
HPD Intervals (95%) : [−1404, −1248]
Credibility Interval (95.00%) : [−1405, −1252]
Taux d'acceptation global : 50.9% (MH : proposal = distribution of calibrated date)

Data : SacA18761
Mode : −1358   Mean : −1307   Std deviation : 58
Q1 : −1353.98   Q2 (Median) : −1310.43   Q3 : −1272.46
HPD Intervals (95%) : [−1400, −1192]
Credibility Interval (95.00%) : [−1393, −1192]
Taux d'acceptation global : 37.2% (MH : proposal = distribution of calibrated date)

Data : SacA15970
Mode : −1358   Mean : −1326   Std deviation : 45
Q1 : −1360.19   Q2 (Median) : −1334.55   Q3 : −1296.54
HPD Intervals (95%) : [−1404, −1243] [−1230, −1226]
Credibility Interval (95.00%) : [−1404, −1243]
Taux d'acceptation global : 49.2% (MH : proposal = distribution of calibrated date)

Data : SacA18762
Mode : −1329   Mean : −1353   Std deviation : 33
Q1 : −1380.57   Q2 (Median) : −1349.86   Q3 : −1327.17
HPD Intervals (95%) : [−1415, −1296]
Credibility Interval (95.00%) : [−1412, −1296]
Taux d'acceptation global : 66.0% (MH : proposal = distribution of calibrated date)

Data : SacA15971
Mode : −1331   Mean : −1348   Std deviation : 36
Q1 : −1375.82   Q2 (Median) : −1347.07   Q3 : −1322.72
HPD Intervals (95%) : [−1415, −1278]
Credibility Interval (95.00%) : [−1413, −1279]
Taux d'acceptation global : 63.0% (MH : proposal = distribution of calibrated date)

Data : SacA18763
Mode : −1429   Mean : −1435   Std deviation : 28
Q1 : −1448.29   Q2 (Median) : −1432.50   Q3 : −1419.92
HPD Intervals (95%) : [−1494, −1394]
Credibility Interval (95.00%) : [−1493, −1395]
Taux d'acceptation global : 70.6% (MH : proposal = distribution of calibrated date)

Data : SacA15972
Mode : 1755   Mean : 1786   Std deviation : 85
Q1 : 1733.85   Q2 (Median) : 1768.14   Q3 : 1840.74
HPD Intervals (95%) : [1653, 1878] [1908, 1961]
Credibility Interval (95.00%) : [1668, 1945]
Taux d'acceptation global : 95.5% (MH : proposal = distribution of calibrated date)

Data : SacA18764
Mode : 1954   Mean : 1905   Std deviation : 68
Q1 : 1896.88   Q2 (Median) : 1906.68   Q3 : 1954.43
HPD Intervals (95%) : [1695, 1730] [1816, 1841] [1869, 1981]
Credibility Interval (95.00%) : [1715, 1955]
Taux d'acceptation global : 97.2% (MH : proposal = distribution of calibrated date)
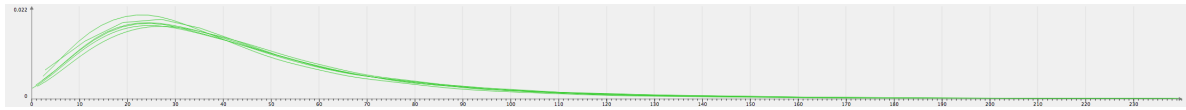
Figure 6.14 – Numerical values related to the modelling of Bouquet 2

Figure 6.15 – Marginal posterior densities of indivudal standard deviations related to the modelling of Bouquet 2
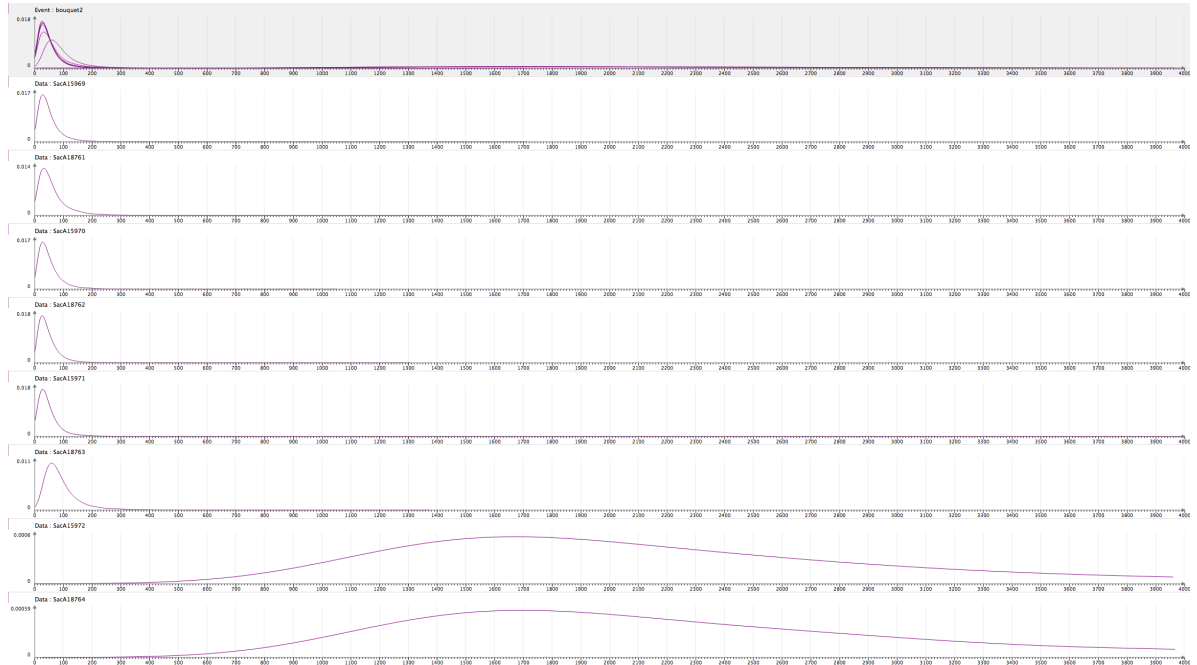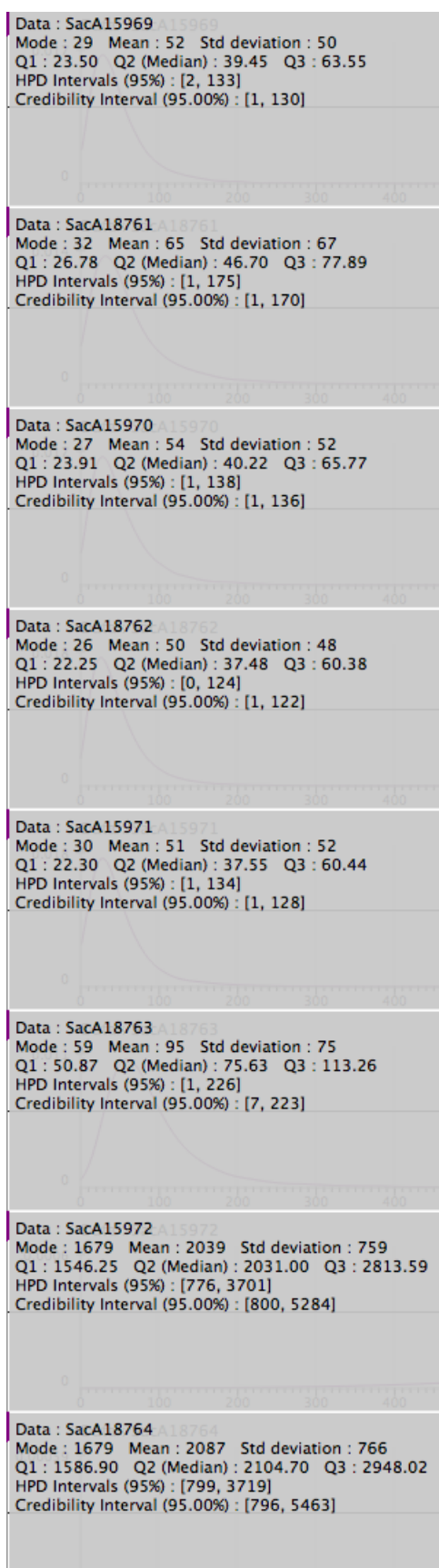
Data : SacA15969
Mode : 29   Mean : 52   Std deviation : 50
Q1 : 23.50   Q2 (Median) : 39.45   Q3 : 63.55
HPD Intervals (95%) : [2, 133]
Credibility Interval (95.00%) : [1, 130]

Data : SacA18761
Mode : 32   Mean : 65   Std deviation : 67
Q1 : 26.78   Q2 (Median) : 46.70   Q3 : 77.89
HPD Intervals (95%) : [1, 175]
Credibility Interval (95.00%) : [1, 170]

Data : SacA15970
Mode : 27   Mean : 54   Std deviation : 52
Q1 : 23.91   Q2 (Median) : 40.22   Q3 : 65.77
HPD Intervals (95%) : [1, 138]
Credibility Interval (95.00%) : [1, 136]

Data : SacA18762
Mode : 26   Mean : 50   Std deviation : 48
Q1 : 22.25   Q2 (Median) : 37.48   Q3 : 60.38
HPD Intervals (95%) : [0, 124]
Credibility Interval (95.00%) : [1, 122]

Data : SacA15971
Mode : 30   Mean : 51   Std deviation : 52
Q1 : 22.30   Q2 (Median) : 37.55   Q3 : 60.44
HPD Intervals (95%) : [1, 134]
Credibility Interval (95.00%) : [1, 128]

Data : SacA18763
Mode : 59   Mean : 95   Std deviation : 75
Q1 : 50.87   Q2 (Median) : 75.63   Q3 : 113.26
HPD Intervals (95%) : [1, 226]
Credibility Interval (95.00%) : [7, 223]

Data : SacA15972
Mode : 1679   Mean : 2039   Std deviation : 759
Q1 : 1546.25   Q2 (Median) : 2031.00   Q3 : 2813.59
HPD Intervals (95%) : [776, 3701]
Credibility Interval (95.00%) : [800, 5284]

Data : SacA18764
Mode : 1679   Mean : 2087   Std deviation : 766
Q1 : 1586.90   Q2 (Median) : 2104.70   Q3 : 2948.02
HPD Intervals (95%) : [799, 3719]
Credibility Interval (95.00%) : [796, 5463]

Figure 6.16 – Numerical values of indivudal standard deviations related to the modelling of Bouquet 2

6.1.3   Modelling of bouquets 1 and 2 simultaneously

Now let's say we want to date Bouquet 1 and Bouquet 2 simultaneously. See bouquets12.chr.

The study period was chosen to start at -2000 and end at 2000 using a step of 1 year. All other parameters are those used for the modelling of Bouquet 1 and of Bouquet 2.

Three different modellings are compared here. In the first modelling, no further constraints are included (See Figure 6.17). In the two next modellings, two bounds are introduced to constrain the beginning and the end of the burial of Sennefer (See Figure 6.18). Indeed, the burial of Sennefer is assumed to have happened between the accession date of Tutankamun and the accession date of Horemheb (See [23]). These accesssion dates are considered as bounds in ChronoModel. There are two different ways to introduce a bound. A bound may be fixed (Accession date of Tutankamun -1356 and Accession date of Horemheb -1312) or a bound may have a uniform distribution (Accession date of Tutankamun : uniform on [-1360; -1352], Accession date of Horemheb : uniform on [-1316;-1308]).

Figure 6.19 displays the marginal posterior densities of both bouquets when the modelling does not include bounds. Figure 6.20 displays the marginal posterior densities of both bouquets when the modelling includes bounds, using fixed bounds (figure on the left hand side) and using uniform bounds (figure on the right hand side). From these results, we can see that the introduction of bounds helps restrain the posterior densities and the HPD interval of both events. However, using fixed bounds or bounds having a uniform distribution with a small period (8 years) lead to similar results. Numerical values are presented in Table 6.1.
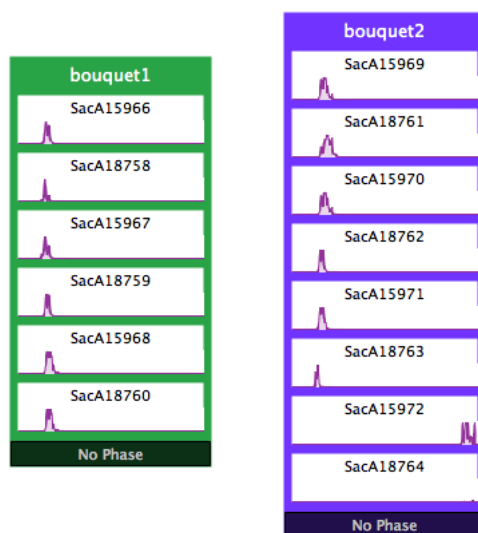
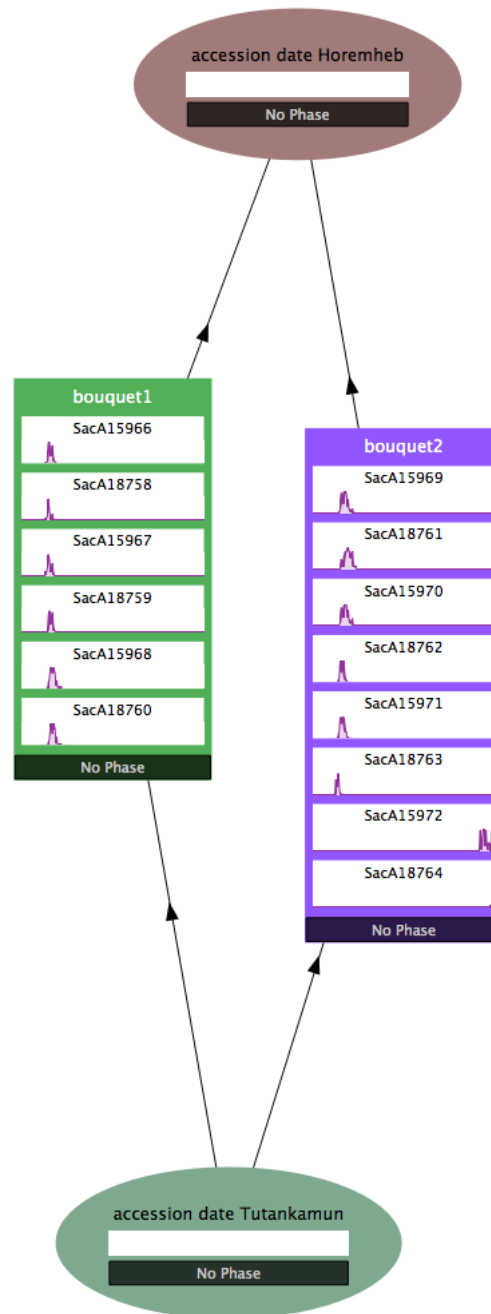Figure 6.17 – Modelling of Bouquets 1 and 2 without bounds

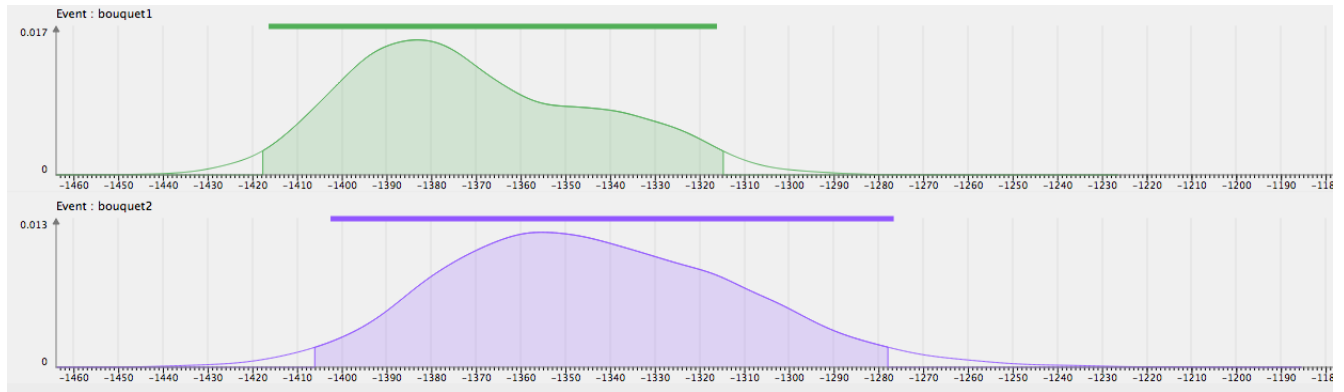Figure 6.18 – Modelling of Bouquets 1 and 2 including bounds

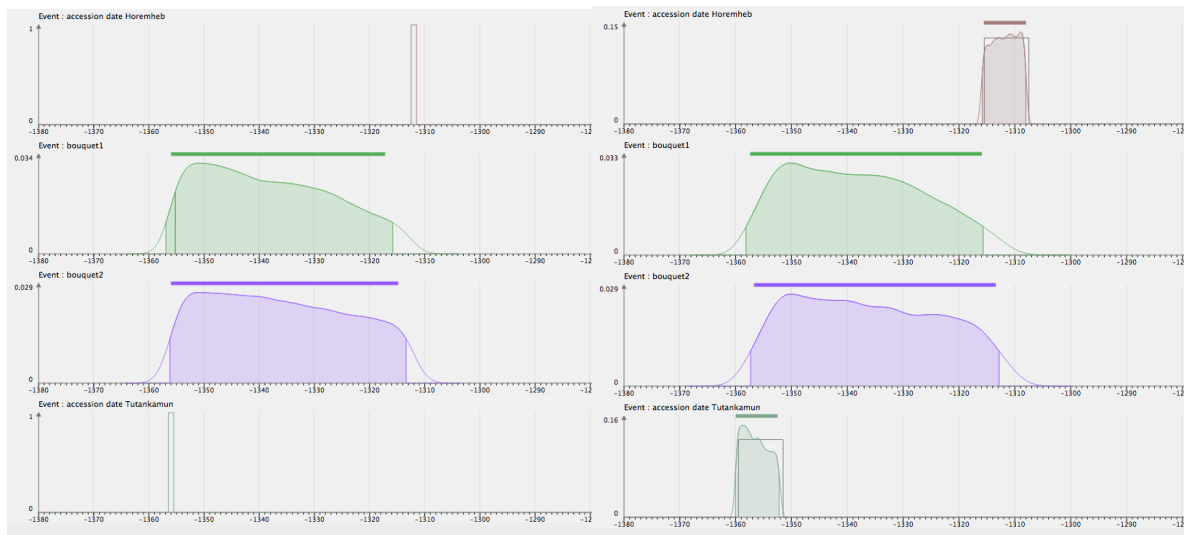Figure 6.19 – Marginal posterior densities related to the modelling of Bouquets 1 and 2 without bounds

Figure 6.20 – Marginal posterior densities related to the modelling of Bouquets 1 and 2 with bounds (fixed bounds on the left handside figure, with uniform bounds on the right handside figure)

	modelling of Bouquets 1 and 2		
	Without bounds	With fixed bounds	With uniform bounds
Event Bouquet 1			
Mean	-1371	-1338	-1338
HPD region	[-1418; -1315]	[-1357; -1316]	[-1358; -1316]
Event Bouquet 2			
Mean	-1344	-1336	-1336
HPD region	[-1406; -1278]	[-1356; -1313]	[-1357; -1313]

Table 6.1 – Numerical values related to the modellings of Bouquets 1 and 2

6.1.4 Modelling of the phase including bouquets 1 and 2 - Estimation of the duration of the phase

Let's say that now we want to estimate the duration of the phase including both bouquets. This phase might be seen as the duration of Sennefer's burial. Three different modellings are possible and presented here in turn. For all of them, the study period is chosen to start at -2000 and end at 2000 using a step of 1 year. All other parameters are those used for the modelling of Bouquet 1 and of Bouquet 2.

6.1.4.1 Phase without constraints

The modelling of the phase including both bouquets is displayed in Figure 6.21. In this modelling, no further constraints are included. The phase's duration is kept unknown.
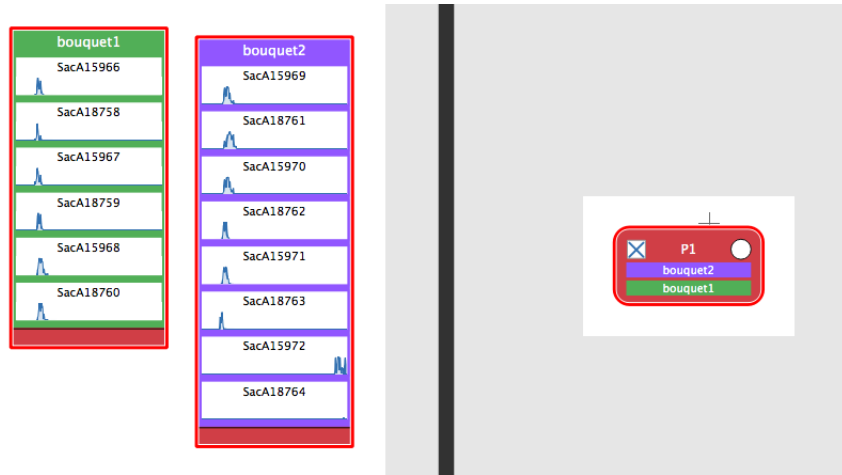

Figure 6.21 – Modelling of the phase including Bouquets 1 and 2 including a phase

Figure 6.22 displays the marginal posterior densities of both events and those of the beginning and the end of the phase. Statistical results regarding Bouquet 1 and Bouquet 2 are unchanged by the introduction of the phase, the results are similar to those presented in the last section when no bounds were introduced (See Table 6.1). In addition, this modelling allows to estimate the mean duration of the phase (35) as well as its credibility interval ([0, 101]).
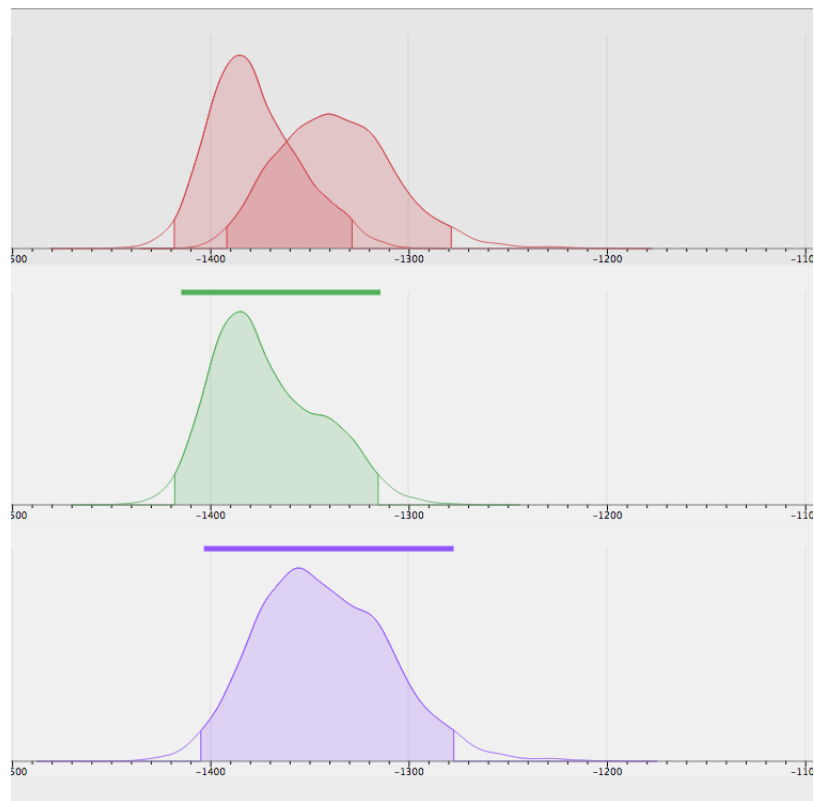
Figure 6.22 – Marginal posterior densities related to the modelling of Bouquets 1 and 2 including a phase and without bounds. The densities of the minimum and the maximum are drawn in red, the density of Bouquet 1 is drawn in green, the density of Bouquet 2 is drawn in purple.

6.1.4.2 Phase with bounds

Now, let's include prior information about the accession dates of Tutankamun and Horemheb. We include the two fixed bounds as detailled in section 6.1.3. Here two modellings are possible using ChronoModel. Figure 6.23 represents the first modelling in which bounds constrain events. Figure 6.24 represents the second modelling in which bounds are included in separated phases, using one phase for each bound, and stratigraphic constraints are placed between phases. However, these two modellings give similar results. The marginal posterior densities of all parameters are presented in Figure 6.25. The mean duration of the bouquets' phase is 12 years associated with a credibility interval of [0, 33] that is smaller than the one estimated without bounds.
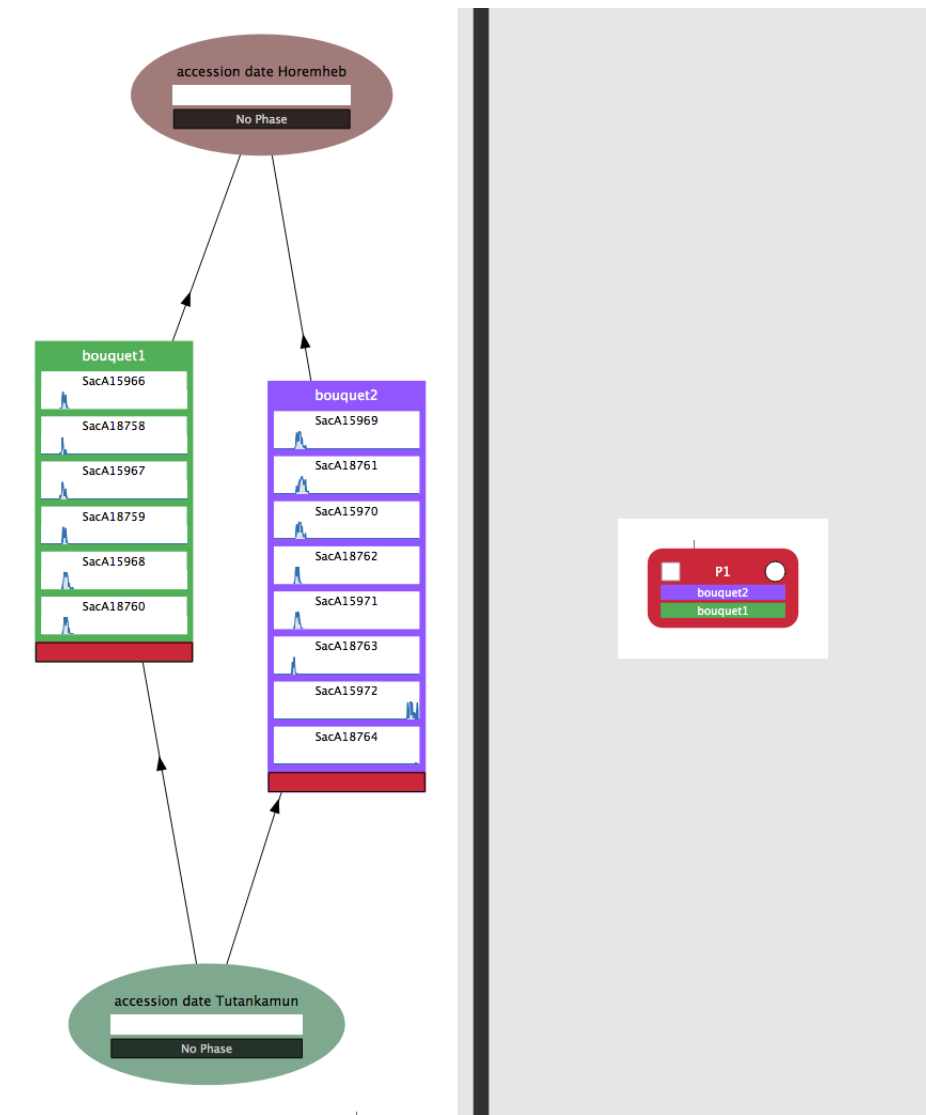

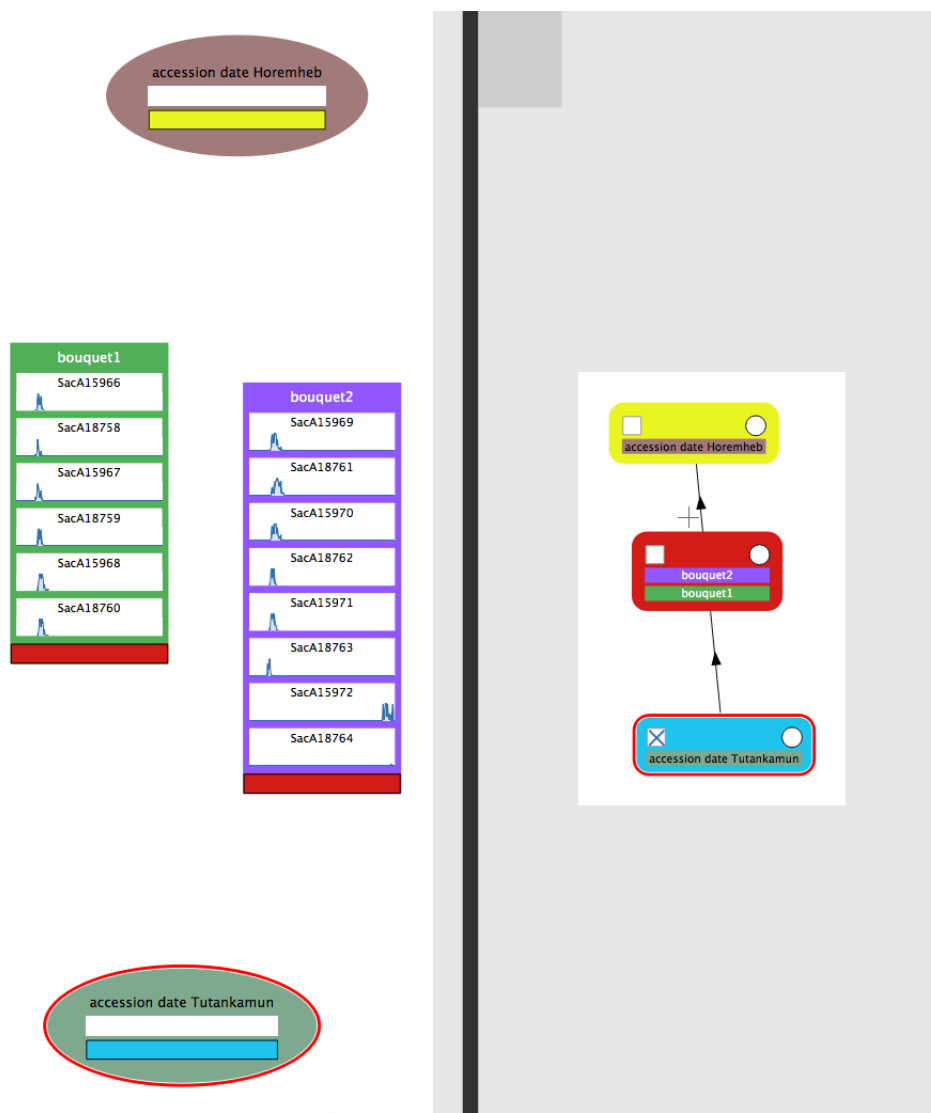Figure 6.23 – First modelling 1 of Bouquets 1 and 2 including a phase and bounds

Figure 6.24 – Second modelling of Bouquets 1 and 2 including a phase and bounds
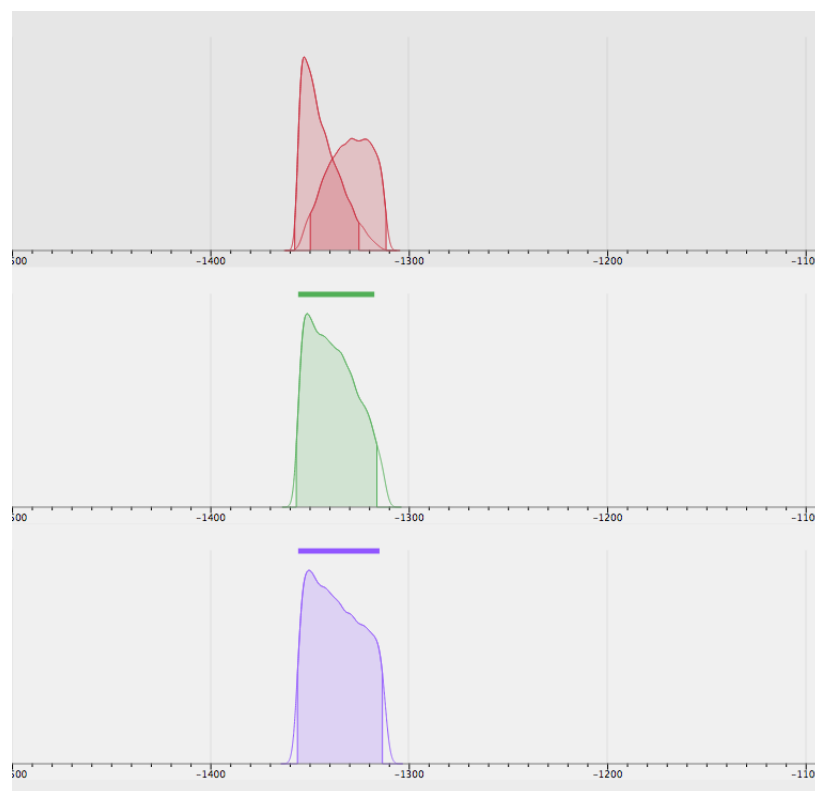
Figure 6.25 – Marginal posterior densities related to the modelling of Bouquets 1 and 2 including a phase and bounds. The densities of the minimum and the maximum are drawn in red, the density of Bouquet 1 is drawn in green, the density of Bouquet 2 is drawn in purple.

6.1.4.3   Phase with fixed duration

The phase of Sennefer's burial is assumed to have happened between the accession date of Tutankamun (-1356) and the accession date of Horemheb (-1312). Hence, the duration of the phase is smaller than 44 years.

In this last modelling, no bounds are included but the maximum duration of the phase is fixed at 44 years, the delay between the accession date of Tutankamun and the accession date of Horemheb.
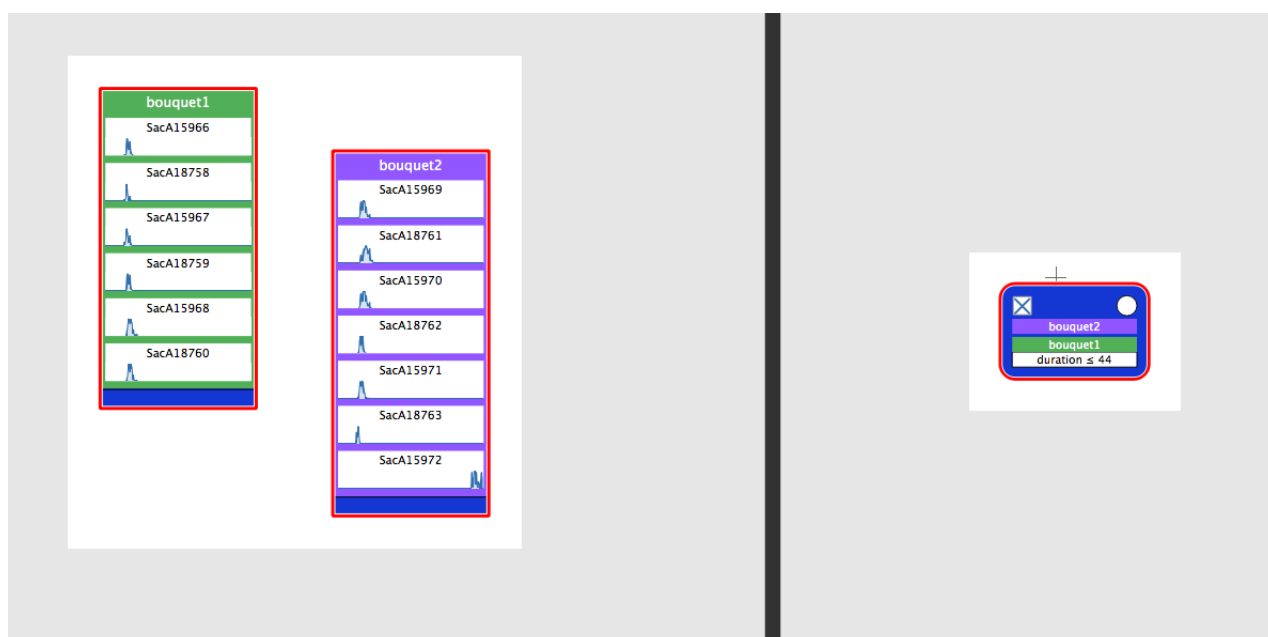

Figure 6.26 – Modelling of the phase including Bouquets 1 and 2 having a fixed duration

This modelling leads to the following results:

Duration of the phase		
Mean	20	
Credibility interval	[0; 41]	
Event	**Bouquet 1**	**Bouquet 2**
Mean	-1364	-1357
HPD region	[-1408; -1317]	[-1403; -1306]

Table 6.2 – Numerical values related to the phase including Bouquets 1 and 2 and having a fixed duration

In that case, although the duration of the phase is reduced compared to the modelling without bounds, however, the estimation of the calendar dates of bouquet 1 and bouquet 2 are less precise than those estimated with the modelling with bounds.
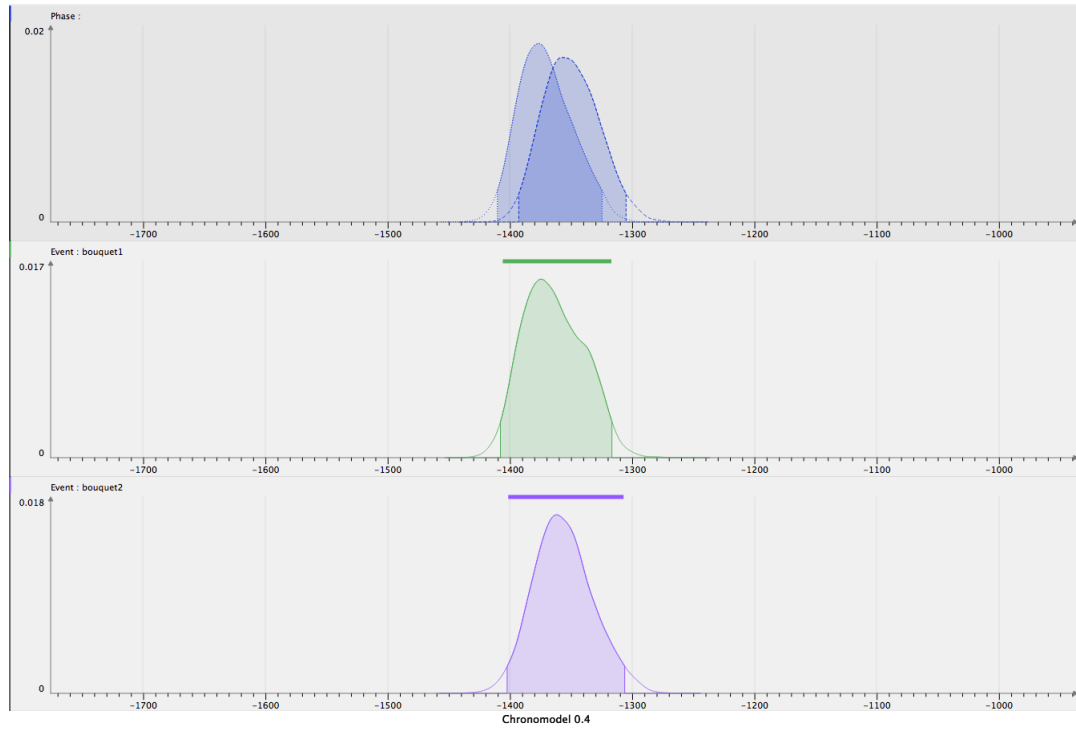
Figure 6.27 – Marginal posterior densities related to the modelling of the phase including Bouquets 1 and 2 having a fixed duration.. The densities of the minimum and the maximum are drawn in red, the density of Bouquet 1 is drawn in green, the density of Bouquet 2 is drawn in purple.

In fact, the prior information about the maximum duration of a phase gives less information than two constraining bounds.

6.2   Toy scenario

In this example, we present a fictitious archaeological excavation with stratigraphy on several structures. This toy scenario is used to give an idea different modellings that may be designed with ChronoModel. Figure 6.28 shows the stratigraphy of the archaeological field. For all modellings, the study period is -2 000 to 2 000.
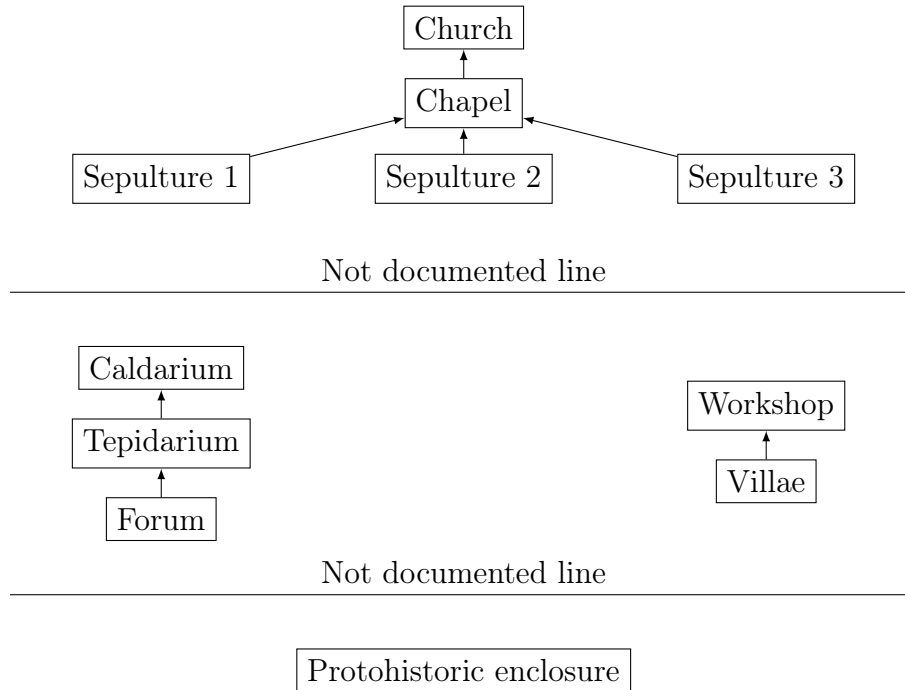

Figure 6.28 – Field model

Table 6.3 listes all the measurements and the corresponding structures. Each structure will be considered as an event in ChronoModel, and so there are only one radiocarbon specimen for each event.

Structure	Ref. 14C	Value and standard error
Enclos	Pr1	2540 ± 50
Villae	GR5	1850 ± 30
Atelier	GR4	1735 ± 30
Forum	GR3	1764 ± 30
Tepidarium	GR2	1760 ± 30
Caldarium	GR1	1734 ± 30
Sepulture 1	M3	1350 ± 35
Sepulture 2	M4	1390 ± 30
Sepulture 3	M5	1370 ± 50
Chapelle	M2	1180 ± 30
Eglise	M1	950 ± 35

Table 6.3 – Radiocarbon datations

6.2.1 A sequential model without phases

In this first modelling, no phase is used. Each measurement is included in one event and stratigraphic constraints between events are added to the modelling. Figure 6.29 shows the first modelling of this toy scenario by ChronoModel.
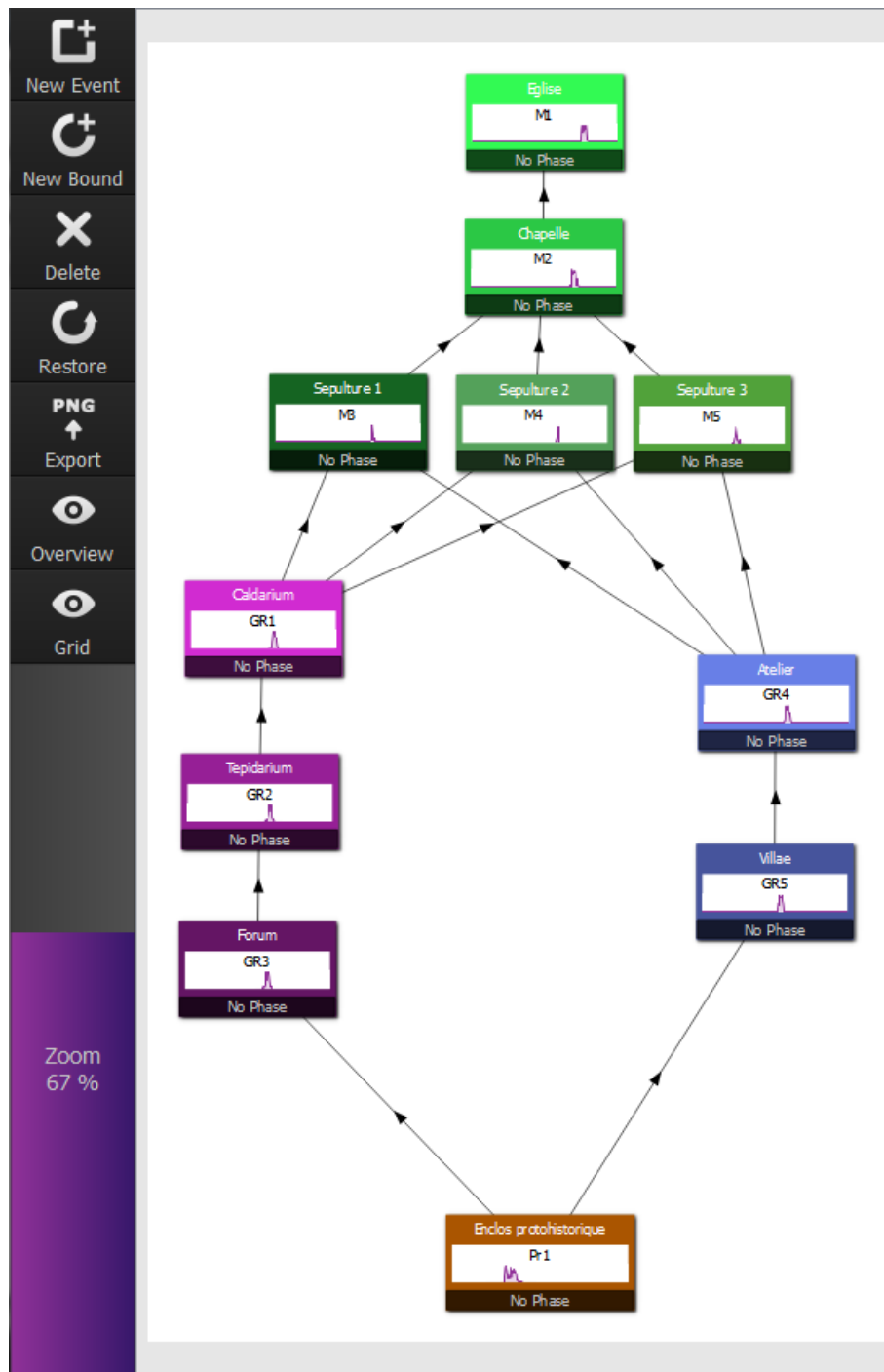

Figure 6.29 – Design of the sequential model by ChronoModel

Markov chains were checked, convergence was reached before the Acquire period, au-

tocorrelation functions felt under the 95% confidence interval after a lag of 2 and acceptance rates were all about 43%. Figure 6.30 shows the marginal posterior densities of the events of this sequential model.
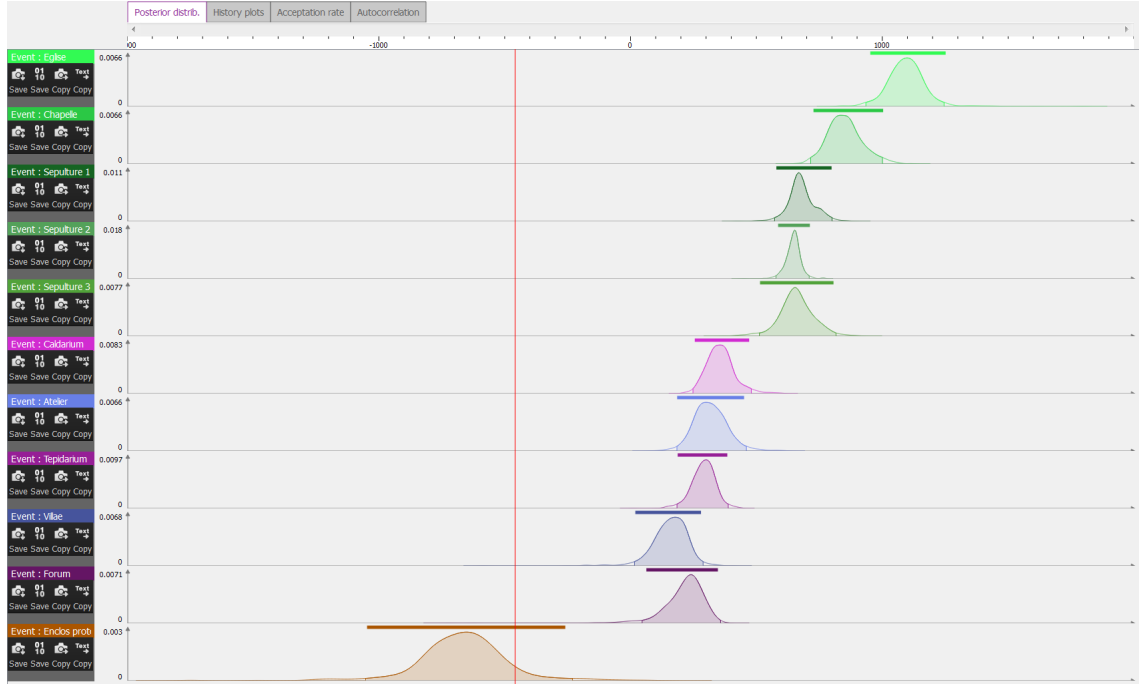

Figure 6.30 – Marginal posterior densities obtained from the sequential model using ChronoModel

6.2.2 Grouping events into phases

Another equivalent way to build the chronology is to introduce phases. In our example, we can see 4 sequences nested in 4 phases. Each phase corresponds to a group structures. Figure 6.31 displays the design of this second model including phases. For this modelling, constraints between events were only kept within phases. Constraints between events of different phases were replaced by constraints between phases. However, both modellings lead to the same results. This second modelling is also a way to simplify the design of the model limiting the number of constraints between events. Figure 6.32 shows the marginal posterior densities computed with the second modelling. These denstiities are similar to those obtained from the first modelling. However, now information about phases : the density if the beginning and the end of each phase may also be seen.
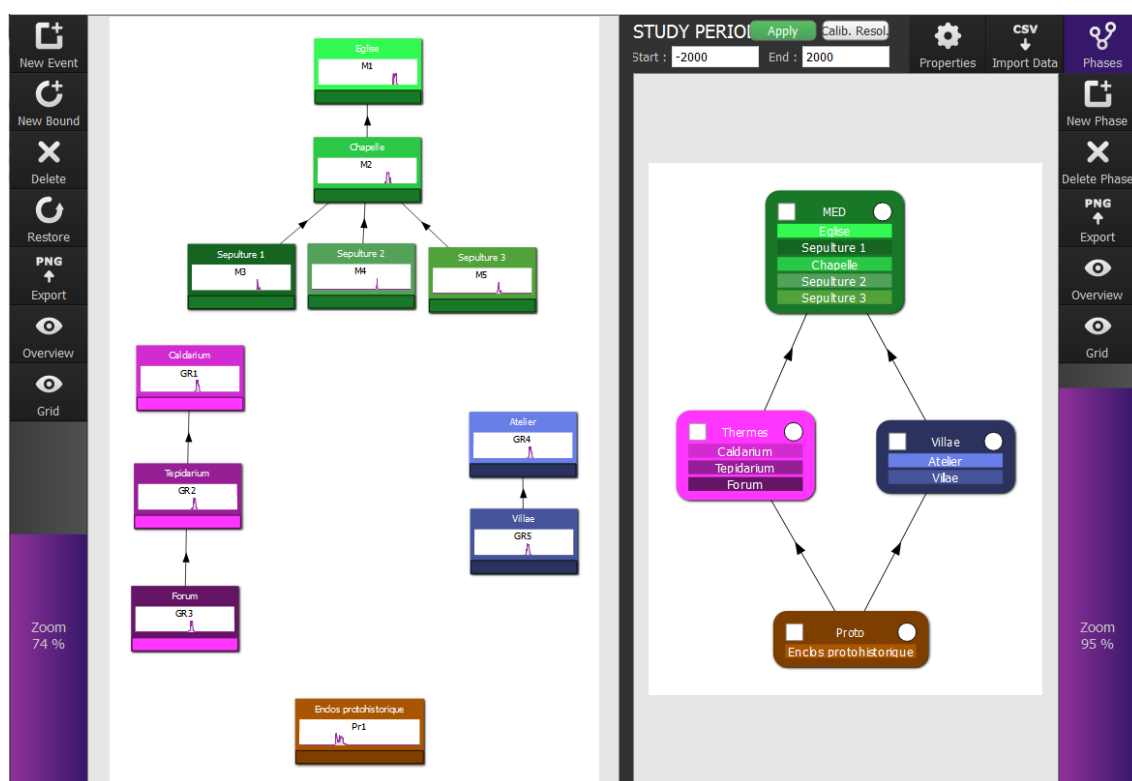
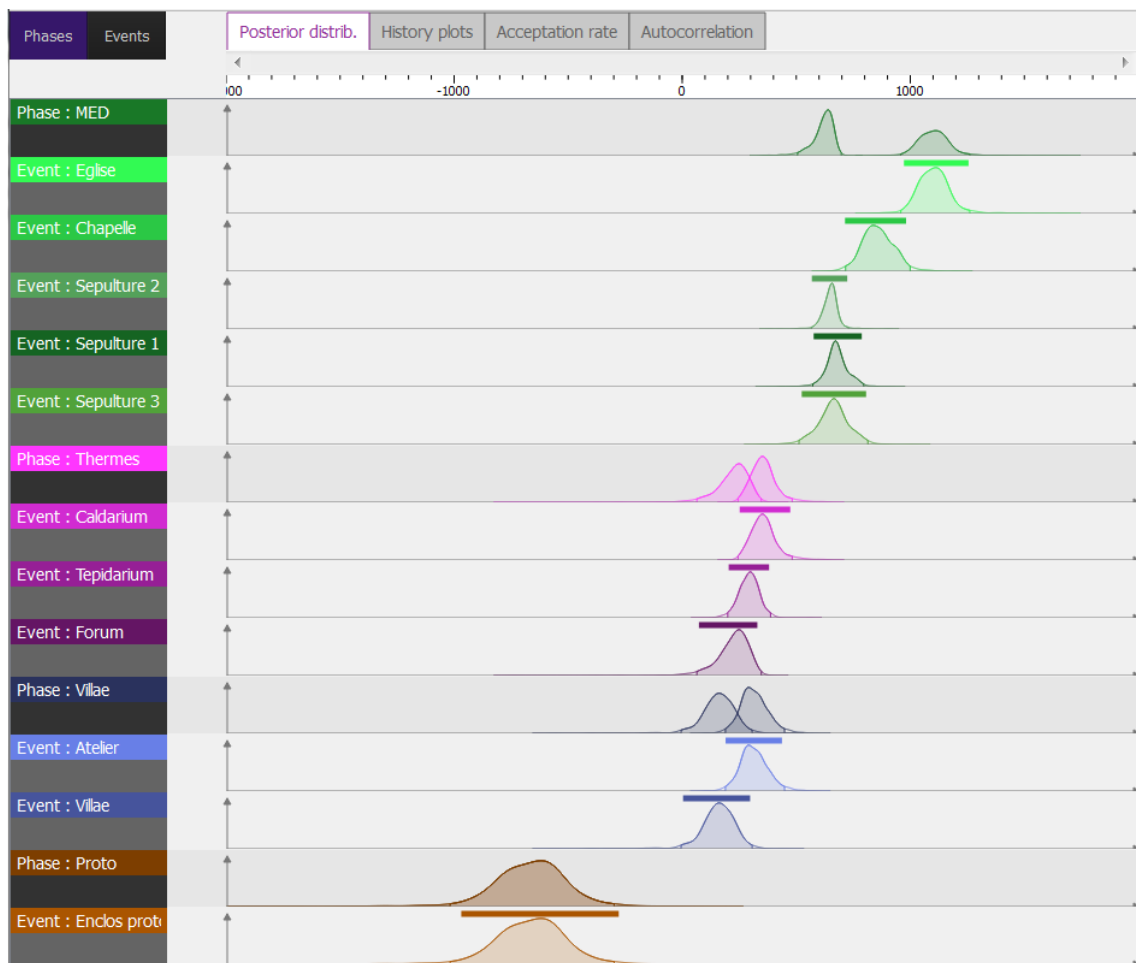Figure 6.31 – Design of the second modelling by ChronoModel : introduction of phases

Figure 6.32 – Marginal posterior densities of all events and phases modelised by the second model

6.2.3   Grouping events into two kinds of phases

Now, keeping the same sequences, we may add new phases corresponding to typo-chronological criteria.
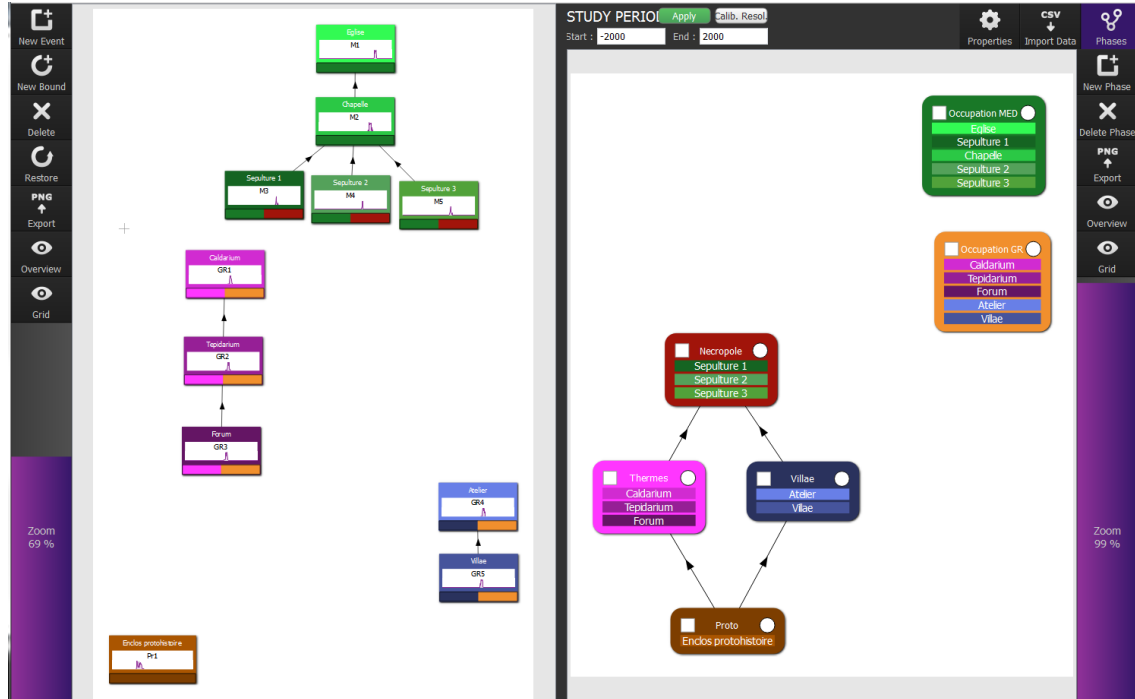

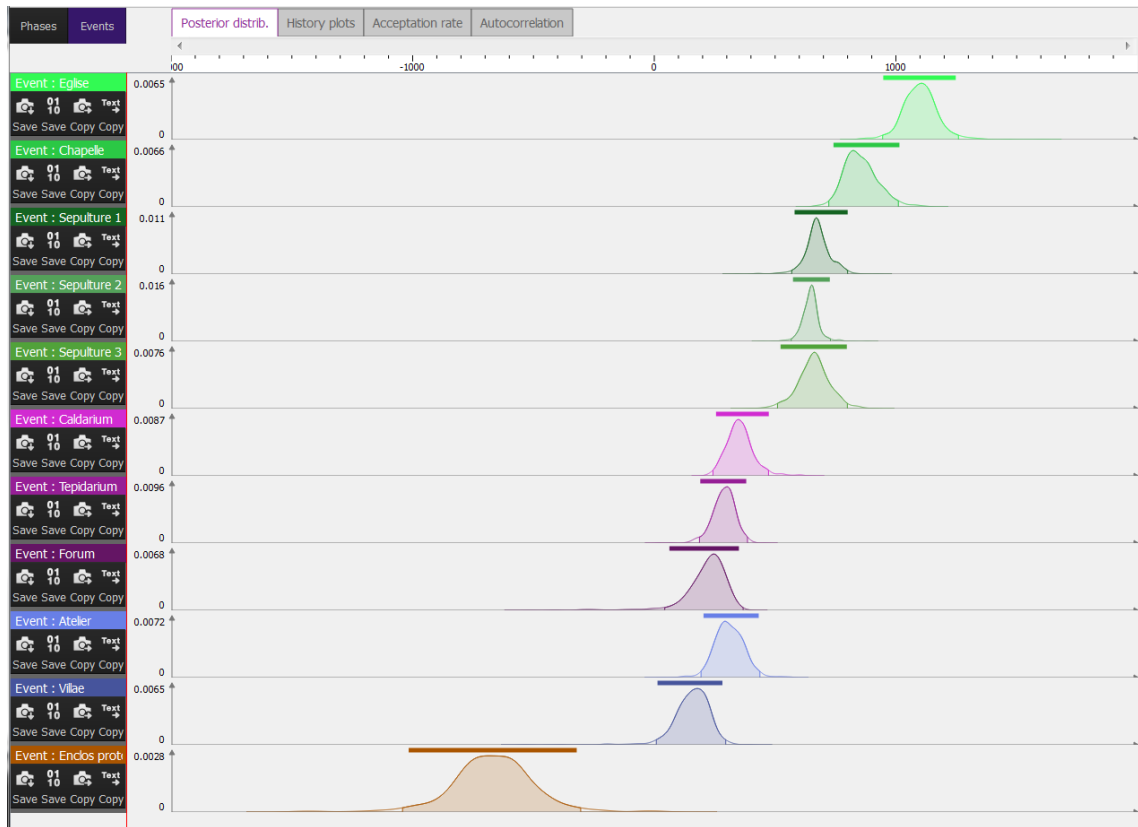Figure 6.33 – Design of the third modelling by ChronoModel

Figure 6.34 – Marginal posterior densities of the events - Third modelling

Figure 6.35 – Marginal posterior densities of the beginning and the end of the phases - Third modelling

Bibliography

[1] Lanos P, Philippe A. Hierarchical Bayesian modeling for combining Dates in archaeological context; 2015.

[2] Christen JA. Summarizing a set of radiocarbon determinations: a robust approach. Applied Statistics. 1994;43(3):489–503.

[3] Lunn D, Jackson C, Spiegelhalter D. The BUGS Book: A Practical Introduction to Bayesian Analysis. Chapman and Hall/CRC; 2012.

[4] Daniels MJ. A prior for the variance in hierarchical models. Canad J Statist. 1999;27(3):567–578.

[5] Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian Approaches to Clinical Trials and Health-Care Evaluation. Chichester: Wiley; 2004.

[6] Ward GK, Wilson SR. Procedures for comparing and combining radiocarbon age determinations: a critique. Archaeometry. 1978;20(1):19–31.

[7] Stuiver M, Reimer P, Bard E, Beck J, Burr G, Hughen K, et al. INTCAL98 radiocarbon age calibration, 24,000-0 cal BP. Radiocarbon. 2006;40(3). Available from: https://journals.uair.arizona.edu/index.php/radiocarbon/article/view/3781.

[8] Reimer P, Baillie M, Bard E, Bayliss A, Beck J, Bertrand C, et al. IntCal04 terrestrial radiocarbon age calibration, 0-26 cal kyr BP. Radiocarbon. 2004;46(3). Available from: https://journals.uair.arizona.edu/index.php/radiocarbon/article/view/4167.

[9] Hughen K, Baillie M, Bard E, Beck J, Bertrand C, Blackwell P, et al. Marine04 marine radiocarbon age calibration, 0-26 cal kyr BP. Radiocarbon. 2004;46(3). Available from: https://journals.uair.arizona.edu/index.php/radiocarbon/article/view/4168.

[10] McCormac FG, Hogg A, Blackwell P, Buck C, Higham T, Reimer P. SHCal04 Southern Hemisphere calibration, 0-11.0 cal kyr BP. Radiocarbon. 2004;46(3). Available from: https://journals.uair.arizona.edu/index.php/radiocarbon/article/view/4169.

[11] Reimer P, Baillie M, Bard E, Bayliss A, Beck J, Blackwell P, et al. IntCal09 and Marine09 Radiocarbon Age Calibration Curves, 0-50,000 Years cal BP. Radiocarbon. 2009;51(4). Available from: https://journals.uair.arizona.edu/index.php/radiocarbon/article/view/3569.

[12] Reimer P, Bard E, Bayliss A, Beck J, Blackwell P, Ramsey CB, et al. IntCal13 and Marine13 Radiocarbon Age Calibration Curves 0–50,000 Years cal BP. Radiocarbon. 2013;55(4). Available from: https://journals.uair.arizona.edu/index.php/radiocarbon/article/view/16947.

[13] Hogg A, Hua Q, Blackwell P, Niu M, Buck C, Guilderson T, et al. SHCal13 Southern Hemisphere Calibration, 0–50,000 Years cal BP. Radiocarbon. 2013;55(4). Available from: https://journals.uair.arizona.edu/index.php/radiocarbon/article/view/16783.

[14] Hervé G, Chauvin A, Lanos P. Geomagnetic field variations in Western Europe from 1500BC to 200AD. Part I: Directional secular variation curve. Physics of the Earth and Planetary Interiors. 2013 May;218:1–13. Available from: http://www.sciencedirect.com/science/article/pii/S0031920113000265.

[15] Hervé G, Chauvin A, Lanos P. Geomagnetic field variations in Western Europe from 1500BC to 200AD. Part II: New intensity secular variation curve. Physics of the Earth and Planetary Interiors. 2013 May;218:51–65. Available from: http://www.sciencedirect.com/science/article/pii/S0031920113000277.

[16] Kovacheva M, Kostadinova-Avramova M, Jordanova N, Lanos P, Boyadzhiev Y. Extended and revised archaeomagnetic database and secular variation curves from Bulgaria for the last eight millennia. Physics of the Earth and Planetary Interiors. 2014;236:79 – 94. Available from: http://www.sciencedirect.com/science/article/pii/S0031920114001605.

[17] Buck CE, Litton CD, Cavanagh W G. The Bayesian Approach to Interpreting Archaeological Data. England: Chichester, J.Wiley and Son; 1996.

[18] Robert CP, Casella G. Monte Carlo Statistical Methods. 2nd ed. Springer; 2004.

[19] Robert CP. Simulation of truncated normal variables. Statistics and Computing. 1995;5(2):121–125. Available from: http://dx.doi.org/10.1007/BF00143942.

[20] Roberts GO, Rosenthal JS. Optimal Scaling for Various {Metropolis-Hastings} Algorithms. Statistical Science. 2001;16(4):351–367. Available from: http://dx.doi.org/10.2307/3182776.

[21] Härdle WK, Müller M, Sperlich S, Werwatz A. Nonparametric and Semiparametric Models; 2012. Available from: https://books.google.com/books?hl=fr{&}lr={&}id=wqX7CAAAQBAJ{&}pgis=1.

[22] Samworth RJ, Wand MP. Asymptotics and optimal bandwidth selection for highest density region estimation. The Annals of Statistics. 2010 jun;38(3):1767–1792. Available from: http://projecteuclid.org/euclid.aos/1269452654.

[23] Quiles A, Aubourg E, Berthier B, Delque-Količ E, Pierrat-Bonnefois G, Dee MW, et al. Bayesian modelling of an absolute chronology for Egypt's 18th Dynasty by astrophysical and radiocarbon methods. Journal of Archaeological Science. 2013;40:423–432.