Centro de Investigación en Matemáticas, A.C.

Master Thesis

F14C Bayesian Calibration: Approach and Open Problems

Author:

Marco Antonio Aquino Lopez

Supervisor:

Dr. J. Andres Christen

CIMAT

TESIS:

Que para obtener el grado de:

Maestro en Ciencias

con orientación en Probabilidad y Estadística

Guanajuato, Guanajuato, Septiembre, 2014

CIMAT
CENTRO DE INVESTIGACION EN MATEMATICAS A. C.

Centro de Investigación en Matemáticas, A.C.

Acta de Examen de Grado

Acta No.: 078

Libro No.: 002

Foja No.: 078

En la Ciudad de Guanajuato, Gto., siendo las 15:00 horas del día 29 de septiembre del año 2014, se reunieron los miembros del jurado integrado por los señores:

DR. ROGELIO RAMOS QUIROGA (CIMAT)
DR. SERGIO PÉREZ ELIZALDE (COLPOS)
DR. JOSÉ ANDRÉS CHRISTEN GRACIA (CIMAT)

bajo la presidencia del primero y con carácter de secretario el segundo, para proceder a efectuar el examen que para obtener el grado de

MAESTRO EN CIENCIAS
CON ESPECIALIDAD EN PROBABILIDAD Y ESTADÍSTICA

sustenta

MARCO ANTONIO AQUINO LÓPEZ

en cumplimiento con lo establecido en los reglamentos y lineamientos de estudios de posgrado del Centro de Investigación en Matemáticas, A.C., mediante la presentación de la tesis

" F14C BAYESIAN CALIBRATION: APPROACH AND OPEN PROBLEMS "

Los miembros del jurado examinaron alternadamente al (la) sustentante y después de deliberar entre sí resolvieron declararlo (a):

APROBADO

DR. ROGELIO RAMOS QUIROGA
Presidente

DR. SERGIO PÉREZ ELIZALDE
Secretario

DR. JOSÉ ANDRÉS CHRISTEN GRACIA
Vocal

CIMAT
DIRECCION GENERAL

DR. JOSÉ ANTONIO STEPHAN DE LA PEÑA MENA
Director General

You must see with eyes unclouded by hate. See the good in that which is evil, and the evil in that which is good. Pledge yourself to neither side, but vow instead to preserve the balance that exists between the two.

Hayao Miyazaki

Centro de Investigación en Matemáticas, A.C.

Abstract

Maestro en Ciencias

F14C Bayesian Calibration: Approach and Open Problems

by Marco Antonio Aquino Lopez

In recent years, a discussion on calibrating radiocarbon data, conventional Radiocarbon Age vs F14C, has taken place amount the users of radiocarbon dating. In this work, the calibration of F14C is discuss and compared to that of the Radiocarbon Age, traditional format, to provide a reliable and comparable calibration in terms of F14C.

Recientemente, la discusión de como calibrar datos de radiocarbono, ya sea usando edad radiocarbonica o F14C , ha tenido lugar en la comunidad de usuarios de las mismas. En este trabajo, se discute la calibración en términos de F14C y es comparada con calibración convencional en términos de edad radiocarbonica, esto con el fin de proporcionar una calibración en F14C que sea confiable.

Agradecimientos

Agradezco sinceramente al Dr. J. Andrés Cristen García por su apoyo y guía en la realización de este trabajo, así como en el apoyo brindado durante mi estancia en Guanajuato. De igual forma agradezco al Dr Maarten Blaauw por su guía y apoyo a la hora de obtener datos que ayudaron en la realización este trabajo.

A mis padres; Filiberta Lopez Santiago y Gildardo Aquino Pérez a quien les agradezco su consejo y apoyo durante toda mi vida y formación académica.

A mis hermanas; Veronica Donaji, Elisa Aurora y Elsa, por estar presentes en los momentos buenos como difíciles de mi vida.

A Vania Xospa Ramirez, por su apoyo en esta etapa de mi vida la cual al principio fue complicada y confusa.

A mis amigos y compañeros con los cuales no solo compartí cursos si no también experiencias de vida.

Agradezco al CONACYT por el apoyo recibido en la realización de mis estudios de maestría. Al CIMAT, A.C. por la gran oportunidad que se me otorgo para desarrollarme profesionalmente.

Dedicatoria

A mis padres:

Filiberta Lopez Santiago y Gildardo Aquino Pérez

Por su apoyo y consejo incondicional en todos los momentos de mi vida.

Contents

Abstract ii

Acknowledgements iii

Contents v

List of Figures vii

List of Tables ix

1 Introduction 1
- 1.1 Outline of the thesis . . . 3

2 F14C Calibration 4
- 2.1 Introduction . . . 4
- 2.2 Data Analysis . . . 4
- 2.3 Calibration Curve . . . 5
- 2.4 Calibration . . . 8
- 2.5 Comparative . . . 10
- 2.6 Discussion . . . 14

3 Models with outliers 16
- 3.1 Introduction . . . 16
- 3.2 Normal model . . . 17
- 3.3 T Model . . . 18
  - 3.3.1 T model justification . . . 18
    - 3.3.1.1 Integrated Likelihood . . . 19
- 3.4 Outliers Models . . . 20
  - 3.4.1 Markov Chain Monte Carlo . . . 22
  - 3.4.2 Normal Model with outliers . . . 23
  - 3.4.3 T Model with outliers . . . 24
- 3.5 Robustness . . . 25
  - 3.5.1 First Test . . . 27
    - 3.5.1.1 Normal model . . . 27

3.5.1.2 T model . . . 27
3.5.1.3 Normal Model with outliers . . . 28
3.5.1.4 T Model with outliers . . . 30
3.5.1.5 Comparison . . . 31
3.5.2 Second Test . . . 32
3.5.2.1 Normal model . . . 32
3.5.2.2 T model . . . 32
3.5.2.3 Normal model with outliers . . . 34
3.5.2.4 T model with outliers . . . 36
3.5.2.5 Comparison . . . 38
3.5.3 Third test . . . 38
3.5.3.1 T model . . . 39
3.5.3.2 Normal model with outliers . . . 40
3.5.3.3 T model with outliers . . . 41
3.5.4 Discussion . . . 43
3.6 Conclusions . . . 43

4 Real Data Calibration **46**
4.1 Queen's University's data . . . 47
4.1.1 Calibration . . . 50
4.2 Comparison Data . . . 50
4.2.1 Data . . . 52
4.2.2 Young Data . . . 53
4.2.3 Mid-age Data . . . 54
4.2.4 Older Data . . . 55
4.2.5 Conclusion . . . 55
4.3 The Shroud of Turin . . . 56
4.3.1 Calibration . . . 57
4.4 Conclusions . . . 60

5 Discussion **61**

A Precision in F14C **64**

Bibliography **66**

List of Figures

1.1 Calibration curve in terms of Radiocarbon age, considering 1 standard deviation around the mean. *the figure only shows up to 3000 years BP, for viewable reasons* 2

2.1 Part of 'Intcal13', Calibration curve in terms of Radiocarbon Age. . 5
2.2 8
2.3 Calibration of data around 750 years BP, from data in table 2.2. . . 10
2.4 Error involved by using wrongly assuming Normality around F14C data. 12
2.5 Comparison with data around 500 years BP. 13
2.6 Comparison with data around 5000 years BP. 13
2.7 Comparison with data around 26000 years BP. 14
2.8 Comparison with data around 36000 years BP. 14

3.1 Calibration of data in table 3.1 using Normal model. 27
3.2 Calibration of data in table 3.1 using T model. 28
3.3 MCMC's time series for the calibration using the Normal model with outliers. 29
3.4 Calibration of data in table 3.1 using Normal model with outliers. . 29
3.5 MCMC's time series for the calibration using the T model with outliers. 30
3.6 Calibration of data in table 3.1 using T model with outliers. 31
3.7 Calibration using normal model; (A)Calibration considering outliers and Normal models using data from table 3.2. (B) Calibration without outliers using the Normal model with outliers using data from table 3.2. 33
3.8 Calibration using T model; (A)Calibration considering outliers and T models using data from table 3.2. (B) Calibration without outliers using the T model with outliers using data from table 3.2. . . . 34
3.9 MCMC's time series for the calibration using the Normal model with outliers. 35
3.10 Calibration of data in table 3.2 using Normal model with outliers. . 35
3.11 MCMC's time series for the calibration using the T model with outliers. 37
3.12 Calibration of data in table 3.2 using T model with outliers. 37
3.13 Calibration using T model. 39

3.14 MCMC's time series for the calibration using the Normal model with outliers. 40
3.15 Calibration of data from table 3.3 using Normal model with outliers. 41
3.16 MCMC's time series for the calibration using the T model with outliers. 42
3.17 Calibration of data from table 3.3 using T model with outliers. . . . 42

4.1 F14C calibration of data in table 4.2. 51
4.2 Radiocarbon age Calibration of data in table 4.1 51
4.3 Calibrations comparison between F14C and Conventional Radiocarbon age. 52
4.4 Calibration of data around 600 years BP. 53
4.5 Calibration of data around 1350 years BP. 53
4.6 Calibration of data around 5500 years BP. 54
4.7 Calibration of data around 24000 years BP. 54
4.8 Calibration of data around 42000 years BP. 55
4.9 A piece of the Shroud of Turin. Wikipedia 2014. 56
4.10 MCMC time series for the calibration of the Shroud of Turin data. 58
4.11 Calibration of the Shroud of Turin data using F14C. 58

A.1 Law of Radiocarbon Decay with precision factor. 64
A.2 Precision in F14C. 65

List of Tables

2.1 F14C calibration curve obtained by Delta Method. 7
2.2 Synthetic data around 750 years BP. 9

3.1 Synthetic data around 2050 years BP in F14C. 26
3.2 Synthetic data around 350 years BP with outliers. 26
3.3 Synthetic data around 1275 years BP with outliers. 26
3.4 Probabilities of a given datum being an outlier. 29
3.5 Probabilities of a given datum been an outlier. 31
3.6 Probabilities of a given datum been an outlier. 36
3.7 Probabilities of a given data being an outlier. 38
3.8 Probabilities of a given datum been an outlier. 41
3.9 Probabilities of a given datum being an outlier. 43

4.1 Queen's University's Radiocarbon age data,(Conventional dates). . 48
4.2 Queen's University's Radiocarbon age data, (F14C data). 49
4.3 Queen's University's Radiocarbon age data. 52
4.4 Radiocarbon determinations for the Shroud of Turin obtained from Damon P. E., et al (1989). 57
4.5 Radiocarbon F14C determinations for the Shroud of Turin, obtained by approximation. 57

Chapter 1

Introduction

Bayesian Statistics has been widely accepted in archeology, paleoecology, and other related areas, mainly due to its contributions to radiocarbon dating, also known as calibration. This thesis focuses on calibration using $F14C$ data, and it compares $F14C$ calibration to the traditional radiocarbon age calibration, generally used.

Radiocarbon dating, also known as carbon dating, is a very popular technique which is used to date organic material. It dates material using the decay of carbon-14 (^{14}C), which has a half-life of 5730 years. This means that after 5730 years, only half of the original ^{14}C remains in the sample; this leads to the law of radiocarbon decay, that is;

$$N(t) = N(0)e^{-\lambda t}, \tag{1.1}$$

where $N(t)$ is the amount of ^{14}C at time t after death, and λ is the decay rate. From equation 1.1, t can be solved and used as an estimate of the time since the organism has died. That is;

$$\begin{aligned} t &= \frac{-1}{\lambda} \ln\left(\frac{N(t)}{N(0)}\right) \\ &= \frac{-1}{\lambda} \ln\left(a^{14}\right). \end{aligned} \tag{1.2}$$

Notice that this transformation is based on the law of radiocarbon decay, but this does not mean that this date is accurate, due to the transformation being based on the following assumptions:

- Time-independent levels of ^{14}C in the atmosphere;
- a half-life of 5730 years;

For this reason, a calibration is necessary. To be able to calibrate the age of the material, a curve was created which took into account the levels of ^{14}C in the atmosphere throughout time, by radiocarbon dating precisely dendrochronology dated wood samples. A piece of the curve is shown in Figure 1.1, and is known as the calibration curve; in this case, 'Intcal13', Reimer et al. (2013):

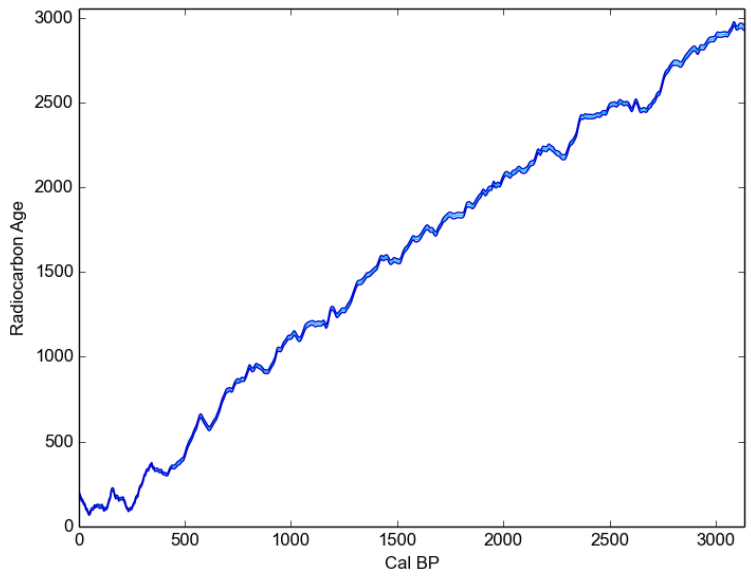


FIGURE 1.1: Calibration curve in terms of Radiocarbon age, considering 1 standard deviation around the mean.
the figure only shows up to 3000 years BP, for viewable reasons

From the curve, it becomes clear that the first assumption is false. The estimate of t obtained previously is not an accurate way to report the date, and does not provide accurate information at all; moreover, this estimate, t, can confuse non-experts. For this reasons, in recent years, users of carbon dating have brought

the idea of using the radiocarbon activity, known as F14C, as the main datum to be used in the calibration instead of the traditional radiocarbon age. This thesis approaches this question, and implements $F14C$ calibration using several models.

1.1 Outline of the thesis

As it was mentioned, this thesis implements the $F14C$ calibration, and also compares models to give an objective opinion about the use of the models in question. In chapter 2, the $F14C$ calibration is implemented and compared to the traditional radiocarbon years calibration. All of this is using the robust model introduced in Christen and Pérez (2009). Also, as part of the implementation, an F14C calibration curve is created in this chapter as well.

In chapter 3, $F14C$ calibration is implemented using the traditional Normal Model and also using the T model, also, both models are used to implement an outlier structure. The four models are compared to verify how robust the models are to outliers. the comparison leads to an objective opinion regarding the use of these models.

Last, in chapter 4, a calibration of data with suspected outliers is used to illustrate the use of the $F14C$ calibration and compares the models in a real scenario. The data used in this example is the data from The Shroud of Turin.

Chapter 2

F14C Calibration

2.1 Introduction

In this Chapter, $F14C$ calibration is justified, implemented and compared to the traditional Radiocarbon age calibration. To justify the use of the $F14C$ a model (i.e. a probability distribution which describes randomness of the phenomenon) is necessary. To assume such a distribution in a correct way, an analysis of the F14C data, is necessary; this analysis can be found on section 2. After the analysis of the data, the necessity of a calibration curve is clear and such curve is created in in section 3 of the chapter, this curve is created using 'Intcal13', Reimer et al. (2013). With a structure of the data and a calibration curve, the calibration is finally introduced in section 4 and tested using simulated data. Last a comparison between Traditional calibration and $F14C$ calibration is made.

2.2 Data Analysis

Each F14C determination is a measurement of the concentration of ^{14}C in the sample. It can be assumed that the observation follows a normal distribution, i.e. F14C follows a normal distribution, since it is a measurement. On the other hand, the traditional assumption is that the radiocarbon age, $t = -\frac{1}{\lambda}\ln(F14C)$, follows

a normal distribution. This thesis is focused on F14C calibration. Considering the model presented in Christen & Pérez (2009), the model used to calibrate F14C data in this work, is the T model. The reasons to favor this model over the traditional normal model are detailed in Chapter 3 of this work.

2.3 Calibration Curve

An important part of carbon dating is the calibration curve. In the case of 'Intcal13', Reimer et al. (2013), the curve gives the estimated levels of ^{14}C on Earth up to 50,000 years before present (1950 AD). The calibration curve is a summary of data obtained from known calendar age, such as tree rings around the globe. The curve is the best description of levels of ^{14}C through time. It is based on analyzing the ^{14}C levels of precisely dated samples (for the first 15,000 years dendrochronology is used to date wood samples)

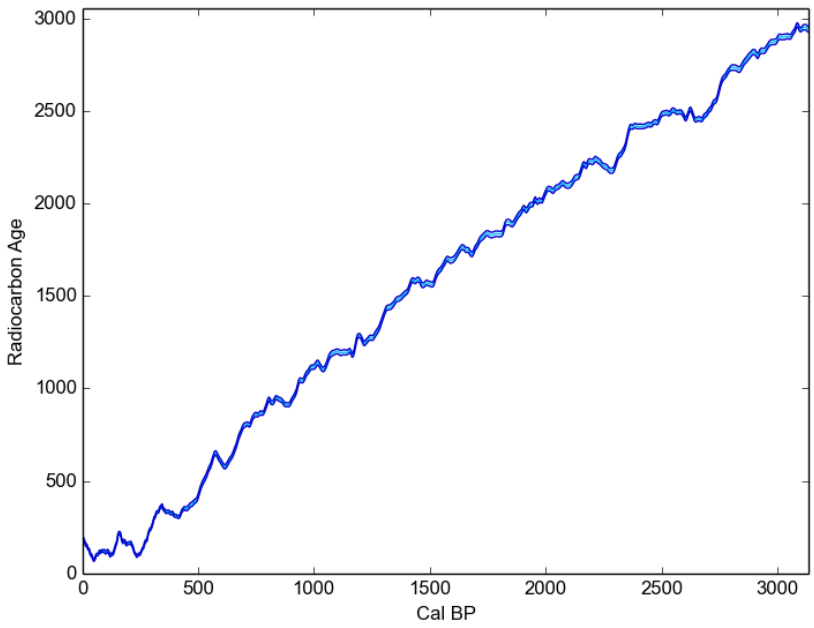


FIGURE 2.1: Part of 'Intcal13', Calibration curve in terms of Radiocarbon Age.

The importance of the calibration curve lies on estimating the age of organic material. This process, called calibration, is crucial to the carbon dating process.

The calibration curve, 'Intcal13' shown in figure 2.1, presents the data in its 'Radiocarbon age' format. This is because nowadays almost all laboratories provide the data in this format, as it was previously mentioned.

To calibrate data in its $F14C$ format, it becomes necessary to have a calibration curve which takes data in this format. In other words, it is necessary to have a $F14C$ calibration curve. Since no such curve exists, one must be estimated using the existing Intcal13 curve.

To be able to obtain a calibration curve in $F14C$ terms it is important to notice that each point in the calibration curve is an estimate of the true 'radiocarbon age' at a given time. If we consider the independence between estimates, we can approach this problem from a statistical point of view, as follows:

Let $\hat{T}_t$ be the 'radiocarbon age' estimate at time t, and $\hat{f}_t$ be the estimate of the radiocarbon activity, F14C, at time t. Now using the law of radiocarbon decay,

$$f_t = exp\left(-\lambda T_t\right).$$

from 'Intcal13', $\hat{T}_t$, and $\hat{\sigma}_t$ are known variables, so it is possible to define the function:

$$f_t = f(T_t) = \exp\left(-\lambda T_t\right). \tag{2.1}$$

To estimate the mean and standard deviation of $f(T_t)$ using estimates $\hat{T}_t$ and $\hat{\sigma}_t$,the Delta Method can be used;

Theorem (Delta Method) 1. If $\tau = g(\theta)$ where g is differentiable and $g'(\theta) \neq 0$ then;

$$\frac{\hat{\tau}_n - \tau}{\hat{se}(\hat{\tau}_n)} \to \mathcal{N}(0,1),$$

where $\hat{\tau}_n = g(\hat{\theta})$ and

$$\hat{se}(\hat{\tau}_n) = |g'(\hat{\theta})|\hat{se}(\hat{\theta}). \tag{2.2}$$

If calendar year t is associated to T, 'radiocarbon age', and $\sigma(t)$, its standard deviation, then f_t =F14C$_t$ and $\sigma(f_t)$ can be estimated using equations 2.1 and 2.2, as follows;

$$\hat{f}(T_t) = \exp(-\lambda \hat{T}_t),$$
$$\hat{\sigma}(f(T_t)) = |-\lambda \exp(-\lambda \hat{T}_t)|\sigma(\hat{T}_t).$$

Using this approximation, the F14C calibration curve was created. A small portion of it is displayed in table 2.1;

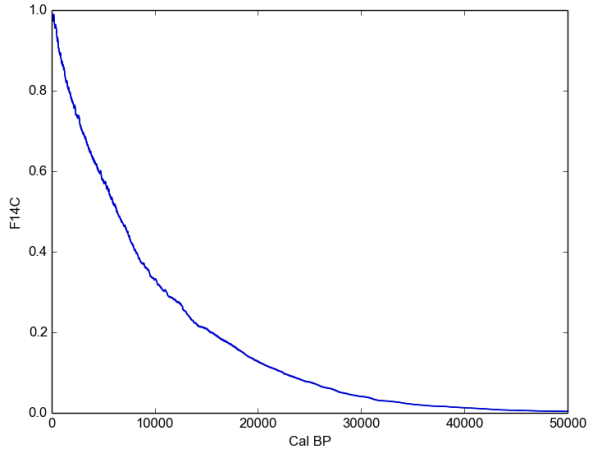
TABLE 2.1: F14C calibration curve obtained by Delta Method.

CAL BP	14C age	YR BP Error	F14C	F14C error
0	199	8.0	0.9755	0.00097
5	188	8.0	0.9768	0.00097
10	172	8.0	0.9788	0.00097
⋮	⋮	⋮	⋮	⋮

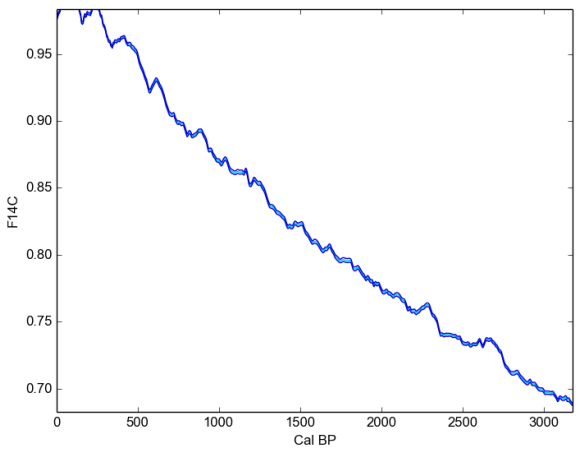
figure 2.2 shows the F14C calibration curve

In table table 2.1 the numerical results obtained with the delta method are observed; in this work only 4 decimals are used, this is because the data provided by Queen's University use this format and using more digits would not provide any additional information for the calibration. It is important to notice that the technique used to obtain the curve in this work provides only an approximation, this is another reason why a calibration curve for F14C should be created.

From figure 2.2, it is possible to observe the similarity in the calibration curve's behavior to that of the radiocarbon decay law, which would be expected from a F14C calibration curve. The F14C calibration curve obtained can be used to calibrate data when reported in F14C terms.



(A) F14C calibration curve obtained by approximating 'Intcal13'.



(B) Zoomed F14C calibration curveobtained by approximating 'Intcal13'.

FIGURE 2.2

2.4 Calibration

Using the calibration curve obtained in the previous section, and the T model described in detail in Chapter 3, a calibration of data is possible. The objective of the calibration is to obtain a posterior distribution given the data. In this case

the data has the following structure:

$$y_j \mid \theta \sim \mathscr{N}\left(\mu(\theta), \alpha\omega_j(\theta)\right),$$

where α follows

$$\pi(\alpha) = InvGa(\alpha, a, b) = \frac{b^a}{\Gamma(a)}\alpha^{-(a+1)} \exp\left(-\frac{b}{\alpha}\right).$$

For the prior distribution for α an Inverse Gamma distribution is used. When modeling the variance of a model it is common to use an inverse gamma distribution since this distribution takes positive values and is flexible enough to describe many different scenarios. Using this structure, the likelihood for this model is

$$\begin{aligned}
\mathscr{L}(\bar{y} \mid \theta, \alpha) &= \prod_{j=1}^{m} f(y_j \mid \theta, \alpha) \\
&= \alpha^{-m/2} \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi\omega(\theta)_j^2}} \exp\left(-\frac{(y_j - (\mu(\theta) + \phi_j\delta_j))^2}{2\alpha\omega(\theta)_j^2}\right) \\
&\propto \alpha^{-m/2} \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right] \exp\left(-\frac{1}{2}\sum_{j=1}^{m} \frac{((y_j - \mu(\theta)) - \phi_j\delta_j)^2)^2}{\alpha\omega(\theta)_j^2}\right). \qquad (2.3)
\end{aligned}$$

and using this likelihood, the posterior distribution has the form;

$$\pi(\theta \mid \bar{y}) \propto \left[b + \sum_{j=q}^{m} \frac{(y_j - \mu(\theta))^2}{2\omega(\theta)_j^2}\right]^{-\frac{m+2a}{2}} \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right]. \qquad (2.4)$$

Now that the posterior distribution has been obtained, calibration is possible. To illustrate the calibration of some data, an artificial set of data is created. This data is around 750 years BP, and is shown in table 2.2,

TABLE 2.2: Synthetic data around 750 years BP.

^{14}Cdetermination	Standard deviation
0.8986	0.0015
0.9026	0.0005
0.8956	0.0013
0.8999	0.0005

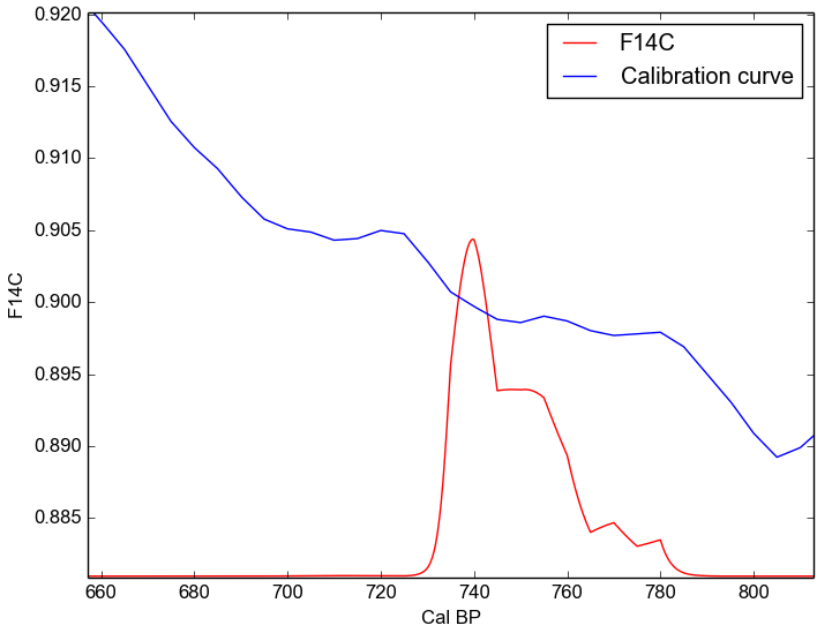


FIGURE 2.3: Calibration of data around 750 years BP, from data in table 2.2.

Figure 2.3 shows the posterior distribution of θ given the observed data. This posterior provides a way to determine the most probable age for the organic material from which the data was obtained. This procedure is called calibration.

It is important to notice that the structure to obtain the calibration is not any different from that of the traditional radiocarbon age calibration. For this, a theoretical justification is necessary.

2.5 Comparative

To compare the two calibrations, it is important to establish the assumptions of each calibration, which are the following;

(a) If the assumption made to calibrate data using F14C data is that F14C is the real data measured in the laboratory, then F14C is normally distributed.

(b) On the other hand, 'radiocarbon age' is assumed to be normally distributed for traditional reasons. This implies false assumptions about the levels of ^{14}C in the atmosphere, among others.

To observe how the traditional assumption of considering 'radiocarbon age' as normally distributed, affects the calibrations, lets recall the law of radiocarbon decay.

$$t(F14C) = -\frac{1}{\lambda}\ln(F14C), \tag{2.5}$$

Since the correct assumption is that $F14C \sim \mathscr{N}(\mu, \sigma^2)$, it is desirable to observe the error when the assumption $t \sim \mathscr{N}(\mu_t, \sigma_t^2)$ is made. For this purpose, the Taylor series can be used.

Theorem (Taylor Series) 1. If a function $g(x)$ has a derivative of order, r, that is if $g^r(x) = \frac{d^r}{dx^r}g(x)$ exists, then for any constant a, the Taylor polynomial of order r about a is:

$$T_r(x) = \sum_{k=0}^{r} \frac{g^{(k)}(a)}{k!}(x-a)^k. \tag{2.6}$$

Using the Taylor series of the first order centered around F14C, $t(x)$ can be written as,

$$t(x) = -\lambda\log(F14C) - \frac{\lambda}{F14C}(x - F14C) + O(\xi) \tag{2.7}$$

where $O(\xi)$ is the error related to the Taylor series approximation. It is important to notice that at the datapoint no error is present, but as the calibration moves away from the point the error increases. This is the error which would affect the calibration, since the data obtained is not exact, but a random variable.

The approximation gives a linear expression therefore the assumption $t \sim \mathscr{N}(\mu_t, \sigma_t^2)$ would now make sense. However, since it is only an approximation, the error of

using such an approximation can be calculated, this error tells us how different the "radiocarbon age" data is from that of F14C as the process moves away from the observed datum, in other words, how the standard deviation is affected by the use of the approximation.

$$\begin{aligned} O(\xi) &= -\lambda \log(x) + \lambda \log(F14C) + \frac{\lambda}{F14C}(x - F14C) \\ &= \lambda \left(\log(F14C) - \log(x) + \frac{1}{F14C}(x - F14C) \right). \end{aligned} \tag{2.8}$$

The error can be seen in figures 2.4.

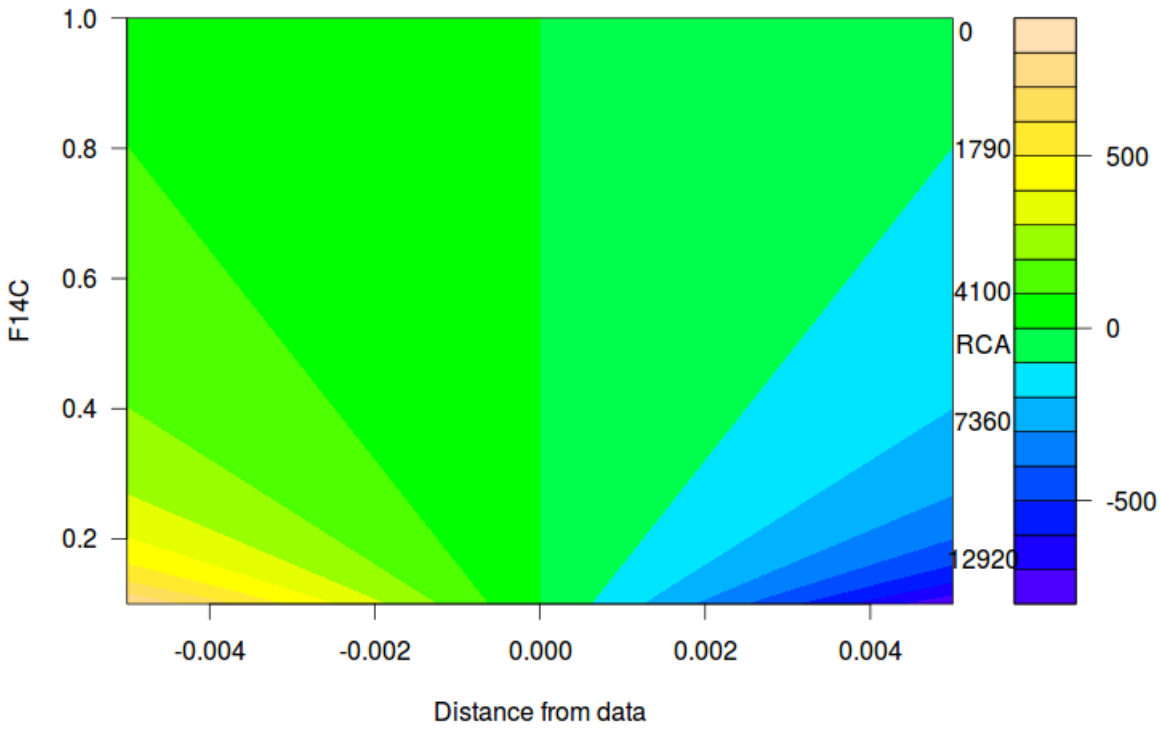


FIGURE 2.4: Error involved by using wrongly assuming Normality around F14C data.

As it is shown in figure 2.4, the error increases as the calibrations are younger.

To better illustrate the effected of the error from the assumptions between F14C and 'radiocarbon years', the following calibrations are introduced.

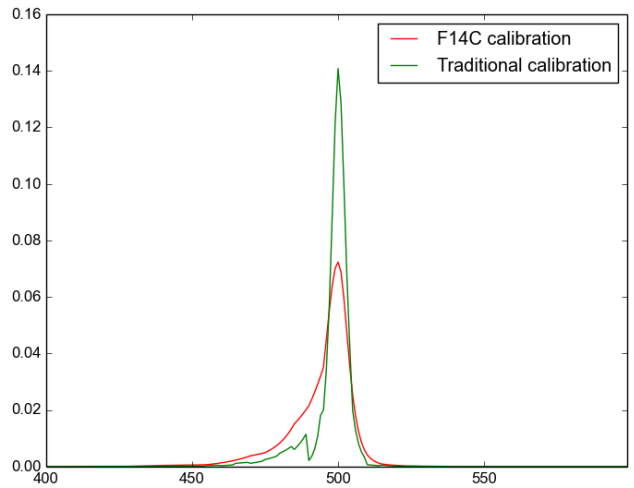


FIGURE 2.5: Comparison with data around 500 years BP.

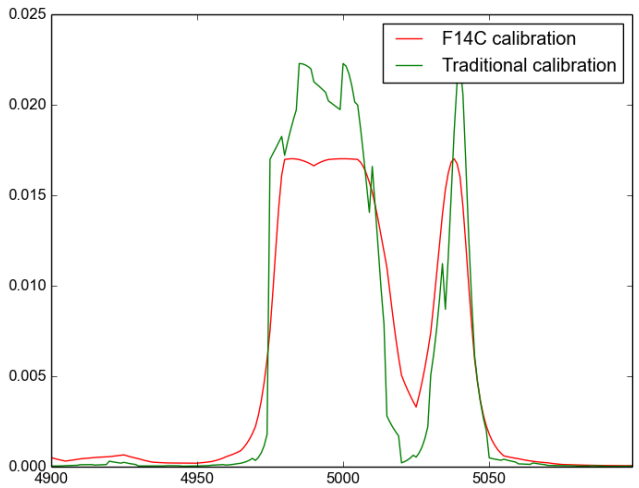


FIGURE 2.6: Comparison with data around 5000 years BP.

It can be seen from these figures that as the calibration becomes older, the discrepancies between the calibrations increases. This occurs because the calibration curve used to calibrate the data is only an approximation.

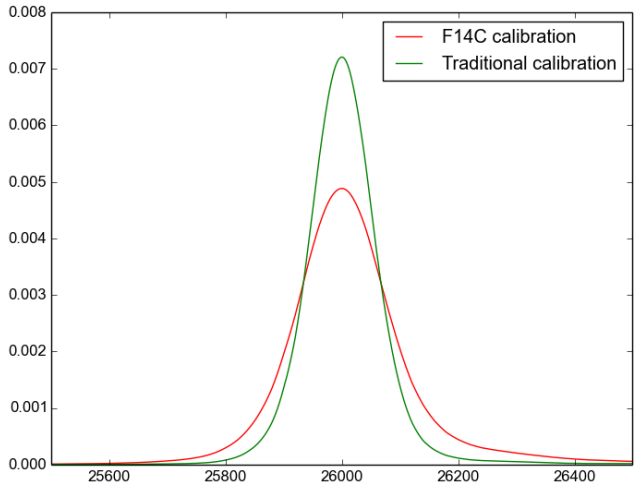


FIGURE 2.7: Comparison with data around 26000 years BP.

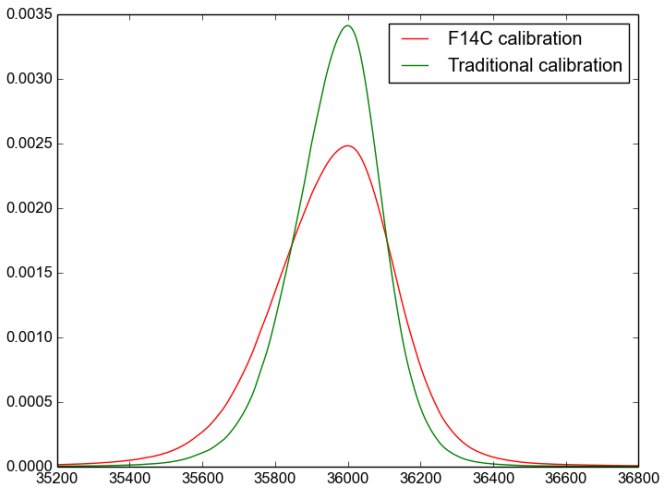


FIGURE 2.8: Comparison with data around 36000 years BP.

2.6 Discussion

In this chapter, calibration of F14C was introduced and compared to the traditional calibration. It is important to remark that the F14C calibration is suggested in papers like Reimer, Brown and Reimer (2004). This work discusses the idea of applying F14C calibration to any set of data regardless of age. Also a calibration curve for F14C data was created, which would help calibrate data when F14C data is presented.

As for the F14C calibration, it is explained in section 2.4 that this calibration does not seem to be modified when using F14C data or 'radiocarbon age'. Since the posterior distribution has the same form, and the only modifications are the data input into the model, the calibration curve has to be used properly. It is for this reason that transitioning from a 'radiocarbon age' to a F14C calibration seems natural and easy, since every model proposed in terms of 'radiocarbon age' works in terms of F14C. In this work, the idea of carrying the calibration to F14C is encouraged since F14C is a more honest and natural data format than the traditional 'radiocarbon age'. giving that, the data is obtained from the laboratory and not just a transformation which misinforms the users. Also, when using F14C, the user sees the necessity to calibrate the data before getting any conclusions. This leads to a more transparent interpretation of the given data.

On the other hand, when working with the calibration curve, several confusions arise. For example, in Reimer et al. (2013) indicates that age-corrected $\Delta^{14}C$ can be found in the curve, 6th colum. On the other hand, Eriksson et al, 2011 describes age-corrected $\Delta^{14}C$ as $\Delta = (F14C \cdot e^{\delta - x} - 1)$. This would create confusion, since it would be desireable to obtain a calibration curve from this value.

It is very important to notice that the calibration curve obtained in this work is an approximation of what would be desirable to obtain. It is a calibration curve in terms of F14C without being previously transformed to 'radiocarbon age'. Since this transformation affects the calibration, it is a transformation that uses incorrect assumptions. Information is lost when the levels of F14C are approximated. For this reason, a natural F14C calibration curve should be created to prevent the calibration in terms of F14C being dependent on approximations.

Chapter 3

Models with outliers

3.1 Introduction

When working with radiocarbon dating, a common problem arises. This problem is the relatively high probability of obtaining outliers in a data set. This is because the procedure to obtain a dating is rather delicate, and it depends on very complicated chemical and physical processes. As mentioned in Christen (1994) *"The quality of the radiocarbon dating can be affected by four main sources of error which can lead to the production of an outlying radiocarbon determination*

(a) *the quality of the care on the archaeological sampling to ensure that the sample taken can realistically provide results for the events that we wish to date (see Bowman(1990), chapter 5);*

(b) *the quality of the care in taking and handling the sample on site to avoid contamination with older or younger material (see Bowman(1990), oages 27-28);*

(c) *the quality of the care in sample handling and preparation in the laboratory to ensure that the sample undergoes appropriate pretreatment and does not become contaminated before analysis for* 14^C *can take place (for information about pretreatment see Bowman (1990), pages 28-30)*;

(d) other non controllable random factors that can appear during the whole process (see Scott at al. (1990a, b))"

As can be seen from the four possible sources of error, radiocarbon dating is a delicate process which is likely to contain outliers. For these reasons, an analysis on models using calibrated data is necessary. It is important to notice that this analysis is valid for F14C calibration and traditional analysis as well, as is the case for regular calibration.

This chapter gives an exploratory analysis on models proposed for traditional calibration and compares their results to give a suggestion on how to deal with radiocarbon data in its F14C form.

3.2 Normal model

To calibrate a radiocarbon determination y_i, in terms of F14C, the normal model has been used traditionally to infer about age θ. This model is reasonable, considering that y_j, F14C, is an observation from a laboratory experiment. The model can be express as;

$$y_j \mid \theta \sim \mathscr{N}\left(\mu(\theta), \omega(\theta)_j^2\right), \tag{3.1}$$

where $\mu(\cdot)$ is the calibration curve, $\omega(\theta)_j^2 = \sigma^2(\theta) + \sigma_j^2$, and σ_j is the reported standard deviation for y_j, is the $\sigma(\theta)$ is the error related to the calibration curve at θ.

If a sample $y = (y_1, y_2, ..., y_m)$ of m ^{14}Cdeterminations for the same age is considered, θ, it is desirable to incorporate this information into the model and obtain a better approximation for θ. This is accomplished by using the likelihood, which

is express as;

$$\begin{aligned}\mathscr{L}(\bar{y} \mid \theta) &= \prod_{j=1}^{m} f(y_j \mid \theta) \\ &= \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi\omega(\theta)_j^2}} \exp\left(-\frac{(y_j - \mu(\theta))^2}{2\omega(\theta)_j^2}\right) \\ &\propto \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right] \exp\left(-\frac{1}{2}\sum_{j=1}^{m} \frac{(y_j - \mu(\theta))^2}{\omega(\theta)_j^2}\right). \end{aligned} \tag{3.2}$$

With equation 3.2 it is possible to obtain the posterior distribution from which the calibration is possible; the posterior distribution in this case is,

$$\pi(\theta \mid \bar{y}) \propto \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right] \exp\left(-\frac{1}{2}\sum_{j=1}^{m} \frac{(y_j - \mu(\theta))^2}{\omega(\theta)_j^2}\right). \tag{3.3}$$

3.3 T Model

As we observed in the previous chapter the normal model is not sufficient to calibrate data due to it possibly containing outliers. For this reason a more robust model (in the sense that it would not be so sensitive to outliers) is needed. In Christen & Pérez (2009) a model which considered the error of the determinations as a random variable is presented, with the objective of capturing more information of the phenomena. This model is now presented and put to test using a set of artificial data containing very clear outliers.

3.3.1 T model justification

It is well known that ^{14}C activity precision depends on age. This means that σ_j is jointly distributed with y_j, therefore, the assumption in the Normal model has faults and we should consider the variability of the precision for the determination.

The variability of σ_j^2 can be expressed by using a variable α which is multiplied by the variable of the observation given θ, i.e.

$$y_j \mid \theta \sim \mathscr{N}\left(\mu(\theta), \alpha\omega_j(\theta)\right). \tag{3.4}$$

now given a random sample of ^{14}C determinations, the likelihood is;

$$\begin{aligned}
\mathscr{L}(\bar{y} \mid \theta, \alpha) &= \prod_{j=1}^{m} f(y_j \mid \theta, \alpha) \\
&= \alpha^{-m/2} \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi\omega(\theta)_j^2}} \exp\left(-\frac{(y_j - \mu(\theta))^2}{2\alpha\omega(\theta)_j^2}\right) \\
&\propto \alpha^{-m/2} \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right] \exp\left(-\frac{1}{2}\sum_{j=1}^{m} \frac{(y_j - \mu(\theta))^2}{\alpha\omega(\theta)_j^2}\right).
\end{aligned} \tag{3.5}$$

Up to now, we have only talked about α in terms of a random variable, but no distribution has been given for this parameter. If one considers that this parameter is a positive random variable and is part of the scale parameter in a normal distribution, it is reasonable to describe this parameter. This means that the prior distribution for α is;

$$\pi(\alpha) = InvGa(\alpha, a, b) = \frac{b^a}{\Gamma(a)}\alpha^{-(a+1)} \exp\left(-\frac{b}{\alpha}\right), \tag{3.6}$$

Furthermore, given θ;

$$\alpha\omega_j \mid \theta \sim InvGa(a, b\omega_j(\theta)). \tag{3.7}$$

3.3.1.1 Integrated Likelihood

At this point, it is important to remember that the variable of interest is the age θ, while α is a nuisance parameter. In Bayesian theory, nuisance parameters are

naturally eliminated by obtaining the marginal distribution, which in this case, would be the integrated likelihood. This can be expressed as;

$$\begin{aligned}
\mathscr{L}_{a,b}(\bar{y} \mid \theta) &= \int_0^\infty \mathscr{L}(\bar{y} \mid \theta, \alpha)\pi_{a,b}(\alpha)d\alpha \\
&\propto \int_0^\infty \alpha^{-m/2} \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right] e^{\left(-\frac{1}{2}\sum_{j=1}^{m} \frac{(y_j-\mu(\theta))^2}{\alpha\omega(\theta)_j^2}\right)} \frac{b^a}{\Gamma(a)} \alpha^{-(a+1)} e^{\left(-\frac{b}{\alpha}\right)} d\alpha \\
&\propto \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right] \left[b + \frac{(y_j-\mu(\theta))^2}{2\omega(\theta)_j^2}\right]^{-\frac{m+2a}{2}}. \qquad (3.8)
\end{aligned}$$

This model takes its name from the fact that the second term of the likelihood is the density of a t distribution with 2a degrees of freedom.

This leads to the posterior distribution being;

$$\pi(\theta \mid \bar{y}) \propto \left[b + \frac{(y_j-\mu(\theta))^2}{2\omega(\theta)_j^2}\right]^{-\frac{m+2a}{2}} \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right]. \qquad (3.9)$$

3.4 Outliers Models

As it was mentioned in the introduction, our model is best described as using a localization-scale model, and in Chapter 2 and 3, the Normal and T models were introduced and tested using noise and abnormal outliers. In this section, the shift and the length of the shift of the outliers are treated as random variables which it is desirable to estimate. Due to this, the model becomes more complex and the use of Markov Chain Monte Carlo, MCMC, becomes necessary.

Since an outlier is considered, in this context, as an observation with a shift in the localization parameter, and using $f(y_j)$ a localization-scale distribution, the

model with a shift in this parameter can be express as;

$$y_j \sim f(y_j \mid \mu + \phi_j \delta_j, \sigma_j^2), \tag{3.10}$$

where $\mu + \phi_j \delta_j$ is the localization parameter and σ_j^2 is the scale parameter of the distribution. Notice that $\phi_j \delta_j$ shifts the distribution. Also ϕ_j and δ_j are variables which one wants to estimate, and can be modeled using prior information. Considering,

$$\phi_j = \begin{cases} 1 & \text{if } y_j \text{ is an outlier,} \\ 0 & \text{otherwise} \end{cases} \tag{3.11}$$

this means that ϕ_j is a Bernoulli distribution. On the other hand, δ_j is the length of the shift, which means it has to be a localization-scale variable. Therefore

$$\delta_j \sim \mathscr{N}\left(0, \sigma^{*2}\right), \tag{3.12}$$

σ^{*2} is a relatively big variance that should capture the length between the outliers and rest of the data. It should be chosen using the prior knowledge of the length of the outliers.

Using this structure, it is possible to write the posterior distribution as follows:

$$\pi(\theta, \delta, \phi \mid \bar{y}) \propto \prod_{j=1}^{m} f(y_j \mid \theta, \delta, \phi_j) p_j^{\phi_j} (1 - p_j)^{\phi_j} \left[f_{\mathscr{N}\left(0, \sigma^{*2}\right)}(\delta_j) \right]^{\phi_j}, \tag{3.13}$$

since this posterior is not conjugated and is multidimensional, a numerical method needs to be implemented. In this case, an MCMC algorithm is used.

3.4.1 Markov Chain Monte Carlo

Markov Chain Monte Carlo, MCMC, is a series of methods used for sampling from a probability distribution. These methods use a Markov Chain to move in the support of the distribution and obtain a sample from the target distribution. In this thesis, a Metropolis-Hastings algorithm and a Gibbs sampler is to be implemented into the MCMC to estimate the desired variables.

In this case, the distribution in equation 3.13 is the target distribution, a Metropolis-Hastings algorithm is used to move θ and δ , and the instrumental distributions are;

$$\theta \text{ is moved using } \mathscr{N}\left(\theta_i, \sigma_1\right) \tag{3.14}$$

$$\delta_j \text{ is moved using } \mathscr{N}\left(\theta_i - y_j, \sigma_2\right). \tag{3.15}$$

On the other hand, ϕ has to be approached differently, since ϕ is a Bernoulli distribution. This means that the real variable of interest is the probability of success. In this case, the probability of a given datum being an outlier. For this reason, a Gibbs algorithm should be used. So one needs to calculate the probability of ϕ, given the rest of the variables;

$$\mathscr{P}(\phi_j = 1 \mid \theta, \delta, \bar{y}, \phi_{i \neq j}) = \mathcal{K}\mathscr{L}(\theta \mid \delta, \bar{y}, \phi) f_{N(0,\sigma^{*2})}(\delta_j) p_j \tag{3.16}$$

$$\mathscr{P}(\phi_j = 0 \mid \theta, \delta, \bar{y}, \phi_{i \neq j}) = \mathcal{K}\mathscr{L}(\theta \mid \delta, \bar{y}, \phi)(1 - p_j). \tag{3.17}$$

where $\mathcal{K}$ is the normalization constant for equations 3.16 and 3.17. It is possible to obtain the normalization constant, and in consequence, obtain the exact

probability for expressions 3.16 and 3.17. One can obtain this constant as follows:

$$\mathscr{P}(\phi_j = 1 \mid \theta, \delta, \bar{y}, \phi_{i\neq j}) = \frac{\mathscr{L}(\theta \mid \delta, \bar{y}, \phi) f_{N(0,\sigma^{*2})}(\delta_j) p_j}{\mathscr{L}(\theta \mid \delta, \bar{y}, \phi) f_{N(0,\sigma^{*2})}(\delta_j) p_j + \mathscr{L}(\theta \mid \delta, \bar{y}, \phi)(1-p_j)}, \tag{3.18}$$

equation 3.18 gives a way to update the probability in the the Gibbs sampling algorithm.

Using the instrumental distribution describe in 3.18, 3.14,and 3.15, the MCMC method can sample from the posterior distribution described in equation 3.13. It is important to notice that the structure only depends on the data being modeled using a localization-scale distribution.

3.4.2 Normal Model with outliers

To apply the methodology described in the previous chapter, an assumption about the distribution of the data has to be made. In this case, the normal model is assumed,

$$y_j \sim \mathscr{N}(\mu(\theta) + \phi_j \delta_j, \omega_j(\theta)^2), \tag{3.19}$$

Using this assumption, the likelihood can be written as,

$$\begin{aligned}
\mathscr{L}(\bar{y} \mid \theta, \phi_j, \delta_j) &= \prod_{j=1}^{m} f(y_j \mid \theta) \\
&= \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi\omega(\theta)_j^2}} \exp\left(-\frac{(y_j - (\mu(\theta) + \phi_j\delta_j))^2}{2\omega(\theta)_j^2}\right) \\
&\propto \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right] \exp\left(-\frac{1}{2}\sum_{j=1}^{m} \frac{((y_j - \phi_j\delta_j) - \mu(\theta))^2}{\omega(\theta)_j^2}\right),
\end{aligned} \tag{3.20}$$

with the likelihood in equation 3.20, the posterior distribution from which it is desirable to simulate from is proportional to;

$$\pi(\theta, \delta, \phi \mid \bar{y}) \propto \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right] \exp\left(-\frac{1}{2}\left[\sum_{j=1}^{m} \frac{((y_j - \phi_j\delta_j) - \mu(\theta))^2}{\omega(\theta)_j^2} + \phi_j \frac{\delta_j^2}{2\sigma^{*2}}\right]\right) p_j^{\phi_j}(1-p_j)^{\phi_j}. \tag{3.21}$$

3.4.3 T Model with outliers

The last model which this thesis considers is the T model incorporating outliers. This means the distributional assumption is as follows;

$$y_j \mid \theta \sim \mathscr{N}\left(\mu(\theta) + \phi_j\delta_j, \alpha\omega_j(\theta)\right). \tag{3.22}$$

The prior;

$$\alpha\omega_j \mid \theta \sim InvGa(a, b\omega_j(\theta)). \tag{3.23}$$

is now given a random sample of ^{14}C determinations. The likelihood is;

$$\begin{aligned}
\mathscr{L}(\bar{y} \mid \theta, \alpha) &= \prod_{j=1}^{m} f(y_j \mid \theta, \alpha) \\
&= \alpha^{-m/2} \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi\omega(\theta)_j^2}} \exp\left(-\frac{(y_j - (\mu(\theta) + \phi_j\delta_j))^2}{2\alpha\omega(\theta)_j^2}\right) \\
&\propto \alpha^{-m/2} \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right] \exp\left(-\frac{1}{2}\sum_{j=1}^{m} \frac{((y_j - \mu(\theta)) - \phi_j\delta_j)^2)^2}{\alpha\omega(\theta)_j^2}\right).
\end{aligned} \tag{3.24}$$

With this likelihood, we can again follow the procedure described, to obtain a likelihood which depends only in on θ, δ, ϕ, i.e.,

$$\mathscr{L}_{a,b}(\bar{y} \mid \theta, \delta, \phi) = \int_0^\infty \mathscr{L}(\bar{y} \mid \theta, \delta, \phi, \alpha)\pi_{a,b}(\alpha)d\alpha$$

$$\propto \int_0^\infty \prod_{j=1}^{m} \left[\frac{\alpha^{-1/2}}{\omega(\theta)_j}\right] e^{\left(-\frac{1}{2}\sum_{j=1}^{m} \frac{((y_j-\mu(\theta))-\phi_j\delta_j)^2)^2}{\alpha\omega(\theta)_j^2}\right)} \frac{b^a}{\Gamma(a)}\alpha^{-(a+1)}e^{\left(-\frac{b}{\alpha}\right)}d\alpha$$

substituting $\mu_S(\theta, \delta, \phi) = \mu(\theta)) + \phi_j\delta_j$ and following the procedure in section integrated likelihood,

$$\propto \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right] \left[b + \frac{(y_j - \mu_S(\theta, \phi_j, \delta_j)))^2}{2\omega(\theta)_j^2}\right]^{-\frac{m+2a}{2}}$$

; substituting $\mu_S(\theta, \delta, \phi) = \mu(\theta)) + \phi_j\delta_j$ back,

$$\propto \left[b + \frac{((y_j - \phi_j\delta_j) - \mu(\theta))^2}{2\omega(\theta)_j^2}\right]^{-\frac{m+2a}{2}} \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right]. \tag{3.25}$$

Using the likelihood 3.25, the posterior distribution for this model can be written as;

$$\pi(\theta, \delta, \phi) \propto \left[b + \frac{((y_j - \phi_j\delta_j) - \mu(\theta))^2}{2\omega(\theta)_j^2}\right]^{-\frac{m+2a}{2}} p_j^{\phi_j}(1-p_j)^{\phi_j} e^{\phi_j \frac{\delta_j^2}{2\sigma^{*2}}} \prod_{j=1}^{m} \left[\frac{1}{\omega(\theta)_j}\right]. \tag{3.26}$$

3.5 Robustness

Now that the posteriors for each model are expressed, it becomes of interest to figure out how robust the model is in the presence of outliers, which are not that uncommon in the field of radiocarbon dating.

To test the robustness of each model, three sets of artificial data are created. The first one, shown in table 3.1, represents a set of data with no outliers and its antiquity is around 2050 years BP.

TABLE 3.1: Synthetic data around 2050 years BP in F14C.

^{14}Cdetermination	Standard deviation
0.7712	0.0029
0.7695	0.0024
0.7718	0.0033
0.7713	0.0023

The second one, shown in table 3.2, is a set of data around 350 years BP. It represents a set of two outliers which are somewhat distant from the regular measurements. It is important to notice that this data provides an unusual set of outliers, that do not occur frequently in a natural process.

TABLE 3.2: Synthetic data around 350 years BP with outliers.

^{14}Cdetermination	Standard deviation
0.9580	0.0034
0.9579	0.0058
0.9580	0.0074
0.9580	0.0047
0.9579	0.0037
0.9580	0.0045
0.8981*	0.0057
0.8980*	0.0051

Lastly, the set of artificial data, shown in table 3.3, provides a set of data, around 1200 years BP, with really extreme outliers, which are almost impossible to obtain in radiocarbon dating. If such abnormal data are obtained, they would surely be discarded from the calibration. The object of this exercise is to see the behavior of the model in a very unusual scenario.

TABLE 3.3: Synthetic data around 1275 years BP with outliers.

^{14}Cdetermination	Standard deviation
0.8498	0.0034
0.8499	0.0058
0.8499	0.0074
0.8499	0.0047
0.8498	0.0037
0.8499	0.0045
0.6500*	0.0057
0.6500*	0.0045

3.5.1 First Test

The object of the first test, using the first data set, is to determine if the models is suitable for data calibration with no outliers present.

3.5.1.1 Normal model

The first model, calibrated using data in table 3.1, is the normal model. Even though, it was discussed that its assumptions were inaccurate. It is desired to see how it behaves to data with and without outliers. Using the software fexplained, the calibration of the first set of data was created and the results are portrayed in figure 3.1 ;

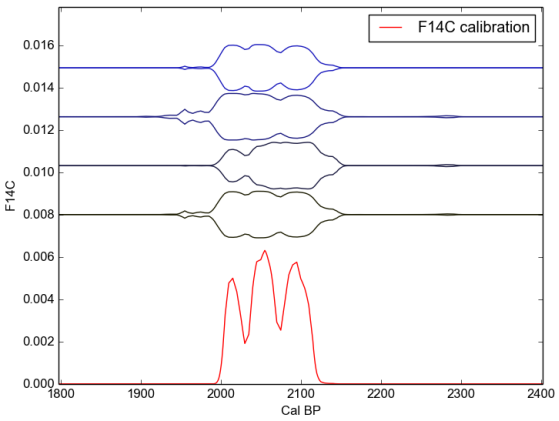


FIGURE 3.1: Calibration of data in table 3.1 using Normal model.

Notice that the calibration includes the real age and gives reasonable intervals around 2000 BP and 2125 BP, which would be a good estimate even if the assumptions are inaccurate.

3.5.1.2 T model

Next, the T model is calibrated with the same data. It is important to notice that this model considers the randomness of the error provided by the laboratory, which is a more robust way to tackle this problem.

Using the software described in chapter 2, the resulting calibration are shown in figure 3.2,

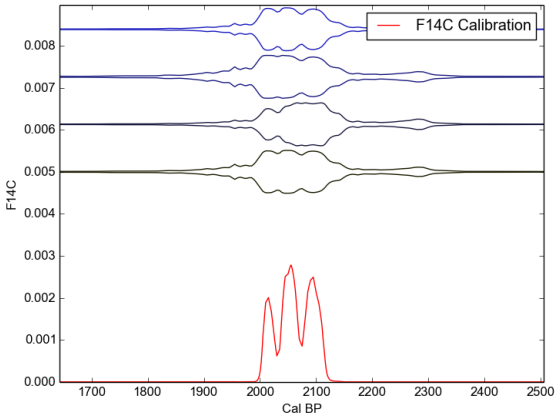


FIGURE 3.2: Calibration of data in table 3.1 using T model.

It can be observed that the calibration looks almost identical and the calibration obtain from the normal model, this is because no outliers are presented.

3.5.1.3 Normal Model with outliers

Even though the Normal Model with outliers is a model specifically built to detect outliers, a calibration is made using this model, with the objective of observing the reaction of the model to these circumstances.

- Burn-in time

  From figure 3.3, it is possible to speculate that the MCMC already converged around 1000 iterations, so a sample from the following iteration would be a proper one.

- Posterior distribution

  From figure 3.4, the sampled posterior can be observed, in black, along with the contribution of each datum, as the colored distributions on top of the

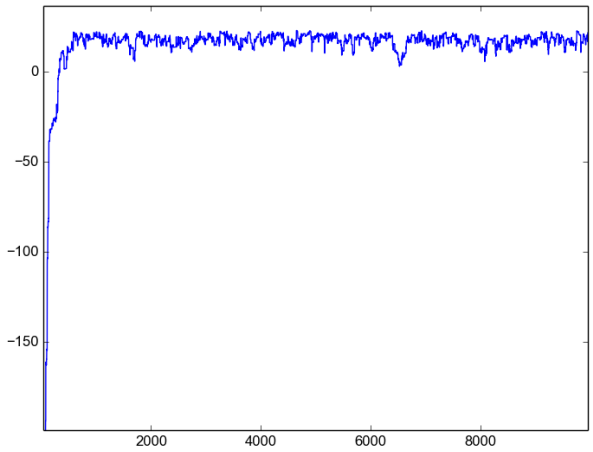


FIGURE 3.3: MCMC's time series for the calibration using the Normal model with outliers.

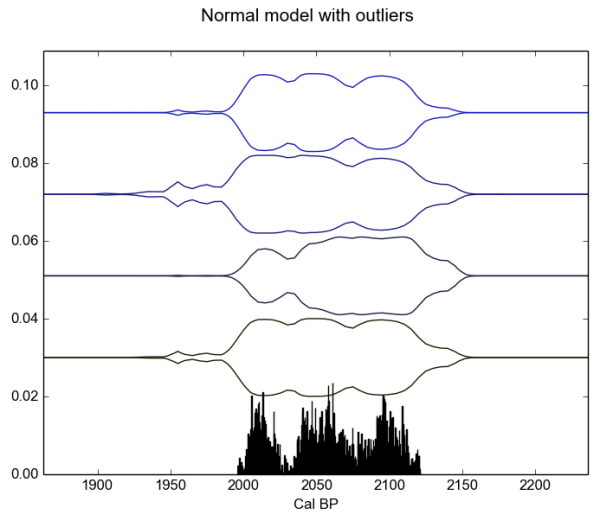


FIGURE 3.4: Calibration of data in table 3.1 using Normal model with outliers.

sample, to the overall posterior. The software also returns the probability of an specific datum been an outlier. In this case, the resulting probabilities are summarized in table 3.4.

TABLE 3.4: Probabilities of a given datum being an outlier.

Data	Probability of being an outlier
0.7712	0
0.7695	0
0.7718	0
0.7713	0

From figure 3.4, it is possible to observe that the distribution is somewhat similar to that of the previous models. However, it is important to remember that only a finite sample from the posterior distribution is observable.

3.5.1.4 T Model with outliers

Just like the normal model with outliers, the calibration with the fist set of data is calibrated as an exercise to observe the reaction of the model to these circumstances.

The software used to obtain the calibration with this model can be found in the appendix.

- Burn-in time

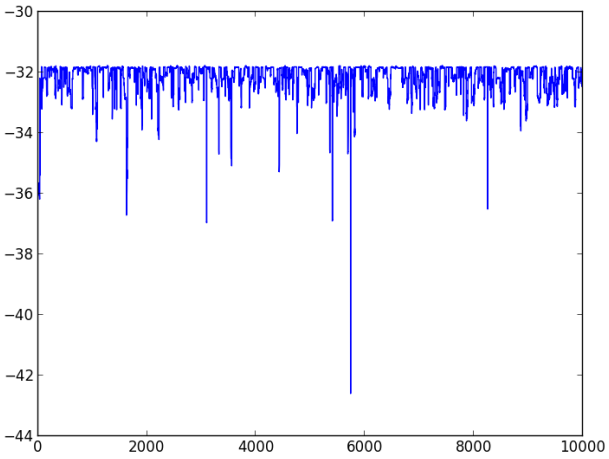


FIGURE 3.5: MCMC's time series for the calibration using the T model with outliers.

From figure 3.5, it is possible to speculate that the MCMC has already converged around 1000 iterations, so a sample from the followed iteration would be a proper one.

- Posterior distribution

From figure 3.6, the sampled posterior can be observed, in black, alone with the contribution of each data, as the colored distributions on top of the

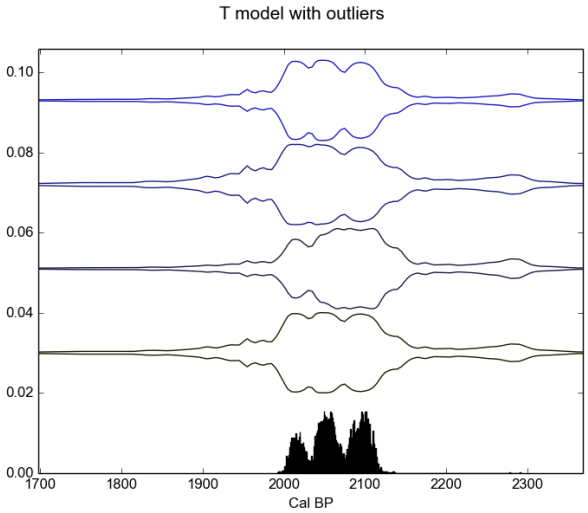


FIGURE 3.6: Calibration of data in table 3.1 using T model with outliers.

sample, to the overall posterior. The software also returns the probability of a specific datum being an outlier. In this case, the resulting probabilities are summarized in table 3.5;

TABLE 3.5: Probabilities of a given datum been an outlier.

Data	Probability of been an outlier
0.7712	0
0.7695	0
0.7718	0
0.7713	0

From figure 3.6 it is possible to observe that the distribution is somewhat similar to that of the previous model. However, just like in the normal model, only a finite sample from the posterior distribution is observed .

3.5.1.5 Comparison

In this exercise, we observed the four models' behavior using data without any real outliers. Although all the models give a good calibration, the last two models became too complex for this task, another problem with these models is the convergence of the method, but most important is the fact that only a sample of the posterior can be observed, which should be avoided if possible. On the

other hand, using the normal and T models, the true posterior distribution can be observed, which is a desirable property. Lastly, it is important to notice that the T model considers the randomness for the error in the calibration, which gives a more realistic approach to the calibration. For this reason, the T model would be the best model to use in a case with data with no suspected outliers.

3.5.2 Second Test

The second test will be performed using the data in table 3.2. This data includes a very clear case of outliers. These kinds of outliers are unusual for a real case but possible, nonetheless. This example provides a way to measure the robustness of the models against outliers which are unusual, but not impossible, in the field.

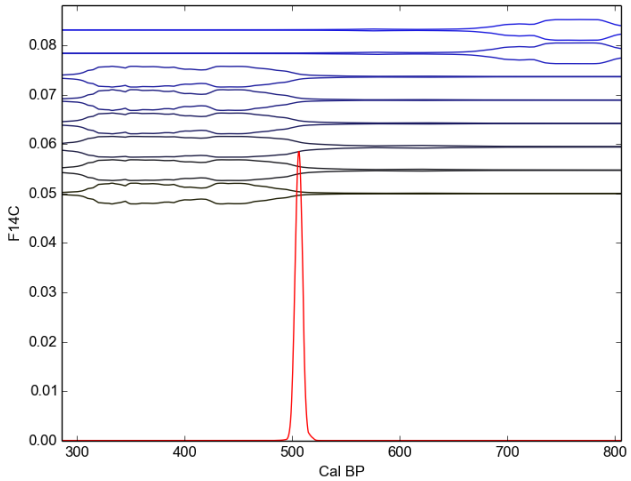
3.5.2.1 Normal model

Even though the normal model does not consider any kind of outliers, the model is tested to observe how big the impact of these outliers is to the calibration. To observe how the outliers affected the calibration, both the calibration considering outliers and without outliers are presented. The resulting calibration is observed in figure 3.7.

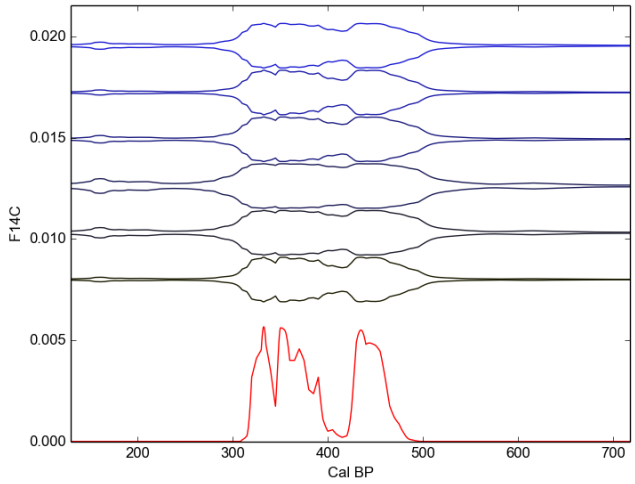
Notice that the calibration is shifted to an older age, and the variance of this posterior is not sufficient to explain the outliers. From this example, it is clear that this model should not be used when any outliers are suspected. Note: since this model did so poorly on this test. There is no point on creating a calibration with even more severe outliers.

3.5.2.2 T model

Since the T model is more robust to outliers, it is expected that the model will behave better under these circumstances than the normal model. The calibration



(A)

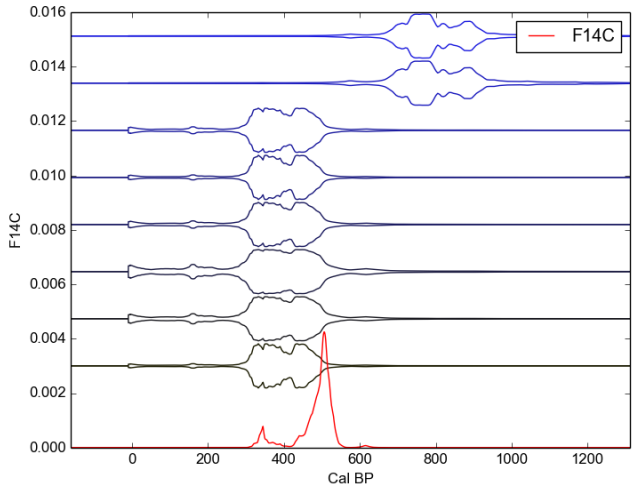


(B)

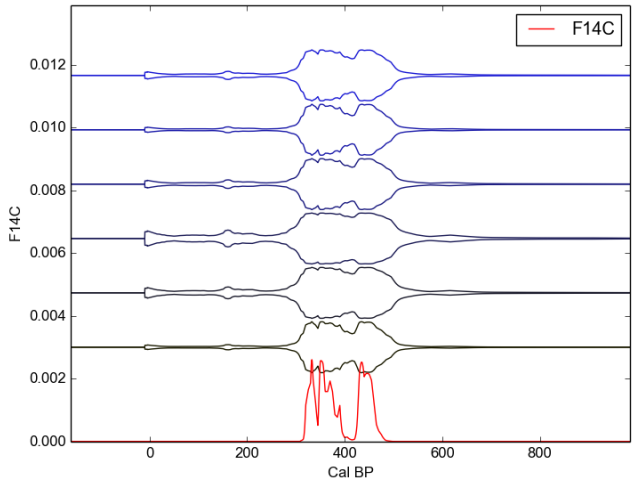
FIGURE 3.7: Calibration using normal model; (A)Calibration considering outliers and Normal models using data from table 3.2. (B) Calibration without outliers using the Normal model with outliers using data from table 3.2.

is shown just like in the normal model, i.e. both calibration with and with no outliers are shown in figure 3.8

It is very important to notice that the distribution for the model does not exclude the true value, and one of its modes lies in this value. Even though the maximum posterior estimation, MAP, is at 505, the probability of the true value is not zero. This example shows how the T model is somewhat robust, even to extreme outliers.



(A)



(B) Calibration without outliers.

FIGURE 3.8: Calibration using T model; (A)Calibration considering outliers and T models using data from table 3.2. (B) Calibration without outliers using the T model with outliers using data from table 3.2.

3.5.2.3 Normal model with outliers

Now, the models with outliers are tested. The first one, the normal model, is calibrated to data with outliers. It is expected to obtain better results. The output of the software using the data are presented next;

- Burn-in time

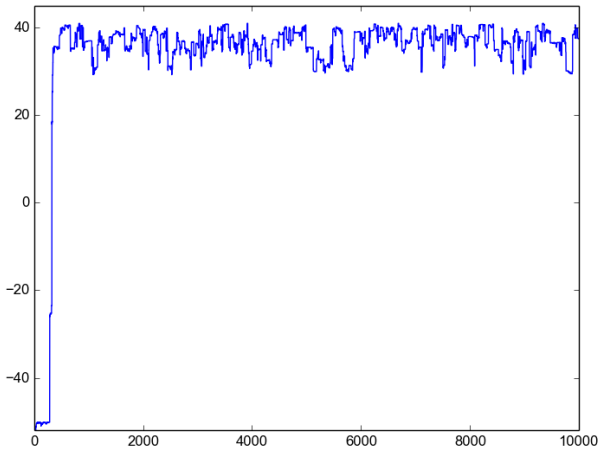


FIGURE 3.9: MCMC's time series for the calibration using the Normal model with outliers.

From figure 3.9, it is possible to speculate that the MCMC converges around 1000 iterations.

- Posterior distribution

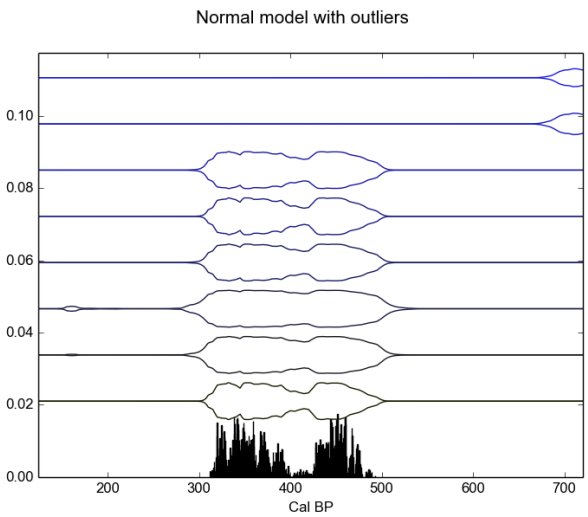


FIGURE 3.10: Calibration of data in table 3.2 using Normal model with outliers.

From figure 3.10, the sampled posterior can be observed, in black, along with the contribution of each data, as the colored distributions on top of the sample, to the overall posterior. The two outliers are shown on the top of the figure. The software also returns the probability of an specific datum

being an outlier. In this case, the resulting probabilities are summarized in table 3.6;

TABLE 3.6: Probabilities of a given datum been an outlier.

Data	Probability of it been an outlier
0.9580	0
0.9579	0
0.9580	0
0.9580	0
0.9579	0
0.9580	0
0.8981*	1
0.8980*	1

From table 3.6 and figure 3.10, it becomes clear that the outliers are identified and corrected. This indicates that the model works as expected.

It is very important to observe that the posterior described in figure 3.10 is very similar to the posterior shown in figure 3.7b. This is the posterior of the data using the normal model without the data suspected of containing outliers. The one in figure 3.8b is the posterior using the T model, excluding the outliers. This shows that the model does correct unusual outliers.

3.5.2.4 T model with outliers

The last model to be used to calibrate the data is the T model with outliers. This model considers two robust approaches; the T model presented previously and the shift of outliers. The results are now presented;

- Burn-in time

  From figure 3.11, it is possible to speculate that the MCMC converges around 3000 iterations.

- Posterior distribution

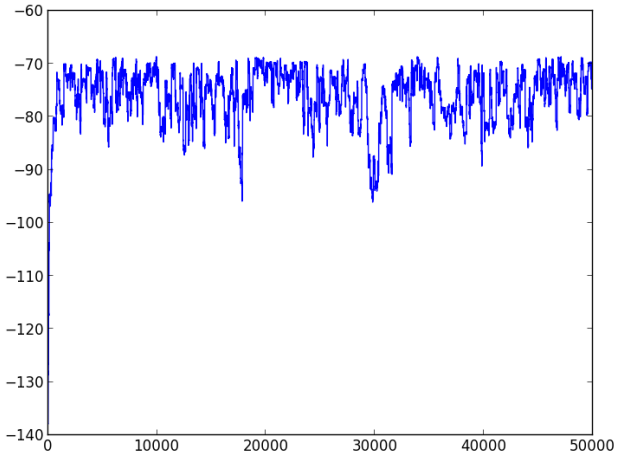


FIGURE 3.11: MCMC's time series for the calibration using the T model with outliers.

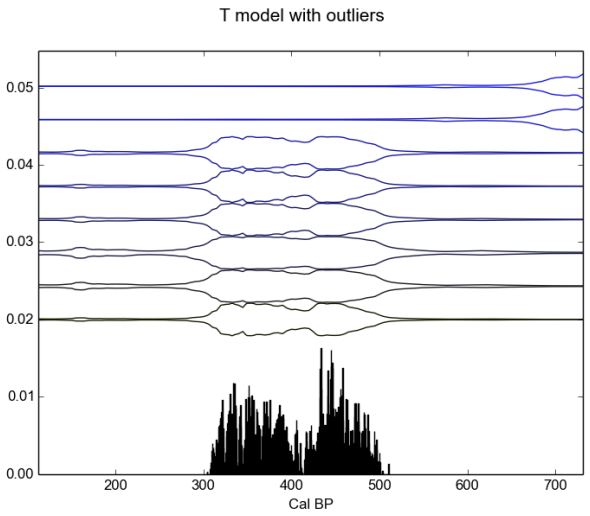


FIGURE 3.12: Calibration of data in table 3.2 using T model with outliers.

From figure 3.12, the sampled posterior can be observed, in black, alone with the contribution of each data, as the colored distributions on top of the sample, to the overall posterior. The two outliers are shown on the top of the figure. The software also returns the probability of a specific datum been an outlier in this case the resulting probabilities are summarized in table 3.9.

As it was expected, the model finds and corrects the outliers and provides a posterior. The posterior looks very similar to when outliers are removed by hand from the sample, but not any better than that of the normal model with outliers. This

Table 3.7: Probabilities of a given data being an outlier.

Data	Probability of it being an outlier
0.8498	0
0.8499	0
0.8499	0
0.8499	0
0.8498	0
0.8499	0
0.6500*	1
0.6500*	1

does not sound reasonable at first sight, since both models, normal and T, give similar results when no outliers are presented.

3.5.2.5 Comparison

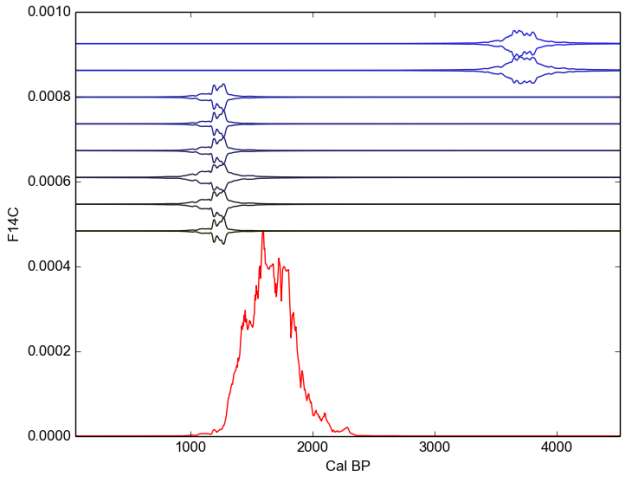
Using data with outliers shows how the models behave in an unusual, but not impossible, environment. To discuss the results, first we take on the normal model. It gives very poor results, and it should not be used when any kind of outliers are suspected. On the other hand, the T model gives a more accurate posterior, but is somewhat affected by the unusual outliers. On the other hand, let's notice that both models find and correct the outliers. This means that the calibration is proportional to that of the model used when no outliers are presented. Since both models, Normal and T, gave similar results when no outliers were presented, and the T model describes the variability of the phenomena in a more realistic way, the T model would give a more realistic calibration with or without the outlier structure.

3.5.3 Third test

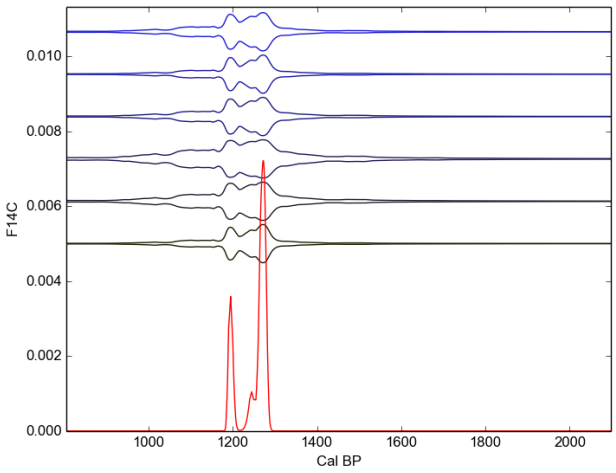
This set of data contains a very particular set of outliers. This comes from the fact that these outliers would be extremely rare to the point of being almost impossible to obtain in a regular scenario.This test will only be an exercise to observe how the models react to data with this kind of outliers.

3.5.3.1 T model

Like in the previous test, the model is compared using the same data, considering outliers and without them. The results are presented in figure 3.13. The distribution in this case is affected by the outliers, but the true value does not have zero probability. Like the previous test, the model does concentrate the probability on a shift mode and the variance becomes quite large to to obtain a reasonable calibration.



(A) Calibration of data in table 3.3 considering outliers, and using the T model.



(B) Calibrationin in table 3.3 without outliers, and using the T model.

FIGURE 3.13: Calibration using T model.

3.5.3.2 Normal model with outliers

Since this model was created to identify and shift the outliers, the expected results would not vary from the previous set of data. It would be assumed that the model would identify the outliers more easily since they are much more different from the rest of the data. The resulting calibrations for this model are the following;

- Burn-in time

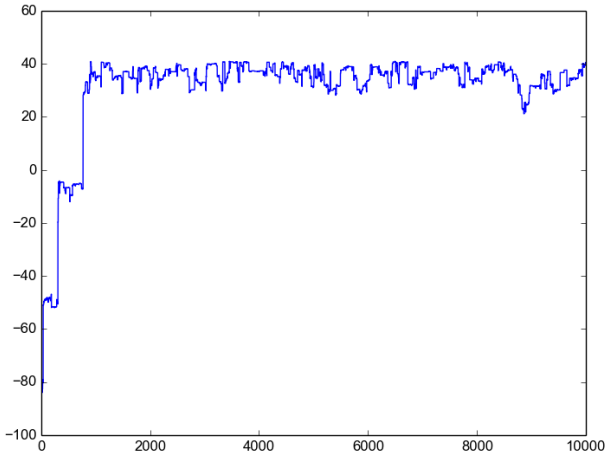


FIGURE 3.14: MCMC's time series for the calibration using the Normal model with outliers.

  From figure 3.14, it is possible to speculate that the MCMC has converged around 1000 iterations.

- Posterior distribution

  From figure 3.15, the sampled posterior can be observed, in black, alone with the contribution of each datum, as the colored distributions on top of the sample, to the overall posterior. The two outliers are shown on the top of the figure. The software also returns the probability of a specific datum being an outlier. In this case, the resulting probabilities are summarized in table 3.8.

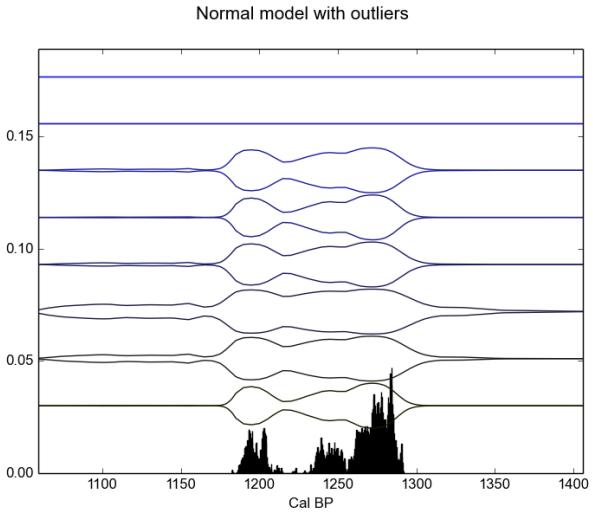


FIGURE 3.15: Calibration of data from table 3.3 using Normal model with outliers.

TABLE 3.8: Probabilities of a given datum been an outlier.

Data	Probability of it been an outlier
0.8498	0
0.8499	0
0.8499	0
0.8499	0
0.8498	0
0.8499	0
0.6500*	1
0.6500*	1

As expected, the model finds and shifts the outliers, even with such extreme outliers.

3.5.3.3 T model with outliers

Finally, the T model is tested using the same data. The results are now shown.

- Burn-in time

  From figure 3.16, it is possible to speculate that the MCMC converges around 3000 iterations.

- Posterior distribution

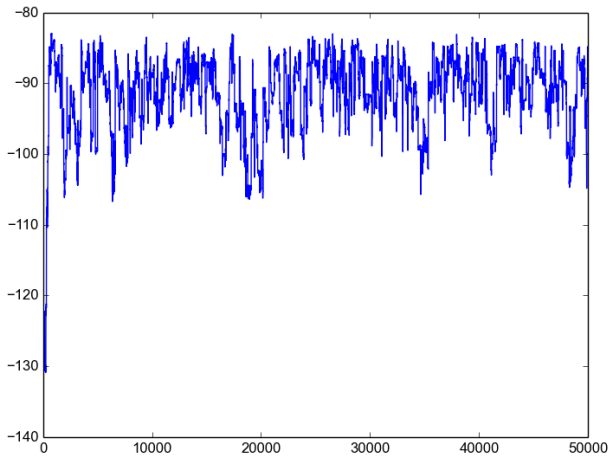


FIGURE 3.16: MCMC's time series for the calibration using the T model with outliers.

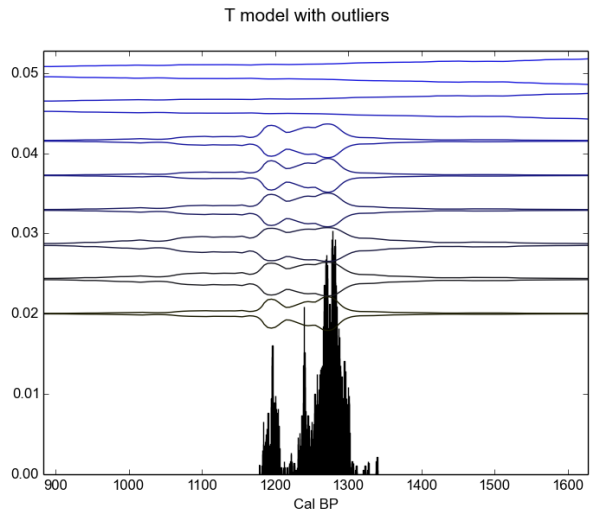


FIGURE 3.17: Calibration of data from table 3.3 using T model with outliers.

From figure 3.17, the sampled posterior can be observed, in black, along with the contribution of each datum, as the colored distributions on top of the sample, to the overall posterior. The two outliers are shown on the top of the figure. The software also returns the probability of a specific datum been an outlier, in this case the resulting probabilities are summarized in table 3.9.

Just like in the previous test, the distribution in this case is very similar to that of the normal model with outliers.

TABLE 3.9: Probabilities of a given datum being an outlier.

Data	Probability of it being an outlier
0.8498	0
0.8499	0
0.8499	0
0.8499	0
0.8498	0
0.8499	0
0.6500*	1
0.6500*	1

3.5.4 Discussion

The results from this test are very similar to that of the previous test. This means that most of the assumptions that are observed in cases with unusual outliers do not change when looking at more extreme outliers. This is a desirable property to have in a robust model.

3.6 Conclusions

This chapter discussed a very common problem in radiocarbon calibration, which is the presence of outliers in the data. The problem was approached using four different models; Normal, T, Normal with outliers, and T with outliers. A series of tests were performed to the models, which depended on the length of the shift per outlier. The data from each test can be found on tables 3.1, 3.2, 3.3.

Using the results from the tests, it becomes clear that three models are remarkable useful; the T model, Normal model with outliers, and T model with outliers. These three models show robustness to outliers. Even if the three models gave favorable results, however, not all models should be used; a proper way to calibrate would be to identify the situation and select a model based on those assumptions, for this an analysis of the situation is described.

First, let us comment on the robustness of the T model. This model, when presented with data, gives a deterministic distribution which may be obtained without

the use of any kind of numerical method. This means faster and more precise results for the calibration. Unfortunately, when presented with unusual cases of outliers, the model provided a posterior which had bigger variance depending on the shift of the outliers. This is not desirable at all. This model was not created with the purpose of discriminating outliers, but to consider the variability of the data; for this reason, the calibration is drastically affected. It is remarkable that this model, even if it was not created with the purpose of considering outliers, does not discard the true value of the data, like in the case of the normal model. For this reason, the T model should be used when presenting data which is not suspected to contain outliers.

Next, the normal and T models with structure to consider outliers, are specifically built to discriminate outliers. Both models identify the outliers very efficiently and shift them in an acceptable way. Both cases proved similar resulting posterior distributions to similar calibrations. This leads to a discussion on which model is more appropriate to use.

As it was mentioned on section 4, the structure to identify and shift the outliers is not dependent on the distribution used to described the data, which leads to the assumption that the distribution which should be used is that which describes the data in a more realistic way. As it was mentioned in Christen and Pérez (2009), and introduced in section 3 of this work, the T model is a model which describes the variability of the data in a more realistic way, for this reason the T model is a better choice to describe the variability of the data and, consequently, the model that should be used to calibrate data with outliers in combination with the outlier structure mentioned in section 4. Another advantage that the T model provides against the normal model is related time to the convergence of the method. Since the T model has heavier tails, the convergence of the method is faster when starting from a farther initial point. For this reason the T model using the outlier structure should be used to calibrate data which contains outliers, in an effort to detect and correct the outliers, which leads us to use the information regarding the variance of the data which this data contains.

Finally, it is important to notice that the models with outliers can only be solved using the numeral method. This leads to a dependence on the convergence of the method. The user has to speculate about the convergence of the method, and also the method only provides a finite sample from the posterior. For these reasons, the models with outliers, if possible, should be avoided. This means that if no outliers are suspected in the data the structure for outliers should not be used, and instead T model should be used to calibrate the data; if not for other reasons simply for the sake of parsimony.

Chapter 4

Real Data Calibration

Up to now the calibration models have been tested using artificial data. In this last chapter, two examples with real data are used to test the calibration reliability. For this purpose, two sets of data are used: the first test comes from 'The School of Geography, Archaeology and Palaeoecology at Queen's University Belfast',(Dr. Ron Raimer, AMS radiocarbon LAB, personal communication), and consist of 85 F14C data obtained from the laboratory in 2013, and 52 Radiocarbon age data obtained in 2014. Both sets of data belong to the same calendar age. The second set of data used is from the Shroud of Turin obtained in 1988, Damon P. E., et al, 1989.

Both sets of data are calibrated using the calibrations obtained in Chapters 2 and 3 respectively, in the case of the Shroud of Turin's data an analysis using outliers is performed, for this reason the T model with outliers is used to calibrate the data. On the other hand, the data from Queen's University is not suspected to contain outliers and for this reason the T model is sufficient to describe the variability of this data, as was justified in chapter 2 of this work.

4.1 Queen's University's data

The data comes from one single sample all dates correspond to the same calendar year. The data consists of 54 Radiocarbon age measurements along with 54 F14C measurements, the data points were obtained on different dates and in the same laboratory, therefore they can be considered independent of each other.

Table 4.1: Queen's University's Radiocarbon age data,(Conventional dates).

age	ageσ	date run	age	ageσ	date run
1378	24	2014-01-17	1443	22	2014-01-17
1451	24	2014-01-22	1434	26	2014-01-22
1465	25	2014-01-25	1435	25	2014-01-25
1512	26	2014-01-30	1478	23	2014-01-30
1365	26	2014-01-31	1453	30	2014-01-31
1452	26	2014-02-07	1560	35	2014-02-07
1449	23	2014-02-20	1449	34	2014-02-20
1442	29	2014-02-21	1415	24	2014-02-21
1433	22	2014-03-04	1433	24	2014-03-04
1526	31	2014-03-07	1464	26	2014-03-07
1415	24	2014-03-10	1444	24	2014-03-10
1503	31	2014-03-14	1442	26	2014-03-14
1390	25	2014-03-24	1383	26	2014-03-24
1435	23	2014-04-09	1505	25	2014-04-09
1453	24	2014-04-10	1466	25	2014-04-10
1414	19	2014-04-29	1404	24	2014-04-29
1362	33	2014-05-02	1460	34	2014-05-02
1410	32	2014-05-07	1474	24	2014-05-07
1433	40	2014-05-13	1443	27	2014-05-13
1427	30	2014-05-16	1448	22	2014-05-16
1406	24	2014-05-22	1423	24	2014-05-22
1396	23	2014-06-18	1395	20	2014-06-18
1392	23	2014-06-20	1511	21	2014-06-20
1413	24	2014-06-26	1399	24	2014-06-26
1411	34	2014-06-27	1399	33	2014-06-27
1448	20	2014-07-02	1425	19	2014-07-02

Table 4.2: Queen's University's Radiocarbon age data, (F14C data).

F14C	F14Cσ	date run	F14C	F14Cσ	date run
0.8371	0.0029	2013-01-21	0.8320	0.0023	2013-02-01
0.8361	0.0027	2013-02-07	0.8434	0.0034	2013-02-07
0.8334	0.0027	2013-03-06	0.8361	0.0027	2013-06-07
0.8334	0.0025	2013-03-06	0.8328	0.0030	2013-03-14
0.8389	0.0027	2013-03-15	0.8368	0.0026	2013-03-15
0.8367	0.0026	2013-05-01	0.8384	0.0028	2013-05-02
0.8439	0.0024	2013-05-13	0.8337	0.0025	2013-05-13
0.8395	0.0024	2013-05-29	0.8317	0.0027	2013-05-31
0.8301	0.0028	2013-05-02	0.8355	0.0028	2013-05-29
0.8393	0.0025	2013-06-08	0.8312	0.0026	2013-06-22
0.8322	0.0025	2013-06-27	0.8258	0.0030	2013-06-29
0.8374	0.0029	2013-07-05	0.8313	0.0040	2013-07-05
0.8354	0.0030	2013-07-09	0.8393	0.0024	2013-07-17
0.8346	0.0024	2013-08-05	0.8379	0.0026	2013-08-09
0.8411	0.0025	2013-08-09	0.8328	0.0024	2013-08-13
0.8362	0.0024	2013-08-15	0.8401	0.0031	2013-08-20
0.8325	0.0024	2013-08-20	0.8328	0.0024	2013-09-19
0.8387	0.0026	2013-09-20	0.8347	0.0029	2013-09-20
0.8382	0.0022	2013-09-21	0.8284	0.0024	2013-09-25
0.8359	0.0022	2013-09-25	0.8349	0.0027	2013-05-31
0.8313	0.0035	2013-11-08	0.8486	0.0026	2013-09-21
0.8240	0.0028	2013-11-14	0.8366	0.0046	2013-11-15
0.8328	0.0024	2013-11-21	0.8385	0.0032	2013-07-09
0.8358	0.0042	2013-11-22	0.8428	0.0026	2013-12-12
0.8276	0.0032	2013-12-12	0.8366	0.0046	2013-12-13
0.8376	0.0022	2013-12-19	0.8427	0.0027	2013-07-17

Using the data shown in tables 4.1 and 4.2 both Radiocarbon age and F14C calibrations are performed to observe the reliability of the F14C calibration in comparison to the traditional calibration.

4.1.1 Calibration

As it is mentioned in chapter 3, the optimal way to calibrate data, is first to calibrate it using the T model, since this model shows robustness to certain outliers, and then if outliers are suspected the calibration with the T model with outliers is recommended.

To calibrate the data first it is important to set a model which will be used for the calibration. In this case no outliers are suspected in the sample and for this reason the T model is sufficient to describe the variability of the samples. After agreeing on the model which is going to be used for the calibration the calibration is performed.

In figures 4.2 and 4.1 the calibration of Radiocarbon age and F14C data can be observed respectively.

Since the calibration is not too old, the calibrations are very similar, as can be expected; furthermore, the calibration for the F14C data presents a higher density around 1325, this can be explained because the F14C data sample is bigger than the Radiocarbon age data, and as more data is collected a higher precision is expected.

To have a better comparison both curves are combined, and are shown in figure 4.3

4.2 Comparison Data

As it was mention in chapter 2, the calibration which is used in this thesis strongly depends on the calibration curve which was created in in chapter 2, also in this chapter, a comparison was used to see the effect that the calibration curve and

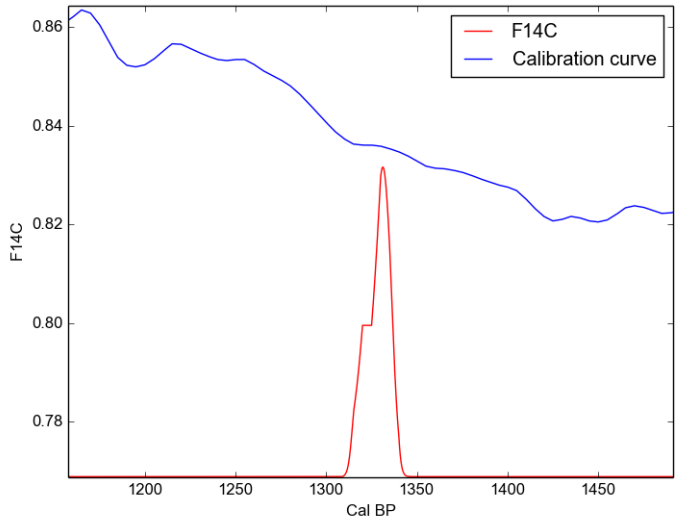


FIGURE 4.1: F14C calibration of data in table 4.2.

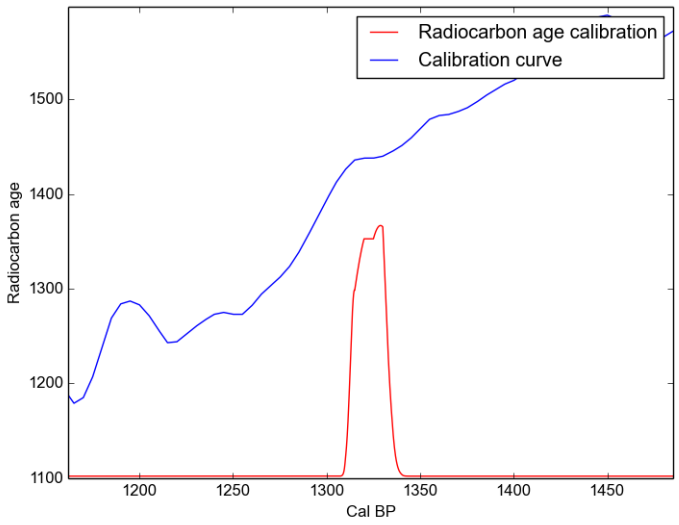


FIGURE 4.2: Radiocarbon age Calibration of data in table 4.1

the incorrect assumption of a normal distribution can have on the calibration. In this section some data obtained from Queen's university is used to confirm the suspicions which arise in chapter 2.

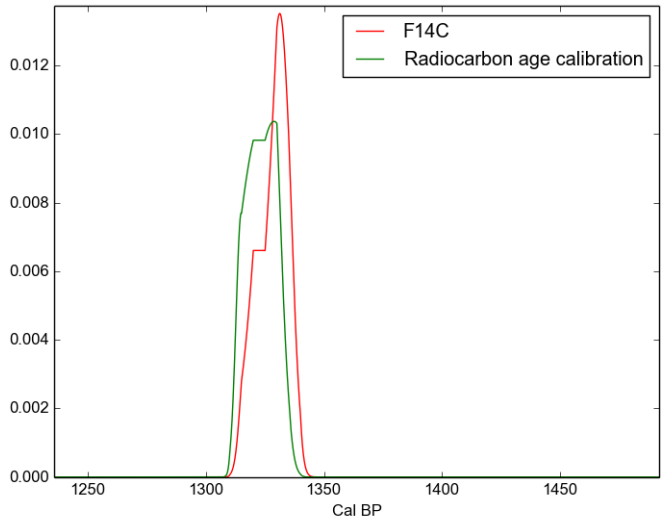


FIGURE 4.3: Calibrations comparison between F14C and Conventional Radiocarbon age.

4.2.1 Data

The data obtained for this section are five samples with different dates, both the F14C and the radiocarbon age are reported and both are from the same sample. This set of data will provide a good idea of the behavior of the calibration around different ages. The data is presented in table 4.3

TABLE 4.3: Queen's University's Radiocarbon age data.

Radiocarbon age	ageσ	f14c	f14cσ
636	30	0.9239	0.0035
1463	32	0.8335	0.0033
4776	39	0.5518	0.0027
19769	74	0.0854	0.0008
38693	632	0.0081	0.0006

To calibrate data, 3 scenarios are used: Data before 1390, data between 1391 and 25000, and data older than 25000. These scenarios are chosen because 'Intcal13', Reimer et al. (2013), has a resolution of 5 years in the first scenario, 10 years in the second, and 20 years in the last. Since both calibration curve follow this resolution, these scenarios seen reasonable.

4.2.2 Young Data

The first scenario to be studied is for data before 1390 BP. The data which have this property are the first 3 samples in table 4.3. The calibrations are presented in the same order as that of the data set.

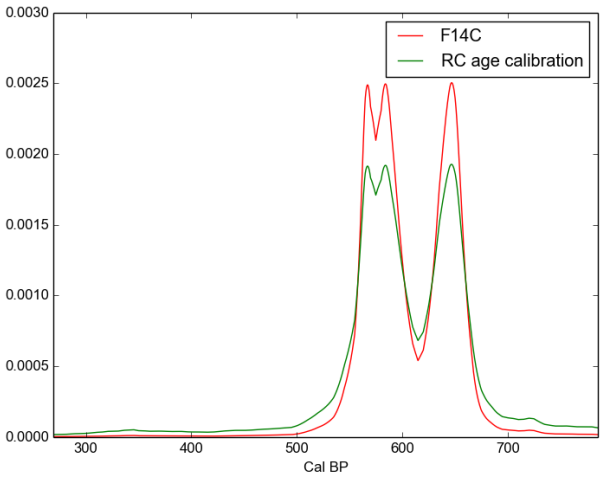


FIGURE 4.4: Calibration of data around 600 years BP.

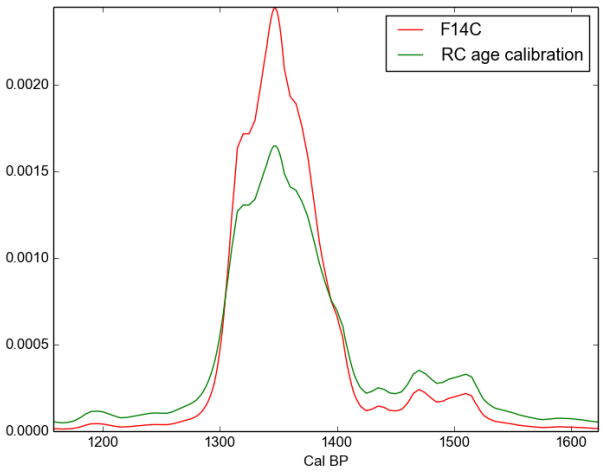


FIGURE 4.5: Calibration of data around 1350 years BP.

From the calibrations in figures 4.4 ,4.5 ,4.6 , it is possible to observe that calibrations are really similar and reach their modes in the same values. This was expected as similar results are obtained in chapter 2.

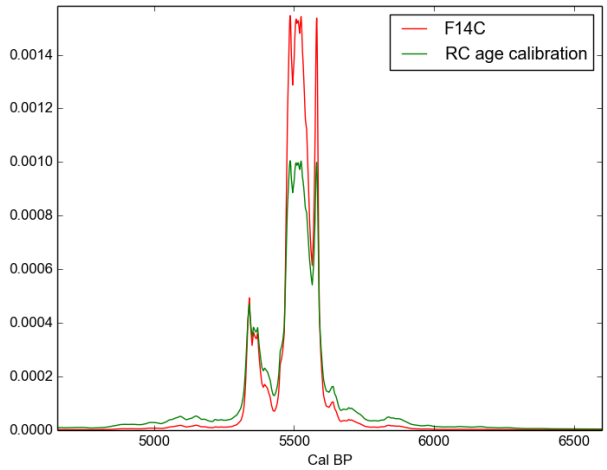


FIGURE 4.6: Calibration of data around 5500 years BP.

4.2.3 Mid-age Data

In this case only one datum follows the condition, $[19769, 74]$ and $[0.0854, 0.0008]$, the calibration is shown in figure 4.7, just like in the previous case, the calibrations are very similar, and reach their mode at the same age, which would indicate that even in mid-age data the calibration still gives a good estimate.

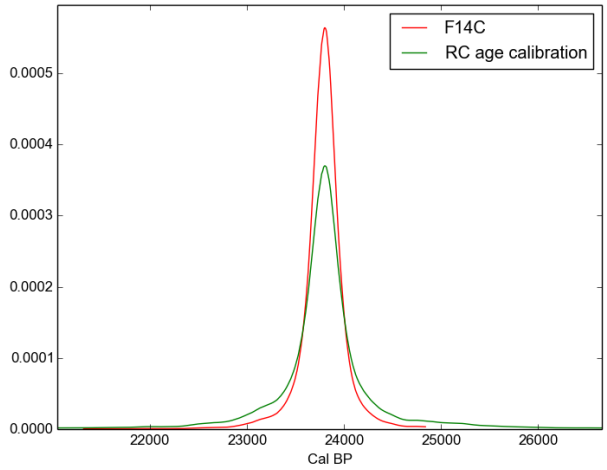


FIGURE 4.7: Calibration of data around 24000 years BP.

4.2.4 Older Data

In the case of the older data one sample is provided $[38693, 632]$ and $[0.00810.0006]$ it is important to notice that these data have a big standard deviation which would be expected in a calibration of this magnitude, the calibration of this data is presented in figure 4.8.

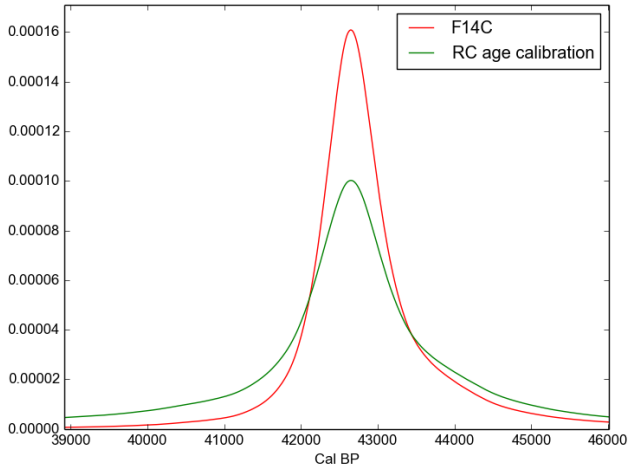


FIGURE 4.8: Calibration of data around 42000 years BP.

Like in the excercise provided in chapter 2, the older calibration differs from that of the traditional calibration more than younger calibrations, this is expected since the calibration curve is only an approximation and would affect the calibration.

4.2.5 Conclusion

As it was expected, both calibrations seem to provide very similar results. It is important to notice that the F14C calibration depends on 'Intcal13'. Even with the dependence on the calibration curve the F14C still gives a good estimate. for the time been, it would be recommended to stick with the traditional calibration, simply because of the dependence in 'Intcal13' in the F14C calibration, but the creation of a calibration curve in terms of F14C is strongly suggested, so the true benefits of this calibration can be explored.

4.3 The Shroud of Turin

"The Shroud of Turin or Turin Shroud (Italian: Sindone di Torino) is a length of linen cloth bearing the image of a man who appears to have suffered physical trauma in a manner consistent with crucifixion." Wikipedia. 2014.

FIGURE 4.9: A piece of the Shroud of Turin.
Wikipedia 2014.

Many people have speculated that the Shroud, see figure 4.9, could be the burial shroud of Jesus of Nazareth. For this reason, the shroud has been the object of an intense debate amount theologians, historians and researchers, until finally in 1988 a sample from the shroud was taken to be dated using radiocarbon dating. For this study, Three radiocarbon dating laboratories were involved, Arizona, Oxford and Zurich. The obtained data is presented in table 4.4, taken from the original publication Damon P.E, et al 1989.

Later on, Christen (1994) used this data to test the normal model with outliers, using the traditional radiocarbon age format. The intention of using this date in this work is to give a real example of calibrating data which may contain outliers using F14C.

It is important to notice that the data was reported originally in 'radiocarbon age'. For this reason, a transformation back to F14C is necessary. Using the same

TABLE 4.4: Radiocarbon determinations for the Shroud of Turin obtained from Damon P. E., et al (1989).

Laboratory	^{14}C age	Standard deviation
Arizona	591	30
	690	35
	606	41
	701	33
Oxford	795	65
	730	45
	745	55
Zurich	733	61
	722	56
	635	57
	639	45
	679	51

transformation described in chapter 2, while working with the F14C calibration curve, the data present in table 4.5 is obtain;

TABLE 4.5: Radiocarbon F14C determinations for the Shroud of Turin, obtained by approximation.

Laboratory	^{14}Cdetermination	F14C	F14Cσ
Arizona	591	0.9290	0.0034
	690	0.9176	0.0074
	606	0.9273	0.0047
	701	0.9164	0.0037
Oxford	795	0.9057	0.0073
	730	0.9131	0.0051
	745	0.9114	0.0062
Zurich	733	0.9127	0.0069
	722	0.9140	0.0063
	635	0.9239	0.0064
	639	0.9235	0.0051
	679	0.9189	0.0058

This data can now be calibrated using the same software as the artificial data in chapter 3.

4.3.1 Calibration

Again to calibrate this data it is important to first select a model. Since the data is suspected to contain outliers the model which is going to be used is the T model

with outliers. To perform the calibration lets remember that this requires MCMC since the posterior can only be obtained with this method. For this reason first the graph of the convergence of the method is presented in figure 4.10

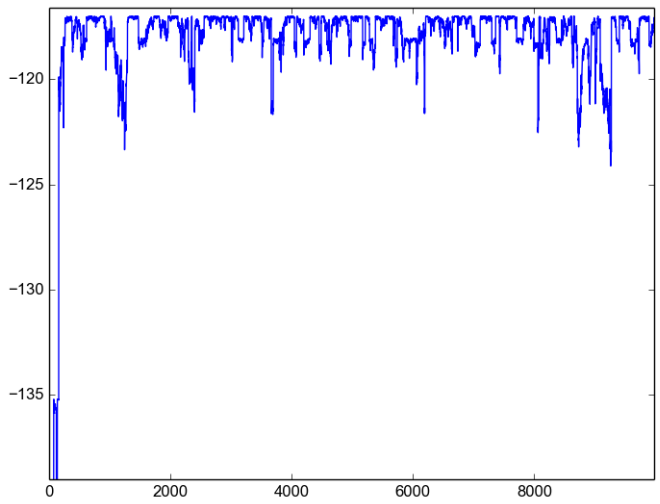


FIGURE 4.10: MCMC time series for the calibration of the Shroud of Turin data.

From figure 4.10 it is possible to speculate that the chain has converged by the 2000th iteration, and the rest of the sample can be taken as a sample from the posterior distribution of interest.

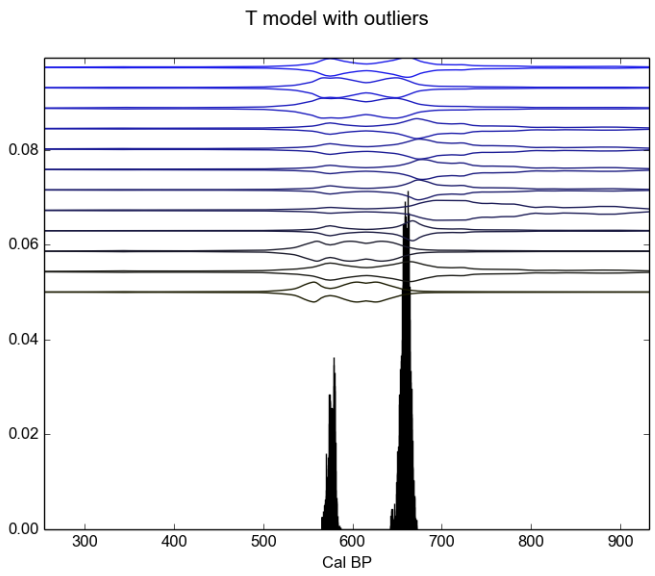


FIGURE 4.11: Calibration of the Shroud of Turin data using F14C.

In this case, the calibration does not identify outliers which would be expected since the T model is robust to this type of outliers, using this model the Shroud is placed between 563 and 673 years before present (1950 AD) which translates to a date between $(1275, 1388)$ AD; furthermore, the calibration is consistent with that of the original study which placed the Shroud around $(1260, 1390)AD$.

4.4 Conclusions

As it was expected, both F14C calibrations return results similar to that of the traditional calibration, this derives from several factors like;

- The calibration is somewhat recent, and as it shown in chapter 2, for recent calibration the error in assuming normality is not noticeable.
- The F14C calibration curve is an approximation obtained from 'intcal13' which leads to the F14C calibration been directly linked to the 'radiocarbon age' calibration.

From these two examples we have shown that the F14C calibration is competitive with traditional calibration even when no calibration curve has been developed for this calibration and an approximation is needed. Also, the F14C calibration is a more proper way to calibrate data, since this is the data obtained from the laboratory and it would not confuse non-experts.

Chapter 5

Discussion

The main objective for this thesis is the comparison of calibration in terms of F14C and Radiocarbon Age. To achieve the comparison; a provisional calibration curve is created alone with a calibration in terms of F14C, based on models used in Radiocarbon Age. The comparison is done using the error obtain from wrongly assuming Normality and also using artificial and real data.

To calibrate the F14C data, a model had to be chosen; in this case the T model is used for its robustness. After choosing a model the necessity of an F14C calibration curve arises, and since no such curve exists at the moment, an approximation of such a curve was created in section 3 of chapter 2. It is important to notice that the obtained curve is an approximation and it is strongly suggested that a true curve, in terms of F14C, is created in the near future. After obtaining an acceptable curve, F14C calibration is introduced in chapter 2 section 4. To justify the use of F14C calibration, the assumptions in traditional calibration are questioned. In section 5, the error in assuming normality in a transformation of F14C is explored. Given the results from these analyses, F14C calibration is encouraged, since this calibration provides a more transparent process, and is free of errors from wrongly assumed normality. In addition to the analysis, the format for reporting F14C is discuss and a subjection on such format is introduce in the append of this work. This last reason is very important because when a laboratory reports data in the traditional radiocarbon age, inexperienced users would not see

a necessity to calibrate the data, and this would lead to wrong interpretations thereof; these incorrect interpretations would no occur when using F14C data from the beginning.

In Chapter 3, Four models are constructed and tested with outliers, since outliers are very common in radiocarbon data these tests are very important to give an objective opinion on which model to use when calibrating data, whether it is F14C or radiocarbon age data. The models are constructed in section 2, and are tested in section 3. The models were tested using three sets of data, where two of them were purposely corrupt with outliers. After testing the four models, the T model performed reasonable well, even when presented with outliers, the other models that give favorable results were the models with a structures for outliers, these models use the normal and T likelihood; both models gave similar results, but it is important to notice that the models with outlier structure only consider a shifting of the data but the model behind it has to represent the variability of the data if this data was not a outlier. For this reason the use of the T model would also be considered to be more appropriate. Furthermore, since the T model has heavier tails, convergence of the MCMC would be faster. In conclusion, the data should be calibrated using the T model when the data is not suspected to contain abnormal outliers. When outliers are suspected, the T model with the outlier structure should be used, but only in cases when outliers are clear so the information of the variance which this data provides would not be lost.

In Chapter 4 uses real data to compared F14C and Radiocarbon Age calibrations. First, two sets of data obtained from the same calendar year, obtained in both F14C and Radiocarbon Age, are used to compare the results from both calibrations in a real scenario. Second, observations, both in F14C and Radiocarbon Age, dating from different dates are used to observe the effects of antiquity in both calibrations. Last, a popular example is use to observe the general performance of F14C calibration. The data used in this example is from P.E Damon, et al (1989) related to The Shroud of Turin. From this real scenarios, it is observed that the F14C calibration is competitive compared to the traditional Radiocarbon Age, obtaining similar results. It is important to remark the fact that F14C is very

dependent on the approximation to the calibration curve obtained in Chapter 2, and the true benefits of using F14C can not be fully explored until a calibration curve for F14C is developed.

In conclusion, this work has analyzed two important aspects of Radiocarbon dating, which are calibration of F14C and the selection of models in different scenarios, along with a provisional F14C calibration curve. It would be desirable that this work inspire the creation of a true F14C calibration curve, which would be a step in the direction of calibrating using the true data obtained in the laboratory and eradicate the confusion which radiocarbon age could generate in the general public.

Appendix A

Precision in F14C

When working with Radiocarbon Age the data is rounded to their nearest unit, i.e. if the laboratory obtain a calibration of 1233.3, 1233 would be reported, it becomes important to have a idea of how many decimals in the F14C calibration, this precision would implied. For this purport a similar figure to figure 2.4 was created. In this case, only old date is taken into account because the F14C would have lower count.

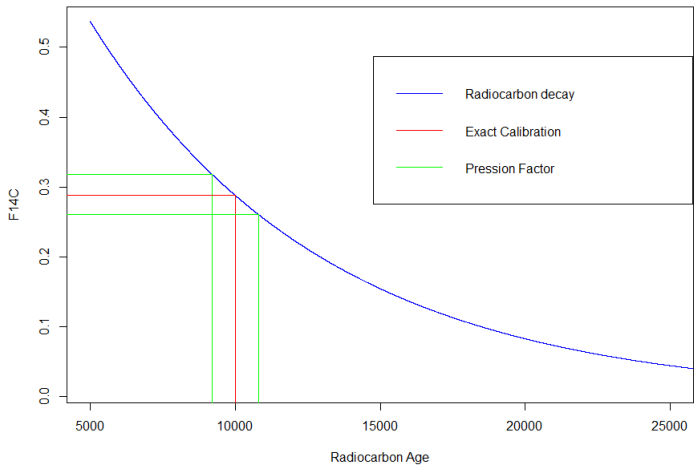


FIGURE A.1: Law of Radiocarbon Decay with precision factor.

In figure A.1 the blue figure represents the Law of Radiocarbon Decay, if x is observed, red in figure, it is desirable to see how many decimals in F14C would be involve when rounding to the nearest unit, in the figure this would be the

difference between the red and green line. It is important to notice that this distance increases while carbon count decreases, for this reason only low counts are observed since newer data would obtain such precision with less decimals.

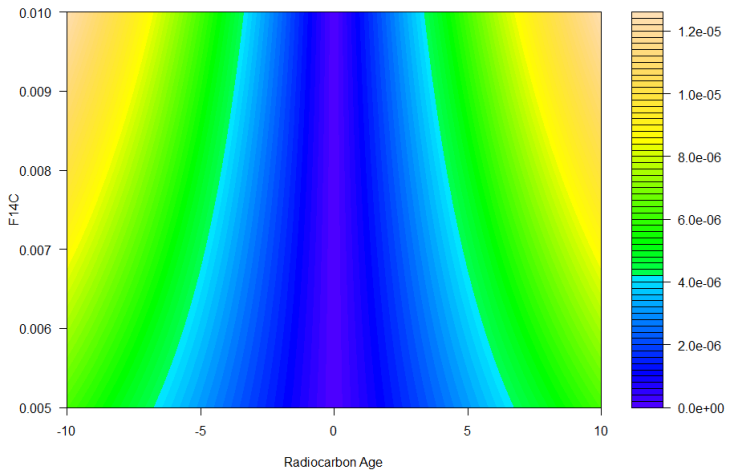


FIGURE A.2: Precision in F14C.

From figure A.2 it is observed that 6 decimals are enough to obtain a equivalent to rounding to the nearest unit, for this reason rounding up to the 7th decimal, in F14C, would be equivalent rounding up to the nearest unit in Radiocarbon Age.

Since seven decimals would not seen elegant when working with F14C, multiplying such number for 10^3 would give a more elegant variable, and only four decimals are needed to obtain the equivalent rounding. This means that if .9734524 is observed, then the new variable would be represented 973.4524, this way the variable would look more elegant and it would not give the false impression of been more precise of what it really is.

Bibliography

1. Christen J A, 1994, Summarizing a Set of Radiocarbon Dertminations : a Robust Approach, Journal of Royal Statistical Society. Series C (Applied Statistics).

2. Christen J A and Pérez S, 2009, A New Robust Statistical Model for Radiocarbon Data, Radiocarbon.

3. Plicht J van der and Hogg A, 2006, A note on reporting radiocarbon, Quaternary Geochronology.

4. Reimer P J et al., 2013,INTCAL13 AND MARINE13 Radiocarbon Age Calibration Curve 0-50000 years Cal BP, Radiocarbon.

5. Reimer P J, Thomas A Brown and Ron W Reimer, 2004, Discussion: Reporting and Calibration of Post-Bomb Data, Radiocarbon.

6. Damon P. E. , et al, 1989, Radiocarbon Dating of the Shroud of Turin, Nature.

7. Wasserman L, 2004, All of Statistics: a concise course in statistical inference, Springer.

8. Spivak M, 2006, Calculus, Publish or Perish.