NTNU – Trondheim
Norwegian University of
Science and Technology

Condition monitoring and failure prediction in complex systems

Gunnhild Hardersen
Presthus

Master of Science in Physics and Mathematics
Submission date: June 2014
Supervisor: Bo Henry Lindqvist, MATH
Co-supervisor: Henrik Madsen, Technical University of Denmark

Norwegian University of Science and Technology
Department of Mathematical Sciences

Problem description

- Description of a motivating case study of the bearing in a wind turbine.
- Probabilistic modeling in a general situation.
- Statistical inference and prediction in the probabilistic model.
- Application to the motivating case study.

Assignment given: January 17, 2014.
Supervisor: Bo Henry Lindqvist (NTNU).
Co-supervisor: Henrik Madsen (DTU).

Preface

This thesis concludes my joint Nordic Five Tech Master's program in Applied and Engineering Mathematics (N5TeAM) and leads to the degrees Master of Science in Applied Physics and Mathematics at the Norwegian University of Science and Technology (NTNU) and Master of Science in Engineering at the Technical University of Denmark (DTU). The work was carried out during the spring semester of 2014.

I would like to thank my supervisor, Bo Lindqvist, at NTNU and my co-supervisor Henrik Madsen at DTU. I am especially grateful for the many enlightning and encouraging discussions with Bo Lindqvist, as well as the invaluable feedback.

I would also like to thank my classmates for the exciting years at NTNU and DTU, and in particular my N5TeAM classmates. To all loved ones, thank you for the support.

Trondheim 2014-06

Gunnhild Hardersen Presthus

Abstract

We consider a monitored system $W(t)$ at time t which is modeled by a stochastic process. Failure of the system is connected to the process overreaching a certain threshold. The system is governed by an unobservable marker process in such a way that the state of the marker process is connected to the technical condition of the system. The system exhibits different traits for the different states of the marker process. The goal is to estimate the first passage time of the critical threshold. The stochastic process under study is the Wiener process, more precisely, a piecewise Wiener process with change points. The change points signify the occurrence of an event which causes a change in parameters of the Wiener process. The change points are governed by the marker process: the time between change points is the time spent in each state for the marker process. This is an unknown quantity, which is estimated by the observable Wiener process. The situation with one change point and two change points and the Wiener process parameters known and unknown are examined and numerical examples are provided for simulated data. The formulas are extended to m change points. A Bayesian approach is used, and Markov Chain Monte Carlo methods are employed to estimate the distribution of the process parameters. In order to predict the hitting time, the hitting time cumulative distribution functions are estimated through simulation of Wiener processes, straightforward calculations and a time transformation approach.

Sammendrag

Vi betrakter et overvåket system $W(t)$ ved tidspunkt t modellert ved en stokastisk prosess. Systemfeil er knyttet til at prosessen krysser en gitt terskel. Prosessen styres av en ikke-observerbar markørprosess, på en slik måte at markørprosessens tilstand er knyttet til systemets tekniske tilstand. Prosessen har ulike egenskaper for de ulike stadiene i markørprosessen. Målet er å estimere tidspunktet for første passasje over den kritiske terskelen. Den stokastiske prosessen som studeres er Wienerprosessen, mer presist, en stykkevis Wienerprosess med endringstidspunkter. Disse tidspunktene inntreffer samtidig med en hendelse som forårsaker en endring i prosessens parametere. Endringstidspunktene styres av markørprosessen: Tiden mellom tidspunktene er tiden markørprosessen tilbringer i hver tilstand. Dette er en ukjent størrelse, som estimeres via den observerbare Wienerprosessen. Situasjonen med ett og to endringstidspunkter studeres grundig og numeriske eksempler er gitt for simulerte data. Bayesisk inferens og Markov Chain Monte Carlo-metoder anvendes for å estimere sannsynlighetsfordelingen til endringstidspunktene, og parameterene i Wienerprosessen. For å predikere tidspunktet for første passasje av den kritiske terskelen, estimeres den kumulative fordelingen for passasjetiden til den kritiske terskelen ved hjelp av simuleringer av Wienerprosesser og beregninger med og uten transformasjon av tiden.

Contents

Chapter 1

Introduction

This thesis will study deterioration processes modeled by an important stochastic process, namely the Wiener process. The model and statistical inference are the main topics, and the goal is to obtain estimates to predict the future behavior of the process, which could be used in connection with condition-based maintenance.

Condition-based maintenance can be defined as (Rausand and Høyland, 2004, p. 363): an approach to maintenance in which maintenance actions are decided based on measurements of variables that are correlated with deterioration. Examples of such variables are temperature, pressure, erosion, vibration and noise levels. The maintenance strategy is sometimes called predictive maintenance. Condition-based maintenance requires a monitoring system for measurements of the variables, as well as a mathematical model that predicts the behavior of the deterioration process. When repair is difficult, involves risk, is costly in itself or leads to costly downtime, condition monitoring may be important to ensure that no production is lost. Examples that could be thought of are offshore structures, such as sub-sea structures or wind turbines. The latter will be the motivating example when studying Wiener processes in this thesis.

Wiener processes have been used in a wide range of subjects, perhaps because of their tractable mathematical properties. One such subject is finance, where the Wiener process without drift and with standard variance is important. This special case of a Wiener process is called standard Brownian motion. One important usage in finance is the Black-Scholes equation, which models stock price under a risk-neutral probability measure. The differential of the Brownian motion enters in the stochastic differential equation in a term which aims to describe the uncertainty in the stock price. The solution to the Black-Scholes equation is known, and the Brownian motion enters as part of the solution.

In degradation modeling it is natural to consider the Wiener process with drift. For the Wiener process with positive drift, the first passage time to a given threshold of interest has inverse Gaussian distribution. In reliability engineering, it has been studied by for example Whitmore: in Whitmore (1986), multiple modes of failure are introduced to a multivariate Brownian motion. The situation with multiple failure modes is often called competing risks. Length of stay in hospital has also been modeled by use of

a Wiener process, in Horrocks and Thompson (2004), in a competing risk situation. The Wiener process with drift represents a health level process and has the multiple outcomes: death in hospital or healthy discharge. A time scale transformation is applied in Whitmore and Schenkelberg (1997) to a Wiener process to predict life time. The time transformation is actually inspired by the time transformation in Doksum and Høyland (1992), which will be examined more thoroughly in section 4.3.2. In Whitmore et al. (1998) a bivariate Wiener process is governed by an unobservable marker process. Many different applications are suggested in a variety of fields, among them marriage failure, where an appropriate marker process is a social estrangement index, AIDS death, where CD4 cell count is an appropriate marker process and metal fatigue failure, with dominant crack length as marker process. The setting with a marker process is the topic for this thesis.

The change point problem for Wiener processes was studied by Shiryaev (1963). The setting is an observed Wiener process with one unknown change point, after which the drift of the process changes from 0 to r for some r. A Bayesian approach equivalent to the approach in this report is used to detect the change point. The multiple change point problem is studied in a frequentistic setting in Hawkins (2001), who provides an exact algorithm for finding maximum likelihood estimates of the change points.

In this thesis, the mathematical model for the behavior of the system will be the Wiener process, which, in our example, will model temperature. The case study from Lindqvist and Slimacek (2013) will be analyzed more thoroughly. Whereas Lindqvist and Slimacek (2013) developed estimates for change points in Wiener processes, this report will examine the cumulative distribution functions of the hitting time of a specified threshold. Wiener processes are used to model temperature in the bearing of a wind turbine, which is governed by the failure development. The failure development itself is modeled by a hidden Markov model. The case study developed by Valland et al. (2012) is presented in section 2. General theory is developed in section 3. Section 4 shows how the theory can be used to model the case study, and finally the theory is applied to simulated data in section 5. Section 6 concludes on the theory developed and its applicability, and suggests further work.

Chapter 2

Case study of wind turbine bearing

In this section a motivating case study for failure prediction in a monitored system will be presented. The setting is an offshore wind turbine, which is monitored in a number of ways. By linking all the monitored variables to an index number, an overview of the network of instruments or structures may easily be obtained. The technical condition index (TCI) aims to do just this: to include all information about the condition of an instrument or structure in one index number. In order to link the monitored variables to a health indication index by a mathematical model, one must have knowledge of the failure development process. The mathematical model will be studied in this thesis.

2.1 Failure development process

The stages of the failure development can be modeled as described in Valland et al. (2012). It is suggested that failure in a wind turbine bearing can develop through the following stages:

0. As good as new
1. Impurities in oil
2. Mechanical wear
3. Micropitting, Pitting
4. Chipping
5. Bearing breakdown
6. Turbine shut down

where the first stage, 0, which signifies normal conditions, has been added for convenience. In Valland et al. (2012), the different health indicators of the failure development were particles in oil, vibrations, abnormal noise, temperature variation and visible signs

of the failure development. Vibrations and temperature variations are observed by machine monitoring. The remaining indicators are observed by more demanding methods, namely oil analysis, human senses and inspection. Impurities in oil is the stage 1 condition and is according to Valland et al. (2012) the only indication which is detectable at stage 1. At stage 2, mechanical wear, it is also possible to detect the failure development through temperature variations, vibrations and visible signs. At stage 4, chipping, the failure development is detectable through all health indicators.

It is suggested in Valland et al. (2012) that the temperature is increasing throughout the failure development as a result of increased bearing friction. This is reflected in Lindqvist and Slimacek (2013) by the drift of the temperature process increasing through the failure development states. This means that the temperature rises at faster rates through failure development stages. Note that after breakdown, states 5 and 6, the temperature is assumed to drop linearly. An example of the temperature process for a process which is moving through the failure development is shown in figure 2.1, from Valland et al. (2012).

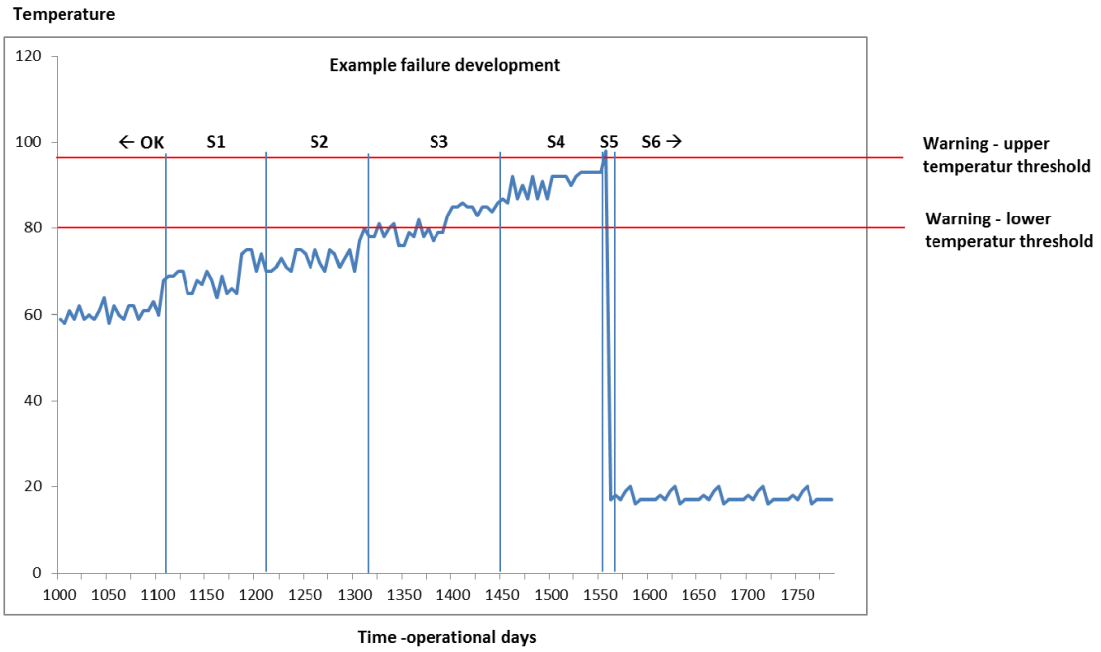

Figure 2.1.1: Example of a typical temperature process through failure development given in Valland et al. (2012). $S1$ denotes stage 1, $S2$ denotes state 2 and so on. Note that state 0 which has been added in this thesis corresponds to OK.

As Valland et al. (2012) points out, one problem with the TCI framework is that it only provides historical information, and no way to predict development in order to plan maintenance. Such predictions is the aim for this report, and we will connect the described failure development to a model which makes it possible to predict the future

development.

Chapter 3

Probabilistic modeling in the general setting

This section gives an introduction to the theory of Markov models and Wiener processes that will later be applied to the model of the motivating case study. Bayesian inference and Markov Chain Monte Carlo methods, which are used to making inference about the case study are also covered.

3.1 Markov model

To model the underlying failure development, we will use a *discrete time Markov chain*, a special stochastic process, which can be defined as in (Ross, 2007, p. 185): Denote the stochastic process Z_n, $\{Z_n, n = 0, 1, 2, \dots\}$, which takes a finite or countable number of values. From now on, we will assume that the number of values is finite. We introduce the notation $Z_n = i$, meaning that the process is in state i at time n. When the process is in state i there is a probability P_{ij} that it will next be in state j, that is

$$P\{Z_{n+1} = j | Z_n = i, Z_{n-1} = i_{n-1}, \dots, Z_1 = i_1, Z_0 = i_0\} = P_{ij},$$

for all $i_0, i_1, \dots, i_{n-1}$. Thus, the conditional distribution of the state Z_{n+1} given all the past states $Z_0, Z_1, \dots, Z_{n-1}$ and the present state Z_n, depends only the present, and is independent of the past, that is, the process is memoryless. The memoryless property is referred to as the *Markov property*. Note that $P_{ij} \geq 0$, $\sum_{j=0}^{\infty} P_{ij} = 1$ for all i, j.

Let $0 = S_0 < S_1 < \dots$ be the times at which transitions occur. Define $\widetilde{U}_i = S_{i+1} - S_i$ as the ith interoccurrence time. Let U_i denote the time spent during a visit to state i. It is known from for example (Rausand and Høyland, 2004, p. 305) that because of the Markov property of the process, the random variable U_i is memoryless, and thus exponentially distributed for a continuous time Markov chain and geometrically distributed for a discrete time Markov chain.

The classification of states of a Markov chain is important when constructing Markov chains for Markov Chain Monte Carlo methods, of which there is a short review in section

3.5. Two of the necessary classifications of states in Markov chains are *irreducible* states and *ergodic* states. A Markov chain is said to be irreducible if all states are accessible from any state. An ergodic state in a finite state space Markov chain is a state which is *recurrent* and *aperiodic*. A state i is recurrent if, starting in state i, the probability of ever returning to state i is 1. A state i is periodic with period k if return to state i must occur in a multiple of k time steps. On the other hand, a state is aperiodic if $k = 1$. If all states in a Markov chain are aperiodic, then the Markov chain itself is said to be aperiodic. Correspondingly, a Markov chain is ergodic if all its states are ergodic.

3.1.1 Hidden Markov model

A hidden Markov chain can be described as in (Givens and Hoeting, 2013, p. 175). Suppose there is a Markov chain of unobservable variables $Z_0, Z_1, \ldots$ indexed by time n. The variables represent the state of a Markov process, thus the chain has the Markov property. Although the states Z_n are unobservable, there is an observable sequence of random variables $X_0, X_1, \ldots$, such that X_n is dependent on the process state at the same time, Z_n. This gives the model

$$X_n \sim f_x(x_n|z_n) \qquad\qquad Z_n \sim f_z(z_n|z_{n-1}),$$

which is known as a hidden Markov model. The hidden Markov model can act as the marker process as in Whitmore et al. (1998), of which several examples were provided in chapter 1. Note that the marker process examples were not necessarily unobservable processes.

Thus, given the state of the hidden Markov chain, we have full knowledge about the distribution of the observable process, X_n. To make inference about X_n, we must thus make inference about Z_n. This will be done by Bayesian inference, which is briefly reviewed in section 3.3.

3.2 The Wiener process and the Inverse Gaussian distribution

A stochastic process $W(t)$ is a Wiener process with drift parameter ν and variance parameter σ^2 if

1. $W(0) = 0$ with probability one,
2. For every $t > 0$, $W(t)$ is normally distributed with mean νt and variance $\sigma^2 t$,
3. $W(t)$ has stationary and independent increments.

If $\nu \neq 0$, we say that the process is a Wiener process with drift. When $\nu = 0$, the expected value of the process $W(t)$ at time t is zero at every time. In fact, when $\nu = 0$, $W(t) = \sigma B(t)$, for a standard Brownian motion process $B(t)$. The standard Brownian motion as a stochastic process is defined as the Wiener process, except that for every

$t > 0$, $B(t)$ is normally distributed with mean 0 and variance t. When the Wiener process $W(t)$ has a non-zero drift ν, the expected value of the process $W(t)$ is increasing with time if $\nu > 0$ and decreasing when $\nu < 0$, which explains the name of ν, the drift coefficient. In fact, given $\nu > 0$, and the times r and s such that $r > s$, $W(r)$ is first-order stochastically dominant over $W(s)$, that is, for all x, $P(W(r) \geq x) \geq P(W(s) \geq x)$.

Define the first passage time T of a threshold value a as the first time the process crosses the threshold value:

$$T = \min_{t>0}(W(t) \geq a).$$

An important and mathematically tractable attribute of the Wiener process, given in for example Aalen and Gjessing (2001), is that if $\nu > 0$, the first passage time to a level $W(t) \geq a > 0$ is inverse Gaussian distributed, with density

$$f(t; \nu, \sigma, a) = \frac{a}{\sqrt{2\pi}\sigma} t^{-\frac{3}{2}} \exp\left(-\frac{(a-\nu t)^2}{2t\sigma^2}\right), \quad t > 0. \tag{3.2.1}$$

The first passage time is sometimes referred to as the *hitting time* of the threshold a. The mean and variance of the hitting time T are given by

$$E[T] = \frac{a}{\nu}, \qquad Var[T] = \frac{a\,\sigma^2}{\nu^3}. \tag{3.2.2}$$

The inverse Gaussian distribution can also be expressed by the two parameters $\mu = a/\nu$ and $\lambda = a^2/\sigma^2$.

$$f_T(t; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi t^3}} \exp\left(-\frac{\lambda}{2\mu^2}\frac{(t-\mu)^2}{t}\right), \quad t > 0.$$

We shall however find it convenient to consider the distribution with three parameters. The survival function, $S(t; \nu, \sigma, a) = P(T > t)$, is displayed in for example Lindqvist and Slimacek (2013):

$$S(t; \nu, \sigma, a) = \Phi\left(\frac{a-\nu t}{\sigma\sqrt{t}}\right) - \exp\left(\frac{2a\nu}{\sigma^2}\right)\Phi\left(\frac{-a-\nu t}{\sigma\sqrt{t}}\right).$$

Thereby, the cumulative distribution $F(t; \nu, \sigma, a) = 1 - S(t; \nu, \sigma, a) = P(T \leq t)$ is also known:

$$F(t; \nu, \sigma, a) = \Phi\left(\frac{\nu t - a}{\sigma\sqrt{t}}\right) + \exp\left(\frac{2a\nu}{\sigma^2}\right)\Phi\left(\frac{-a-\nu t}{\sigma\sqrt{t}}\right). \tag{3.2.3}$$

3.3 Bayesian inference

Contrary to the traditional, frequentist approach, where the parameters are regarded as constants, the Bayesian approach regards the parameters as stochastic variables, with a probability density. This probability density function is called a *prior*, and holds our believes of how the parameter is distributed in the parameter space, prior to collecting

data. Thus, Bayesian inference aims to include our knowledge of the quantities we want to estimate. The knowledge can stem from own experience or previous experiments, and is subjective. Data is then collected, and the prior is updated with the information from the sample. The updated distribution is called the *posterior distribution.* The update is done by Bayes rule:

Theorem 1. (Bayes' Rule) *Let $A_1, A_2, \ldots$ be a partition of the sample space, and let B be any set. Then for each $i = 1, 2, \ldots,$*

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)P(A_j)}.$$

Following this, suppose that θ is the parameter we want to make inference about, given the observations $\mathbf{x}$, whose sampling distribution we denote by $f(\mathbf{x}|\theta)$. Let $\pi(\theta)$ be the prior distribution of the parameters, and $\pi(\theta|\mathbf{x})$ be the posterior distribution of the parameters given the observations. Then the posterior distribution of the parameter, θ, given the sample, $\mathbf{x}$, is

$$\pi(\theta|\mathbf{x}) = \frac{f(\mathbf{x}|\theta)\pi(\theta)}{\int f(\mathbf{x}|\theta')\pi(\theta')d\theta'}.$$

Note that the denominator, $\int f(\mathbf{x}|\theta')\pi(\theta')d\theta'$ is the marginal distribution of $\mathbf{x}$. Since $\mathbf{x}$ is observed, the marginal distribution is a constant, thus the posterior distribution is proportional to $f(\mathbf{x}|\theta)\pi(\theta)$, that is

$$\pi(\theta|\mathbf{x}) \propto f(\mathbf{x}|\theta)\pi(\theta).$$

This implies that

$$\pi(\theta|\mathbf{x}) \propto L(\theta|\mathbf{x})\pi(\theta)$$

where $L(\theta|\mathbf{x})$ is the likelihood of θ given the observation $\mathbf{x}$.

The choice of priors is crucial. It is important that the priors tail behavior is similar to the posterior tail behavior, and specifically that the support of the prior covers the support of the posterior. One should make sure that the prior does not influence the posterior too much, unless the prior information is certain in some way.

3.4 Monte Carlo integration

Monte Carlo integration is the statistical evaluation of an integral using the evaluations of the integrand at a set of points drawn randomly from a distribution which has support over the entire range of integration. One of the important applications is evaluating and expectation $E[t(x)]$, shown in (Givens and Hoeting, 2013, p. 151). Let $f(x)$ be the density of the random variable X, and let μ denote the expectation of $t(x)$ with respect to f. Given an i.i.d. sample, that is, a sample of independent, identically distributed $X_1, \ldots, X_n$ from f, μ can be approximated by the Monte Carlo (MC) estimator $\hat{u}_{\text{MC}}$

$$\hat{\mu}_{\text{MC}} = \frac{1}{n}\sum_{i=1}^{n} t(X_i) \to \int t(x)f(x)dx = \mu.$$

MC integration has slow convergence, but its strengths are considerable. The technique is very simple. Furthermore, other integration techniques are severely punished by increasing dimensions, but this effect is not as severe for MC integration, and the simplicity of implementation is preserved. Note that if we only wish to evaluate the integral over some region Ω which has probability mass $\int_\Omega f(x)dx = V$, the estimate, $\hat{\mu}_{\text{MC}}$, must be scaled with V. The technique for a discrete random variable is equivalent. Let this be termed MC summation. The MC estimator $\hat{\mu}_{\text{MC}}$, for a discrete random variable X is

$$\hat{\mu}_{\text{MC}} = \frac{1}{n}\sum_{i=1}^{n} t(X_i) \rightarrow \sum_{x_i \in \Omega} t(x_i)p(x_i) = \mu,$$

where Ω denotes the sample space of X and $p(x_i) = P(X = x_i)$.

3.5 Markov Chain Monte Carlo methods

Markov Chain Monte Carlo (MCMC) methods aim at generating a sample from a target distribution which can be evaluated easily, but from which it is difficult to sample. It can be employed as a specific strategy for MC integration. An introduction is given in (Givens and Hoeting, 2013, chapter 7). An example can be the posterior distribution of a random variable, which may not be in a known class of distributions. The sampling strategy in MCMC methods is to create an irreducible, aperiodic Markov chain for which the stationary distribution equals the target distribution. The distribution of the Markov chain $X^{(t)}$ converges to the limiting stationary distribution when the chain is irreducible and aperiodic. Thus, after some time, the Markov chain $X^{(t)}$ converges to the target distribution f. The construction of the Markov chain is crucial, as well as diagnosing whether the chain has converged.

3.5.1 Metropolis-Hastings algorithm

The Metropolis-Hastings algorithm is a general method for constructing a Markov chain, given in for example (Givens and Hoeting, 2013, chapter 7.1). The algorithm samples from a proposal distribution $g(x|x^{(t)})$.

There are different classes of proposals to chose from, one is the independence chains, where the proposal distribution g is such that $g(x^*|x^{(t)}) = g(x^*)$. The ratio becomes

$$R(x^{(t)}, X^*) = \frac{f(X^*)g(x^{(t)})}{f(x^{(t)})g(X^*)}.$$

A Markov chain with such a proposal is irreducible and aperiodic, thus converges to the limiting stationary distribution, if $g(x) > 0$ whenever $f(x) > 0$. One simple choice in Bayesian inference is to use the prior as proposal, $g(x) = \pi(\theta)$ in an independence chain. The posterior distribution $\pi(\theta|y) = p(\theta)L(\theta|y)$ is the target distribution. Thus,

$$R(\theta^{(t)}, \theta^*) = \frac{p(\theta^*)L(\theta^*|y)p(\theta^{(t)})}{p(\theta^{(t)})L(\theta^{(t)}|y)p(\theta^*)} = \frac{L(\theta^*|y)}{L(\theta^{(t)}|y)}$$

that is, Metropolis-Hastings ratio simplifies to the likelihood ratio.

3.5.2 Diagnostics of output

Although the Markov chain generated by the described approach has the target distribution as limiting distribution, it is necessary to investigate whether the chain has run sufficiently long, such that the output represents the target distribution. Thus, a number of tests are suggested in the literature to check for convergence of the chain, as well as the mixing of the chain, that is, how far apart in a sequence the samples must be to be considered approximately independent. In the following, a small collection of popular diagnostics will be shortly reviewed. It is important to remember that no test is completely reliable.

To investigate the mixing properties of a chain, one can inspect the *sample path*, often called the *trace plot*. The sample path is a plot of the realizations of $X^{(t)}$ versus the iteration number t. Poor mixing manifests as the chain staying at or near the same value for a long time, where as good mixing results in the chain moving quickly around in the entire support region of the target distribution f.

A plot of the *autocorrelation* of the chain can also provide indications of the mixing properties of the chain. The autocorrelation at *lag* i is the correlation between iterates that are i iterations apart. If the chain is mixing poorly, the autocorrelation will be decaying slowly as a function of i. On the other hand, if the chain is mixing well, the autocorrelation will decay relatively fast. In higher dimensional problems, cross-correlations between parameters may also be worth investigating, as high cross-correlations could be a symptom of poor mixing. It seems thinning of the output to reduce autocorrelation is often practiced, as found by Link and Eaton (2012). If we want to thin by k, we only use every kth iteration of the output. However, Link and Eaton (2012) advise using thinning only to reduce the size of a data set to facilitate computations. Rather than thinning by k to lessen autocorrelation, one should rather sample for k times longer. As an effect, the autocorrelation effects should wash out in the large sample size.

For an independence sampler, the acceptance rate, that is, the relative number of times when the proposal is accepted as the new iteration, should be high. A high acceptance rate is crucial for fast convergence of the chain. However, in a target density which is not unimodal, the chain can get stuck in areas of high posterior density. This could give a very large acceptance rate, and result in poor mixing.

An important point of the diagnostic of a Markov chain is to assess the *burn-in*. It is only the limiting distribution of the chain for which $X^{(t)} \sim f$, and the starting point will affect the first iterates. The dependence on the starting point may be strong, and to lessen this effect, the first D iterates, the burn-in period, are usually discarded. Whether burn-in period and run length are reasonable can be found by comparison of within-chain and between-chain variance. The run length should be increased if the within-chain variance is considerably smaller than the between-chain variance. This is estimated from M runs of the MCMC algorithm to create separate chains of equal length. The starting values of the M chains should be spread over the support of f. In fact, Brooks and Gelman (1998) advises that the starting points should be over-dispersed with respect to the target distribution, to ensure absence of falsely diagnosed convergence. A formula for calculating the distance between the within-chain and between-chain variance

in multidimensional chains is given by Brooks and Gelman (1998). Let W be the within-chain variance and B/n be the between-chain variance, where n is the number of runs in the chain. Let $x_{jt}^{(i)}$ denote the ith element of the p-dimensional parameter vector in chain j at time t. Let $x_{j\cdot}$ denote the mean over time, $\frac{1}{n}\sum_{t=1}^{n} x_{jt}$, and $x_{\cdot\cdot}$ denote the mean over time and over all chains, $\frac{1}{m}\sum_{j=1}^{m} x_{j\cdot}$. The between-chain variance B/n can be found by

$$B/n = \frac{1}{m-1}\sum_{j=1}^{m}(x_{j\cdot} - x_{\cdot\cdot})(x_{j\cdot} - x_{\cdot\cdot})',$$

and the within-chain variance W can be found by

$$W = \frac{1}{m(n-1)}\sum_{j=1}^{m}\sum_{t=1}^{n}(x_{jt} - x_{j\cdot})((x_{jt} - x_{j\cdot})').$$

Finally, a scalar measure of the distance between the two variances is

$$\widehat{R}^p = \frac{n-1}{n} + \frac{m+1}{m}\lambda_1,$$

where λ_1 is the largest eigenvalue of the symmetric positive definite matrix $W^{-1}B/n$. The distance measure $\widehat{R}^p$ converges to 1 as $n \to \infty$.

The measure is in fact an estimator of the squared scale reduction factor, which is a ratio of the pooled posterior variance and the within-chain variance. If it is large, it suggests that the variance in the parameters can be decreased by increasing the number of simulations, or that the simulated sequences have not fully toured the target distribution, such that W will increase by further simulations. If the ratio is close to 1, each of the m sets of n simulated observations is close to the target distribution.

Algorithm 1 The Metropolis-Hastings algorithm

Require: $x^{(0)}$ such that $f\left(x^{(0)}\right) > 0$

for $t = 0$ to m **do**

Sample X^* from the proposal distribution $g\left(\cdot|x^{(t)}\right)$

Compute the Metropolis-Hastings ratio $R\left(x^{(t)}, X^*\right)$ where

$$R(u, v) = \frac{f(v)g(u|v)}{f(u)g(v|u)}$$

Sample a value for $X^{(t+1)}$ according to the following rule

$$X^{(t+1)} = \begin{cases} X^*, \text{ with probability } \min\left\{R\left(x^{(t)}, X^*\right), 1\right\} \\ x^{(t)}, \text{ otherwise} \end{cases}$$

end for

Chapter 4

Modeling the case study

In this section, the set of equations and approximations necessary for calculating the cumulative distribution function (CDF) of the hitting time T of a critical threshold a will be presented. Different approaches will be used, and the setting with one and two change points will be studied in detail. Generalizations for $d > 2$ change points are given.

4.1 Modeling the case study

We proceed by following the approach of Lindqvist and Slimacek (2013), by modeling the failure development stage of the system, which was described in section 2, as a hidden Markov model. Note that we now assume that the initial state in failure development, state 1, is detectable through temperature variations when the methods which will be presented are applied, as opposed to the situation in section Valland et al. (2012) where temperature variations were detectable from stage 2 and onward.

We will model the failure in the wind turbine by a hidden Markov model: the state in the failure development is the hidden Markov process, and the observable random variable is the temperature. Note that we assume discrete observations of $W(t)$, such that the process $W(t)$ is recorded at some $t_1, t_2, \ldots, t_n$. We assume that, for each state $j = 0, \ldots, 6$ in the failure development, there is a probability q_i to stay in the state and a probability $1-q_j$ to leave the state and jump to state $j+1$. Thus $p_{jj} = 1-q_j, p_{j,j+1} = q_j$ and $p_{j,i} = 0$ for $(j,i) \neq (j,j)$ or $(j, j+1)$. This models that the failure development cannot repair itself, and that it does not skip a stage, so to speak. The initial state, state 0, models some kind of stationarity in the temperature. The hidden Markov process is sketched in figure 4.1.1.

As for the temperature, $W(t)$, we assume a piecewise Wiener process. The process has change points τ_j such that the temperature is a Wiener process in each time interval $[\tau_j, \tau_{j+1})$, with parameters ν_j and σ_j^2. We may have $\sigma_j^2 = \sigma^2$ such that the variance parameter is the same throughout the failure development. The drift parameter is assumed to start at zero and increase throughout the states until breakdown, that is $0 = \nu_0 < \nu_1 \leq \cdots \leq \nu_4$. In this report, we will only model the temperature until state

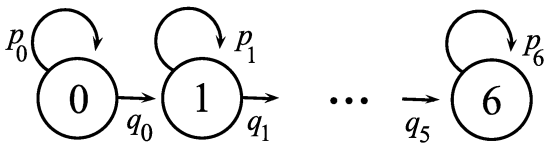

Figure 4.1.1: A schematic sketch of the discrete time Markov model for the state of the failure development. In each state j , q_j is the probability of jumping to state $j+1$, and $p_j = 1 - q_j$ is the probability of staying in state j, shown with arrows. As stated, it is only possible to go from state j to $j+1$ for $j = 1, \ldots, 6$.

4, after which it is assumed by Valland et al. (2012) that the temperature is linearly decreasing.

A realization of a piecewise Wiener process with increasing drift parameter is shown in figure 4.1.2. The process has been simulated in the time interval $[0, t_n] = [0, 1000]$. The process has two change points, $\tau_1 = 500$ and $\tau_2 = 750$, and was simulated with the following parameters: $\nu_0 = 0, \nu_1 = 0.0015, \nu_2 = 0.0035, \sigma_0 = \sigma_1 = \sigma_2 = 0.05$. In a failure development setting, the plot can be interpreted as the process being in the normal state for $t \in [0, \tau_1)$, in an initial state of failure development for $t \in [\tau_1, \tau_2)$, and in a more severe state of failure development for $t \in [\tau_2, t_n)$.

Let X_i denote the temperature increment at time step i in a state, that is $X_i = W(t_i) - W(t_{i-1})$. Then the increments in each state, conditioned on the state Z_i, are normally distributed with the state parameters

$$X_i | Z_i = j \sim \mathcal{N}\left((t_i - t_{i-1})\nu_j, (t_i - t_{i-1})\sigma_j^2\right),$$

that is,

$$f(x_i | z_i = j) = \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}\sigma_j} \exp\left\{-\frac{(x_i - (t_i - t_{i-1})\nu_j)^2}{2(t_i - t_{i-1})\sigma_j^2}\right\}.$$

Note that this is only valid if the latent Markov process is in state j both in time t_i and t_{i-1}.

Furthermore, consider the situation of having observed a Wiener process from time 0 to some time t_n, such that $W(0) = 0$ and $W(t_n) = w_n$, and having set a critical threshold, a. In order to find the distribution of the hitting time T from the "now" point in time t_n and the "now" temperature w_n, we regard the process such that the last observation point, (t_n, w_n) is the origin. Denote the new process, which starts at (t_n, w_n), W_2. Let s be the new time, $s = t - t_n$ and $W_2(s) = W(t) - w_n$. Given that the process has not already crossed the threshold, it is known that $T > t_n$. The hitting time and the threshold a can be transformed according to the information we hold from having observed the process. Note that the process starting in the new origin is a Wiener process, such that the hitting time is inverse Gaussian. Let $S = T - t_n$ be the transformed hitting time and note that $S \sim \mathcal{IG}(\nu, \sigma, a - w_n)$. In the situation of having observed the Wiener process until time $t = t_n$, it is natural to view the transformed hitting time as the remaining time until first passage of the threshold a.

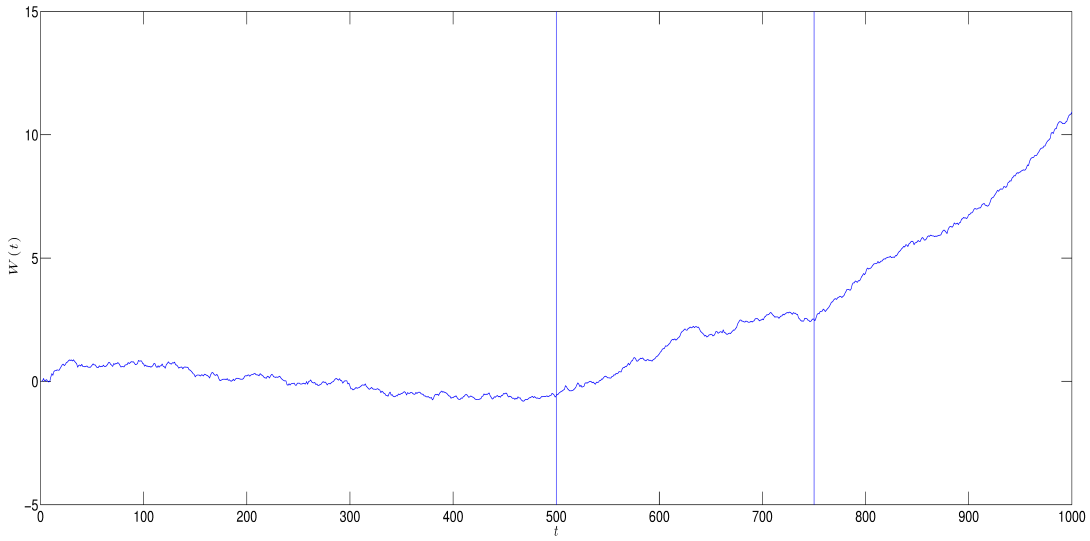

Figure 4.1.2: A realization of a Wiener process in the time interval [0, 1000] with two change points: $\tau_1 = 500, \tau_2 = 750$, marked with vertical lines. The process was simulated with the following parameters: $\nu_0 = 0, \nu_1 = 0.015, \nu_2 = 0.035$ and $\sigma_i = \sigma = 0.05, i = 0, 1, 2$.

The probability model of the marker process, Z_i, follows from its Markov model characteristics, the memorylessness. This implies an exponential probability distribution for the time spent in each state in continuous time and a geometric probability distribution for the time spent in each state in discrete time. Thus, the time spent in state Z_i, U_i, is geometrically distributed

$$U_i \sim \mathcal{G}eom\,(q_i)\,,$$

that is,

$$f(u_i) = q_i(1-q_i)^{u_i-1}, \quad u_i = 1, 2, \ldots$$

The physical interpretation is that each time unit, for example each day, it is equally likely that the chain moves from state i to state $i+1$, and there is independence from day to day. In fact, the probability of the chain moving from state i to state $i+1$ is q_i. Note that in stage 0, this means that there is an equal probability of particles entering the oil each time unit, for example, each day.

Figure 4.1.3 shows the connection between the development through stages S_i and the temperature process.

Throughout the discussion, we will continue the assumption of Lindqvist and Slimacek (2013), that $T > \tau_1$, that is, that the threshold cannot be crossed before the first change point has occurred. This is because the threshold temperature is so much larger than the temperature in the normal condition that it will not be crossed by the

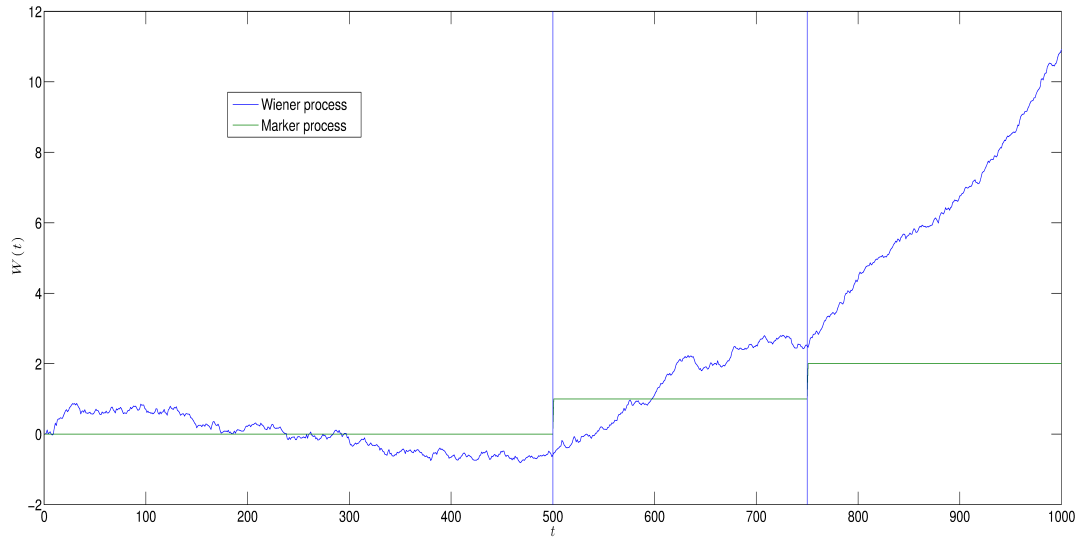

Figure 4.1.3: The simulated piecewise Wiener process which was shown in figure 4.1.2 where the marker process is the latent Markov model, which was sketched in figure 4.1.1.

natural fluctuations of temperature. We also assume that $T > t_n$, since if $W(t_i) \geq a$, for any $i = 1, \ldots, n$, $P(T \leq t_n) = 1$ and the hitting time $T = t_i$ can be found from the observations. For the remainder of the report, we will assume equidistant measurements, such that, $t_i = i\Delta$ for $i = 1, \ldots, n$. Then the temperature increment at time t_i, X_i, given the state $Z_i = j$ is normally distributed with mean $\Delta\nu_j$ and variance $\Delta\sigma_j^2$: $X_i|Z_i = j \sim \mathcal{N}\left(\Delta\nu_j, \Delta\sigma_j^2\right)$. We will also assume that the system is in the normal condition in the starting point of observations, that is, $Z_0 = 0$.

4.2 One unknown change point, τ

We now assume that the process has one change point, τ_1, as was the situation in Lindqvist and Slimacek (2013). Remember that the parameters of the piecewise Wiener process in the time interval $[0, \tau_1)$ are $0 = \nu_0$ and σ_0, and in the time interval $[\tau_1, \infty)$ the parameters are $\nu_1 > \nu_0$ and σ_1. We will assume that $\sigma_0 = \sigma_1 = \sigma$ and denote $\nu = \nu_1$ and $\tau = \tau_1$.

Firstly, we shall consider ν and σ known. As Lindqvist and Slimacek (2013) state, it may be reasonable to assume that ν and σ are known from expert judgment or statistical analysis of past data. Two different strategies will be used to find the cumulative distribution function (CDF) of the hitting time T. Both strategies employ the Bayesian posterior distribution of the change point given the temperature increments, $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$. The two strategies differ in how the conditional distribution of T given τ and the temperature increments $x_1, \ldots, x_n$, $P(T \leq t|\tau, x_1, \ldots, x_n; \nu, \sigma, a)$ is found: the first, presented in section 4.2.1 will employ the inverse Gaussian distribution formulae, while the second, presented in section 4.2.2 rely on simulation of Wiener processes to find the cumulative hitting time distribution. Having found the conditional distribution, the marginal distribution of T given the temperature increments $x_1, \ldots, x_n$ is found by summation over the posterior distribution:

$$P(T \leq t|x_1, \ldots, x_n; \nu, \sigma, a) = \sum_{\tau \in \Omega} P(T \leq t|\tau, x_1, \ldots, x_n; \nu, \sigma, a)\pi(\tau|x_1, \ldots, x_n; \nu, \sigma) \tag{4.2.1}$$

where Ω is the range of τ. In practice, the summation will be done by drawing random values from the posterior distribution, as in MC summation, in both strategies.

Given $P(T \leq t|x_1, \ldots, x_n; \nu, \sigma, a)$ one can predict the future behavior of the process by considering $P(T \leq t|x_1, \ldots, x_n; \nu, \sigma, a)$ for many values t. Then, one can decide on the risk level one is willing to take regarding the probability that the hitting time has already been reached. That, is, choose the estimator $\widehat{T}_\alpha$ as the t for which $P(T \leq t|x_1, \ldots, x_n; \nu, \sigma, a) = \alpha$. For a low risk, one should schedule maintenance for a low α, such as $P(T \leq t|x_1, \ldots, x_n; \nu, \sigma, a) = 0.01$. One could also look at the median, the t for which $P(T \leq t|x_1, \ldots, x_n; \nu, \sigma, a) = 0.5$ for multiple systems, which will result in a higher risk.

Note that the hitting time T and the change point τ is restricted to the same time grid as the measurements.

The posterior distribution of $\tau, \pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, was found by Lindqvist and Slimacek (2013), and will be shown for completeness. The likelihood of the change point, $L(\tau|x_1, \ldots, x_2; \nu, \sigma)$, is

$$
\begin{aligned}
L(\tau|x_1,\ldots,x_2;\nu,\sigma) &= \prod_{i=1}^{n} f(x_i|\tau) \\
&= \prod_{i=\Delta}^{\tau/\Delta-1} \frac{1}{\sqrt{2\pi\Delta}\sigma} \exp\left(-\frac{x_i^2}{2\Delta\sigma^2}\right) \prod_{i=\tau/\Delta}^{t_n/\Delta} \frac{1}{\sqrt{2\pi\Delta}\sigma} \exp\left(-\frac{(x_i-\Delta\nu)^2}{2\Delta\sigma^2}\right) \\
&\propto \exp\left(-\frac{1}{2\Delta\sigma^2}\left[\sum_{i=\Delta}^{(\tau-1)/\Delta} x_i^2 + \sum_{i=\tau/\Delta}^{t_n/\Delta} (x_i-\Delta\nu)^2\right]\right) \\
&\propto \exp\left(-\frac{1}{2\Delta\sigma^2}\left[2\Delta\nu \sum_{i=\tau/\Delta}^{t_n/\Delta} (x_i) + (n-\tau+1)\,\Delta^2\nu^2\right]\right)
\end{aligned}
$$

if $\tau \le t_n$ and $L(\tau|x_1,\ldots,x_2;\nu,\sigma) \propto 1$ if $\tau > t_n$, that is

$$
\begin{aligned}
&L(\tau|x_1,\ldots,x_n;\nu,\sigma) \propto \\
&\begin{cases}
\exp\left(-\frac{1}{2\Delta\sigma^2}\left[2\Delta\nu\sum_{i=\tau/\Delta}^{t_n/\Delta}(x_i) + (n-\tau+1)\,\Delta^2\nu^2\right]\right) & \text{if } \tau \le t_n \\
1 & \text{if } \tau > t_n.
\end{cases}
\end{aligned} \tag{4.2.2}
$$

As the time spent in each state is modeled by a hidden Markov process, it is natural to assume a geometric prior for the change points:

$$\pi(\tau) = \lambda(1-\lambda)^{\tau-1}, \qquad \tau = 1, 2, \ldots$$

as noted in section 4.1. Thus the posterior distribution of τ is

$$
\begin{aligned}
&\pi(\tau|x_1,\ldots,x_n;\nu,\sigma) \propto L(\tau|x_1,\ldots,x_2;\nu,\sigma)\pi(\tau) \propto \\
&\begin{cases}
\lambda(1-\lambda)^{\tau-1}\exp\left(-\frac{1}{2\Delta\sigma^2}\left[2\Delta\nu\sum_{i=\tau/\Delta}^{t_n/\Delta}(x_i) + (n-\tau+1)\,\Delta^2\nu^2\right]\right) & \text{if } \tau \le t_n \\
\lambda(1-\lambda)^{\tau-1} & \text{if } \tau > t_n.
\end{cases}
\end{aligned} \tag{4.2.3}
$$

4.2.1 Formulaic approach

The formulaic approach will use that, conditioned on τ and given that τ is passed, the distribution of T is known. The estimator will be denoted $\widehat{P}_f(T \le t|x_1,\ldots,x_n;\nu,\sigma,a)$. We continue the assumption that the threshold temperature a will not we crossed before the change point has occurred, $T > \tau$. This implies $P(T \le \tau) = 0$. Furthermore, we have assumed that the process starts in the normal condition. Thus $P(T \le 0) = 0$.

Consider first the situation where $\tau < t_n$. In this, case, we will use that the shifted hitting time, $S = T - t_n$ is shifted as described in section 4.1, such that $S \sim \mathcal{IG}(\nu,\sigma,a - W(t_n)$. Thus,

$$
\begin{aligned}
P(T \le t|\tau,x_1,\ldots,x_n;\nu,\sigma,a) &= P(T-t_n \le t-t_n|\tau,x_1,\ldots,x_n;\nu,\sigma,a)) \\
&= P(S \le t-t_n|\tau,x_1,\ldots,x_n;\nu,\sigma,a-W(t_n))) \\
&= F(t-t_n;\nu,\sigma,a-W(t_n))
\end{aligned}
$$

if $\tau < t_n$, where $F(\cdot;\beta,\gamma,\delta)$ is the inverse Gaussian cumulative distribution function given in equation (3.2.1) with drift parameter β, variance parameter γ and threshold parameter δ.

In the situation where $\tau > t_n$, because of the assumption that the threshold is not crossed before τ, $P(T < \tau) = 0$, the process should be shifted to the new point $(\tau, W(\tau))$, from which the shifted threshold time is inverse Gaussian, $S = T - \tau \sim \mathcal{IG}(\nu, \sigma, a - W(\tau))$. However, $W(\tau)$ is in the future, thus unknown. Conditioned on $W(\tau)$, the distribution is known:

$$\begin{aligned} P(T \leq t|\tau, W(\tau), x_1, \ldots, x_n; \nu, \sigma, a) &= P(T - t_n \leq t - t_n|\tau, W(\tau), x_1, \ldots, x_n; \nu, \sigma, a)) \\ &= P(S \leq t - t_n|\tau, W(\tau), x_1, \ldots, x_n; \nu, \sigma, a - W(\tau)) \\ &= F(t - t_n; \nu, \sigma, a - W(\tau)). \end{aligned}$$

We can marginalize over $W(\tau)$. Let $X_{\tau - t_n}$ denote the temperature increment from t_n to τ. Remember that $X_{\tau - t_n}$ is normally distributed with mean 0 and variance $(\tau - t_n)\sigma^2$, that is

$$X_{\tau - t_n} \sim \mathcal{N}(0, (\tau - t_n)\sigma^2).$$

Thus,

$$P(T \leq t|\tau, x_1, \ldots, x_n; \nu, \sigma, a) = \int_{-\infty}^{a} F(t - t_n; \nu, \sigma, a - (W(t_n) + x)) f_{X_{\tau - t_n}}(x) dx$$

where $f_{X_{\tau - t_n}}(x)$ is the normal pdf with mean 0 and variance $(\tau - t_n)\sigma^2$.

In practice, the integration may be tedious. Two solutions will be suggested in the following. The first is to employ MC summation by drawing random increments from $\mathcal{N}(0, (\tau - t_n)\sigma^2)$.

The second solution is a simplification, approximating $W(\tau) \approx W(t_n)$. Note that given $W(t_n)$, the expected value of $W(\tau)$ is $E[W(\tau)|W(t_n)] = W(t_n)$.

Thus, using MC summation to draw M values of $W(\tau)$ for each $\tau, w_1, \ldots, w_M$ from the given normal distribution and N values of $\tau, \tau_1, \ldots, \tau_N$, from the posterior distribution the marginal distribution of T given the measurements $x_1, \ldots, x_n$ can be found as

$$\begin{aligned} P(T \leq & t|x_1, \ldots, x_n; \nu, \sigma, a) = \\ & \frac{1}{N} \sum_{i=1}^{N} P(T|\tau_i, x_1, \ldots, x_n; \nu, \sigma, a - W(t_n)) I_{\tau_i \leq t_n} + \\ & \frac{1}{N} \frac{1}{M} \sum_{i=1}^{N} \sum_{m=1}^{M} P(T|\tau_i, x_1, \ldots, x_n; \nu, \sigma, a - w_m) I_{\tau_i > t_n}. \end{aligned} \tag{4.2.4}$$

where I_a is the indicator function, which is one of a is true, and 0 if a is false. Using the simplifying approximation $W(\tau) \approx W(t_n)$, we find the marginal distribution of T given the measurements $x_1, \ldots, x_n$ as

$$P(T \leq t|x_1, \ldots, x_n; \nu, \sigma, a) = \frac{1}{N} \sum_{i=1}^{N} P(T|\tau_i, x_1, \ldots, x_n; \nu, \sigma, a - W(t_n)). \tag{4.2.5}$$

Note that this approximation will cause less variation in the results than using the MC summation in equation (4.2.4) above.

Thus, finally, we find formulaic estimate of the marginal distribution of T given the temperature measurements by summing over τ using an MC summation: sampling over the posterior distribution of τ:

$$\widehat{P}_f(T \leq t|x_1, \ldots, x_n; \nu, \sigma, a) = \frac{1}{N} \sum_{i=1}^{N} P(T|\tau_i^*, x_1, \ldots, x_n; \nu, \sigma, a),$$

where the τ_i^* are a sample from the posterior distribution of τ, $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$. The pseudo-algorithm for the formulaic approach is given in algorithm 2.

Algorithm 2 Formulaic approach for estimating the hitting time CDF of $t = t'$ for piecewise Wiener process with one unknown change point τ.

Require: $m, \{W(t_i), i = 1, \ldots, n\}, \nu, \sigma, a > W(t_n), t'$
  $\Delta = t_n/n$
  $x_i = W((i+1) \cdot \Delta) - W(i \cdot \Delta)$
  Find posterior distribution of τ: $\pi(\tau|x_1, \ldots, x_n)$
  for $k = 1$ to m **do**
    Sample τ_k^* from $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$
    Calculate $P(T \leq t'|\tau_k^*, x_1, \ldots, x_n; \nu, \sigma, a)$ from equation (4.2.4) or (4.2.5).
  end for
  return $\widehat{P}_f(T \leq t'|x_1, \ldots, x_n; \nu, \sigma, a) = \frac{1}{m} \sum_{k=1}^{m} P(T \leq t'|\tau_k^*, x_1, \ldots, x_n; \nu, \sigma, a)$

4.2.2 Simulation approach

The simulation estimator of $P(T \leq t|x_1, \ldots, x_n; \nu, \sigma, a)$ will be denoted $\widehat{P}_S(T \leq t|x_1, \ldots, x_n; \nu, \sigma, a)$. The simulation approach estimates $P(T \leq t|\tau, x_1, \ldots, x_n; \nu, \sigma, a)$ by simulating a number, say K, of Wiener processes from the last observation, at time t_n and temperature $W(t_n)$, until the processes cross the threshold temperature a, that is, until time T_i and temperature a. This is done by drawing temperature increments X_j for each time point, $t = j\Delta$, $j = n+1, n+2, \ldots$ until the process reaches temperature a, such that the time, say $l\Delta$ is the hitting time T_i. As long as $\tau < \infty$, T has a proper distribution, and the hitting time will occur with probability 1, hence, each simulation of a Wiener process will result in a hitting time. Note that conditioned on τ, the distribution of X_j is known:

$$X_j|\tau \sim \begin{cases} \mathcal{N}(0, \Delta\sigma^2) & \text{if } \tau > j\Delta. \\ \mathcal{N}(\Delta\nu, \Delta\sigma^2) & \text{if } \tau \leq j\Delta, \end{cases}$$

thus, the above gives the rule for how the temperature increments are sampled.

The respective hitting times $T_1, \ldots, T_k$ are stored. The hitting times T_k are i.i.d. from the hitting time distribution. Thus $P(T \leq t|\tau, x_1, \ldots, x_n; \nu, \sigma, a)$ can be estimated

as the empirical CDF, $\widehat{F}_{K,\tau_j^*}(t)$,

$$\widehat{F}_{K,\tau_j^*}(t) = \frac{1}{K}\sum_{i=1}^{K} I_{T_i \leq t}. \tag{4.2.6}$$

The simulated solution does not require the assumption that $T > \tau$. Thereby, the approach simultaneously provides a simple way to check the assumption, as we can estimate $P(T \leq \tau | \tau, x_1, \ldots, x_n; \nu, \sigma, a)$ by $\widehat{F}_{K,\tau_j^*}(\tau_j^*)$.

Note that in the formulaic approach, we must calculate $P(T \leq t | \tau, x_1, \ldots, x_n; \nu, \sigma, a)$ for each t, but in the simulation approach, can obtain the estimates for all t from the same data set. Also note that in the formulaic approach, $P(T \leq t | \tau, x_1, \ldots, x_n; \nu, \sigma, a)$ is readily available, but in the simulation approach, the simulations give a sample from the population of $P(T \leq t | \tau, x_1, \ldots, x_n; \nu, \sigma, a)$.

Finally, the marginalized distribution of T given $x_1, \ldots, x_n$, $P(T \leq t | x_1, \ldots, x_n; \nu, \sigma, a)$ is found as for the formulaic approach, by drawing a large number N of change point realizations τ_i^* from the posterior distribution, and let the mean denote the simulation estimator $\widehat{P}_S(T \leq t | x_1, \ldots, x_n; \nu, \sigma, a)$ for the probability $P(T \leq t | x_1, \ldots, x_n; \nu, \sigma, a)$:

$$\widehat{P}_S(T \leq t | x_1, \ldots, x_n; \nu, \sigma, a) = \frac{1}{N}\sum_{i=1}^{N} \widehat{F}_{K,\tau_i^*}(t). \tag{4.2.7}$$

The simulation is described in algorithm 3. The description is schematic, and implementation can of course be done much more efficient.

4.2.3 Evaluating estimates

In both approaches, for each t, the estimate of $P(T \leq t | x_1, \ldots, x_n; \nu, \sigma, a)$ is found by calculating the mean of M values $P(T \leq t | \tau_i^*, x_1, \ldots, x_n, \nu, \sigma, a)$ for a M values of τ_i^*. Thus, a $1 - \alpha$ credible interval is given by the $1 - \alpha$ and α percentiles of the M values $P(T \leq t | \tau_i^*, x_1, \ldots, x_n, \nu, \sigma, a)$.

4.2.4 τ and ν unknown, σ known

If ν is an unknown parameter, we introduce a prior for ν in order to obtain the joint posterior distribution $\pi(\tau, \nu, | x_1, \ldots, x_n; \sigma)$. We assume that ν is independent of the change point τ, such that the joint prior of the two parameters $\pi(\tau, \nu)$ is the product of their marginal priors:

$$\pi(\tau, \nu) = \pi(\nu)\pi(\tau).$$

Since we require $\nu > 0$, a natural prior distribution is the gamma distribution, which has support on the interval $(0, \infty)$, such that

$$\nu \sim \mathcal{G}am(\alpha, \beta).$$

Algorithm 3 Simulation to find hitting time CDF for piecewise Wiener process with one unknown change point τ, ν and σ known.

Require: $J, \{W(t_i), i = 1, \ldots, n\}, \nu, \sigma, a > W(t_n), K$
  $\Delta = t_n/n$
  $x_i = W(i \cdot \Delta) - W((i-1) \cdot \Delta)$
  Find posterior distribution of τ: $\pi(\tau | x_1, \ldots, x_n; \nu, \sigma)$
  $t = t_n$
  for $j = 1$ to J **do**
    Sample τ_j^* from $\pi(\tau | x_1, \ldots, x_n; \nu, \sigma)$
    for $k = 1$ to K **do**
      $t = t + \Delta$
      while $W(t - \Delta) \leq a$ **do**
        if $\tau_j^* \leq t_n$ **then**
          Draw increment
          $X(t) \sim \mathcal{N}(\nu \cdot \Delta, \sigma^2 \cdot \Delta)$
        else
          Draw increment
          $X(t) \sim \mathcal{N}(0, \sigma^2 \cdot \Delta)$
        end if
        $W(t) = W(t - \Delta) + X(t)$
        $t = t + \Delta$
      end while
      $T_k = t$ is the k-th hitting time given τ_j^*
    end for
    $\widehat{F}_{K,\tau_j^*}(t) = \frac{1}{K} \sum_{k=1}^{K} I_{T_k} \leq t$ is the estimate of $P(T \leq t | \tau_j^*, x_1, \ldots, x_n; \nu, \sigma, a)$.
  end for
  Estimate $P(T \leq t | x_1, \ldots, x_n; \nu, \sigma, a)$ by

$$\widehat{P}_S(T \leq t | x_1, \ldots, x_n; \nu, \sigma, a) = \frac{1}{J} \sum_{j=1}^{J} \widehat{F}_{K,\tau_j^*}(t).$$

  return $\widehat{P}_S(T \leq t | x_1, \ldots, x_n; \nu, \sigma, a)$.

That is, the probability density function, pdf, of ν, $f(\nu)$ is given by

$$f(\nu) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \nu^{\alpha-1} \exp\left(-\frac{\nu}{\beta}\right).$$

where α and β are prior parameters which must be tuned such that the chain can move quickly across the entire domain. Note that the gamma distribution is sometimes represented by the rate parameter $1/\beta$.

It is often recommended to have an uninformative prior for parameters θ, such as the uniform distribution, which has density

$$f_\theta(\theta) = \frac{1}{b-a} \quad \text{if } \theta \in [a, b]$$

and 0 otherwise, or Jeffreys prior, which has density

$$f_\theta(\theta) \propto \sqrt{\det(\mathcal{I})}$$

where $\mathcal{I}$ is the Fisher information,

$$\mathcal{I}(\theta) = E_\theta \left[\left(\frac{\partial}{\partial \theta} \log f(X; \theta) \right)^2 \right]$$

where $E_\theta[\cdot]$ signifies the expectation with respect to the parameter θ, and $f(X; \theta)$ is the density of the data X given the parameter θ. However, in multidimensional cases it can be hard to keep track of the implications of the marginal uninformative prior on the joint prior distribution, which may be the case if we choose to continue the geometric prior on τ, as is implied by the Markov model. Note also that in the formulaic approach, we use the inverse Gaussian CDF,

$$F(t; \nu, \sigma, a) = \Phi\left(\frac{\nu t - a}{\sigma\sqrt{t}}\right) + \exp\left(\frac{2a\nu}{\sigma^2}\right) \Phi\left(\frac{-a - \nu t}{\sigma\sqrt{t}}\right)$$

from equation (3.2.1). Note that in the MCMC procedure, if a large value of $\nu^{(t)}$ is drawn and accepted, the exponential term blows up and cause numerical problems in the calculations. Thus, it may be necessary to choose priors that limit the occurrences of such events.

The likelihood from equation (4.2.2) remains the same, except it is now a function of both τ and ν :

$$\begin{aligned} &L(\tau, \nu | x_1, \ldots, x_n; \sigma) \propto \\ &\quad \begin{cases} \exp\left(-\frac{1}{2\Delta\sigma^2}\left[2\Delta\nu \sum_{i=\tau/\Delta}^{t_n/\Delta}(x_i) + (n - \tau + 1)\,\Delta^2\nu^2\right]\right) & \text{if } \tau \le t_n \\ 1 & \text{if } \tau > t_n. \end{cases} \end{aligned} \tag{4.2.8}$$

Thus the posterior is simply the posterior in equation (4.2.3) multiplied by the prior distribution for ν, $\pi(\nu)$, $\pi(\tau, \nu | x_1, \ldots, x_n; \sigma) \propto L(\tau, \nu | x_1, \ldots, x_n; \sigma)\pi(\tau)\pi(\nu)$.

With ν unknown, the posterior distribution, as the likelihood, becomes two-dimensional, $\pi(\tau, \nu|x_1, \ldots, x_n; \sigma)$. However, with the prior $\pi(\tau, \nu) = \pi(\tau)\pi(\nu)$ as the proposal, the Metropolis-Hastings ratio $R([\tau, \nu]^{(t)}, [\tau, \nu]^*)$ remains the likelihood ratio, now with both ν and τ as variables.

Note that both the simulation and formulaic approach will be conceptually the same with the new unknown parameter. Both will now sample pairs of (τ_i^*, ν_i^*) realizations from the posterior distribution. The simulation approach now relies on drawing pairs of (τ_i^*, ν_i^*) realizations from the posterior distribution, and thereafter simulating temperature increments according to the rule

$$X_j|\tau, \nu \sim \begin{cases} \mathcal{N}(0, \Delta\sigma^2) & \text{if } \tau_i^* > j\Delta \\ \mathcal{N}(\Delta\nu_i^*, \Delta\sigma^2) & \text{if } \tau_i^* \leq j\Delta \end{cases}$$

The simulation for each parameter pair (τ_i^*, ν_i^*) is stopped when the threshold temperature a is reached at the hitting time T_i. Thus, the simulation approach becomes:

$$\widehat{P}_S(T \leq t|x_1, \ldots, x_n; a, \nu, \sigma) = \frac{1}{N}\sum_{i=1}^{N} \widehat{F}_{K,\tau_i^*,\nu_i^*}(t), \tag{4.2.9}$$

where $\widehat{F}_{K,\tau_i^*,\nu_i^*}(t)$ is the empirical CDF given by

$$\widehat{F}_{K,\tau_i^*,\nu_i^*}(t) = \frac{1}{K}\sum_{i=1}^{K} I_{T_i \leq t} \tag{4.2.10}$$

where the K pairs of values (τ_i^*, ν_i^*) have been sampled from the posterior distribution $\pi(\tau, \nu|x_1, \ldots, x_n; \sigma)$.

The difference in the formulaic approach, apart from drawing parameter pairs (τ_i^*, ν_i^*) from the two-dimensional posterior distribution $\pi(\tau, \nu|x_1, \ldots, x_n; \sigma)$, is merely that ν_i^* is the drift parameter in the expression for the inverse Gaussian CDF,

$$P(T \leq t|\tau_i^*, \nu_i^*, x_1, \ldots, x_n; \sigma, a) = F(T - \max(\tau_i^*, t_n); \nu_i^*, \sigma, a - W(\max\{\tau_i^*, t_n\})),$$

where, if $\tau_i^* > t_n, W(\tau_i^*)$ is resolved by MC integration or simplified by $W(\tau_i^*) \approx W(t_n)$ as discussed. which is then used in to calculate the estimator,

$$\widehat{P}_f(T \leq t|x_1, \ldots, x_n; \sigma, a) = \frac{1}{m}\sum_{k=1}^{m} P(T \leq t|\tau_k^*, x_1, \ldots, x_n; \nu_k^*, \sigma, a). \tag{4.2.11}$$

Evaluating estimators

Note that as with only one unknown parameter, both the simulation approach estimate, $P_S(T \leq t|x_1, \ldots, x_n; \sigma, a)$, and the formulaic approach estimate, $P_f(T \leq t|x_1, \ldots, x_n; \sigma, a)$, are given as means of vectors. Thus, the same method of finding a credible interval is used: a $1 - \alpha$ credible interval is given by the $1 - \alpha$ and α percentiles of the M values $P(T \leq t|\tau_i^*, x_1, \ldots, x_n; \nu_i^*, \sigma, a)$ in the summation for both methods.

4.2.5 τ, ν and σ unknown

Consider now the case where the change point τ, the drift parameter ν and the variance parameter σ is unknown. In this case the likelihood is changed:

$$\begin{aligned}
L(\tau,\nu,\sigma|x_1,\ldots,x_n) &= \prod_{i=1}^{n} f(x_i|\tau,\nu,\sigma) \\
&= \prod_{i=1}^{\tau-1} \frac{1}{\sqrt{2\pi\Delta}\sigma} \exp\left(-\frac{x_i^2}{2\Delta\sigma^2}\right) \prod_{i=\tau}^{n} \frac{1}{\sqrt{2\pi\Delta}\sigma} \exp\left(-\frac{(x_i-\Delta\nu)^2}{2\Delta\sigma^2}\right) \\
&\propto \frac{1}{\sigma^n} \exp\left(-\frac{1}{2\Delta\sigma^2}\left[\sum_{i=1}^{\tau-1} x_i^2 + \sum_{i=\tau}^{n}(x_i-\Delta\nu)^2\right]\right) \\
&\propto \frac{1}{\sigma^n} \exp\left(-\frac{1}{2\Delta\sigma^2}\left[\left(\sum_{i=1}^{n} x_i^2\right) + 2\Delta\nu\left(\sum_{i=\tau}^{n} x_i\right) + (n-\tau+1)\,\Delta^2\nu^2\right]\right).
\end{aligned}$$

if $\tau \leq t_n$ and

$$\begin{aligned}
L(\tau,\nu,\sigma|x_1,\ldots,x_n) &= \prod_{i=1}^{n} f(x_i|\tau,\nu,\sigma) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\Delta}\sigma} \exp\left(-\frac{x_i^2}{2\Delta\sigma^2}\right) \\
&\propto \frac{1}{\sigma^n} \exp\left(-\frac{1}{2\Delta\sigma^2}\left[\sum_{i=1}^{n} x_i^2\right]\right).
\end{aligned}$$

if $\tau > t_n$. Putting it together, we obtain

$$L(\tau,\nu,\sigma|x_1,\ldots,x_n) \propto \begin{cases} \frac{1}{\sigma^n}\exp\left(-\frac{1}{2\Delta\sigma^2}\left[\left(\sum_{i=1}^{n} x_i^2\right) + 2\Delta\nu\left(\sum_{i=\tau}^{n} x_i\right) + (n-\tau+1)\,\Delta^2\nu^2\right]\right) & \text{if } \tau \leq t_n \\ \frac{1}{\sigma^n}\exp\left(-\frac{1}{2\Delta\sigma^2}\left[\sum_{i=1}^{n} x_i^2\right]\right) & \text{if } \tau > t_n. \end{cases} \tag{4.2.12}$$

We propose to continue the gamma prior on ν, and add a gamma prior on σ, assuming independence of all the three parameters,

$$\pi(\tau,\nu,\sigma) = \pi(\tau)\pi(\nu)\pi(\sigma).$$

where $\tau \sim \mathcal{G}eom(\lambda)$ as before, $\nu \sim \mathcal{G}am(\nu_\alpha,\nu_\beta)$ as before and $\sigma \sim \mathcal{G}am(\sigma_\alpha,\sigma_\beta)$. We make that note that the numerical problems with the exponential term in the inverse Gaussian CDF, see equation (3.2.3), may worsen now that $\sigma^{(t)}$ may be a small value in addition to $\nu^{(t)}$ being large.

The parameters σ_α and σ_β should be tuned such that they do not influence the prior too much, unless the information available is very strong, and such that the Markov chain is allowed to move across the entire domain quickly.

Note that both the simulation and formulaic approach will be conceptually the same with the third unknown parameter. Both will now sample triplets of $(\tau_i^*, \nu_i^*, \sigma_i^*)$ realizations from the posterior distribution. The simulation approach now relies on drawing triplets of $(\tau_i^*, \nu_i^*, \sigma_i^*)$ realizations from the posterior distribution, and thereafter simulating temperature increments according to the rule

$$X_j \sim \begin{cases} \mathcal{N}(0, \Delta(\sigma_i^*)^2) & \text{if } \tau_i^* > j\Delta \\ \mathcal{N}(\Delta\nu_i^*, \Delta(\sigma_i^*)^2) & \text{if } \tau_i^* \leq j\Delta \end{cases}$$

The simulation for each parameter triplet realization $(\tau_i^*, \nu_i^*, \sigma_i^*)$ is stopped when the threshold temperature a is reached at the hitting time T_i. Thus, the simulation approach becomes:

$$\widehat{P}_S(T \leq t|x_1, \ldots, x_n; a) = \frac{1}{N}\sum_{i=1}^{N} \widehat{F}_{K,\tau_i^*,\nu_i^*,\sigma_i^*}(t), \tag{4.2.13}$$

where $\widehat{F}_{K,\tau_i^*,\nu_i^*,\sigma_i^*}(t)$ is the empirical CDF given by

$$\widehat{F}_{K,\tau_i^*,\nu_i^*,\sigma_i^*}(t) = \frac{1}{K}\sum_{i=1}^{K} I_{T_i \leq t} \tag{4.2.14}$$

where the K triplets of realizations $(\tau_i^*, \nu_i^*, \sigma_i^*)$ have been sampled from the posterior distribution $\pi(\tau, \nu, \sigma|x_1, \ldots, x_n)$.

The difference in the formulaic approach, apart from drawing parameter triplet realizations $(\tau_i^*, \nu_i^*, \sigma_i^*)$ from the three-dimensional posterior distribution $\pi(\tau, \nu, \sigma|x_1, \ldots, x_n)$, is that ν_i^* is the drift parameter and σ_i^* is the variance parameter in the expression for the inverse Gaussian CDF,

$$P(T \leq t|\tau_i^*, \nu_i^*, \sigma_i^*, x_1, \ldots, x_n) = F(T - \max(\tau_i^*, t_n); \nu_i^*, \sigma_i^*, a).$$

which is then used in to calculate the estimator,

$$\widehat{P}_f(T \leq t|x_1, \ldots, x_n; a) = \frac{1}{m}\sum_{k=1}^{m} P(T \leq t|\tau_k^*, x_1, \ldots, x_n; \nu_k^*, \sigma_k^*, a). \tag{4.2.15}$$

Evaluating estimators

Again, as with one and two unknown parameters, both the simulation approach estimate, $P_S(T \leq t|x_1, \ldots, x_n; a)$, and the formulaic approach estimate, $P_f(T \leq t|x_1, \ldots, x_n, a)$, are given as means of vectors. Thus, the same method of finding a credible interval is used: a $1-\alpha$ credible interval is given by the $1-\alpha$ and α percentiles of the M values $P(T \leq t|\tau_i^*, x_1, \ldots, x_n; \nu_i^*, \sigma_i^*, a)$ for both methods.

4.3 Two unknown change points, τ_1 and τ_2

We now assume that there are two change points, τ_1 and τ_2 with corresponding parameters ν_1 and σ_1, ν_2 and σ_2. As before, $0 = \nu_0$ and σ_0 are the parameters of the initial process. Three different strategies will be outlined to find the hitting time CDF. The first, in section 4.3.1, is equivalent to the formulaic approach in the process with one change point, described in section 4.2.1. The third approach, in section 4.3.3, is equivalent to the simulation approach in the process with one change point, described in section 4.2.2. The approach presented in section 4.3.2 elegantly simplifies the formulaic approach by a time transformation and is applicable to the process with two or more change points. In the formulaic approaches, with and without the time transformation, the hitting time CDF will be found by conditioning on the two change points τ_1 and τ_2, and then the conditioned variables will be marginalized out by MC summation. In the simulation approach, the method will be similar: the simulations are done conditioned on the change points, which are marginalized out by MC summation. Again, having found the hitting time CDF, one can choose the estimator $\widehat{T}_\alpha$ as the t for which $P(T \leq t|x_1, \ldots, x_n; \nu, \sigma, a) = \alpha$, where α varies with the level of risk one is willing to take that the threshold has already been crossed.

Thus, we need the distribution of the change points. As in section 4.2 we will use the Bayesian posterior distribution.

The likelihood of the data is, for $\tau_1 < \tau_2 \leq n$:

$$
\begin{aligned}
L(\tau_1, \tau_2|x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2) &= \prod_{i=1}^{n} f(x_i|\tau_1, \tau_2; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2) \\
&= \prod_{i=1}^{\tau_1/\Delta-1} \frac{1}{\sqrt{2\pi\Delta}\sigma_0} \exp\left(-\frac{x_i^2}{2\Delta\sigma_0^2}\right) \times \prod_{i=\tau_1/\Delta}^{\tau_2/\Delta-1} \frac{1}{\sqrt{2\pi\Delta}\sigma_1} \exp\left(-\frac{(x_i - \Delta\nu_1)^2}{2\Delta\sigma_1^2}\right) \times \\
&\quad \prod_{i=\tau_2/\Delta}^{t_n/\Delta} \frac{1}{\sqrt{2\pi\Delta}\sigma_2} \exp\left(-\frac{(x_i - \Delta\nu_2)^2}{2\Delta\sigma_2^2}\right) \\
&\propto \frac{1}{\sigma_0^{\tau_1-1}} \frac{1}{\sigma_1^{\tau_2-\tau_1}} \frac{1}{\sigma_2^{n-\tau_2+1}} \times \\
&\quad \exp\left(-\frac{1}{2\Delta\sigma_0^2} \sum_{i=1}^{\tau_1/\Delta-1} x_i^2 - \frac{1}{2\Delta\sigma_1^2} \sum_{i=\tau_1/\Delta}^{\tau_2/\Delta-1} (x_i - \Delta\nu_1)^2 - \frac{1}{2\Delta\sigma_2^2} \sum_{i=\tau_2/\Delta}^{t_n/\Delta} (x_i - \Delta\nu_2)^2\right).
\end{aligned}
$$

For $\tau_1 \leq n < \tau_2$ the likelihood is:

$$\begin{aligned}
L(\tau_1,\tau_2|x_1,\ldots,x_n;\nu_1,\nu_2,\sigma_0,\sigma_1,\sigma_2) &= \prod_{i=1}^{n} f(x_i|\tau_1,\tau_2;\nu_1,\nu_2,\sigma_0,\sigma_1,\sigma_2) \\
&= \prod_{i=1}^{\tau_1/\Delta-1} \frac{1}{\sqrt{2\pi\Delta}\sigma_0}\exp\left(-\frac{x_i^2}{2\Delta\sigma_0^2}\right) \times \prod_{i=\tau_1/\Delta}^{t_n/\Delta} \frac{1}{\sqrt{2\pi\Delta}\sigma_1}\exp\left(-\frac{(x_i-\Delta\nu_1)^2}{2\Delta\sigma_1^2}\right) \\
&\propto \frac{1}{\sigma_0^{\tau_1-1}}\frac{1}{\sigma_1^{n-\tau_1+1}} \times \exp\left(-\frac{1}{2\Delta\sigma_0^2}\sum_{i=1}^{\tau_1/\Delta-1} x_i^2 - \frac{1}{2\Delta\sigma_1^2}\sum_{i=\tau_1/\Delta}^{t_n/\Delta}(x_i-\Delta\nu_1)^2\right).
\end{aligned}$$

For $n < \tau_1 < \tau_2$, we have as for one change point

$$\begin{aligned}
L(\tau_1,\tau_2|x_1,\ldots,x_n;\nu_1,\nu_2,\sigma_0,\sigma_1,\sigma_2) &= \prod_{i=1}^{n} f(x_i|\tau_1,\tau_2;\nu_1,\nu_2,\sigma_0,\sigma_1,\sigma_2) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\Delta}\sigma_0}\exp\left(-\frac{x_i^2}{2\Delta\sigma_0^2}\right) \\
&= \frac{1}{\sigma_0^n}\exp\left(-\frac{1}{2\Delta\sigma_0^2}\sum_{i=1}^{n} x_i^2\right)
\end{aligned}$$

The marginal prior distribution of τ_1 is geometric as before:

$$\pi(\tau_1) = \lambda_1(1-\lambda_1)^{\tau_1-1}, \;\; \tau_1 = 1,2,\ldots$$

and 0 otherwise. The time spent in each state is independent, so the prior for τ_2 conditioned on τ_1 is also geometric:

$$\pi(\tau_2|\tau_1) = \lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}, \;\; \tau_1 = 1,2,\ldots;\; \tau_2 = \tau_1+1, \tau_1+2,\ldots$$

and 0 otherwise. Thus the joint prior of τ_1, τ_2 is

$$\begin{aligned}
\pi(\tau_1,\tau_2) &= \pi(\tau_2|\tau_1)\pi(\tau_1) \\
&= \lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}, \;\; \tau_1 = 1,2,\ldots;\; \tau_2 = \tau_1+1, \tau_2+2,\ldots,
\end{aligned} \tag{4.3.1}$$

and 0 otherwise.

Thus the posterior distribution of τ_1, τ_2 given the temperature increment measurements $x_1,\ldots,x_n$ is

$$\pi(\tau_1,\tau_2|x_1,\ldots,x_n;\nu_1,\nu_2,\sigma_0,\sigma_1,\sigma_2) \propto L(\tau_1,\tau_2|x_1,\ldots,x_n;\nu_1,\nu_2,\sigma_0,\sigma_1,\sigma_2)\pi(\tau_1,\tau_2) \propto$$

$$\begin{cases}
\frac{\lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}}{\sigma_0^{\tau_1-1}\sigma_1^{\tau_2-\tau_1}\sigma_2^{n-\tau_2+1}}\times \\
\quad \exp\left(-\sum_{i=1}^{\tau_1/\Delta-1}\frac{x_i^2}{2\Delta\sigma_0^2} - \sum_{i=\tau_1/\Delta}^{\tau_2/\Delta-1}\frac{(x_i-\Delta\nu_1)^2}{2\Delta\sigma_1^2} - \sum_{i=\tau_2/\Delta}^{n}\frac{(x_i-\Delta\nu_2)^2}{2\Delta\sigma_2^2}\right), & \tau_1 < \tau_2 \leq t_n \\
\frac{\lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}}{\sigma_0^{\tau_1-1}\sigma_1^{n-\tau_1+1}}\times \\
\quad \exp\left(-\sum_{i=1}^{\tau_1/\Delta-1}\frac{x_i^2}{2\Delta\sigma_0^2} - \sum_{i=\tau_1/\Delta}^{n}\frac{(x_i-\Delta\nu_1)^2}{2\Delta\sigma_1^2}\right), & \tau_1 \leq t_n < \tau_2. \\
\frac{\lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}}{\sigma_0^n}\exp\left(-\frac{1}{2\Delta\sigma_0^2}\sum_{i=1}^{n} x_i^2\right), & t_n < \tau_1 < \tau_2.
\end{cases}$$

4.3.1 Formulaic approach

The (τ_1, τ_2) plane is sectioned into five regions, for which we will develop different expressions for the hitting time distribution conditioned on the change points. The five cases are illustrated in figure 4.3.1, and each region is bounded by the black lines. We continue the assumption from Lindqvist and Slimacek (2013) that the threshold temperature a will not be crossed before the first change point τ_1 has occurred, that is $P(T \leq \tau_1) = 0$. Following the assumption, in each case, we consider the distribution of the shifted hitting time: $S = T - \max\{t_n, \tau_1\}$. Note also that $P(T \leq t|\tau_1 \geq \tau_2) = 0$, since we have assumed $\tau_1 < \tau_2$. Now, we can go on to establish the expressions for the

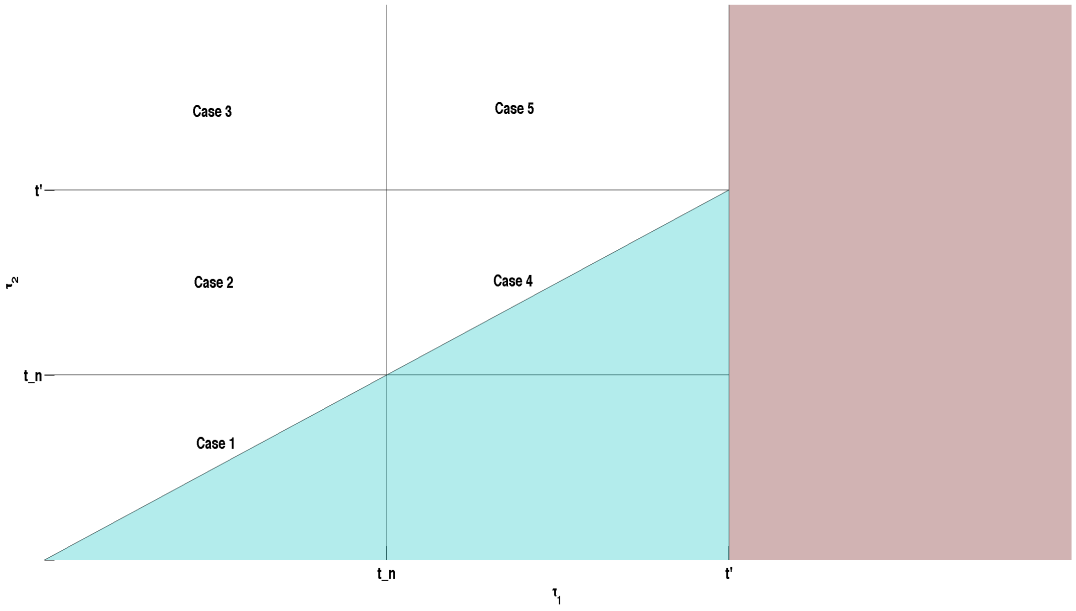

Figure 4.3.1: Illustration of the different regions in the (τ_1, τ_2) area. The light blue region has probability zero since $\tau_1 > \tau_2$, which by assumption cannot happen. The light red region has probability zero because of the assumption that the threshold will not be crossed unless the first change point is crossed, $P(T \leq \tau_1) = 0$.

conditional CDF in each of the five cases. In the following, we will simplify notation of $P(T \leq t|\tau_1, \tau_2, x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2)$ to $P(T \leq t)$, that is, it is implicitly understood that we are conditioning on τ_1, τ_2 in the current region.

Case 1: $\tau_1 < \tau_2 \leq t_n < t'$

In this case, the shifted hitting time, $S = T - t_n$, is simply inverse Gaussian distributed, $S \sim IG(\nu_2, \sigma_2, a - W(t_n))$. Thus,

$$P(T \leq t') = P(T - t_n \leq t' - t_n) = P(S \leq t' - t_n) = F(t' - t_n; \nu_2, \sigma_2, a - W(t_n)) \quad (4.3.2)$$

is straightforwardly found by the inverse Gaussian CDF. As before, $F(\cdot;\nu,\sigma,a)$ denotes the inverse Gaussian CDF with drift parameter ν variance parameter σ and threshold parameter a as given in equation (3.2.3).

Case 2: $\tau_1 \leq t_n < \tau_2 \leq t'$

In this case, only the first change point has occurred. However, τ_2 is in the future, and thus $W(\tau_2)$ is unknown. This is analogous to τ in the future in the one change point setting. Hence, we condition on the unknown $W(\tau_2)$:

$$P(T \leq t') = \int_{-\infty}^{\infty} P(T \leq t' | W(\tau_2) = r) f_{W(\tau_2)}(r) dr$$

where $f_{W(\tau_2)}(r)$ is the normal pdf, $W(\tau_2) \sim \mathcal{N}(W(t_n) + (\tau_2 - t_n)\nu_1, (\tau_2 - t_n)\sigma_1^2)$.

We can split the conditional probability in two:

$$P(T \leq t' | W(\tau_2) = r) = P(T \leq \tau_2 | W(\tau_2) = r) + P(\tau_2 < T \leq t' | W(\tau_2) = r).$$

Consider the term $P(T \leq \tau_2 | W(\tau_2) = r)$, when $T \leq \tau_2$, then $T = s$ for some $s \leq \tau_2$, and $W(\tau_2) \geq a$. Hence, $W(\tau_2) - W(s) = r - a$. Thus

$$\begin{aligned} P(T \leq \tau_2 | W(\tau_2) = r) &= \frac{P(T \leq \tau_2, W(\tau_2) = r)}{P(W(\tau_2) = r)} \\ &= \frac{\int_0^{\tau_2 - t_n} P(T \leq \tau_2, W(\tau_2) = r | T = s) f_T(s) ds}{f_{W(\tau_2)}(r)} \\ &= \frac{\int_0^{\tau_2 - t_n} f_{W(\tau_2)-W(s)}(r-a) f_T(s) ds}{f_{W(\tau_2)}(r)}, \end{aligned}$$

where $f_{W(\tau_2)}(r)$ is the normal pdf as before, $f_{W(\tau_2)-W(s)}(r-a)$ is the pdf of the increment $W(\tau_2) - W(s)$, which is normally distributed: $W(\tau_2) - W(s) \sim \mathcal{N}((\tau_2 - s)\nu_1, (\tau_2 - s)\sigma_1^2)$, and $f_T(s)$ is the shifted inverse Gaussian density $T - t_n \sim \mathcal{IG}(\nu_1, \sigma_1, a - W(t_n))$. The shift explains the why integration limits are 0 and $\tau_2 - t_n$ instead of t_n and τ_2.

Consider now the latter term, $P(\tau_2 < T \leq t' | W(\tau_2) = r)$. Note that here we only need consider $W(\tau_2) = r < a$, since if $r \geq a, P(\tau_2 < T) = 0$. We can write

$$\begin{aligned} P(\tau_2 < T \leq t' | W(\tau_2 = r)) &= P(\tau_2 < T, T \leq t' | W(\tau_2) = r) \\ &= P(T \leq t' | \tau_2 < T, W(\tau_2) = r) P(\tau_2 < T | W(\tau_2) = r) \end{aligned}$$

Note that given $\tau_2 < T$ and $W(\tau_2) = r$, the shifted $S = T - t_n$ is inverse Gaussian CDF, $S \sim \mathcal{IG}(\nu_2, \sigma_2, a - r)$. Hence,

$$\begin{aligned} P(\tau_2 < T \leq t' | W(\tau_2 = r)) &= F(t' - t_n; \nu_2, \sigma_2, a - r) P(\tau_2 < T | W(\tau_2) = r) \\ &= F(t' - t_n; \nu_2, \sigma_2, a - r) \left(1 - P(T \leq \tau_2 | W(\tau_2) = r)\right) \end{aligned}$$

where $P(T \leq \tau_2 | W(\tau_2) = r)$ was found above.

Finally, we can put everything together:

$$
\begin{aligned}
&P(T \le t') = \int_{-\infty}^{\infty} P(T \le t' | W(\tau_2) = r) f_{W(\tau_2)}(r) dr \\
&= \int_{-\infty}^{\infty} \left(P(T \le \tau_2 | W(\tau_2) = r) + P(\tau_2 < T \le t' | W(\tau_2) = r) \right) f_{W(\tau_2)}(r) dr \\
&= \int_{-\infty}^{\infty} \int_{0}^{\tau_2 - t_n} f_{W(\tau_2)-W(s)}(r-a) f_T(s) ds dr + \\
&\int_{-\infty}^{a} F(t - \tau_2; \nu_2, \sigma_2, a - r) \left(1 - \frac{\int_0^{\tau_2 - t_n} f_{W(\tau_2)-W(s)}(r-a) f_T(s) ds}{f_{W(\tau_2)}(r)} \right) f_{W(\tau_2)}(r) dr,
\end{aligned}
\tag{4.3.3}
$$

where $f_{W(\tau_2)}(r)$ is the normal pdf as before, $f_{W(\tau_2)-W(s)}(r-a)$ is the normal pdf, $W(\tau_2) - W(s) \sim \mathcal{N}((\tau_2 - s)\nu_1, (\tau_2 - s)\sigma_1^2)$, and $f_T(s)$ is the shifted inverse Gaussian density $T - t_n \sim \mathcal{IG}(\nu_1, \sigma_1, a - W(t_n))$, and $F(\cdot)$ is the inverse Gaussian CDF. Note that in the second term, where $r < a$, $P(T \le t | W(\tau_2) = r)$ is probably small since the process has a positive drift, and the expression can be simplified to neglect $P(T \le t | W(\tau_2) = r)$.

Case 3: $\tau_1 \le t_n < t' < \tau_2$

In this case, only the first change point has occurred, and since the second change point does not occur prior to the time in question, t', the hitting time distribution of $S = T - t_n$ is given as $S \sim \mathcal{IG}(\nu_1, \sigma_1, a - W(t_n))$, thus we have

$$
P(T \le t') = P(S \le t' - t_n) = F(t' - t_n; \nu_1, \sigma_1, a - W(t_n)). \tag{4.3.4}
$$

Case 4 : $t_n < \tau_1 < \tau_2 \le t'$

This case is similar to case 2, except here we need to take into account that also τ_1 is in the future. Thus $W(\tau_1)$ is unknown, and the shifted hitting time is $S = T - \tau_1$. Note that

$$
W(\tau_1) - W(t_n) \sim \mathcal{N}\left(0, (\tau_1 - t_n)\sigma_0^2\right), \quad W(\tau_2) - W(\tau_1) \sim \mathcal{N}\left((\tau_2 - \tau_1)\nu_1, (\tau_2 - \tau_1)\sigma_1^2\right).
$$

Compared to equation (4.3.3), we need to alter the expression for $P(T \le \tau_2 | W(\tau_2) = r)$, which enters in both terms of the equation. Now we must integrate the shifted hitting time s over 0 to $\tau_2 - \tau_1$, and $f_T(s)$ is inverse Gaussian distribution with drift parameter ν_1 and variance parameter σ_1 as before, but the threshold parameter is $a - W(\tau_1)$. Thus

we integrate over $f_{W(\tau_1)}$:

$$
\begin{aligned}
P(T \le t') &= \int_{-\infty}^{\infty} P(T \le t' | W(\tau_2) = r) f_{W(\tau_2)}(r) dr \\
&= \int_{-\infty}^{\infty} \left(P(T \le \tau_2 | W(\tau_2) = r) + P(\tau_2 < T \le t' | W(\tau_2) = r) \right) f_{W(\tau_2)}(r) dr \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{0}^{\tau_2 - \tau_1} f_{W(\tau_2)-W(s)}(r-a) f_T(s) f_{W(\tau_1)}(q) dq ds dr + \\
&\quad \int_{-\infty}^{a} \int_{-\infty}^{\infty} F(t - \tau_2; \nu_2, \sigma_2, a - r) \Bigg(1 - \\
&\frac{\int_{-\infty}^{\infty} \int_{0}^{\tau_2 - \tau_1} f_{W(\tau_2)-W(s)}(r-a) f_T(s) f_{W(\tau_1)}(q) dq ds}{f_{W(\tau_2)}(r)} \Bigg) f_{W(\tau_2)}(r) f_{W(\tau_1)}(q) dq dr, \qquad (4.3.5)
\end{aligned}
$$

where $f_{W(\tau_2)}(r)$ is the normal pdf, $W(\tau_2) \sim \mathcal{N}\left(W(\tau_1) + (\tau_2 - \tau_1)\nu_1, (\tau_2 - \tau_1)\sigma_1^2\right)$, $f_{W(\tau_2)-W(s)}(r-a)$ is the normal pdf, $W(\tau_2) - W(s) \sim \mathcal{N}((\tau_2 - s)\nu_1, (\tau_2 - s)\sigma_1^2)$, and $f_T(s)$ is the shifted inverse Gaussian density $T - \tau_1 \sim \mathcal{IG}(\nu_1, \sigma_1, a - W(\tau_1))$. Again, in the second term, since $r < a$, $P(\tau_2 \le T | W(\tau_2) = r)$ is probably small and can be neglected. As before, a simplification of the integration over $W(\tau_1)$ is to set $W(\tau_1) = E[W(\tau_1)|W(t_n)] = W(t_n)$ instead of integrating over possible values of $W(\tau_1)$.

Case 5: $t_n < \tau_1 \le t' < \tau_2$

In this case we need only consider the process $W_1(t)$, which has parameters ν_1 and σ_1, but the time interval of interest is $(\tau_1, t']$. The corresponding hitting time, $S_1 = T - \tau_1$, is inverse Gaussian distributed $S_1 \sim \mathcal{IG}(\nu_1, \sigma_1, a - W(\tau_1))$. The increment $W(\tau_1) - W(t_n)$ is normally distributed,

$$W(\tau_1) - W(t_n) \sim \mathcal{N}\left(0, (\tau_1 - t_n)\sigma_0^2\right)$$

Given this increment, the distribution of the hitting time is given as

$$P(T \le t' | W(\tau_1)) = P(S_1 \le t' - \tau_1) = F(t' - \tau_1; \nu_1, \sigma_1, a - W(\tau_1)).$$

Thus, we can condition on the increment and integrate out :

$$P(T \le t') = \int_{-\infty}^{\infty} F(t' - \tau_1; \nu_1, \sigma_1, a - W(\tau_1)) f_{W(\tau_1)-W(t_n)}(r) dr. \qquad (4.3.6)$$

As before, a simplification of this is to set $W(\tau_1) = E[W(\tau_1)|W(t_n)] = W(t_n)$ instead of integrating over possible values of $W(\tau_1)$.

This concludes all the five cases in figure 4.3.1, thus, to find the hitting time CDF, $P(T \le t)$ when the change points τ_1 and τ_2 are unknown, we can apply MC summation to draw m pairs of realizations of τ_1, $\tau_{1,k}^*$, and τ_2, $\tau_{2,k}^*$ from $\pi(\tau_1, \tau_2 | x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2)$, to find the approximation $\widehat{P}_n(T \le t)$ to $P(T \le t)$:

$$\widehat{P}_n(T \leq t' | x_1, \ldots, x_n; ; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a) = \frac{1}{m} \sum_{k=1}^{m} P(T \leq t' | \tau_{1,k}^*, \tau_{2,k}^*, x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a)$$

where $P(T \leq t' | \tau_{1,k}^*, \tau_{2,k}^*, x_1, \ldots, x_n; a, \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2)$ is found by consulting the expression for the case which τ_1, τ_2 belong to. The approach is summarized in algorithm 4.

Algorithm 4 Formulaic approach for estimating the hitting time CDF of $t = t'$ for piecewise Wiener process with two unknown change points τ_1 and τ_2.

Require: $m, \{W(t), 0 \leq t \leq t_n\}, \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a > W(t_n), t'$
  $\Delta = t_n / n$
  $x_i = W((i+1) \cdot \Delta) - W(i \cdot \Delta)$
  Find posterior distribution of τ: $\pi(\tau_1, \tau_2 | x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a)$
  for $k = 1$ to m **do**
    Sample $\tau_{1,k}^*, \tau_{2,k}^*$ from $\pi(\tau_1, \tau_2 | x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a)$
    if $\tau_1, \tau_2 \in$ case 1 **then**
      Calculate $P(T \leq t' | \tau_{1,k}^*, \tau_{2,k}^*, x_1, \ldots, x_n; a, \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2)$ from equation (4.3.2)
    else if $\tau_1, \tau_2 \in$ case 2 **then**
      Calculate $P(T \leq t' | \tau_{1,k}^*, \tau_{2,k}^*, x_1, \ldots, x_n; a, \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2)$ from equation (4.3.3)
    else if $\tau_1, \tau_2 \in$ case 3 **then**
      Calculate $P(T \leq t' | \tau_{1,k}^*, \tau_{2,k}^*, x_1, \ldots, x_n; a, \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2)$ from equation (4.3.4)
    else if $\tau_1, \tau_2 \in$ case 4 **then**
      Calculate $P(T \leq t' | \tau_{1,k}^*, \tau_{2,k}^*, x_1, \ldots, x_n; a, \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2)$ from equation (4.3.5)
    else if $\tau_1, \tau_2 \in$ case 5 **then**
      Calculate $P(T \leq t' | \tau_{1,k}^*, \tau_{2,k}^*, x_1, \ldots, x_n; a, \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2)$ from equation (4.3.6)
    end if
  end for
  return

$$\widehat{P}_n(T \leq t' | x_1, \ldots, x_n; ; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a) = \frac{1}{m} \sum_{k=1}^{m} P(T \leq t' | \tau_{1,k}^*, \tau_{2,k}^*, x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a)$$

4.3.2 Formulaic approach with time transformation

The time transformation approach was studied by Doksum and Høyland (1992) in relation to accelerated life testing. In their setting, the process started with a non-zero drift, hence, the starting point corresponds to our τ_1. At a known time τ_2 the drift was intensified. Thus, their setting is a bit different. As the change points are not known to us, we condition on the change points to find the hitting time CDF given the change points, and marginalize to find the hitting time CDF, just as before.

In accordance with Doksum and Høyland (1992), consider the process as follows: The initial process, $W_0(t)$ starting at $t = 0$ has mean $\nu = 0$ and variance parameter σ_0^2. At the first change point $t = \tau_1$ the process $W_1(t)$ begins, with starting temperature $W_0(\tau_1)$, drift parameter ν_1 and variance parameter σ_1^2. At the second change point $t = \tau_2$ the process $W_2(t)$ begins with starting temperature $W_1(\tau_2)$, drift parameter ν_2 and variance parameter σ_2^2. Furthermore, assume that $W_2(t) = W_1(\tau_2 + \alpha(t - \tau_2))$. This puts the following constraints on ν_2 and σ_2^2:

$$\nu_2 = \alpha\nu_1, \qquad \sigma_2^2 = \alpha^2\sigma_1^2. \tag{4.3.7}$$

Thus, we employ the simplifying notation $\nu = \nu_1, \sigma = \sigma_1$ and will write $\alpha\nu$ for ν_2 and $\alpha\sigma$ for σ_2. The following time transformation $\xi(t)$ is proposed by Doksum and Høyland (1992):

$$\xi(t) = \begin{cases} t, & \text{if } \tau_1 < t \leq \tau_2 \\ \tau_2 + \alpha(t - \tau_2), & \text{if } t > \tau_2. \end{cases} \tag{4.3.8}$$

It allows us to express the piecewise Wiener process as $W(\xi(t))$ for $t > \tau_1$, and, more conveniently, the CDF of the hitting time is inverse Gaussian of the time transformation:

$$P(T \leq t|\tau_1, \tau_2; \alpha, \nu, \sigma, a) = F(\xi(t)|\tau_1, \tau_2; \alpha, \nu, \sigma, a), \tag{4.3.9}$$

for $\tau_1 < t$, where $F(\cdot)$ denotes the inverse Gaussian CDF as before, with drift parameter ν, variance parameter σ and threshold parameter a. The parameter α applies only to the transformation $\xi(t)$.

As noted in section 4.3 the posterior distribution of τ_1, τ_2 given the temperature increment measurements $x_1, \ldots, x_n$ is

$$\pi(\tau_1, \tau_2|x_1, \ldots, x_n) \propto L(\tau_1, \tau_2|x_1, \ldots, x_n)\pi(\tau_1, \tau_2) \propto$$
$$\begin{cases} \frac{\lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}}{\sigma_0^{\tau_1-1}\sigma_1^{\tau_2-\tau_1}\sigma_2^{n-\tau_2+1}} \times \\ \quad \exp\left(-\sum_{i=1}^{\tau_1/\Delta-1}\frac{x_i^2}{2\Delta\sigma_0^2} - \sum_{i=\tau_1/\Delta}^{\tau_2/\Delta-1}\frac{(x_i-\Delta\nu_1)^2}{2\Delta\sigma_1^2} - \sum_{i=\tau_2/\Delta}^{n}\frac{(x_i-\Delta\nu_2)^2}{2\Delta\sigma_2^2}\right), & \tau_1 < \tau_2 \leq t_n \\ \frac{\lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}}{\sigma_0^{\tau_1-1}\sigma_1^{n-\tau_1+1}} \times \\ \quad \exp\left(-\sum_{i=1}^{\tau_1/\Delta-1}\frac{x_i^2}{2\Delta\sigma_0^2} - \sum_{i=\tau_1/\Delta}^{n}\frac{(x_i-\Delta\nu_1)^2}{2\Delta\sigma_1^2}\right), & \tau_1 \leq t_n < \tau_2. \\ \frac{\lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}}{\sigma_0^n}\exp\left(-\frac{1}{2\Delta\sigma_0^2}\sum_{i=1}^{n}x_i^2\right), & t_n < \tau_1 < \tau_2. \end{cases}$$

In this setting, with the restrictions given in (4.3.7), we obtain

$$\pi(\tau_1,\tau_2|x_1,\ldots,x_n) \propto L(x_1,\ldots,x_n|\tau_1,\tau_2)\pi(\tau_1,\tau_2) \propto$$
$$\begin{cases} \frac{\lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}}{\sigma_0^{\tau_1-1}\sigma_1^{n-\tau_1-1}\alpha^{n-\tau_2}}\times \\ \quad \exp\left(-\sum_{i=1}^{\tau_1/\Delta-1}\frac{x_i^2}{2\Delta\sigma_0^2}-\sum_{i=\tau_1/\Delta}^{\tau_2/\Delta-1}\frac{(x_i-\Delta\nu)^2}{2\Delta\sigma_1^2}-\sum_{i=\tau_2/\Delta}^{n}\frac{(x_i-\alpha\Delta\nu)^2}{2\alpha^2\Delta\sigma_1^2}\right), & \tau_1<\tau_2\le t_n \\ \frac{\lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}}{\sigma_0^{\tau_1-1}\sigma_1^{n-\tau_1-1}}\times \\ \quad \exp\left(-\sum_{i=1}^{\tau_1-1}\frac{x_i^2}{2\Delta\sigma_0^2}-\sum_{i=\tau_1}^{n}\frac{(x_i-\Delta\nu)^2}{2\Delta\sigma_1^2}\right), & \tau_1\le t_n<\tau_2. \\ \frac{\lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}}{\sigma_0^{n}}\exp\left(-\frac{1}{2\Delta\sigma_0^2}\sum_{i=1}^{n}x_i^2\right), & t_n<\tau_1<\tau_2. \end{cases}$$

If $\sigma_0=\sigma_1=\sigma$, the posterior simplifies to

$$\pi(\tau_1,\tau_2|x_1,\ldots,x_n) \propto L(x_1,\ldots,x_n|\tau_1,\tau_2)\pi(\tau_1,\tau_2) \propto$$
$$\begin{cases} \frac{\lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}}{\alpha^{n-\tau_2}}\times \\ \quad \exp\left(-\frac{1}{2\Delta\sigma^2}\left[\sum_{i=1}^{\tau_1/\Delta-1}x_i^2+\sum_{i=\tau_1/\Delta}^{\tau_2/\Delta-1}(x_i-\Delta\nu)^2+\sum_{i=\tau_2/\Delta}^{n}\frac{(x_i-\alpha\Delta\nu)^2}{\alpha^2}\right]\right), & \tau_1<\tau_2\le t_n \\ \lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}\times \\ \quad \exp\left(-\frac{1}{2\Delta\sigma^2}\left[\sum_{i=1}^{\tau_1/\Delta-1}x_i^2+\sum_{i=\tau_1/\Delta}^{n}(x_i-\Delta\nu)^2\right]\right), & \tau_1\le t_n<\tau_2. \\ \lambda_1(1-\lambda_1)^{\tau_1-1}\lambda_2(1-\lambda_2)^{\tau_2-\tau_1-1}\times\exp\left(-\frac{1}{2\Delta\sigma^2}\sum_{i=1}^{n}x_i^2\right), & t_n<\tau_1<\tau_2. \end{cases}$$

As for one change point, the hitting time CDF is found by MC summation over the posterior distribution of τ_1 and τ_2, $\pi(\tau_1,\tau_2|x_1,\ldots,x_n;\alpha,\nu,\sigma)$, by drawing m pairs of realizations of τ_1, $\tau_{1,k}^*$, and τ_2, $\tau_{2,k}^*$ from $\pi(\tau_1,\tau_2|x_1,\ldots,x_n;\alpha,\nu,\sigma)$. Denote the formulaic time transformed estimate of the hitting time CDF $\widehat{P}_f(T\le t|x_1,\ldots,x_n;\alpha,\nu,\sigma,a)$, and note that as in section 4.3.1, the five different regions in figure 4.3.1 lead to different expressions for $\xi(t)$ in equation (4.3.8) for the shifted hitting time:

$$\xi(t)=\begin{cases} \alpha(t-t_n), & \text{if } \tau_1<\tau_2\le t_n, \text{ Case 1} \\ \tau_2-t_n+\alpha(t-\tau_2), & \text{if } \tau_1<t_n<\tau_2\le t, \text{ Case 2} \\ t-t_n, & \text{if } \tau_1\le t_n<t<\tau_2, \text{ Case 3} \\ \tau_2-\tau_1+\alpha(t-\tau_2) & \text{if } t_n<\tau_1<\tau_2\le t, \text{ Case 4} \\ t-\tau_1 & \text{if } t_n<\tau_1\le t<\tau_2, \text{ Case 5.} \end{cases} \tag{4.3.10}$$

Thus formulaic time transformed estimate of the hitting time CDF is given as

$$\begin{aligned} \widehat{P}_f(T\le t|x_1,\ldots,x_n;\alpha,\nu,\sigma,a) &= \frac{1}{m}\sum_{k=1}^{m}P(T\le t|\tau_{1,k}^*,\tau_{2,k}^*;\alpha,a,\nu,\sigma) \\ &= \frac{1}{m}\sum_{k=1}^{m}F(\xi(t)|\tau_{1,k}^*,\tau_{2,k}^*;\alpha,\nu,\sigma,a), \end{aligned} \tag{4.3.11}$$

where $\xi(t)$ is given in equation (4.3.10). Note that in case 1 and case 3, $\xi(t)$ is independent of τ_1 and τ_2, such that $F(\xi(t)|\tau_1,\tau_2;a,\nu,\sigma)$ can be taken out of the summation for computational savings. Algorithm 5 summarizes the approach.

Algorithm 5 Formulaic time transformation approach for estimating the hitting time CDF of $t = t'$ for piecewise Wiener process with two unknown change points τ_1 and τ_2.

Require: $m, \{W(t_i), i = 1, \ldots, n\}, \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a > W(t_n), t'$

$\Delta = t_n/n$

$x_i = W((i+1) \cdot \Delta) - W(i \cdot \Delta)$

Find posterior distribution of τ: $\pi(\tau_1, \tau_2 | x_1, \ldots, x_n; \alpha, \nu, \sigma)$

for $k = 1$ to m **do**

Sample $\tau^*_{1,k}, \tau^*_{2,k}$ from $\pi(\tau_1, \tau_2 | x_1, \ldots, x_n; \alpha, \nu, \sigma)$

Calculate $\xi_k(t')$ according to equation (4.3.10)

$$P(T \leq t' | \tau^*_{1,k}, \tau^*_{2,k}, x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a) = F(\xi_k(t'); \nu_1, \sigma_1, a),$$

where $F(\cdot\,; \nu_1, \sigma_1, a)$ denotes the inverse Gaussian CDF with drift parameter ν_1 , variance parameter σ_1 and threshold parameter a.

end for

return

$$\widehat{P}_f(T \leq t' | x_1, \ldots, x_n; \nu, \sigma, a) = \frac{1}{m} \sum_{k=1}^{m} F(\xi_k(t'); \nu_1, \sigma_1, a)$$

4.3.3 Simulation approach

The simulation approach is conceptually the same for two change points as it was for one. As for one change point, the simulation approach finds the hitting time CDF conditioned on the change points, $P(T \leq t | \tau_1, \tau_2, x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a)$, by simulating a number, say K, of Wiener processes from the last observation, at time t_n and temperature $W(t_n)$, until the processes cross the threshold temperature a, that is, until time T_i and temperature a. Note that conditioned on τ_1 and τ_2, the distribution of X_i is now:

$$X_j | \tau_1, \tau_2 \sim \begin{cases} \mathcal{N}(0, \Delta\sigma_0^2) & \text{if } j\Delta < \tau_1 \\ \mathcal{N}(\Delta\nu_1, \Delta\sigma_1^2) & \text{if } \tau_1 \leq j\Delta < \tau_2 \\ \mathcal{N}(\Delta\nu_2, \Delta\sigma_2^2) & \text{if } \tau_2 \leq j\Delta \end{cases}$$

Thus, the above gives the rule for how the temperature increments are sampled.

As for one change point, $P(T \leq t | \tau, x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a)$ can be estimated as the empirical CDF, $\widehat{F}_{K, \tau^*_{1,j}, \tau^*_{2,j}}(t)$,

$$\widehat{F}_{K, \tau^*_{1,j}, \tau^*_{2,j}}(t) = \frac{1}{K} \sum_{i=1}^{K} I_{T_i \leq t}.$$

The simulation approach does not require the assumption that $T > \tau$, neither the assumption of Doksum and Høyland (1992) that $W_2(t) = W_1(\tau_2 + \alpha(t - \tau_2))$ nor the constraints on ν_2 and σ_2 that followed.

Finally, the marginalized distribution of the hitting time T given the temperature increments $x_1, \ldots, x_n$, $P(T \leq t|x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a)$ is found as for the formulaic approach, by drawing a large number N of change point realizations $\tau^*_{1,i}, \tau^*_{2,i}$ from the posterior distribution, and let the mean denote the simulation estimator $\widehat{P}_S(T \leq t|x_1, \ldots, x_n;; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a)$ for the probability $P(T \leq t|x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a)$:

$$\widehat{P}_S(T \leq t|x_1, \ldots, x_n;; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a) = \frac{1}{N} \sum_{j=1}^{N} \widehat{F}_{K,\tau^*_{1,j},\tau^*_{2,j}}(t). \tag{4.3.12}$$

The simulation is described in algorithm 6. The description is schematic, and implementation can be done much more efficient.

4.3.4 τ_1, τ_2 and ν_1 unknown

Consider now the situation of Doksum and Høyland (1992) where the parameters can be described as in 4.3.7, and further consider the case where ν_1 is unknown, but α and σ_1 is known. Note that this implies that ν_2 is unknown, since $\nu_2 = \alpha\nu_1$. However, σ_2 is known since $\sigma_2 = \alpha\sigma_1$. Thus, we will estimate only ν_1, but our estimates will take into account both ν_1 and $\nu_2 = \alpha\nu_1$.

We consider $\nu_1 > 0$ in the failure development setting, and thus a gamma prior on ν_1 as in section 4.2 is intuitive. We assume that ν_1 is independent of τ_1 and τ_2, and hence let the joint prior be

$$\pi(\tau_1, \tau_2, \nu_1) = \pi(\tau_1, \tau_2)\pi(\nu_1)$$

where $\pi(\tau_1, \tau_2)$ is the joint prior of τ_1 and τ_2, given in equation (4.3.1). Thus, when using the Metropolis-Hastings algorithm with the prior distribution as the proposal, the Metropolis-Hastings ratio stays the likelihood ratio, but now a function of τ_1, τ_2 and ν_1. The extension to ν_1 unknown in the formulaic and simulation approach is analogous to the extension to ν unknown for the process with one change point.

Algorithm 6 Simulation to find hitting time distribution for piecewise Wiener process with one unknown change point τ.

Require: $m, \{W(t_i), i = 1, \ldots, n\}, \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a > W(t_n), K, t$
  $\Delta = t_n/n$
  $x_i = W(i \cdot \Delta) - W((i-1) \cdot \Delta)$
  Find posterior distribution of τ: $\pi(\tau | x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2)$
  $t = t_n$
  for $j = 1$ to J **do**
    Sample $\tau^*_{1,j}, \tau^*_{2,j}$ from $\pi(\tau_1, \tau_2 | x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2)$
    for $k = 1$ to K **do**
      $t = t + \Delta$
      while $W(t - \Delta) \leq a$ **do**
        if $\tau^*_{2,j} \leq t$ **then**
          Draw increment
          $X(t) \sim \mathcal{N}(\nu_2 \cdot \Delta, \sigma_2^2 \cdot \Delta)$
        else if $\tau^*_{1,j} \leq t$ **then**
          Draw increment
          $X(t) \sim \mathcal{N}(\nu_1 \cdot \Delta, \sigma_1^2 \cdot \Delta)$
        else
          Draw increment
          $X(t) \sim \mathcal{N}(0, \sigma_0^2 \cdot \Delta)$
        end if
        $W(t) = W(t - \Delta) + X(t)$
        $t = t + \Delta$
      end while
      $T_k = t$ is the k-th hitting time given $\tau^*_{1,j}, \tau^*_{2,j}$
    end for
    $\widehat{F}_{K,\tau^*_{1,j},\tau^*_{2,j}}(t) = \frac{1}{K} \sum_{k=1}^{K} I_{T_k} \leq t$ is the estimate of
    $P(T \leq t | \tau^*_{1,j}, \tau^*_{2,j}, x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a)$.
  end for
  Estimate $P(T \leq t | \tau^*_{1,j}, \tau^*_{2,j}, x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a)$ by

$$\widehat{P}_S(T \leq t | x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a) = \frac{1}{J} \sum_{i=1}^{J} \widehat{F}_{K,\tau^*_{1,j},\tau^*_{2,j}}(t).$$

  return $\widehat{P}_S(T \leq t | x_1, \ldots, x_n; \nu_1, \nu_2, \sigma_0, \sigma_1, \sigma_2, a)$.

4.4 m change points

Consider now the setting with m change points and corresponding parameters $0 = \nu_0$ and σ_0, for the initial process, as before, and $\nu_1, \ldots, \nu_m$ and $\sigma_1, \ldots, \sigma_m$. Let $\vec{\nu} = [\nu_0, \nu_1, \ldots, \nu_m]$ and $\vec{\sigma} = [\sigma_0, \sigma_1, \ldots, \sigma_m]$. Let $\vec{\nu}, \vec{\sigma}$ be known. We make the remark that, as before, $\tau_1 < \tau_2 < \cdots < \tau_m$. The likelihood becomes

$$L(\tau_1, \ldots, \tau_m | x_1, \ldots, x_n; \vec{\nu}, \vec{\sigma}) = \begin{cases} L_m(\tau_1, \ldots, \tau_m | x_1, \ldots, x_n; \vec{\nu}, \vec{\sigma}) & \text{if } \tau_m \leq t_n, \\ L_{m-1}(\tau_1, \ldots, \tau_m | x_1, \ldots, x_n; \vec{\nu}, \vec{\sigma}) & \text{if } \tau_{m-1} \leq t_n < \tau_m, \\ \quad \vdots & \\ L_1(\tau_1, \ldots, \tau_m | x_1, \ldots, x_n; \vec{\nu}, \vec{\sigma}) & \text{if } \tau_1 \leq t_n < \tau_2, \\ L_0(\tau_1, \ldots, \tau_m | x_1, \ldots, x_n; \vec{\nu}, \vec{\sigma}) & \text{if } t_n < \tau_1. \end{cases} \tag{4.4.1}$$

where $L_m(\tau_1, \ldots, \tau_m | x_1, \ldots, x_n; \vec{\nu}, \vec{\sigma})$ for $m = 1, 2, \ldots$, is the likelihood if $\tau_m \leq n$:

$$\begin{aligned} L_m(\tau_1, \ldots, \tau_m | x_1, \ldots, x_n; \vec{\nu}, \vec{\sigma}) &= \prod_{i=1}^{n} f(x_i | \tau_1, \ldots, \tau_n; \vec{\nu}, \vec{\sigma}) \\ &= \prod_{i=1}^{\tau_1/\Delta - 1} \frac{1}{\sqrt{2\pi}\Delta\sigma_0} \exp\left\{ -\frac{x_i^2}{2\Delta\sigma_0^2} \right\} \times \prod_{i=\tau_1/\Delta}^{\tau_2/\Delta - 1} \frac{1}{\sqrt{2\pi}\Delta\sigma_1} \exp\left\{ -\frac{(x_i - \Delta\nu_1)^2}{2\Delta\sigma_1^2} \right\} \times \ldots \\ &\quad \times \prod_{i=\tau_m/\Delta}^{n} \frac{1}{\sqrt{2\pi}\Delta\sigma_m} \exp\left\{ -\frac{(x_i - \Delta\nu_m)^2}{2\Delta\sigma_m^2} \right\} \\ &\propto \frac{1}{\sigma_0^{\tau_1}} \times \frac{1}{\sigma_1^{\tau_2 - \tau_1}} \times \cdots \times \frac{1}{\sigma_m^{n - \tau_m + 1}} \times \\ &\quad \exp\left\{ -\frac{1}{2\Delta} \left[\sum_{i=1}^{\tau_1/\Delta - 1} \frac{x_i^2}{\sigma_0^2} + \sum_{i=\tau_1/\Delta}^{\tau_2/\Delta - 1} \frac{(x_i - \Delta\nu_1)^2}{\sigma_1^2} + \cdots + \sum_{i=\tau_m/\Delta}^{n} \frac{(x_i - \Delta\nu_m)^2}{\sigma_m^2} \right] \right\} \end{aligned}$$

and $L_0(\tau_1, \ldots, \tau_m | x_1, \ldots, x_n; \vec{\nu}, \vec{\sigma}) \propto \frac{1}{\sigma_0^n} \exp\left\{ -\frac{1}{2\Delta} \left[\sum_{i=1}^{n} \frac{x_i^2}{\sigma_0^2} \right] \right\}$. When $\vec{\sigma}$ is known, the products of fractions is obsolete and need not be included in the likelihood.

The joint prior distribution of the change points $\tau_1, \ldots, \tau_m$ can be found as before by realizing that the time spent in each state is independent and geometrically distributed, as noted in section 4.3. By defining $\tau_0 = 0$, we can express the joint prior probability as

$$\pi(\tau_1, \ldots, \tau_m) = \prod_{i=1}^{m} \lambda_i (1 - \lambda_i)^{\tau_i - \tau_{i-1} - 1}$$

for $0 < \tau_1 < \tau_2 < \cdots < \tau_m$ with all τ_i integer valued, and 0 otherwise.

The posterior is as before

$$\pi(\tau_1, \ldots, \tau_m | x_1, \ldots, x_n; \vec{\nu}, \vec{\sigma}) \propto \pi(\tau_1, \ldots, \tau_m) \times L(\tau_1, \ldots, \tau_m | x_1, \ldots, x_n; \vec{\nu}, \vec{\sigma}).$$

To find the hitting time CDF, we suggest using either the simulation approach or the formulaic time transformed approach, which are both easily generalized to m change points, as the formulae for the original time scaled problems become tedious to work out.

The simulation approach is naturally extended. A set of $\tau_1, \ldots, \tau_m$ realizations are drawn from the posterior distribution, and the processes are simulated accordingly. The difference from 6 is that more cases on the size of τ_i must be checked in order to draw increments from the correct distribution. The CDF is found by the empirical estimate $\widehat{F}_l(t)$ as before, where l denotes the number of simulated Wiener processes.

The transformed time approach is outlined in Doksum and Høyland (1992). The restrictions on ν_i and σ_i are extended:

$$\nu_i = \alpha\nu_{i-1} \qquad\qquad \sigma_i^2 = \alpha^2\sigma_{i-1}^2$$

for $i = 2, 3, \ldots, m$.

Define the model as for two change points, with the Wiener process $W_0(t), t \in [0, \tau_1)$ with drift parameter $\nu_0 = 0$ and variance parameter σ_0. At τ_1, the process changes to the Wiener process $W_1(t), t \in [\tau_1, \tau_2)$, with parameters ν_1, σ_1. In this case, the process changes again at the m change points until $W_m(t), t \in [\tau_m, \infty)$. Thus, for $t > \tau_1$, we can define the process

$$W_i(t) = W_{i-1}(\tau_i + \alpha_i[t - \tau_i]), \quad t \in [\tau_i, \tau_{i+1}),$$

where $\tau_{k+1} = \infty$. Thus, the drift is multiplied with the factor α_i each time t crosses the change point τ_i. Again, the process can be re-expressed in terms of the process $W_1(t)$ by introducing β_i:

$$\beta_i = \prod_{r=0}^{i} \alpha_i,$$

by defining $\alpha_0 = 1$. Furthermore, define the multiplicative factor $\beta(t)$:

$$\beta(t) = \beta_i, \;\; t \in [\tau_i, \tau_{i+1}), i = 1, \ldots, k.$$

The transformation can be defined as

$$\xi(t) = \begin{cases} t, & \text{if } \tau_1 < t \leq \tau_2 \\ \sum_{r=0}^{i-1} \beta_r(\tau_{r+1} - \tau_r) + \beta_i(t - \tau_i) & \text{if } t \in [\tau_i, \tau_{i+1}), i = 1, \ldots, k. \end{cases} \tag{4.4.2}$$

Thus, we can write $W(t) = W_1(\xi(t)), t > \tau_1$. According to Doksum and Høyland (1992), the CDF is

$$P(T \leq t) = F(\xi(t)), t > \tau_1$$

and 0 otherwise, where $F(\cdot)$ denotes the inverse Gaussian CDF.

Chapter 5

Application to the motivating case study

5.1 One change point

In this section, we will have one particular temperature Wiener process as our starting point. The temperature process $W(t)$ is shown in figure 5.1.1. The process was simulated with the following parameters:

$$\nu = 0.003 \qquad \sigma = 0.02 \qquad \tau = 1500$$

in the time interval $[0, t_n] = [0, 3000]$. The resolution was $t_n/n = \Delta = 1$. For this process we will consider the hitting time of the threshold temperature value $a = [W(t)]+10 = 14$, where $[\cdot]$ denotes rounding to the nearest integer.

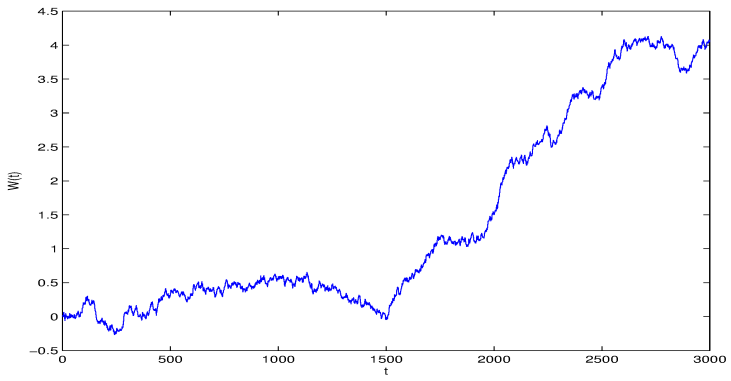

Figure 5.1.1: Piecewise Wiener process simulated in the time interval $[0, t_n] = [0, 3000]$. with one change point $\tau = 1500$. The drift parameter before the change point is $\nu_0 = 0$ and after the change point $\nu = 0.003$. The variance parameter is the same for both pieces, $\sigma = 0.02$. For this process, we will consider the threshold $a = 14$.

In this section, numerical examples will be given for three different cases, first τ unknown and ν and σ known, secondly both τ and ν unknown and σ known, and finally all three parameters τ, ν and σ unknown. In each case, we will also consider different observation lengths t_n of the temperature process, that is, we will consider the observations $\{W(t_i), i = 1, \ldots, n\}$ for different n. Note that the corresponding t_n is the endpoint in time of observation. This is to mimic the situation of online monitoring.

For each of the three cases, the different observations length are $t_n = 1\ 400$, $t_n = 1\ 700$ and $t_n = 3\ 000$. Note that in with the first observation length $t_n = 1\ 400$, the change point has not occurred. The second observation length, $t_n = 1\ 700$, is shortly after the change point. The third change point, $t_n = 3\ 000$ is long after the change point, and should give good information.

In the entire chapter, the approach used to deal with $W(\tau)$ for $\tau > t_n$ was to set $W(\tau) = W(t_n)$. This was done after some comparison of calculations with the MC summation solution, and the approaches were found to give comparable results.

5.1.1 τ unknown, ν and σ known

Firstly in this section, we will consider the MCMC diagnostics of the samples from the posterior distribution $\pi(\tau|x_1, \ldots, x_n)$ for the different t_n. Finally, we will examine the hitting time distributions which we obtain for the different posterior distribution by the formulaic approach, $\widehat{P}_f(T \leq t)$ and the simulation approach, $P_S(T \leq t)$.

In all cases, a geometric prior was put on τ with parameter $\lambda = 1/3000$.

Observation until $t_n = 1\ 400$

Figure 5.1.2 shows the prior distribution of τ together with the posterior distribution, estimated by kernel smoothing density methods, as estimated by available software, based on the sample of τ by MCMC methods. Further diagnostic figures are shown in appendix A.1.1: A trace plot is given in figure A.1.1, a histogram of all values after burn-in is shown in figure A.1.2 and finally, the autocorrelation function, ACF, of the τ iterates is shown in figure A.1.3. The trace plot seems to show that the Markov chain is covering the entire domain of the posterior distribution $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, and the chain is moving fast. The autocorrelation is rapidly decreasing. The acceptance rate was 0.632. Together with the trace plot and the ACF, this points to good mixing. The Rubin-Gelman statistic was calculated with five chains of length $2 \cdot 10^6$ and a burn-in of 10^4 to $\widehat{R}^1 = 1$. Thus, it is assumed that the chain has converged to the posterior distribution.

However, looking at figure 5.1.2, it seems like the prior distribution is still influential. This is natural when we know that the change point has not happened yet. In fact, looking at the posterior distribution, we see that it is very flat until t_n, unlike the prior distribution, $\pi(\tau)$. After the last observation point, we can only rely on the prior distribution for information about τ. This gives rise to the long 95 % credible interval estimates, shown in table 5.1 along with point estimates of the MCMC sampled population of τ from the posterior distribution. The interval estimates are the 2.5 % and 97.5

% percentiles of the posterior distribution along with the 95 % highest posterior density region, HPD region, as estimated by available software.

Observation until t_n = 1 700

In this case, the change point has occurred, and we have collected some data following the parameters ν and σ. Figure 5.1.3 shows the prior distribution of τ along with the posterior distribution estimated by kernel smoothing density methods based on the sample of τ obtained by MCMC methods. The trace plot is shown in figure A.1.4, the ACF is shown in figure A.1.6 and the histogram is shown in figure A.1.5, all in appendix A.1.1. The acceptance rate was 0.428. The lower acceptance rate than at $t_n = 1\ 400$ manifests in all three plots, by the sharper histogram, the less vigorously moving trace plot and the slower decline in the ACF as function of the lag. However, the chain is still moving freely and fast around the range of the posterior distribution, and does not show signs of getting stuck or moving in a specific pattern. The Rubin-Gelman statistic was calculated with five chains of length $2 \cdot 10^6$ and a burn-in of 10^4 to $\widehat{R}^1 = 1$. Thus, it is assumed that the chain has converged to the posterior distribution.

Figure 5.1.3 shows that there is a steep peak in the posterior distribution of τ, $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, around $\tau = 1500$. Table 5.2 shows the statistics of the sample. Even though the mean of τ has decreased significantly compared to $t_n = 1\ 400$, the difference between the mean and the median is large. This reflects the heavy tail of $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, which can be seen in figure 5.1.3. One can note that the interval estimates have narrowed compared to $t_n = 1\ 400$, but are still wide.

Observation until t_n = 3 000

In this case, we have an abundance of measurements after the change point, which allows us to estimate the change point very well. Figure 5.1.4 shows the prior and posterior distributions of τ, given the observations up until $t_n = 3\ 000$. Notice that $\pi(\tau|x_1, \ldots, x_n)$ is so steep that $\pi(\tau)$ is barely visible. The peakedness of the posterior distribution is reflected in the diagnostics plots in appendix A.1.1 as well, the trace plot in figure A.1.7 shows that fewer proposals are accepted, in fact the acceptance rate is 0.038. The ACF in figure A.1.9 is decaying at an even slower rate than at $t_n = 1\ 700$, which is a result of the low acceptance rate. However, looking at longer stretches of samples, the chain is not moving sequentially or otherwise displaying unstationary behavior. The Rubin-Gelman statistic was calculated with five chains of length $2 \cdot 10^6$ and a burn-in of 10^4 to $\widehat{R}^1 = 1$. Thus, it is assumed that the chain has converged to the posterior distribution. It is also assumed that with the large MCMC sample from the posterior distribution, the autocorrelation effects wash out.

Table 5.3 show the statistics of the MCMC sample of τ. Now, the mean and median are close to each other and close to the true $\tau = 1\ 500$. The 2.5th and 97.5th percentiles and the 95% HPD are much narrower than at the earlier t_n, as is natural.

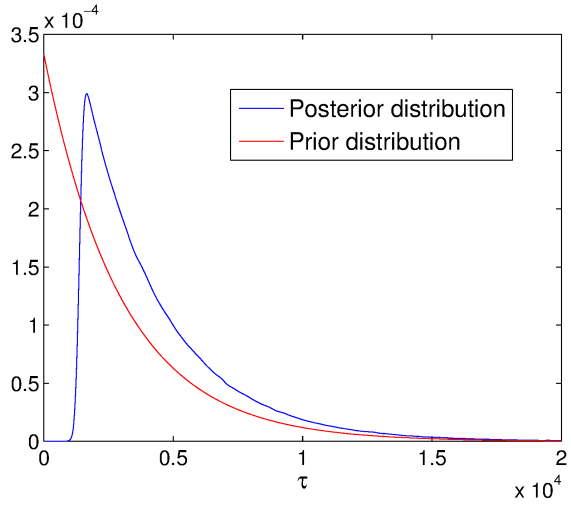

Figure 5.1.2: Prior distribution of τ, $\pi(\tau)$ and posterior distribution of τ, $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, given the observed temperature $W(t)$ up until $t_n = 1\ 400$. Here, ν and σ were considered known.

Table 5.1: Statistics of the posterior distribution of τ, $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, as estimated by MCMC methods. Here, ν and σ were considered known, and the endpoint of observation was $t_n = 1\ 400$.

Point estimates	Mean	4 378
	Median	3 459
Interval estimates	2.5 and 97.5 percentile	[1 454, 12 448]
	95 % HPD	[1 292, 10 389]

Table 5.2: Statistics of the posterior distribution of τ, $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, as estimated by MCMC methods. Here, ν and σ were considered known, and the endpoint of observation was t_n =1 700.

Point estimates	Mean	3 352
	Median	2 149
Interval estimates	2.5 and 97.5 percentile	[1 398, 11 125]
	95 % HPD	[1 254, 9 138]

Table 5.3: Statistics of the posterior distribution of τ, $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, as estimated by MCMC methods. Here, ν and σ were considered known, and the endpoint of observation was $t_n = 3\ 000$.

Point estimates	Mean	1 507
	Median	1 497
Interval estimates	2.5 and 97.5 percentile	[1 357, 1 857]
	95 % HPD	[1 292, 1 682]

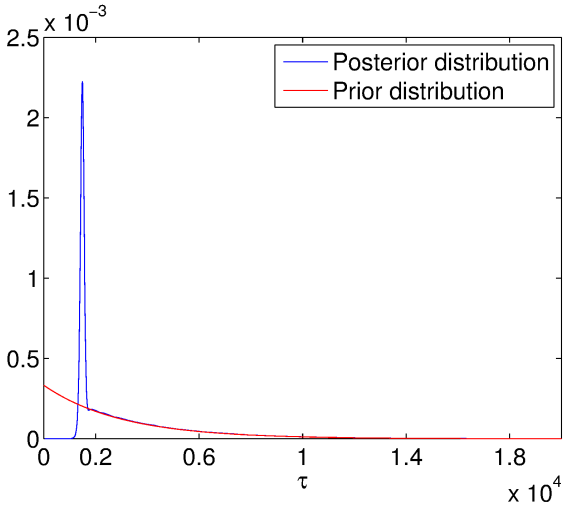

Figure 5.1.3: Prior distribution of τ, $\pi(\tau)$ and posterior distribution of τ, $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, given the observed temperature $W(t)$ up until $t_n = 1\ 700$. Here, ν and σ were considered known.

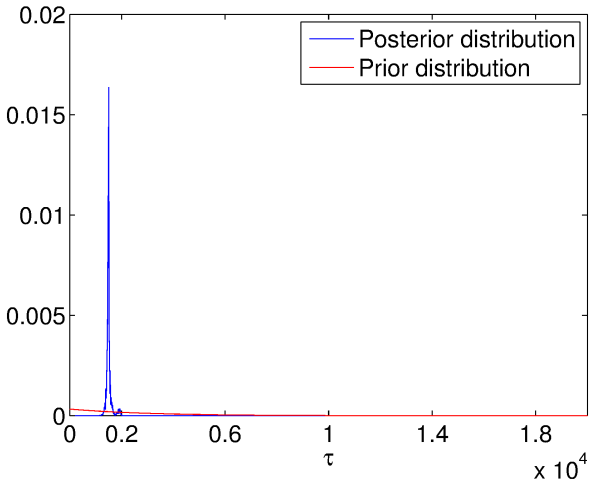

Figure 5.1.4: Prior distribution of τ, $\pi(\tau)$ and posterior distribution of τ, $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, given the observed temperature $W(t)$ up until $t_n = 3\ 000$. Here, ν and σ were considered known.

Hitting time CDF estimates

The hitting time CDF estimates of the temperature threshold $a = 14$ will now be presented. Let it be implicitly understood that the hitting time CDF estimates are given the data and the known parameters. That is, we write $\widehat{P}_f(T \leq t)$ for $P_f(T \leq t; x_1, \ldots, x_n; \nu, \sigma)$. The estimates for all three values of t_n by the formulaic approach $P_f(T \leq t)$, are shown in figure 5.1.5. The estimate $\widehat{P}_f(T \leq t)$ was calculated with $m = 5 \cdot 10^5$ draws from the posterior distribution of τ for each value of t. Figure 5.1.6 shows the hitting time CDF estimate by the simulation solution, $P_S(T \leq t)$, with the credible intervals and true distribution. In the simulation approach, for each realization of τ^* from the posterior distribution, 100 simulations of Wiener processes were run to estimate $P(T \leq t|\tau^*)$. In total, there were $m = 10^4$ realizations τ^* drawn from $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$. Figure 5.1.7 compares the two approaches, $P_f(T \leq t)$, and $P_S(T \leq t)$, and their respective credible intervals. Finally, some selected percentiles of $\widehat{P}_f(T \leq t)$ and $P_S(T \leq t)$ as well as their respective credible intervals are given in table 5.4. The percentiles of the true distribution is shown for comparison.

The estimates are plotted with the so-called *true* distribution, which is what we term the distribution given the true value of τ and the temperature measurements up to $t_n = 3\ 000$. That is, the true distribution is the inverse Gaussian CDF with the true drift parameter ν, true variance parameter σ and threshold parameter $a - W(3000)$, in other words, the distribution given the maximum amount of data in our case. However, the process' path from $t_n = 1\ 400$ to $t_n = 3\ 000$ is just one possible realization, thus the true distribution shows the deviance only for this case.

Note that the shape of the curve $P(T \leq t)$ is determined by ν, σ and $a - W(\max\{t_n, \tau\})$. Thus, in every case where $\tau \leq t_n$, the estimated CDF of T is the same. Thus, when nearly all realizations of τ^* in the MC summation of the formulaic solution $P_f(T \leq t)$, are such that $\tau^* < t_n$, the 95 % credible interval is exactly the same as the estimate itself. This happens for $t_n = 3\ 000$, where the curves of the estimate, the credible interval and the true distribution completely overlap in the figure. Table 5.4 show that the values are in fact exactly the same, as should be expected, since all values of τ^* in MCMC sampling are below $t_n = 3\ 000$, thus giving identical, and correct, estimates.

The spread of the estimates for $t_n = 1\ 400$ and $t_n = 1\ 700$ are of course larger, with the spread of $t_n = 1\ 400$ being the largest, as is natural. Note that the upper limit of the credible interval is much closer to the true distribution than the lower limit of the credible interval. For $t_n = 1\ 700$, the true change point $\tau = 1\ 500$ is passed, so the process has a drift. Thus, as long as $\tau^* \leq t_n$, all estimates of the hitting time CDF will be the same, unless the temperature $W(t_n)$ gives rise to a higher expected temperature at $t = 3\ 000$. From the plot of the temperature process, seen in figure 5.1.1, it is clear that this may be the case, since the process is flattening out at the very end of the time interval, say the time interval [2 500, 3000]. At $t_n = 1\ 400$, there is another possible way for the hitting time CDF estimate to elevate above the true distribution, namely that $\tau^* < 1\ 500$, the true value of τ can be drawn, and when drawn will affect the threshold parameter $a - W(\max\{t_n, \tau\})$. However, every value of $\tau^* > t_n$ will result in a lower hitting time CDF estimate, and remembering figures 5.1.2 and 5.1.3 we know that there

are many such τ^* which can be drawn from $\pi(\tau|x_1, \ldots, x_n)$.

Note also that for both $t_n = 1\ 700$ and $t_n = 1\ 400$, the lower credible interval exceeds the hitting time CDF estimate itself. This is because the number of τ^* values that are lower than the estimate itself are very few, but in the mean, which is the hitting time CDF estimate they contribute with a value at or near zero, and thus influence the mean to a great extent.

It can also be noted that the lower percentiles of the hitting time CDF estimate at $t_n = 1\ 700$ is closer to the true distribution, but for the higher percentiles, the behavior is more similar to that of the hitting time estimate at $t_n = 1\ 400$. This is because of the sharp peak in the posterior distribution of τ and the following heavy tail. This also causes the jaggedness of the lower credible interval. As it is selected as the 2.5th percentile, it is not strictly increasing when different τ^* are drawn for each t.

Note that the smaller number of realizations from the posterior distribution gives broader credible intervals than $\widehat{P}_f(T \leq t)$, but the estimates themselves are very close, see figure 5.1.7. However, in these examples, the number of simulated processes and draws from $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$ were restricted because of computational limitations.

Regardless of the poorer accuracy of the simulation solution in this example, the behavior is the same. The upper credible interval is closer to the true distribution, the credible intervals narrow as t_n grows and at $t_n = 3\ 000$, there is overlap of the true distribution and the simulation estimate $P_S(T \leq t)$. Also, the lower percentiles of $P_S(T \leq t)$ at $t_n = 1\ 700$ are notably steeper than the higher percentiles, reflecting the peak and tail of the posterior distribution. The upper and lower credible interval also reflect the typical staircase-shape of the empirical CDF. This is smoothed out in the mean, $P_S(T \leq t)$, and it is believed that a higher number of simulations and draws from the posterior distribution will smooth out the credible intervals as well.

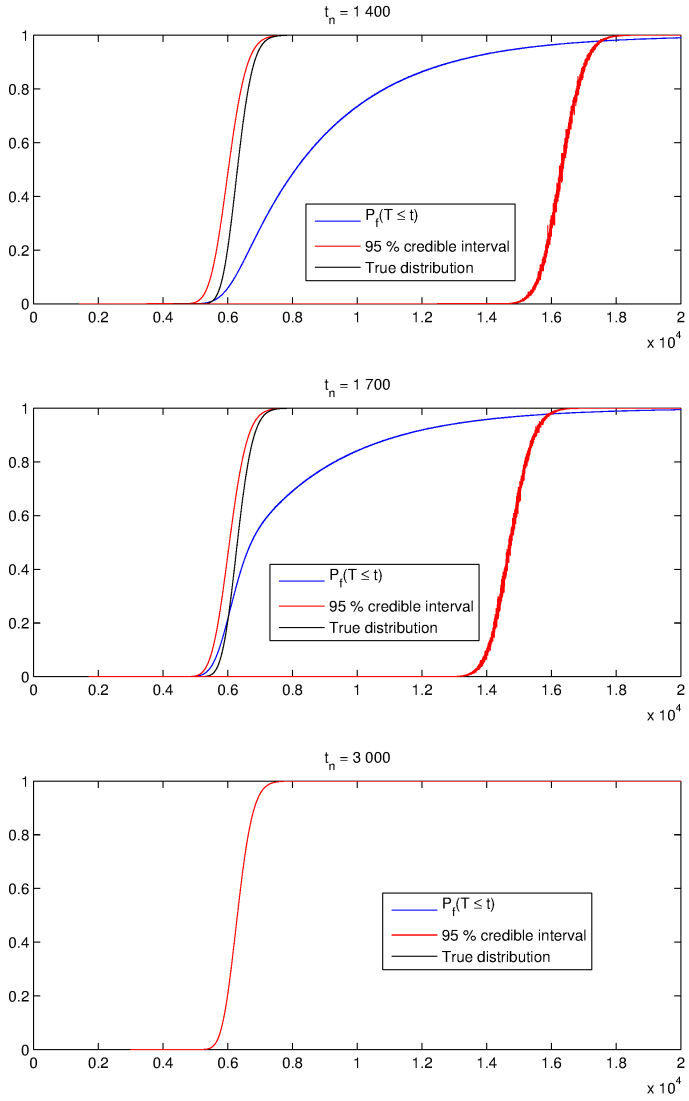

Figure 5.1.5: Formulaic approach to estimating CDF, $\widehat{P}_f(T \leq t)$ for different observation lengths, t_n, shown with 95 % credible intervals and the true distribution. The x-axis is time t while the y-axis is the probability. Top panel: $t_n = 1\ 400$. Middle panel: $t_n = 1\ 700$. Bottom panel: $t_n = 3\ 000$. Here, ν and σ are considered known parameters, while τ is unknown.

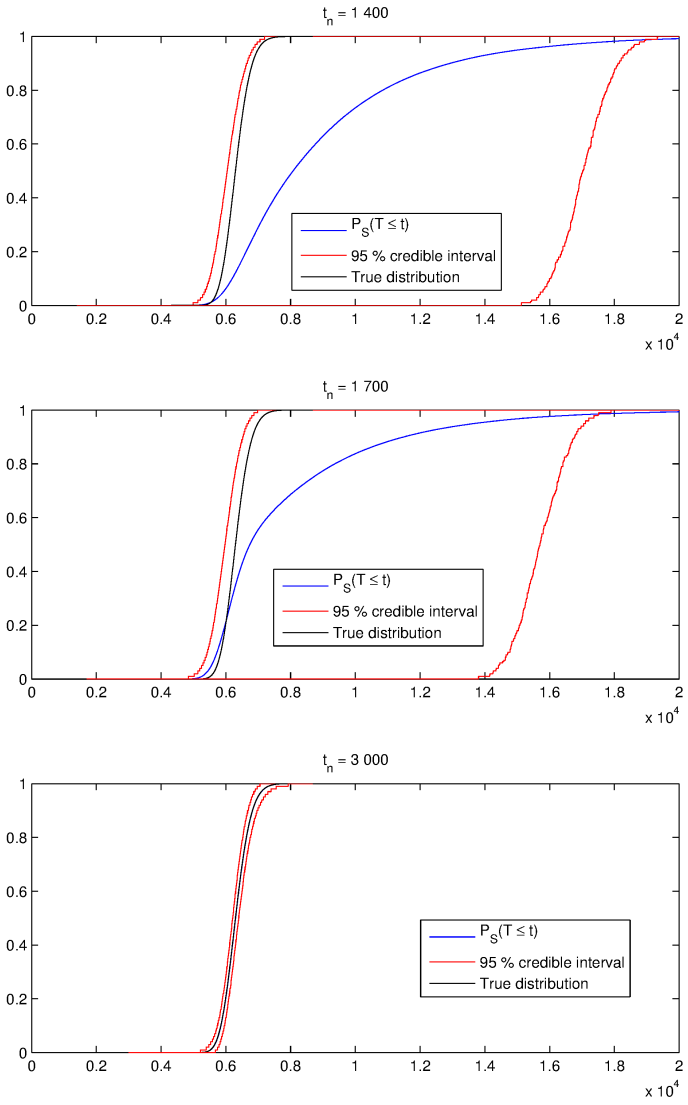

Figure 5.1.6: Simulation approach to estimating CDF, $\widehat{P}_S(T \leq t)$ for different observation lengths, t_n, shown with 95 % credible intervals and the true distribution. The x-axis is time t while the y-axis is the probability. Top panel: $t_n = 1\ 400$. Middle panel: $t_n = 1\ 700$. Bottom panel: $t_n = 3\ 000$.Here, ν and σ are considered known parameters, while τ is unknown.

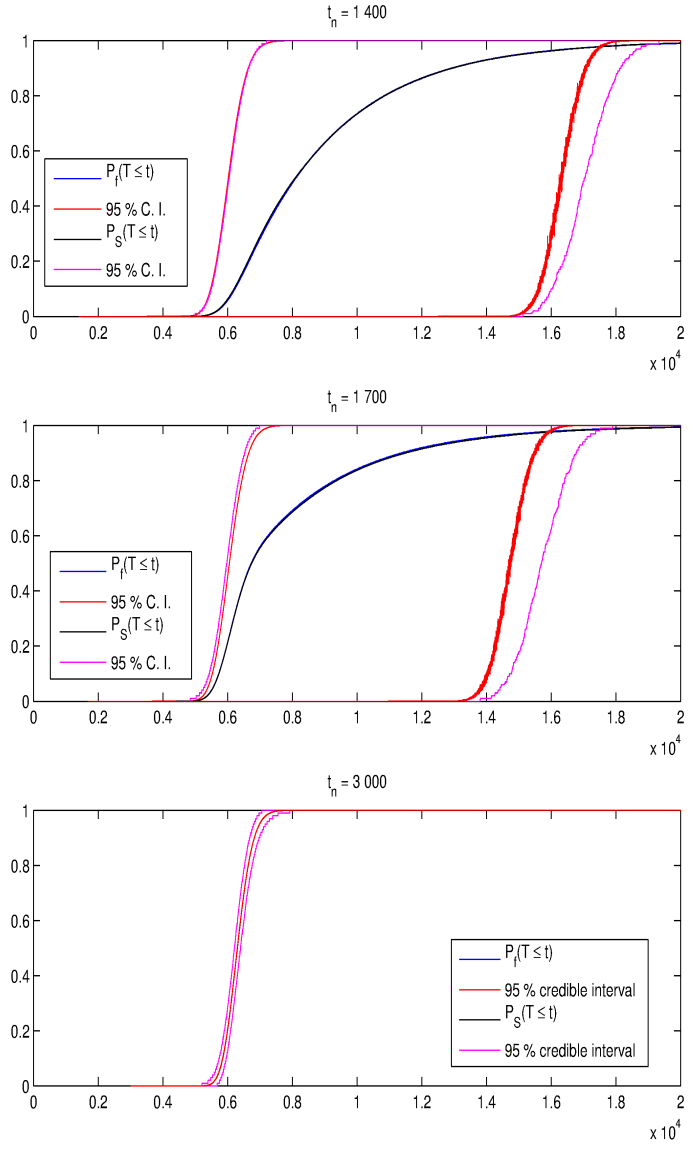

Figure 5.1.7: Comparison of simulation approach to estimating the hitting time CDF, $\widehat{P}_S(T \leq t)$ and formulaic approach to estimated the hitting time CDF, $\widehat{P}_f(T \leq t)$, with their respective 95 % credible intervals for different observation lengths, t_n. The x-axis is time t while the y-axis is the probability. Top panel: $t_n = 1\ 400$. Middle panel: $t_n = 1\ 700$. Bottom panel: $t_n = 3\ 000$. Here, ν and σ are considered known parameters, while τ is unknown.

Table 5.4: Percentiles of the estimated hitting time CDF with τ unknown, and ν and σ known.

Percentile	2.5	5	10	25	50	75	90	95	97.5
Formulaic solution									
$t_n = 1\ 400$									
Upper 95 %	4 212	5 230	5 474	5 722	6 019	6 333	6 634	6 823	6 993
Estimated CDF	5 740	5 946	6 226	6 865	8 093	10 170	12 901	14 965	17 060
Lower 2.5 %	15 200	15 372	15 580	15 885	16 242	16 661	17 029	17 255	17 461
$t_n = 1\ 700$									
Upper 95 %	5 275	5 390	5 527	5 767	6 053	6 359	6 652	6 836	7 001
Estimated CDF	5 410	5 549	5 724	6 078	6 728	8 631	11 370	13 448	15 494
Lower 95 %	13 633	13 796	13 997	14 294	14 684	15 100	15 453	15 683	15 863
$t_n = 3\ 000$									
Upper 95 %	5 631	5 727	5 844	6 050	6 297	6 564	6 823	6 986	7 133
Estimated CDF	5 631	5 727	5 844	6 050	6 297	6 564	6 823	6 986	7 133
Lower 95 %	5 631	5 727	5 844	6 050	6 297	6 564	6 823	6 986	7 133
Simulation solution									
$t_n = 1\ 400$									
Upper 95 %	5 220	5 367	5 507	5 751	6 050	6 366	6 663	6 851	6 967
Estimated CDF	5 717	5 921	6 197	6 830	8 075	10 179	12 925	15 044	17 125
Lower 95 %	15 603	15 796	16 040	16 528	17 038	17 614	18 155	18 514	18 778
$t_n = 1\ 700$									
Upper 95 %	5 132	5 295	5 442	5 691	5 976	6 269	6 537	6 697	6 799
Estimated CDF	5 413	5 553	5 728	6 082	6 743	8 698	11 497	13 681	15 841
Lower 95 %	14 261	14 445	14 777	15 190	15 705	16 300	16 784	17 109	17 373
$t_n = 3\ 000$									
Upper 95 %	5 477	5 624	5 749	5 967	6 214	6 465	6 694	6 826	6 908
Estimated CDF	5 635	5 732	5 848	6 054	6 301	6 568	6 825	6 989	7 137
Lower 95 %	5 787	5 870	5 965	6 155	6 402	6 687	6 995	7 222	7 395
True distribution									
$t_n = 3\ 000$	5 631	5 727	5 844	6 050	6 297	6 564	6 823	6 986	7 133

5.1.2 τ and ν unknown, σ known

In this section, we will examine the situation where both the drift parameter ν and the change point τ are unknown parameters. The variance parameter σ will be considered known. As above, we will examine the MCMC diagnostics of the three cases $t_n = 1\ 400$, $t_n = 1\ 700$ and $t_n = 3\ 000$, before looking at the hitting time CDF estimates by the formulaic and simulation approach. Note that the estimates has been obtained by the same temperature process as before, shown in figure 5.1.1.

The geometric prior on τ was continued such that $\tau \sim \mathcal{G}eom(1/3000)$ and the prior on ν was chosen $\nu \sim \mathcal{G}am(4, 0.0006)$, which was thought to sample from a reasonable range for ν, while avoiding the numerical difficulties discussed in section 4.2.4.

Observation until $t_n = 1\ 400$

Figure 5.1.8 shows the marginal prior distributions for τ and ν as well as the marginal posterior distributions of τ and ν, $\pi(\tau|x_1, \ldots, x_n)$ and $\pi(\nu|x_1, \ldots, x_n)$ respectively, as obtained by kernel smoothing density estimates. Note how $\pi(\nu|x_1, \ldots, x_n)$ is very close to $\pi(\nu)$. Remember that the true change point has not yet occurred, thus in reality we have no temperature increments x_i drawn with expectation ν. As before, the posterior distribution of τ, $\pi(\tau|x_1, \ldots, x_n)$ is quite flat until $\tau = 1\ 400$, and thereafter, it rises sharply, due to the prior distribution $\pi(\tau)$. Note however that compared to figure 5.1.2, there is more density before $\tau = 1\ 400$. This can be explained by the variability of ν. If we consult the contour plot of the joint posterior density, $\pi(\tau, \nu|x_1, \ldots, x_n)$, shown in figure A.1.12 in appendix A.1.2, we see that these smaller values of τ correspond to near-zero drift, which the increments until t_n are likely to have since $E[x_i] = 0$. That is, the situation that the change point has occurred, but the drift is very small. Thus, it is to be expected that $\pi(\nu)$ is very influential of $\pi(\tau|x_1, \ldots, x_n)$, in the situation where $\pi(\tau|x_1, \ldots, x_n; \nu)$ is near zero for $\tau < t_n$. Further diagnostics plots are shown in appendix A.1.2: the marginal trace plots of τ and ν are shown in figure A.1.10, the ACF of the MCMC iterations of τ and ν as function of lag are shown in figure A.1.13 and the histogram of the samples of τ and ν are shown in A.1.11. There were some small cross-correlations between τ and ν, which are not shown. These were assumed to stem from the natural dependency when estimating the two variables which were seen in the contour plot - that if ν is very small, τ can be very small, and if τ is large, ν can be anything. The trace plots and the ACF suggest that the chain is moving rapidly and freely and not getting stuck. This, combined with the calculated Rubin-Gelman statistic $\widehat{R}^2 = 1$, obtained with five chains of run length $2 \cdot 10^6$ and burn-in of 10^4, points to convergence of the chain.

Some sample statistics of τ and ν are given in table 5.5, showing the mean and median as well as the 2.5th and 97.5th percentile and the 95 % HPD. The interval estimates are quite large for the two random variables. The most notable is the difference between the median and mean of τ, as before, caused by the heavy tail. The mean has been increased compared to the prior distribution, as in the situation with one change point and ν known, because the posterior is near zero for $\tau < t_n$.

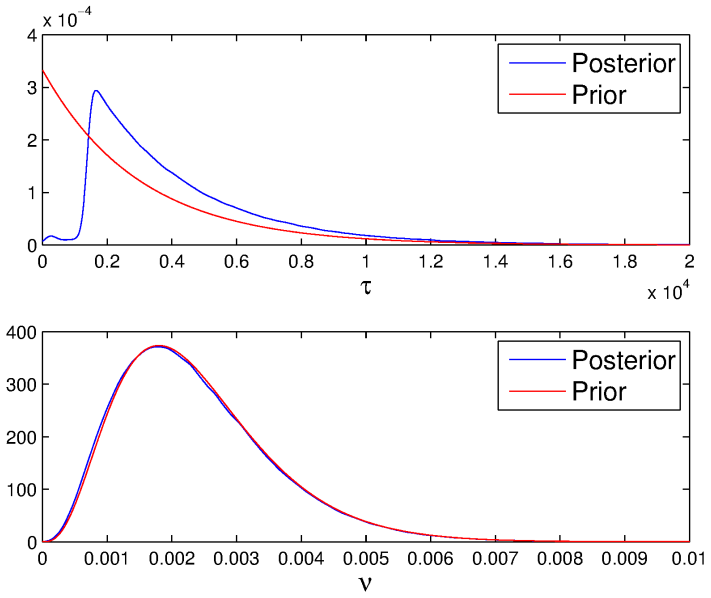

Figure 5.1.8: Prior distribution of τ, $\pi(\tau)$ and ν, $\pi(\nu)$ and marginal posterior distributions of τ and ν, $\pi(\tau, |x_1, \ldots, x_n; \sigma)$ and $\pi(\nu, |x_1, \ldots, x_n; \sigma)$ given the observed temperature $W(t)$ up until $t_n = 1\ 400$. Here, σ is considered known.

Table 5.5: Statistics of the joint posterior distribution of τ and ν, $\pi(\tau, \nu | x_1, \ldots, x_n; \sigma)$, as estimated by MCMC methods. Here, σ was considered known, and the endpoint of observation was t_n =1 400.

	τ	
Point estimates	Mean	4 304
	Median	3 399
Interval estimates	2.5 and 97.5 percentile	[1 397, 12 380]
	95 % HPD	[1 140, 11 309]
	ν	
Point estimates	Mean	0.0024
	Median	0.0022
Interval estimates	2.5 and 97.5 percentile	[0.0006, 0.0052]
	95 % HPD	[0.0004, 0.0047]

Observation until t_n = 1 700

Figure 5.1.9 shows the marginal prior distributions for τ and ν as well as the marginal posterior distributions of τ and ν, $\pi(\tau|x_1,\ldots,x_n)$ and $\pi(\nu|x_1,\ldots,x_n)$ respectively, as obtained by kernel smoothing density estimates. Compared to figure 5.1.8, the change is striking in $\pi(\tau|x_1,\ldots,x_n)$, as was the case when ν was known. However, the change in $\pi(\nu|x_1,\ldots,x_n)$ is not as remarkable, although the distribution is shifted towards larger values of ν. This may stem from the heavy tail of $\pi(\tau|x_1,\ldots,x_n)$, because if $\tau^* > t_n$, there are no temperature increments with drift ν, thus the only information about the drift is the prior distribution.

Further diagnostics plots are shown in appendix A.1.2, the trace plot of the marginal posterior distribution $\pi(\tau|x_1,\ldots,x_n)$ and $\pi(\nu|x_1,\ldots,x_n)$ in figure A.1.14, the histogram in figure A.1.15, ACF for the MCMC iterations from $\pi(\tau|x_1,\ldots,x_n)$ and $\pi(\nu|x_1,\ldots,x_n)$ as function of the lag in figure A.1.17 and finally a contour plot of the joint posterior density $\pi(\tau,\nu|x_1,\ldots,x_n)$ is shown in figure A.1.16. The trace plots show that the acceptance rate has gone down as the posterior density has become more peaked. This also results in the ACF decreasing at a slower rate. Also here, there were some small cross-correlations between τ and ν, which are not shown. However, the chain is moving fast and freely, and even though it gets stuck for smaller amounts of time, it is seems to have good mixing overall. In this case as well, the Rubin-Gelman statistic was calculated $\widehat{R}^2 = 1$, obtained with five chains of run length $2 \cdot 10^6$ and burn-in of 10^4, pointing to convergence of the chain.

Some sample statistics of the Markov chain is show in table 5.6. Note that the mean an median of ν has both risen, as is natural since the change point is now passed, and the cases with small change points τ and near-zero drift ν should be less probable. The mean and median of τ is now decreased, as was the case with ν known. The intervals for τ are decreased as is natural when there is more information. Notice that the intervals for ν does not decrease as much, which can be traced to the heavy tail $\pi(\tau|x_1,\ldots,x_n)$.

Observation until t_n = 3 000

Finally, we consider the situation with observations up until $t_n = 3\,000$, which provides an abundance of measurements of temperature increments X_i after the true change point $\tau = 1\,500$. Figure 5.1.10 shows the marginal prior distributions for τ and ν as well as the marginal posterior distributions of τ and ν, $\pi(\tau|x_1,\ldots,x_n)$ and $\pi(\nu|x_1,\ldots,x_n)$ respectively, as obtained by kernel smoothing density estimates. The peakedness of $\pi(\tau|x_1,\ldots,x_n)$ is almost as sharp as it was with ν known. It is also clear that $\pi(\nu|x_1,\ldots,x_n)$ is more peaked and less influenced by the prior distribution. However, it is centered to the left of the true value of ν, 0.003, which is clear from the statistics shown in table 5.7. It can be noted that when considering the measurements from the true $\tau = 1\,500$ to $t_n = 3\,000$, the mean of the temperature increments, $\bar{X}_i = 0.027$, which is smaller than the true value. Thus it is natural that $\pi(\nu|x_1,\ldots,x_n)$ is shifted towards smaller values of ν.

Further diagnostics plots are shown in appendix A.1.2: figure A.1.18 shows the marginal trace plots of τ and ν, figure A.1.19 shows a histogram of the sample val-

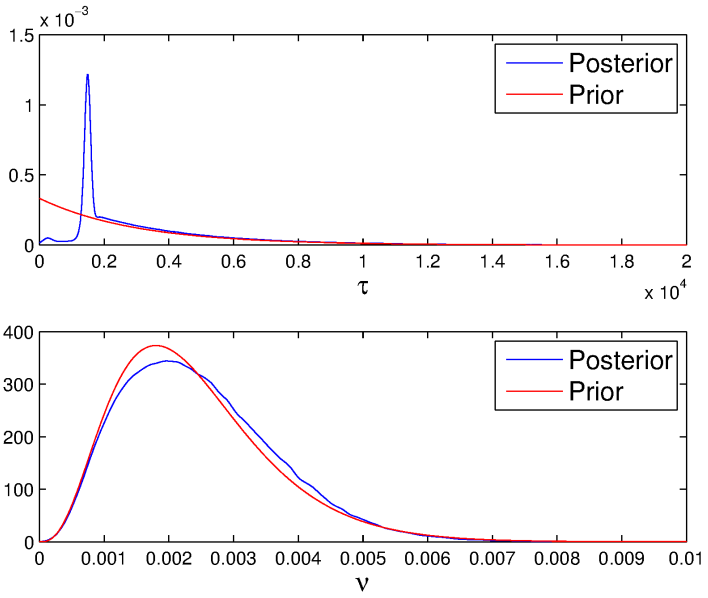

Figure 5.1.9: Prior distribution of τ, $\pi(\tau)$ and ν, $\pi(\nu)$ and marginal posterior distributions of τ and ν, $\pi(\tau, |x_1, \ldots, x_n; \sigma)$ and $\pi(\nu, |x_1, \ldots, x_n; \sigma)$ given the observed temperature $W(t)$ up until $t_n = 1\ 700$. Here, σ is considered known.

Table 5.6: Statistics of the joint posterior distribution of τ and ν, $\pi(\tau, \nu | x_1, \ldots, x_n; \sigma)$, as estimated by MCMC methods. Here, σ was considered known, and the endpoint of observation was t_n =1 700.

	τ	
Point estimates	Mean	3 501
	Median	2 440
Interval estimates	2.5 and 97.5 percentile	[694, 11 400]
	95 % HPD	[1, 9 326]
	ν	
Point estimates	Mean	0.0025
	Median	0.0023
Interval estimates	2.5 and 97.5 percentile	[0.0007, 0.0052]
	95 % HPD	[0.0005, 0.0048]

ues of $\pi(\tau|x_1,\ldots,x_n)$ and $\pi(\nu|x_1,\ldots,x_n)$, figure A.1.20 shows the contour plot of the joint posterior distribution $\pi(\tau,\nu|x_1,\ldots,x_n)$, and finally, the ACF of MCMC iterations of τ and ν as function of the lag are shown in figure A.1.21. It is now clear that the acceptance rate has dropped significantly, which manifests in the trace plots as well as the ACF plots, which are decreasing at a slower rate. However, the decrease is evident, and looking at trace plots of longer stretches, there behavior is not sequential. Noting that the number of measurements is large, it is natural to assume that the likelihood is considerably more peaked than for $t_n = 1\ 700$, for example. The Rubin-Gelman statistic was calculated $\widehat{R}^2 = 1$, obtained with five chains of run length $2 \cdot 10^6$ and burn-in of 10^4, where special care was taken to over-disperse the starting points of the chain. Thus, it is assumed that the Markov chain has converged to the posterior distribution.

Table 5.7 also shows that the interval estimates of τ are now significantly decreased and that the median and mean of τ are close to each other and to the true value of τ. The mean and median of ν is not as close to the true value, and the interval estimates center at the mean (and median). Although the true value of ν is inside both intervals, the intervals are shifted towards smaller values of ν. This is assumed to stem from smaller realized means of the temperature increments as remarked above, and the fact that with $\tau^* < t_n$, a smaller value of ν^* is to be expected.

Hitting time CDF estimates

Finally, the resulting hitting time CDF estimates of the temperature level $a = 14$ for each of the three t_n are shown. The formulaic approach estimates, $\widehat{P}_f(T \leq t)$, is shown in figure 5.1.11, and the simulation approach estimates $P_S(T \leq t)$ is shown in figure 5.1.12, with their respective 95 % credible intervals and the true distribution. Note the different end values of the time axis in the two figures. The comparison of the two approaches in shown in figure 5.1.13. Some percentiles of the hitting time CDF estimates and the credible intervals are shown in table 5.8, with the true distribution. The estimate $\widehat{P}_f(T \leq t)$ was calculated with $m = 5 \cdot 10^5$ draws from the joint posterior distribution of τ and ν, $\pi(\tau,\nu|x_1,\ldots,x_n;\sigma)$, for each value of t. In the simulation approach, for each realization of τ^* and ν^* from the joint posterior distribution, 100 simulations of Wiener processes were run to estimate $P(T \leq t|\tau^*)$. In total, there were $m = 10^4$ realizations τ^* drawn from $\pi(\tau,\nu|x_1,\ldots,x_n;\sigma)$.

The true distribution shown for comparison in figures 5.1.11 and 5.1.12 is the same as with ν known, the inverse Gaussian distribution with the true drift parameter ν, true variance parameter σ and threshold parameter $a - W(3000)$, in other words, the distribution given the maximum amount of data in our case.

Note now that in both approaches, the upper credible interval limit is higher for all observations lengths t_n compared to the situation where ν was known. This is partly because of the increased variation that naturally follows when an unknown parameter is added, and partly because there are now more ways that the hitting time T can become higher: as before, we can obtain $\tau^* < 1\ 500$, which influences only the case $t_n = 1\ 400$, and with ν unknown, we have seen from the plot of $\pi(\tau|x_1,\ldots,x_n)$ that this is more likely. We can also obtain $W(\tau^*) > W(t_n)$, which gives higher values of $P(T \leq t|\tau^*,\nu^*)$

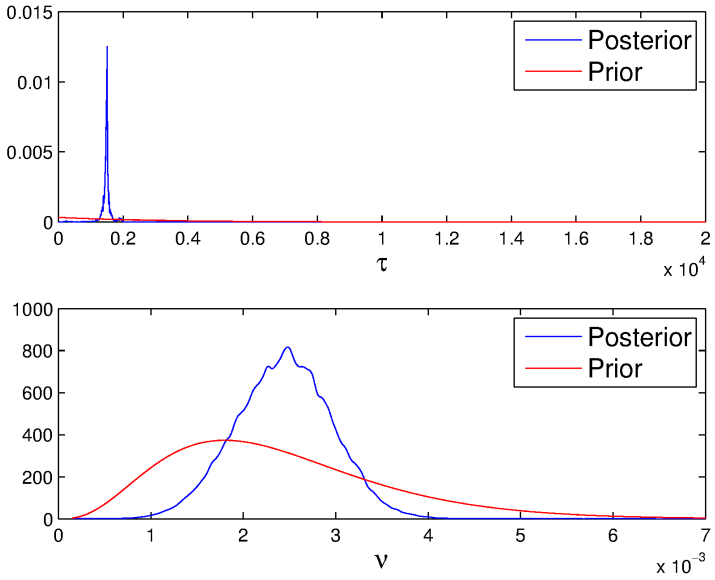

Figure 5.1.10: Prior distribution of τ, $\pi(\tau)$ and ν, $\pi(\nu)$ and marginal posterior distributions of τ and ν, $\pi(\tau, |x_1, \ldots, x_n; \sigma)$ and $\pi(\nu, |x_1, \ldots, x_n; \sigma)$ given the observed temperature $W(t)$ up until $t_n = 3\ 000$. Here, σ is considered known.

Table 5.7: Statistics of the joint posterior distribution of τ and ν, $\pi(\tau, \nu | x_1, \ldots, x_n; \sigma)$, as estimated by MCMC methods. Here, σ was considered known, and the endpoint of observation was t_n =3 000.

	τ	
Point estimates	Mean	1 467
	Median	1 489
Interval estimates	2.5 and 97.5 percentile	[981, 1 861]
	95 % HPD	[1 238, 1 952]
	ν	
Point estimates	Mean	0.0024
	Median	0.0024
Interval estimates	2.5 and 97.5 percentile	[0.0014, 0.0034]
	95 % HPD	[0.0014, 0.0034]

for all t. Furthermore, which is new for the situation with ν unknown, we can obtain $\nu^* > \nu$, which gives higher values of $P(T \leq t|\tau^*, \nu^*)$ for all t. The lower credible interval is also lower, which stems from the fact that when $\nu^* < \nu$, $P(T \leq t|\tau^*, \nu^*)$ decreases compared to the true distribution.

The steepening of the lower percentiles of the hitting time CDF estimates, which was accredited to the peak in the posterior distribution, at $t_n = 1\ 700$ is evident also now, but less so than when ν was known.

Also note that $t_n = 3\ 000$ does not give as narrow credible intervals as before. This shows the greater effect the variability in the parameter ν has on the inverse Gaussian CDF. Neither is the estimate as close to the true distribution as when ν is known, which is natural. It is thought that if the estimate of ν were higher, it would be possible that the hitting time CDF estimate lay between the upper credible interval and the true distribution.

When comparing the two approaches, we see now that the credible intervals of the simulation approach $P_S(T \leq t)$ are narrower than those of the formulaic approach, $P_f(T \leq t)$. It can be confirmed in table 5.8. This is assumed to stem from the fact that fewer values from the posterior distribution was drawn to calculate $P_S(T \leq t)$, thus the entire variability of the distribution may not have been captured. It is assumed that if a larger sample from the posterior distribution was drawn, the two credible intervals would be nearly equal. The fact that the intervals approach each other for increasing t_n strengthens this belief, as the posterior distributions narrows when t_n grows, such that fewer draws are needed to capture the variance.

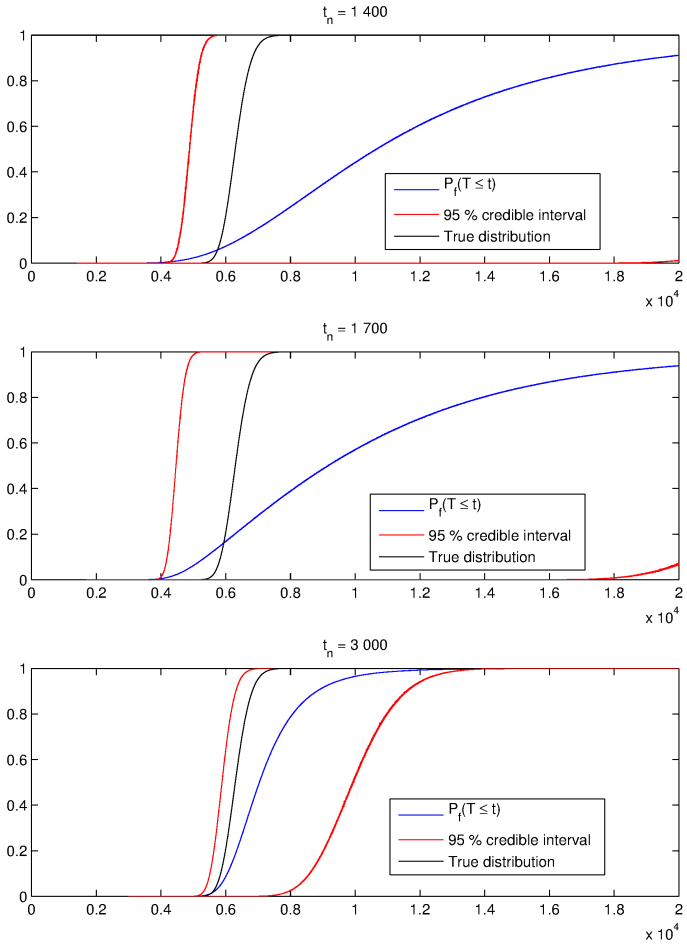

Figure 5.1.11: Formulaic approach to estimating CDF, $\widehat{P}_f(T \leq t)$ for different observation lengths, t_n, shown with 95 % credible intervals and the true distribution. The x-axis is time t while the y-axis is the probability. Top panel: $t_n = 1\ 400$. Middle panel: $t_n = 1\ 700$. Bottom panel: $t_n = 3\ 000$. Here, σ is considered known.

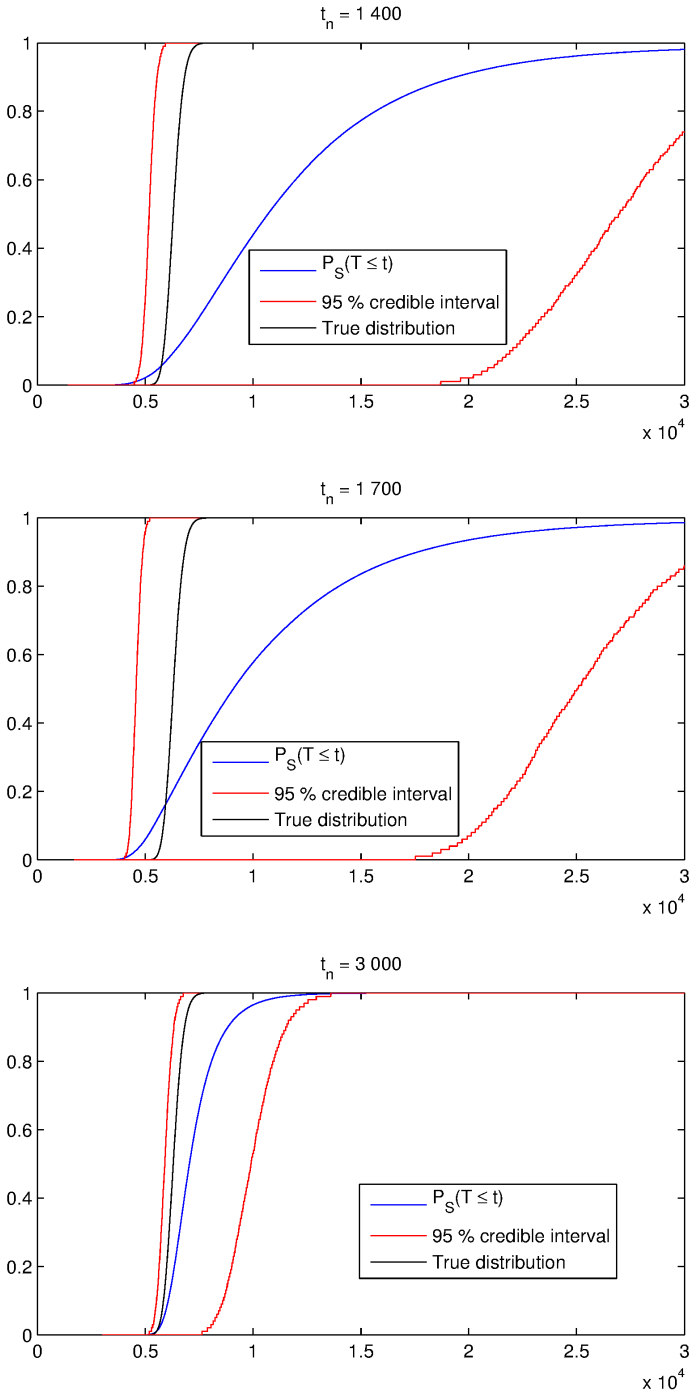

Figure 5.1.12: Simulation approach to estimating CDF, $\widehat{P}_S(T \leq t)$ for different observation lengths, t_n, shown with 95 % credible intervals and the true distribution. The x-axis is time t while the y-axis is the probability. Top panel: $t_n = 1\ 400$. Middle panel: $t_n = 1\ 700$. Bottom panel: $t_n = 3\ 000$. Here, σ is considered known.

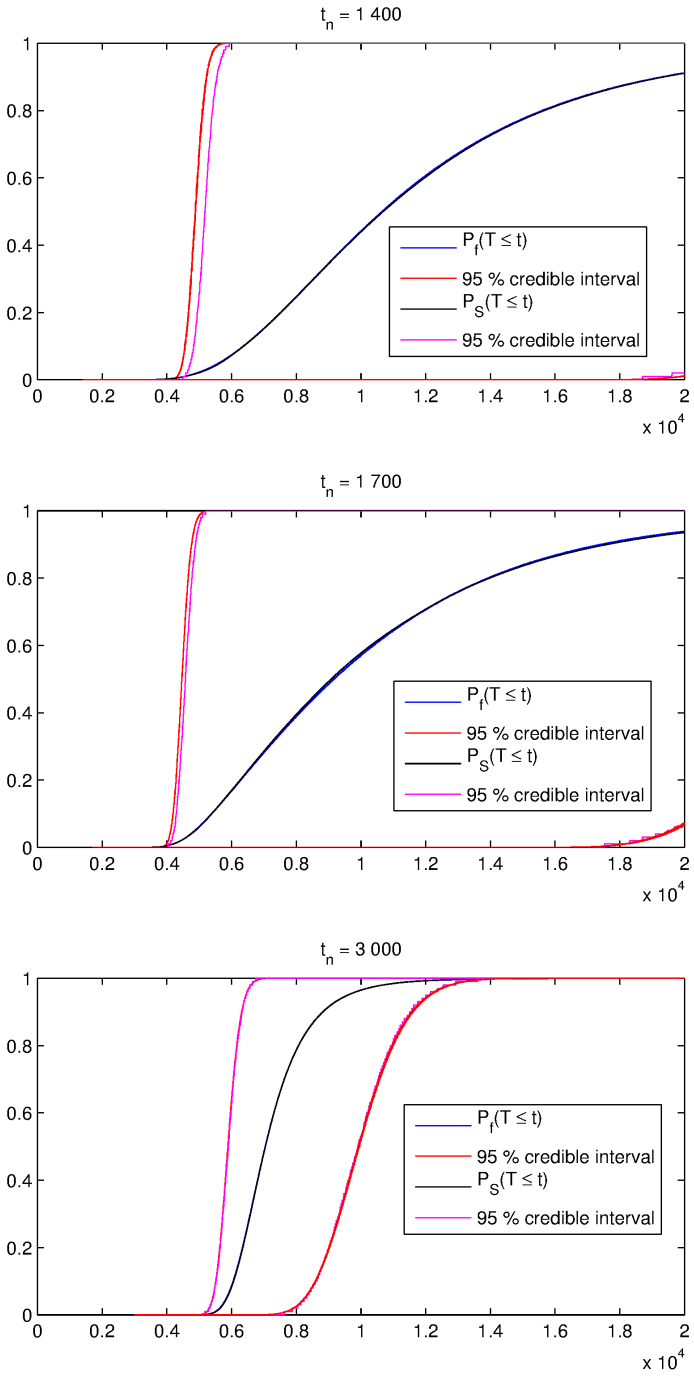

Figure 5.1.13: Comparison of the simulation and formulaic approaches, $P_S(T \leq t)$ and $\widehat{P}_f(T \leq t)$ and their respective 95 % credible intervals with τ and ν as unknown parameters, and σ considered known, for three different observation lengths t_n. The x-axis is time t while the y-axis is the probability. Top panel: $t_n = 1\ 400$. Middle panel: $t_n = 1\ 700$.

Table 5.8: Percentiles of the estimated CDF, τ and ν unknown, σ known.

Percentile	2.5	5	10	25	50	75	90	95	97.5
Formulaic solution									
$t_n = 1\ 400$									
Upper 95 %	4 398	4 478	4 555	4 708	4 878	5 048	5 227	5 335	5 414
Estimated CDF	5 064	5 612	6 366	8 035	10 655	14 428	19 293	$> 20\ 000$	
Lower 2.5 %	$> 20\ 000$								
$t_n = 1\ 700$									
Upper 95 %	4 059	4 117	4 193	4 318	4 464	4 619	4 765	4 854	4 937
Estimated CDF	4 520	4 866	5 387	6 723	9 161	12 805	17 407	$> 20\ 000$	
Lower 95 %	$> 20\ 000$								
$t_n = 3\ 000$									
Upper 95 %	5 337	5 416	5 513	5 681	5 886	6 104	6 313	6 444	6 566
Estimated CDF	5 680	5 842	6 051	6 460	7 042	7 829	8 821	9 597	10 409
Lower 95 %	7 996	8 259	8 583	9 183	9 924	10 759	11 585	12 108	12 615
Simulation solution									
$t_n = 1\ 400$									
Upper 95 %	4 647	4 739	4 834	4 997	5 185	5 369	5 554	5 686	5 760
Estimated CDF	5 116	5 655	6 371	8 035	10 692	14 487	19 380	23 318	27 922
Lower 95 %	20 214	21 169	22 053	24 157	26 996	$> 30\ 000$			
$t_n = 1\ 700$									
Upper 95 %	4 162	4 231	4 297	4 419	4 576	4 739	4 894	4 986	5 952
Estimated CDF	4 527	4 888	5 394	6 694	9 083	12 823	17 632	21 491	25 813
Lower 95 %	18 704	19 689	20 593	22 618	25 064	28 183	$> 30\ 000$		
$t_n = 3\ 000$									
Upper 95 %	5 341	5441	5 526	5 697	5 906	6 120	6 335	6 472	6 571
Estimated CDF	5 695	5 856	6 064	6 470	7 047	7 833	8 830	9 605	10 412
Lower 95 %	8 049	8 320	8 636	9 175	9 981	10 744	11 570	12 149	12 548
True distribution									
$t_n = 3\ 000$	5 631	5 727	5 844	6 050	6 297	6 564	6 823	6 986	7 133

5.1.3 τ, ν and σ unknown

Finally, we let σ as well be an unknown parameter. Again, we will examine the MCMC diagnostics of the three cases $t_n = 1\ 400$, $t_n = 1\ 700$ and $t_n = 3\ 000$, before looking at the hitting time CDF estimates by the formulaic and simulation approach, $\widehat{P}_f(T \leq t)$ and $\widehat{P}_S(T \leq t)$ respectively. Note that the estimates has been obtained by the same temperature process as before, shown in figure 5.1.1.

The geometric prior on τ was continued and the gamma prior on ν were continued such that $\tau \sim \mathcal{G}eom(1/3000)$ and $\nu \sim \mathcal{G}am(4, 0.0006)$.The parameter is given a gamma prior, such that $\sigma \sim \mathcal{G}am(4, 0.01)$ which is thought to be a suitable range for the variance parameter.

Observation until $t_n = 1\ 400$

For the situation where the change point has not yet occurred, $t_n = 1\ 400$, the marginal priors of τ, ν and σ are shown in 5.1.14 along with the marginal posterior distributions $\pi(\tau|x_1, \ldots, x_n), \pi(\nu|x_1, \ldots, x_n)$ and $\pi(\sigma|x_1, \ldots, x_n)$.

Further diagnostics plots are given appendix A.1.3: the marginal trace plots of τ, ν and σ are shown in figure A.1.22, the histogram of the sample values of each parameter is shown in figure A.1.23, the ACF as function of the lag for each of the three parameter iterations is shown in figure A.1.25, and finally, three pairwise contour plots of the pairwise joint posterior distributions, $\pi(\tau, \nu|x_1, \ldots, x_n), \pi(\nu, \sigma|x_1, \ldots, x_n)$ and $\pi(\tau, \sigma|x_1, \ldots, x_n)$ are shown in figure A.1.24. The sample statistics are summarized in table 5.9.

With the third unknown parameter, it is evident that the acceptance rate has decreased greatly, compared to $t_n = 1\ 400$ and σ known. This manifests in the trace plot and in the plot of the ACF. Compared to when σ was known, we double the run length to $4 \cdot 10^6$. This is done to wash out the autocorrelation effects and ensure that the chain is moving across the entire domain. With five chains with over-dispersed starting points, the Rubin-Gelman statistic was calculated, $\widehat{R}^3 = 1$. Thus, with the long run length, it was thought that the Markov chain had converged to the posterior distribution, and was able to traverse the entire range of the parameters.

Note in figure 5.1.14, the posterior distribution $\pi(\sigma|x_1, \ldots, x_n)$ of σ is more unaffected of its prior compared to the two others. This is because all temperature increments x_i are sampled with the same variance parameter σ, thus at $t_n = 1\ 400$, which allows the posterior distribution $\pi(\sigma|x_1, \ldots, x_n)$ to become very narrow and wash out the effect of the prior distribution $\pi(\sigma)$. This is also evident in table 5.9. The interval estimates of τ now show a larger portion of $\tau^* < t_n$, compared to σ known, and hence, and the mean and median of ν has decreased, compared to the situation with σ known. Note that for $\{W(t_i), i = 1, \ldots, n\}$, the empirical standard deviation of the temperature increments x_i, is 0.0203, which is thought to explain why the main portion of the mass of $\pi(\sigma|x_1, \ldots, x_n)$ lies above the true value 0.02. In addition, it is believed that the iterations for which $\tau^* < t_n$ and ν^* is near zero give require larger σ^* to be accepted in the MCMC procedure.

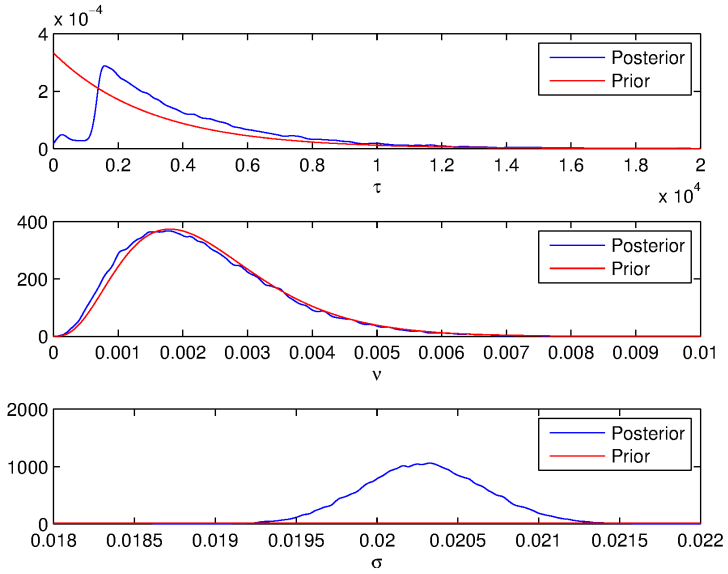

Figure 5.1.14: Prior distributions of τ, $\pi(\tau)$, ν, $\pi(\nu)$ and σ, $\pi(\sigma)$ and posterior distribution of τ, $\pi(\tau|x_1,\ldots,x_n;\nu,\sigma)$, ν, $\pi(\nu|x_1,\ldots,x_n;\nu,\sigma)$ and σ, $\pi(\sigma|x_1,\ldots,x_n;\nu,\sigma)$ given the observed temperature $W(t)$ up until $t_n = 1\ 400$. Here, τ, ν and σ were considered unknown.

Table 5.9: Statistics of the posterior distribution of τ, ν and σ, $\pi(\tau,\nu,\sigma|x_1,\ldots,x_n)$, as estimated by MCMC methods. Here, τ, ν and σ were considered unknown, and the endpoint of observation was t_n =1 400.

	τ	
Point estimates	Mean	4 149
	Median	3 260
Interval estimates	2.5 and 97.5 percentile	[614, 12 130]
	95 % HPD	[7, 10 094]
	ν	
Point estimates	Mean	0.0023
	Median	0.0021
Interval estimates	2.5 and 97.5 percentile	[0.0006, 0.0051]
	95 % HPD	[0.0004, 0.0047]
	σ	
Point estimates	Mean	0.020
	Median	0.020
Interval estimates	2.5 and 97.5 percentile	[0.0196, 0.0211]
	95 % HPD	[0.0195, 0.0210]

Observation until t_n = 1 700

In this situation the "now" point in time, t_n is close after the change point. The marginal priors of τ, ν and σ, $\pi(\tau), \pi(\nu)$ and $\pi(\sigma)$, are shown in 5.1.15 along with the marginal posterior distributions $\pi(\tau|x_1, \ldots, x_n), \pi(\nu|x_1, \ldots, x_n)$ and $\pi(\sigma|x_1, \ldots, x_n)$.

Further diagnostics plots are given appendix A.1.3: the marginal trace plots of τ, ν and σ are shown in figure A.1.26, the histogram of the sample values of each parameter is shown in figure A.1.27, the ACF as function of the lag for each of the three parameters is shown in figure A.1.29, and finally, three pairwise contour plots of the pairwise joint posterior distributions, $\pi(\tau, \nu|x_1, \ldots, x_n), \pi(\nu, \sigma|x_1, \ldots, x_n)$ and $\pi(\tau, \sigma|x_1, \ldots, x_n)$ are shown in figure A.1.28. The sample statistics are summarized in table 5.10.

The acceptance rate decreases further from $t_n = 1\,400$, and we see that the $\pi(\nu|x_1, \ldots, x_n)$ and $\pi(\sigma|x_1, \ldots, x_n)$ in figure 5.1.15 become more jagged. The peak in $\pi(\tau|x_1, \ldots, x_n)$ which arose at $t_n = 1\ 700$ in both cases where σ was known is evident again. The run length was set to $4 \cdot 10^6$ in the case where $t_n = 1\ 700$ as well. This is done to wash out the autocorrelation effects and ensure that the chain is moving across the entire domain. With five chains with over-dispersed starting points, the Rubin-Gelman statistic was calculated, again $\widehat{R}^3 = 1$. Thus, with the long run length, it was believed that the Markov chain had converged to the posterior distribution, and was able to traverse the entire range of the parameters.

As before, the 2.5th percentile and the 97.5th percentile, as well as the 95 % HPD region narrow compared to $t_n = 1\ 400$, as is natural when we obtain more data. Note that also the interval estimates for σ narrow. The estimates for τ are especially notable. Note that both the median and mean are closer to the true value than the estimates for τ when σ was known, see tables 5.6 and 5.2. It is also seen as the peak in figure 5.1.15 being higher than the peak of 5.1.9 and 5.1.3. It may be that when τ, ν and σ are all unknown, there are a larger number of proposals which will lead to a small Metropolis-Hastings ratio, thus a smaller probability of acceptance.

Observation until t_n = 3 000

Finally, we consider the case when we have an abundance of measurements after the change point, $t_n = 3\,000$. The marginal priors of τ, ν and σ are shown in 5.1.16 along with the marginal posterior distributions $\pi(\tau|x_1, \ldots, x_n), \pi(\nu|x_1, \ldots, x_n)$ and $\pi(\sigma|x_1, \ldots, x_n)$.

Further diagnostics plots are given appendix A.1.3: the marginal trace plots of τ, ν and σ, $\pi(\tau), \pi(\nu)$ and $\pi(\sigma)$, are shown in figure A.1.30, the histogram of the sample values of each parameter is shown in figure A.1.31, the ACF as function of the lag for each of the three parameter iterations is shown in figure A.1.33, and finally, three pairwise contour plots of the pairwise joint posterior distributions, $\pi(\tau, \nu|x_1, \ldots, x_n), \pi(\nu, \sigma|x_1, \ldots, x_n)$ and $\pi(\tau, \sigma|x_1, \ldots, x_n)$ are shown in figure A.1.32. The sample statistics are summarized in table 5.11. The severity of the low acceptance rate which manifests in the trace plot is now especially evident in the plot of the ACF. However, the Rubin-Gelman statistic with over-dispersed starting points for 5 chains of run length $4 \cdot 10^6$ and burn-in of 10^4 gives $\widehat{R}^3 = 1$. Considering trace plots for longer runs of time does not show sequential

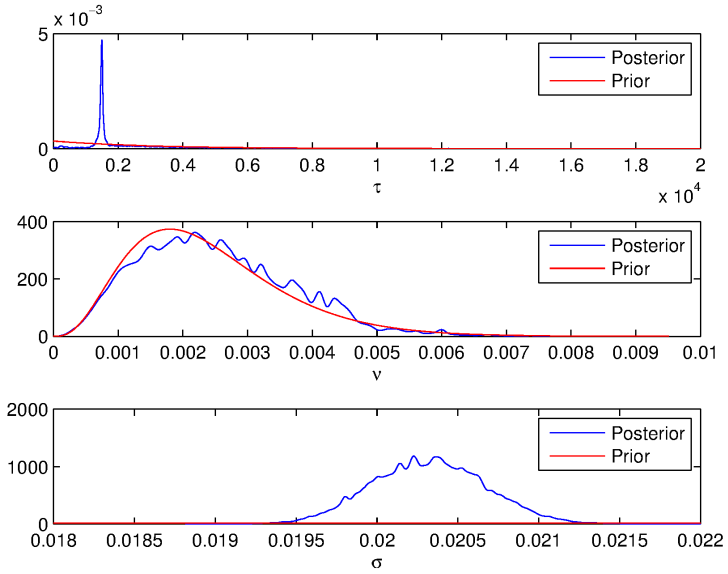

Figure 5.1.15: Prior distributions of τ, $\pi(\tau)$, ν, $\pi(\nu)$ and σ, $\pi(\sigma)$ and posterior distribution of τ, $\pi(\tau|x_1,\ldots,x_n;\nu,\sigma)$, ν, $\pi(\nu|x_1,\ldots,x_n;\nu,\sigma)$ and σ, $\pi(\sigma|x_1,\ldots,x_n;\nu,\sigma)$ given the observed temperature $W(t)$ up until $t_n = 1\ 700$. Here, τ, ν and σ were considered unknown.

Table 5.10: Statistics of the posterior distribution of τ, ν and σ, $\pi(\tau,\nu,\sigma|x_1,\ldots,x_n)$, as estimated by MCMC methods. Here, τ, ν and σ were considered unknown, and the endpoint of observation was t_n =1 700.

	τ	
Point estimates	Mean	2 729
	Median	1 531
Interval estimates	2.5 and 97.5 percentile	[344, 10 002]
	95 % HPD	[1, 7 929]
	ν	
Point estimates	Mean	0.0025
	Median	0.0024
Interval estimates	2.5 and 97.5 percentile	[0.0007, 0.0050]
	95 % HPD	[0.0005, 0.0046]
	σ	
Point estimates	Mean	0.0020
	Median	0.0020
Interval estimates	2.5 and 97.5 percentile	[0.0197, 0.00210]
	95 % HPD	[0.0196, 0.0210]

or unstationary behavior. Thus it is believed that the run length is sufficient for the Markov chain to traverse the range of the three parameters τ, ν and σ, as well as to wash out the autocorrelation effects.

The sample statistics shown in table 5.11 are similar to the comparable statistics in table 5.7, except the interval estimates for τ are wider, as one expects when another unknown parameter is introduced in the model. The statistics for ν are equal to the shown decimals. It can be noted that the sample standard deviation of the temperature increments X_i from time $\tau = 1\ 500$ to $t_n = 3\ 000$ is 0.0209, even higher than that for the temperature increments X_i from time 0 to $\tau = 1\ 500$. This, as well as the uncertainty in ν, mainly, but also in τ, is thought to be the cause of $\pi(\sigma|x_1, \ldots, x_n)$ having the majority of mass located at $\sigma > 0.02$, the true value.

It may be that the prior for σ, which was chosen to be wide and more objective, resulted in a lower acceptance rate than desirable, and that a more subjective prior for σ would have been more appropriate.

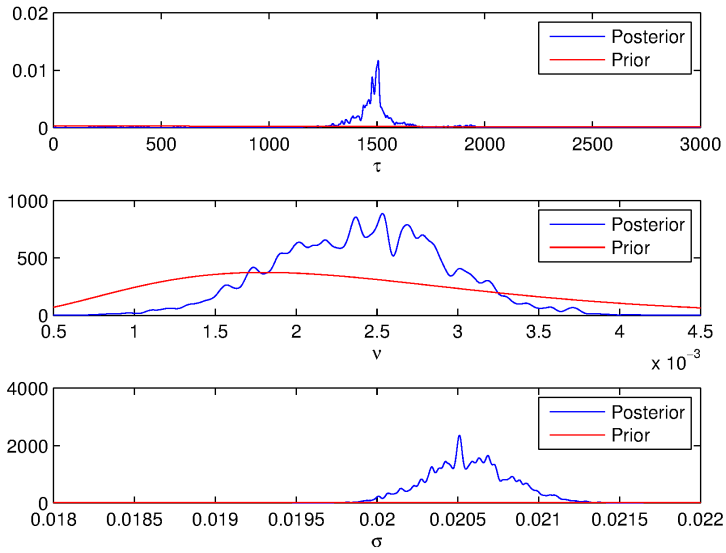

Figure 5.1.16: Prior distributions of τ, $\pi(\tau)$, ν, $\pi(\nu)$ and σ, $\pi(\sigma)$ and posterior distribution of τ, $\pi(\tau|x_1,\ldots,x_n;\nu,\sigma)$, ν, $\pi(\nu|x_1,\ldots,x_n;\nu,\sigma)$ and σ, $\pi(\sigma|x_1,\ldots,x_n;\nu,\sigma)$ given the observed temperature $W(t)$ up until $t_n = 3\ 000$. Here, τ, ν and σ were considered unknown.

Table 5.11: Statistics of the posterior distribution of τ, ν and σ, $\pi(\tau,\nu,\sigma|x_1,\ldots,x_n)$, as estimated by MCMC methods. Here, τ, ν and σ were considered unknown, and the endpoint of observation was $t_n = 3\ 000$.

	τ	
Point estimates	Mean	1 463
	Median	1 487
Interval estimates	2.5 and 97.5 percentile	[844, 1 885]
	95 % HPD	[1 206, 1 957]
	ν	
Point estimates	Mean	0.0024
	Median	0.0024
Interval estimates	2.5 and 97.5 percentile	[0.0014, 0.0034]
	95 % HPD	[0.0014, 0.0034]
	σ	
Point estimates	Mean	0.0021
	Median	0.0021
Interval estimates	2.5 and 97.5 percentile	[0.020, 0.00211]
	95 % HPD	[0.020, 0.0211]

Hitting time CDF estimates

Finally, the resulting hitting time CDF estimates of the temperature level $a = 14$ for each of the three t_n are shown. The formulaic approach estimates, $\widehat{P}_f(T \leq t)$, is shown in figure 5.1.17 with the 95 % credible interval, and the simulation approach estimates $\widehat{P}_S(T \leq t)$ with the 95 % credible interval is shown in figure 5.1.18. Note that the end point value of the time axis is different in the two figures. The comparison of the two approaches and the 95 % credible interval in shown in figure 5.1.19. Some percentiles of the hitting time CDF estimates and the credible intervals are shown in table 5.12, with the true distribution. The estimate $\widehat{P}_f(T \leq t)$ was calculated with $m = 5 \cdot 10^5$ draws from the posterior distribution of τ for each value of t. In the simulation approach, for each realization of τ^*, ν^*, σ^* from the posterior distribution, 100 simulations of Wiener processes were run to estimate $P(T \leq t|\tau^*, \nu^*, \sigma^*)$. In total, there were $m = 10^4$ realizations τ^*, ν^*, σ^* drawn from $\pi(\tau, \nu, \sigma|x_1, \ldots, x_n)$. The true distribution is the same as with ν and σ known, the inverse Gaussian distribution with true drift parameter ν, true variance parameter σ and threshold parameter $a - W(3000)$, in other words, the distribution given the maximum amount of data in our case.

For the formulaic solution, $\widehat{P}_f(T \leq t)$, the upper limit is so to speak unchanged from the situation with σ known, and τ and ν unknown, see table 5.8. The estimate itself is slightly lower, but the lower credible interval at $t_n = 1\ 700$ is slightly higher. This is thought to originate from $\pi(\sigma|x_1, \ldots, x_n)$ which is centered at a higher value than the true parameter 0.02, as well as the fact that the sample of τ was narrower when σ was unknown. The higher values of σ will lead to a less steep hitting time CDF estimate, and the narrow distribution of τ may lead to less extreme tail behavior, thus increasing the lower credible interval at $t_n = 1\ 700$. At $t_n = 3\ 000$, the estimates are very similar to those in table 5.8, where τ and ν was unknown, but σ was known. The somewhat reduced steepness of $P(T \leq t)$ is attributed to the larger values of σ than for $t_n = 1\ 700$.

For the simulation solution as well, $\widehat{P}_S(T \leq t)$, the upper credible limits are so to say unchanged, the estimate itself $\widehat{P}_S(T \leq t)$, is a little less steep, and the lower credible interval at $t_n = 1\ 700$ is also higher than compared to those in table 5.8, where τ and ν was unknown, but σ was known. It is believed to be caused by the narrower posterior distribution of τ and the posterior distribution of σ being shifted to the right of the true value, as for the formulaic approach.

The comparison in figure 5.1.19 shows, as for the situation where τ and ν was unknown, but σ was known, that the simulation approach credible intervals are narrower than the formulaic approach. This is again thought to stem from the computational limitations by drawing fewer values of (τ, ν, σ), and it is again assumed that if a larger number of values were drawn, the intervals of the formulaic and simulation approach would be closer.

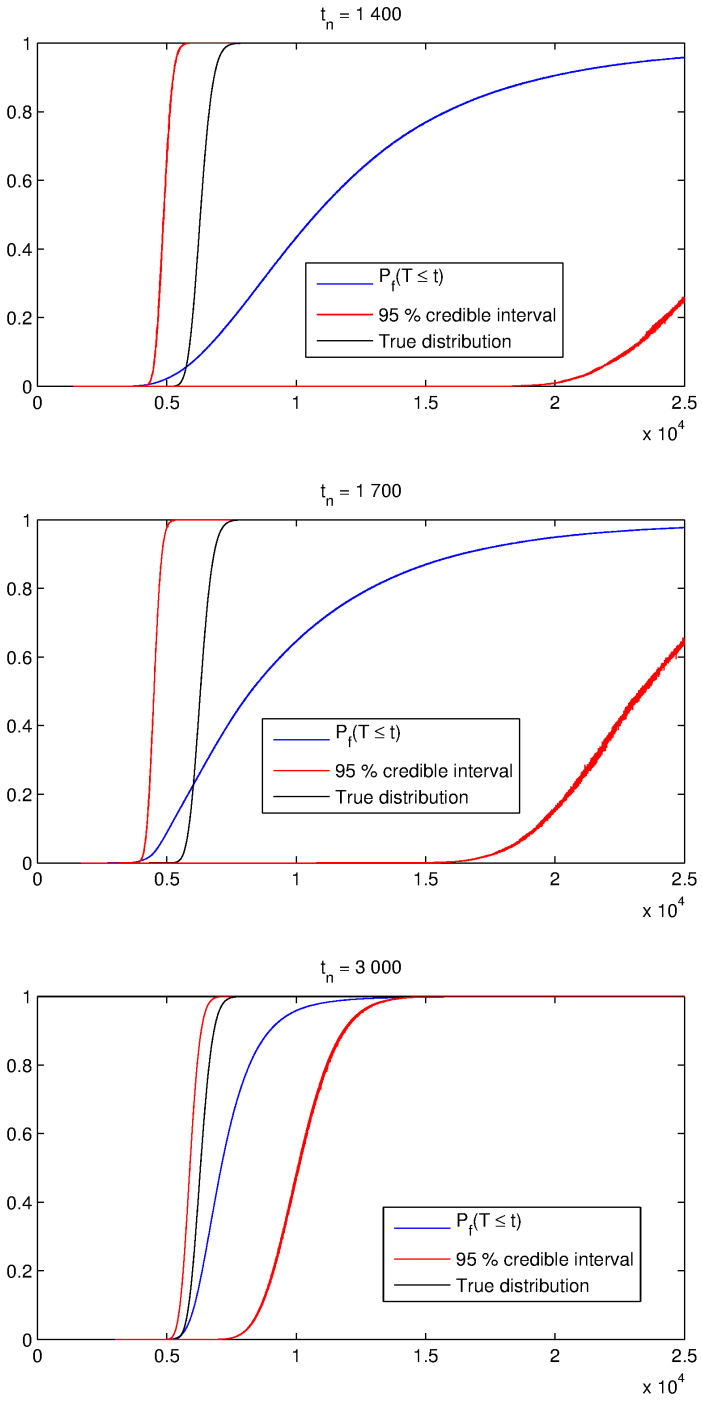

Figure 5.1.17: Formulaic approach to estimating CDF, $\widehat{P}_f(T \leq t)$ for different observation lengths, t_n, shown with 95 % credible intervals and the true distribution. The x-axis is time t while the y-axis is the probability. Top panel: $t_n = 1\ 400$. Middle panel: $t_n = 1\ 700$. Bottom panel: $t_n = 3\ 000$. Here, τ, ν and σ are unknown parameters.

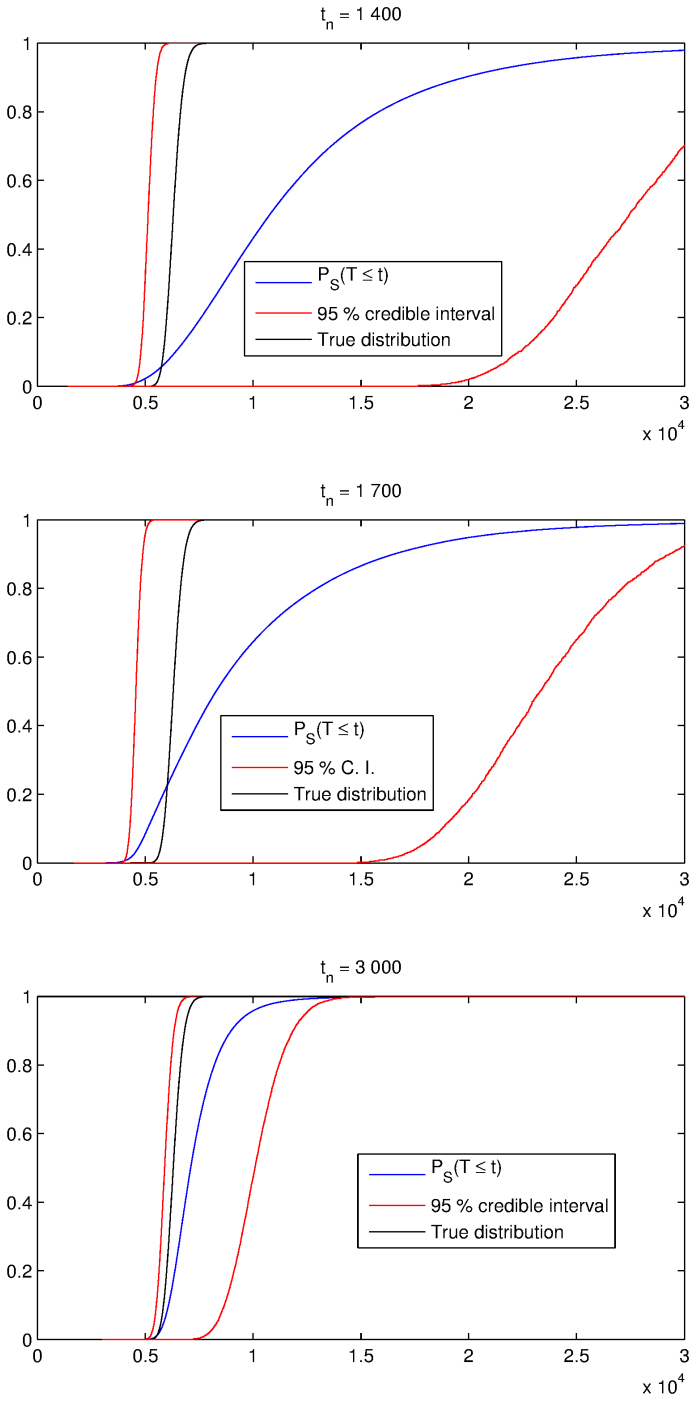

Figure 5.1.18: Simulation approach to estimating CDF, $\widehat{P}_S(T \leq t)$ for different observation lengths, t_n, shown with 95 % credible intervals and the true distribution. The x-axis is time t while the y-axis is the probability. Top panel: $t_n = 1\ 400$. Middle panel: $t_n = 1\ 700$. Bottom panel: $t_n = 3\ 000$.Here, τ, ν and σ are unknown parameters.

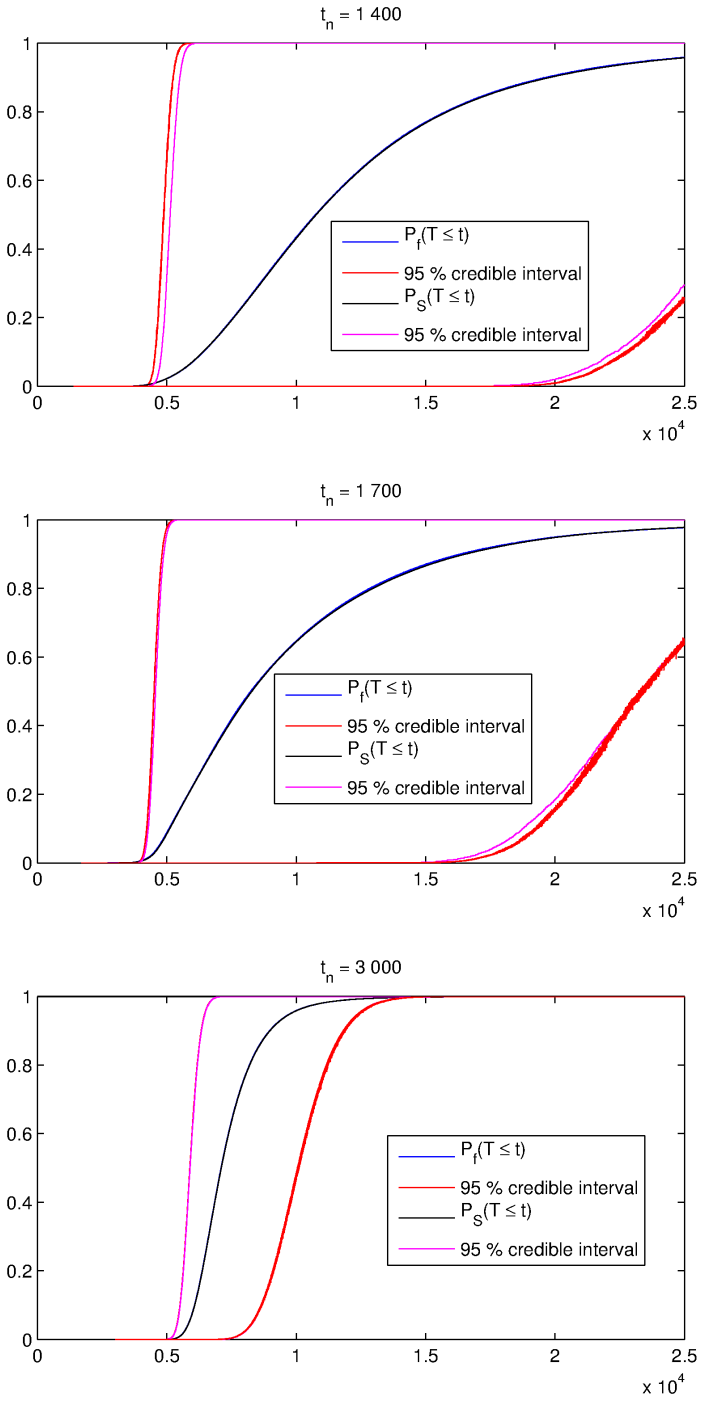

Figure 5.1.19: Comparison of the simulation and formulaic approaches, $\widehat{P}_S(T \leq t)$ and $\widehat{P}_f(T \leq t)$ with τ, ν and σ as unknown parameters, for three different observation lengths t_n. Top panel: $t_n = 1\ 400$. Middle panel: $t_n = 1\ 700$. Bottom panel: $t_n = 3\ 000$. The x-axis is time t while the y-axis is the probability.

Table 5.12: Percentiles of the estimated CDF, τ, ν and σunknown.

Percentile	2.5	5	10	25	50	75	90	95	97.5
Formulaic solution									
$t_n = 1\ 400$									
Upper 95 %	4 401	4 473	4 549	4 692	4 872	5 064	5 232	5 346	5 428
Estimated CDF	5 064	5 635	6 395	8 093	10 743	14 570	19 629	23 771	$> 25\ 000$
Lower 2.5 %	20 776	21 555	22 683	24 882	$> 25\ 000$				
$t_n = 1\ 700$									
Upper 95 %	4 083	4 144	4 217	4 343	4 505	4 663	4 813	4 906	4 989
Estimated CDF	4 441	4 701	5 084	6 189	8 246	11 733	16 342	20 022	24 287
Lower 95 %	17 578	18 312	19 225	20 946	23 235	$> 25\ 000$			
$t_n = 3\ 000$									
Upper 95 %	5 325	5 407	5 506	5 680	5888	6 112	6 329	6 465	6 588
Estimated CDF	5 687	5 855	6 069	6 489	7 095	7 935	8 966	9 769	10 634
Lower 95 %	8 018	8 295	8 643	9 266	10 045	10 929	11 817	12 395	12 922
Simulation solution									
$t_n = 1\ 400$									
Upper 95 %	4 631	4 710	4 799	4 950	5 129	5 325	5 502	5 612	5 707
Estimated CDF	5 073	5 631	6 398	8 114	10 770	14 615	19 802	23 994	28 674
Lower 95 %	20 253	21 158	22 309	24 501	27 451	$>30\ 000$			
$t_n = 1\ 700$									
Upper 95 %	4 140	4 203	4 281	4 410	4 563	4 726	4 883	4 981	5 060
Estimated CDF	4 470	4 728	5 112	6 201	8 284	11 812	16 541	20 188	24 222
Lower 95 %	17 015	17 802	18 766	20 781	23 311	26 331	29 337	$> 30\ 000$	
$t_n = 3\ 000$									
Upper 95 %	5 330	5 419	5 514	5 689	5 896	6 123	6 340	6 476	6 601
Estimated CDF	5 692	5 861	6 076	6 497	7 103	7 945	8 983	9 787	10 655
Lower 95 %	8 044	8 388	8 655	9 307	10 083	10 926	11 863	12 453	12 919
True distribution									
$t_n = 3\ 000$	5 631	5 727	5 844	6 050	6 297	6 564	6 823	6 986	7 133

5.2 Two change points

In this section, we will consider one particular temperature process, which is a piecewise Wiener process, with two change points, τ_1 and τ_2 simulated in the interval $[0, t_n] = [0, 2000]$. The process is shown in figure 5.2.1. The process was simulated with the following parameters:

$$\nu_1 = 0.003 \qquad \nu_2 = 0.0039$$
$$\sigma_0 = \sigma_1 = 0.05 \qquad \sigma_2 = 0.065$$
$$\tau_1 = 800 \qquad \tau_2 = 1400$$

The resolution in the simulation was $t_n/n = \Delta = 1$. For this process, we will consider the threshold temperature $a = [W(t_n)] + 10 = 13$. Thus, the formulaic approach described in section 4.3.2 is applicable with $\alpha = 1.3$. Therefore, we denote ν_1 by ν and ν_2 by $\alpha\nu$.

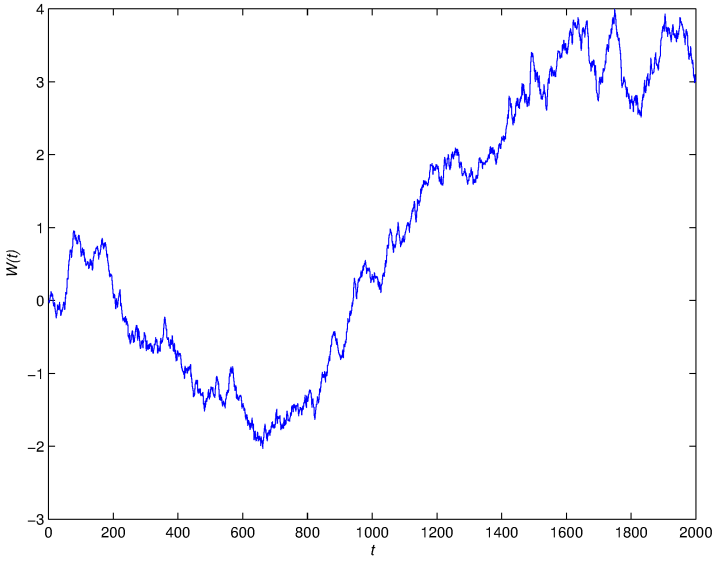

Figure 5.2.1: Piecewise Wiener process with two change points $\tau_1 = 800$ and $\tau_2 = 1400$ simulated in the time interval $[0, t_n] = [0, 2000]$. The initial parameters, in the time interval $[0, \tau_1)$, of the process were $\nu_0 = 0$ and $\sigma_0 = 0.05$. In the time interval $[\tau_1, \tau_2)$, the parameters were $\nu_1 = 0.03$ and $\sigma_1 = 0.05$. The parameters of the last piecewise process, in the time interval $[\tau_2, t_n)$, were $\nu_2 = 0.0039$ and $\sigma_2 = 0.065$. For this process, we will consider the hitting time of the threshold temperature $a = 13$.

Likewise, we write $\sigma = \sigma_1 = \sigma_0$ and $\alpha\sigma = \sigma_2$.

We will estimate the hitting time CDF by the formulaic and simulation approach both in the case where ν, σ and α is known, and in the case there ν is unknown in

addition to the change points. For each case, four observations intervals of different lengths, that is $\{W(t_i), i = 1, \ldots, n\}$, for four different t_n. The different values are $t_n = 700$, $t_n = 1\ 100$, $t_n = 1\ 450$, $t_n = 2\ 000$. For $t_n = 700$, none of the change points have occurred. For $t_n = 1\ 100$, the first change point has occurred, but not the second change point. For the third $t_n = 1\ 450$, the second change point has just occurred, and the last $t_n = 2\ 000$, we have an abundance of measurements after the second change point.

5.2.1 τ_1, τ_2 unknown, ν, σ, α known

In this section, we will first consider the MCMC diagnostics and the posterior distributions of the parameters will be discussed for all t_n, before the hitting time CDF estimates are shown and discussed.

The prior put on τ_1 was geometric, $\tau_1 \sim \mathcal{G}eom(1/2000)$. The prior put on the time spent in the second state, $\rho_2 = \tau_2 - \tau_1$ was also geometric, $\rho_2 \sim \mathcal{G}eom(1/1000)$. Note that this is because we expect the failure development process to spend more time in state 1 than state 2.

Observation until $t_n = 700$

The prior distributions of τ_1 and $\rho_2 = \tau_2 - \tau_1$, $\pi(\tau_1)$ and $\pi(\tau_2 - \tau_1)$ are shown with the respective posterior distributions $\pi(\tau_1|x_1, \ldots, x_n)$ and $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n)$ in figure 5.2.2. The figure also shows the posterior distribution of $\tau_2, \pi(\tau_2|x_1, \ldots, x_n)$. The mean, median, 2.5th and 97.5th percentiles as well as the HPD region are given in table 5.13.

Further diagnostics plots are shown in appendix A.2.1: the marginal trace plots of τ_1 and τ_2 is given in figure A.2.1, the histograms of $\tau_1, \tau_2 - \tau_1$ and τ_2 are given in figure A.2.2, the contour plot of the joint posterior density $\pi(\tau_1, \tau_2|x_1, \ldots, x_n)$ is given in figure A.2.3 and finally, the ACF of the Markov chain iterates are shown in the figure A.2.4. The plots all point to good mixing and fast exploration of the range of τ_1 and τ_2, with the trace plot moving vigorously, the ACF decaying rapidly and smooth histograms and posterior distributions. The acceptance rate was 0.77. The Rubin-Gelman statistic was calculated for five chains with run length $2 \cdot 10^6$ and burn-in 10^4 to $\widehat{R}^2 = 1$. Thus, all signs point to the Markov chain has converged to the posterior distribution, and is sampling from the entire distribution.

Note that the interval estimates in table 5.13 are very wide. The posterior distribution of $\tau_1, \pi(\tau_1|x_1, \ldots, x_n)$ shows that there is little belief that τ_1 has already occurred, as the distribution is peaking after $t_n = 700$. The posterior distribution of $\tau_2, \pi(\tau_2|x_1, \ldots, x_n)$, shows that it is not expected that τ_2 has occurred yet at time $t_n = 700$.

Observation until $t_n = 1\ 100$

The prior distributions of τ_1 and $\rho_2 = \tau_2 - \tau_1$, $\pi(\tau_1)$ and $\pi(\tau_2 - \tau_1)$ are shown with the respective posterior distributions $\pi(\tau_1|x_1, \ldots, x_n)$ and $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n)$ in figure 5.2.3 together with the posterior distribution of $\tau_2, \pi(\tau_2|x_1, \ldots, x_n)$. The mean, median,

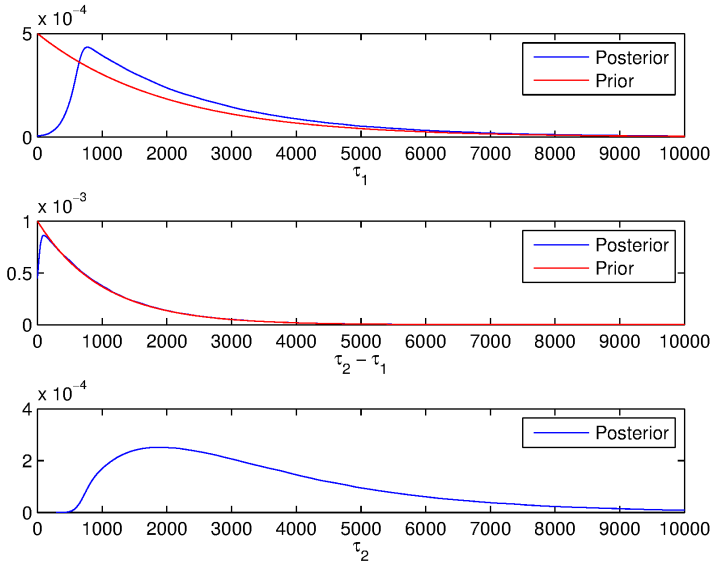

Figure 5.2.2: Prior distribution of τ_1, $\pi(\tau_1)$ and $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1)$, and posterior distributions of τ_1,$\pi(\tau_1|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n; \alpha, \nu, \sigma)$ and τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$ given the observed temperature $W(t)$ up until $t_n = 700$.

Table 5.13: Statistics of the posterior distribution of τ_1, and τ_2 $\pi(\tau_1, \tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, as estimated by MCMC methods. Here, α, ν and σ were considered known, and the endpoint of observation was $t_n = 700$.

	τ_1	
Point estimates	Mean	2 520
	Median	1 910
Interval estimates	2.5 and 97.5 percentile	[512, 7 912]
	95 % HPD	[223, 6 604]
	τ_2	
Point estimates	Mean	3 532
	Median	2 983
Interval estimates	2.5 and 97.5 percentile	[908, 9 280]
	95 % HPD	[689, 7 893]

2.5th and 97.5th percentiles as well as the highest posterior density region are given in table 5.14.

Further diagnostics plots are shown in appendix A.2.1: the marginal trace plots of τ_1 and τ_2 is given in figure A.2.5, the histograms of $\tau_1, \tau_2 - \tau_1$ and τ_2 are given in figure A.2.6, the contour plot of the joint posterior density $\pi(\tau_1, \tau_2 | x_1, \ldots, x_n)$ is given in figure A.2.7 and finally, the ACF of the Markov chain iterations as function of the lag are shown in the figure A.2.8. The trace plots are still showing rigorous exploration of the range of the posterior distribution $\pi(\tau_1 | x_1, \ldots, x_n), \pi(\tau_2 | x_1, \ldots, x_n)$. The ACF is decaying at a slower rate, but still with reasonable speed. The histograms are all smooth. The acceptance rate was 0.463. The Rubin-Gelman statistic was calculated for five chains with run length $2 \cdot 10^6$ and burn-in 10^4 to $\widehat{R}^2 = 1$. It is again assumed that the Markov chain has converged to the posterior distribution, and thus that the sample is representative of the posterior distribution.

The posterior distribution of $\tau_1, \pi(\tau_1 | x_1, \ldots, x_n)$ now has a marked peak around $\tau_1 = 800$, pushing the posterior distribution of $\tau_2, \pi(\tau_2 | x_1, \ldots, x_n)$ further to the right, as it is seen in all statistics for τ_2 shown in table 5.14. There is a drop in the mean and median estimates of τ_1, and the interval estimates narrow compared to table 5.13.

Observation until $t_n = 1\ 450$

The prior distributions of τ_1 and $\rho_2 = \tau_2 - \tau_1$, $\pi(\tau_1)$ and $\pi(\tau_2 - \tau_1)$ are shown with the respective posterior distributions $\pi(\tau_1 | x_1, \ldots, x_n)$ and $\pi(\tau_2 - \tau_1 | x_1, \ldots, x_n)$ in figure 5.2.4 together with the posterior distribution of $\tau_2, \pi(\tau_2 | x_1, \ldots, x_n)$. The mean, median, 2.5th and 97.5th percentiles as well as the highest posterior density region are given in table 5.15.

Further diagnostics plots are shown in appendix A.2.1: the marginal trace plots of τ_1 and τ_2 is given in figure A.2.9, the histograms of $\tau_1, \tau_2 - \tau_1$ and τ_2 are given in figure A.2.10, the contour plot of the joint posterior density $\pi(\tau_1, \tau_2 | x_1, \ldots, x_n)$ is shown in figure A.2.11 and finally, the ACF of the Markov chain iterations are shown in the figure A.2.12. It is evident that the acceptance rate has decreased from the trace plot, where there are longer stretches where the Markov chain is stuck at one value. The acceptance rate was 0.127. However, it is evident that these values are from the region of higher posterior density, and the chain is still moving rigorously between stretches of standing still. This causes the ACF to decay at a slower rate. It is especially the ACF for τ_2 which is decaying slowly. The Rubin-Gelman statistic was calculated for five chains with run length $2 \cdot 10^6$ and burn-in 10^4 to $\widehat{R}^2 = 1$.

Looking at figure 5.2.4, it is evident that $\pi(\tau_2 | x_1, \ldots, x_n)$ has a narrow peak near the true τ_2, even though t_n is very close to the true value of $\tau_2 = 1\ 400$. However, $\pi(\tau_2 | x_1, \ldots, x_n)$ still has some density mass for higher values of τ_2, which is evident also in the difference between the mean and median of τ_2, shown in table 5.15, and the interval estimates. It may seem like the chain gets stuck around $\tau_2 = 1450$. Remember that the our prior information was about ρ_2, the time spent in state 2, $\rho_2 = \tau_2 - \tau_1$, and note that we are sampling independently from the distributions of τ_1 and ρ_2. From figure 5.2.4, we can see that the spread of ρ_2 is as wide or even wider than the spread

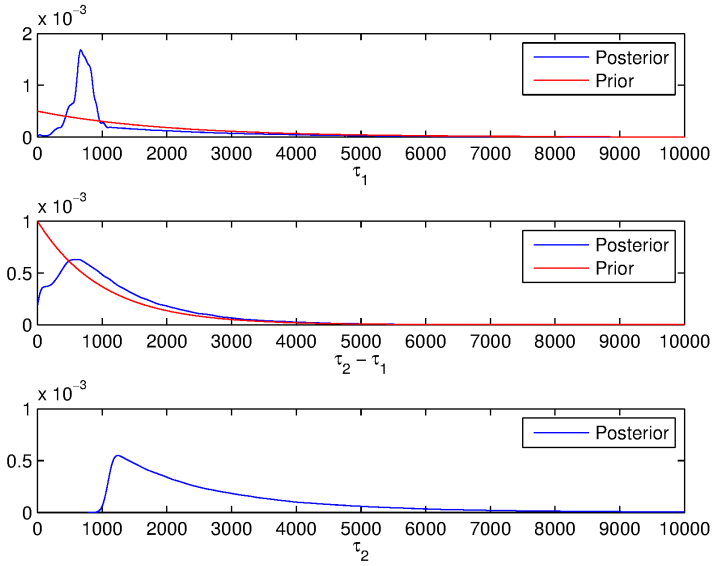

Figure 5.2.3: Prior distribution of τ_1, $\pi(\tau_1)$ and $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1)$, and posterior distributions of τ_1,$\pi(\tau_1|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n; \alpha, \nu, \sigma)$ and τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$ given the observed temperature $W(t)$ up until $t_n = 1\ 100$.

Table 5.14: Statistics of the posterior distribution of τ_1, and τ_2 $\pi(\tau_1, \tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, as estimated by MCMC methods. Here, α, ν and σ were considered known, and the endpoint of observation was $t_n = 1\ 100$.

τ_1		
Point estimates	Mean	1 609
	Median	834
Interval estimates	2.5 and 97.5 percentile	[339, 6 556]
	95 % HPD	[1, 5 159]
τ_2		
Point estimates	Mean	2 853
	Median	2 252
Interval estimates	2.5 and 97.5 percentile	[1 133, 7 963]
	95 % HPD	[1 069, 6 601]

in τ_1. Thus, it seems that the Markov chain is only accepting τ_2^* proposals close to the true change point because the likelihood is much larger. This may stem from the fact that the variance parameter changes, which is does not for τ_1 and did not in the setting with only one change point, τ. It is found that in situations with $\alpha = 2$, for example, the acceptance rate decreases strikingly, and in situations where $\alpha = 1$, the acceptance rate increases equally strikingly, resulting in a posterior density of τ_2, $\pi(\tau_2|x_1, \ldots, x_n)$ which is as wide as the posterior density of τ_1, $\pi(\tau_1|x_1, \ldots, x_n)$. Thus, it is assumed that the increasing ACF and decreasing acceptance rate is due to the parameters of the problem, and not to a fault in the implementation or weakness in the algorithm of the MCMC methods. Hence, it is assumed that the Markov chain has converged to the joint posterior distribution $\pi(\tau_1, \tau_2|x_1, \ldots, x_n)$.

Note also that the peak in $\pi(\tau_1|x_1, \ldots, x_n)$ is covering a larger range of τ_1 values, and that the interval estimates are significantly reduced. Note that $\tau_1 < \tau_2$, such that a sharp definition of τ_2 gives an upper limit of τ_1.

Observation until t_n = 2 000

The prior distributions of τ_1 and $\rho_2 = \tau_2 - \tau_1$, $\pi(\tau_1)$ and $\pi(\tau_2 - \tau_1)$ are shown with the respective posterior distributions $\pi(\tau_1|x_1, \ldots, x_n)$ and $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n)$ in figure 5.2.5 together with the posterior distribution of $\tau_2, \pi(\tau_2|x_1, \ldots, x_n)$. The mean, median, 2.5th and 97.5th percentiles as well as the highest posterior density region are given in table 5.16.

Further diagnostics plots are shown in appendix A.2.1: the marginal trace plots of τ_1 and τ_2 is given in figure A.2.13, the histograms of $\tau_1, \tau_2 - \tau_1$ and τ_2 are given in figure A.2.14, the contour plot of the joint posterior density $\pi(\tau_1, \tau_2|x_1, \ldots, x_n)$ is shown in figure A.2.15 and finally, the ACF of the Markov chain iterates are shown in the figure A.2.16. It is clear that the bad signs at $t_n = 1\ 450$ have gotten worse. The trace plot shows that the acceptance rate has decreased further, which causes the ACF to decay even slower than at $t_n = 1\ 450$. The acceptance rate was 0.005. The histogram shows that the distribution of τ_2 is more spiked, but note that the scale is different. The Rubin-Gelman statistic was calculated for five chains with run length $2 \cdot 10^6$ and burn-in 10^4 to $\widehat{R}^2 = 1$. It is believed that the long run length of $2 \cdot 10^6$ will wash out the autocorrelation effects and allow the entire range of (τ_1, τ_2) to be traversed.

Note in table 5.16 that the interval estimates of τ_2 are very narrow. Note also that the estimated mean and median of τ_2 are very close to the true value and that the interval estimates contain the true value of $\tau_2 = 1\ 400$. Note that the interval estimates of τ_1 has narrowed, mainly from above. This is because $\tau_1 < \tau_2$, as for $t_n = 1\ 450$, thus the narrow distribution of τ_2 is pushing the distribution of τ_1 towards smaller τ_1. However, the median is not very different from the estimated median of τ_1 at $t_n = 1\ 450$. Thus, it seems like the it is only the larger values of τ_1 that is affected, and that the sampling from the smaller values of τ_1 is as for $t_n = 1\ 450$.

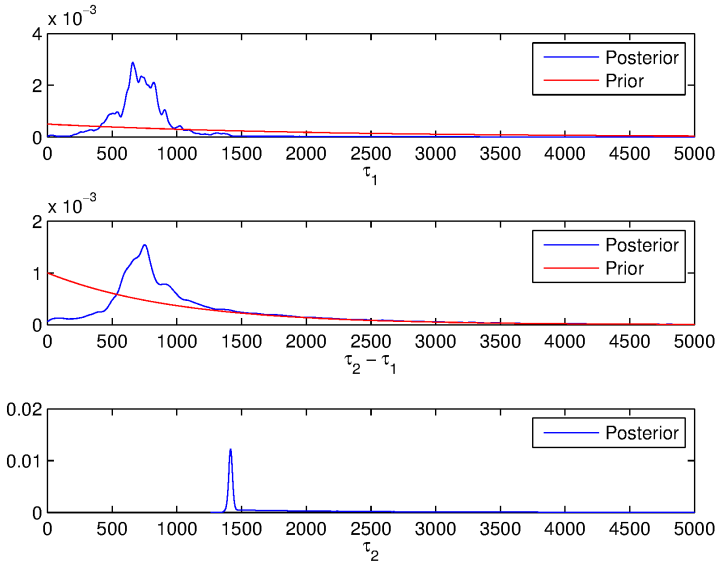

Figure 5.2.4: Prior distribution of τ_1, $\pi(\tau_1)$ and $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1)$, and posterior distributions of τ_1,$\pi(\tau_1|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n; \alpha, \nu, \sigma)$ and τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$ given the observed temperature $W(t)$ up until $t_n = 1\ 450$.

Table 5.15: Statistics of the posterior distribution of τ_1, and τ_2 $\pi(\tau_1, \tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, as estimated by MCMC methods. Here, α, ν and σ were considered known, and the endpoint of observation was $t_n = 1\ 450$.

τ_1		
Point estimates	Mean	884
	Median	725
Interval estimates	2.5 and 97.5 percentile	[297, 3 230]
	95 % HPD	[1, 1 849]
τ_2		
Point estimates	Mean	2 099
	Median	1 543
Interval estimates	2.5 and 97.5 percentile	[1 397, 5 386]
	95 % HPD	[1 356, 4 397]

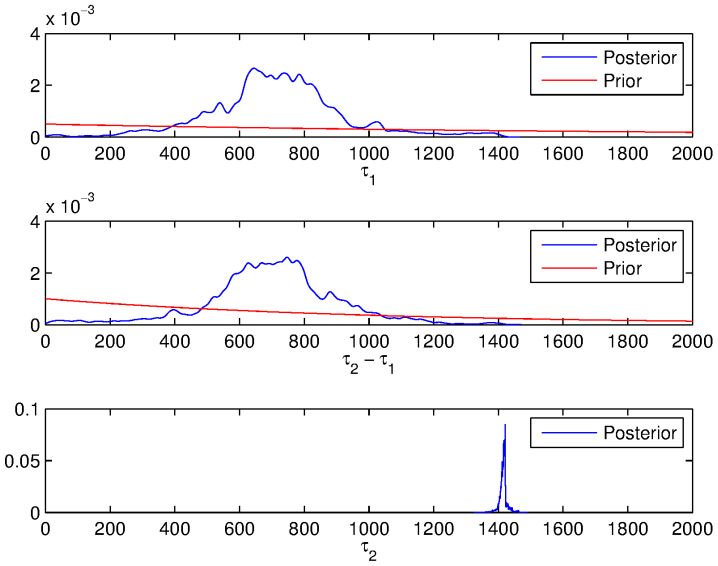

Figure 5.2.5: Prior distribution of τ_1, $\pi(\tau_1)$ and $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1)$, and posterior distributions of τ_1,$\pi(\tau_1|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n; \alpha, \nu, \sigma)$ and τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$ given the observed temperature $W(t)$ up until $t_n = 2\ 000$.

Table 5.16: Statistics of the posterior distribution of τ_1, and τ_2 $\pi(\tau_1, \tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, as estimated by MCMC methods. Here, α, ν and σ were considered known, and the endpoint of observation was $t_n = 2\ 000$.

	τ_1	
Point estimates	Mean	716
	Median	711
Interval estimates	2.5 and 97.5 percentile	[288, 1 225]
	95 % HPD	[244, 1 153]
	τ_2	
Point estimates	Mean	1 416
	Median	1 416
Interval estimates	2.5 and 97.5 percentile	[1 388, 1 443]
	95 % HPD	[1 388, 1 443]

Hitting time CDF estimates

In the formulaic approach, $\widehat{P}_f(T \leq t)$, there were $m = 500\ 000$ draws from the posterior distribution in the MC summation. In the simulation approach, $\widehat{P}_S(T \leq t)$, there were 10 000 draws from the posterior distribution and 100 simulations of Wiener processes for each draw from the posterior distribution. Note that the difference in number of draws from the posterior distribution is done because of computational limitations. The estimated hitting time CDF by the formulaic approach, $\widehat{P}_f(T \leq t)$, is shown in figure 5.2.6 along with 95 % credible intervals and the true distribution, and selected percentiles of $\widehat{P}_f(T \leq t)$ and the corresponding 95 % credible intervals are given in table 5.17. The estimated hitting time CDF by the simulation approach, $\widehat{P}_S(T \leq t)$, is shown in figure 5.2.7 with 95 % credible intervals and the true distribution. Table 5.18 displays selected percentiles of $\widehat{P}_S(T \leq t)$ and the corresponding credible intervals. The two approaches are compared in figure 5.2.8.

The true distribution shown in the figures is calculated as for one change: with the true parameters τ_1 and τ_2 and observations until $t_n = 2\ 000$. This means that the true hitting time distribution is inverse Gaussian with drift parameter $\alpha\nu$, variance parameter $\alpha\sigma$ and threshold parameter $a - W(2000)$. Note that as for one change point, the true distribution is only one possible realization of the development when $t_n < 2\ 000$, thus it only shows the deviance in this particular process.

The overall behavior of the two estimation methods is similar. At $t_n = 700$, before the first change point, the MCMC sampling gives the belief that the first change point has probably not occurred, and the second change point has definitely not occurred. Thus, we are left with the prior information of when in the future the change points will occur. Note that since $\lambda_1 = 1/2000$, prior to observing the process, the expectance of τ_1 is $E[\tau_1] = 2000$. This explains why the true distribution is partly outside the credible interval of the estimated hitting time CDFs.

At $t_n = 1\ 100$, the lower percentiles of the estimated distributions have become steeper. This is similar to the behavior in the case of one change point, at $t_n = 1$ 700, which was shortly after the (only) change point. The distribution has a heavy tail, which is thought to derive from the fact that there are no firm beliefs on when the second change point will occur. However, it is clear that the credible intervals have narrowed, and we make the remark that the true distribution is inside the credible interval, thus our beliefs about the future development is more similar to the actual future development for this particular process.

At $t_n = 1\ 450$, we see that both estimates $\widehat{P}_f(T \leq t)$ and $\widehat{P}_S(T \leq t)$ are above the true distribution. Observe that when standing in time $t_n = 1\ 450$, and believing that $\tau_2 < t_n$, the expectation of $W(2000), E[W(2000)] = W(t_n) + (2000 - t_n)\alpha\nu = 4.9$, but the realized temperature is $W(2000) = 3.0$. We notice, as is natural, that the credible intervals of the estimates have narrowed further.

As for the differences in the estimates, we see that at $t_n = 2\ 000$, there formulaic estimate $\widehat{P}_f(T \leq t)$ and the respective credible interval limits are equal. This is not true for $\widehat{P}_S(T \leq t)$ and its credible interval curves, but this is thought to be due to the

smaller number of draws from the posterior distribution. That is, it is believed that the curves would overlap if the number of draws from the posterior distribution and number of simulated piecewise Wiener processes for each draw were larger.

We notice that at $t_n = 2\ 000$, the formulaic estimate is actually steeper than the true distribution. This is difficult to explain by other factors than statistical variations, due to the fact that we have only one observed temperature process. One could suspect that the smaller variation in the distribution comes from the fact that we are using the expected value of $W(\tau_1)$ in cases where $\tau_1 > t_n$, but his will not affect the particular situation when $t_n = 2\ 000$, because we know from 5.16 that the number of draws where $\tau_2 > t_n$ is negligible, thus the effect this has on the variation is assumed to be very small. However, it may be that the calculation of the mean $P_f(T \leq t) = 1/m \sum_{i=1}^{m} P(T \leq t|\tau_i^*)$, when nearly all terms in the sum are equal, is sensitive to round-off error.

We also remark that the credible intervals of the simulation approach, $\widehat{P}_S(T \leq t)$ are not only broader, but have a gentler slope than the formulaic approach, $\widehat{P}_f(T \leq t)$, for $t_n = 700$ and $t_n = 1\ 100$. However, this becomes less distinct for $t_n = 1\ 450$ and $t_n = 2\ 000$. As the two latter values of t_n correspond to narrower posterior distributions, the gentler slope in the two former values of t_n is thought to be caused by the smaller number of draws from the posterior distribution, which disables the simulation approach from accounting for the same amount of variation as the formulation approach. Having said that, it may also be yet another manifestation of the steepness of the formulaic approach relative to the true distribution. Nevertheless, the approaches give estimates that quite close for all the different t_n and the estimates for t_n are close to the true distribution when both change points have occurred.

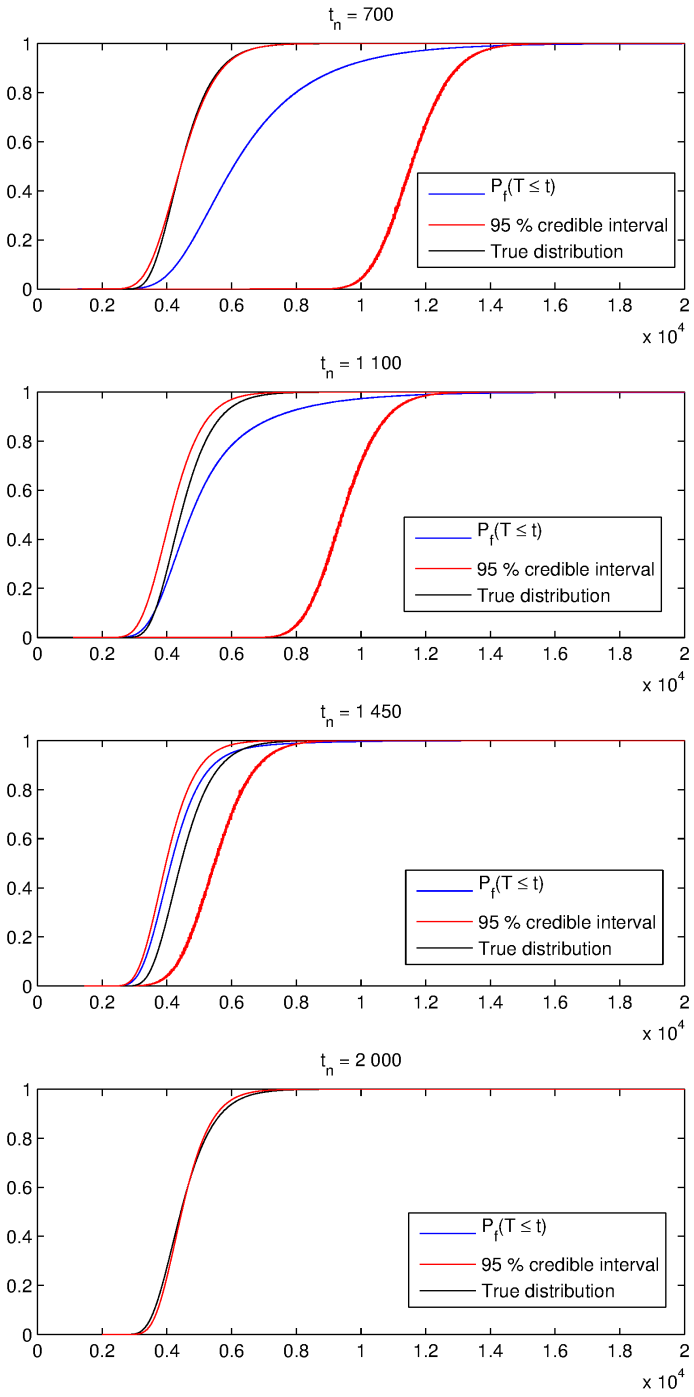

Figure 5.2.6: Formulaic approach to estimating CDF, $\widehat{P}_f(T \leq t)$ shown with 95 % credible intervals for different observation lengths, t_n. Shown with true distribution. Here, τ_1 and τ_2 are unknown parameters, ν, σ and α are known parameters. Top panel: $t_n = 700$. Second panel: $t_n = 1\ 100$. Third panel: $t_n = 1\ 450$. Bottom panel: $t_n = 2\ 000$. The x-axis is time and y-axis is probability.

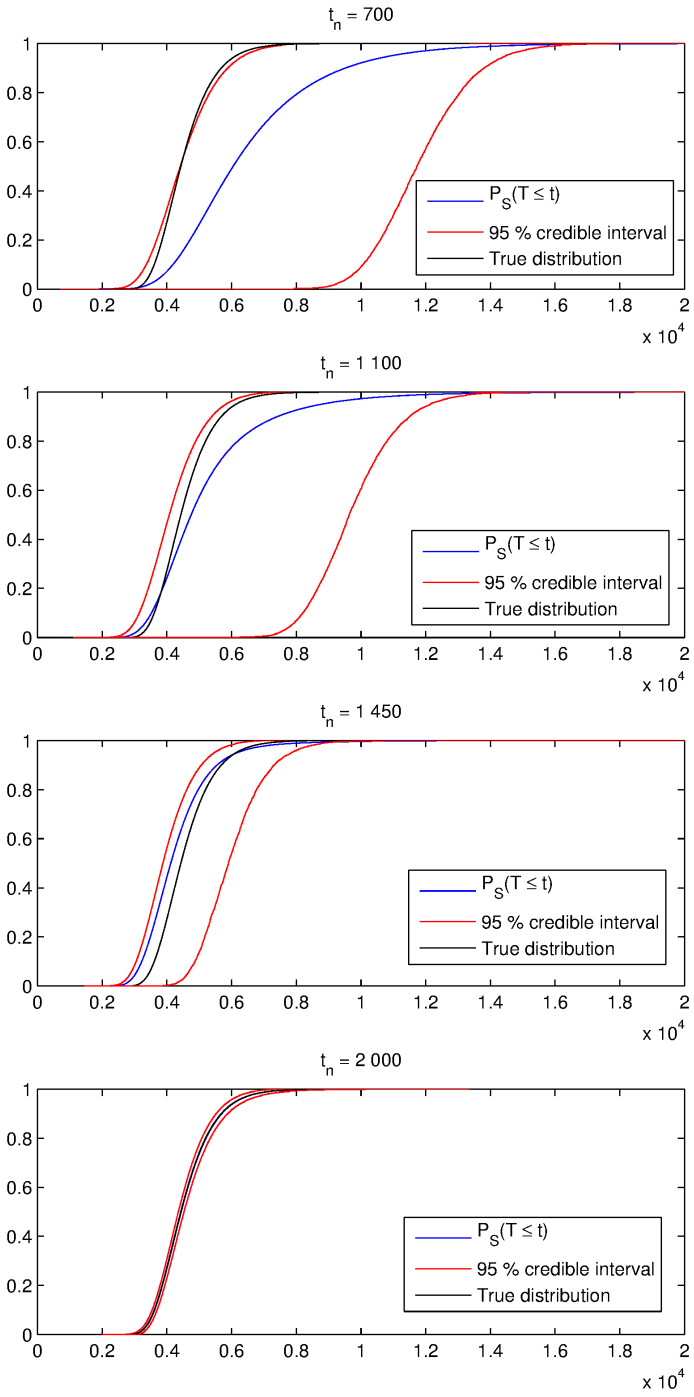

Figure 5.2.7: Simulation approach to estimating CDF, $\widehat{P}_S(T \leq t)$ shown with 95 % credible intervals for different observation lengths, t_n. Shown with true distribution. Here, τ_1 and τ_2 are unknown parameters, ν, σ and α are known parameters. Top panel: $t_n = 700$. Second panel: $t_n = 1\ 100$. Third panel: $t_n = 1\ 450$. Bottom panel: $t_n = 2\ 000$. The x-axis is time and y-axis is probability.

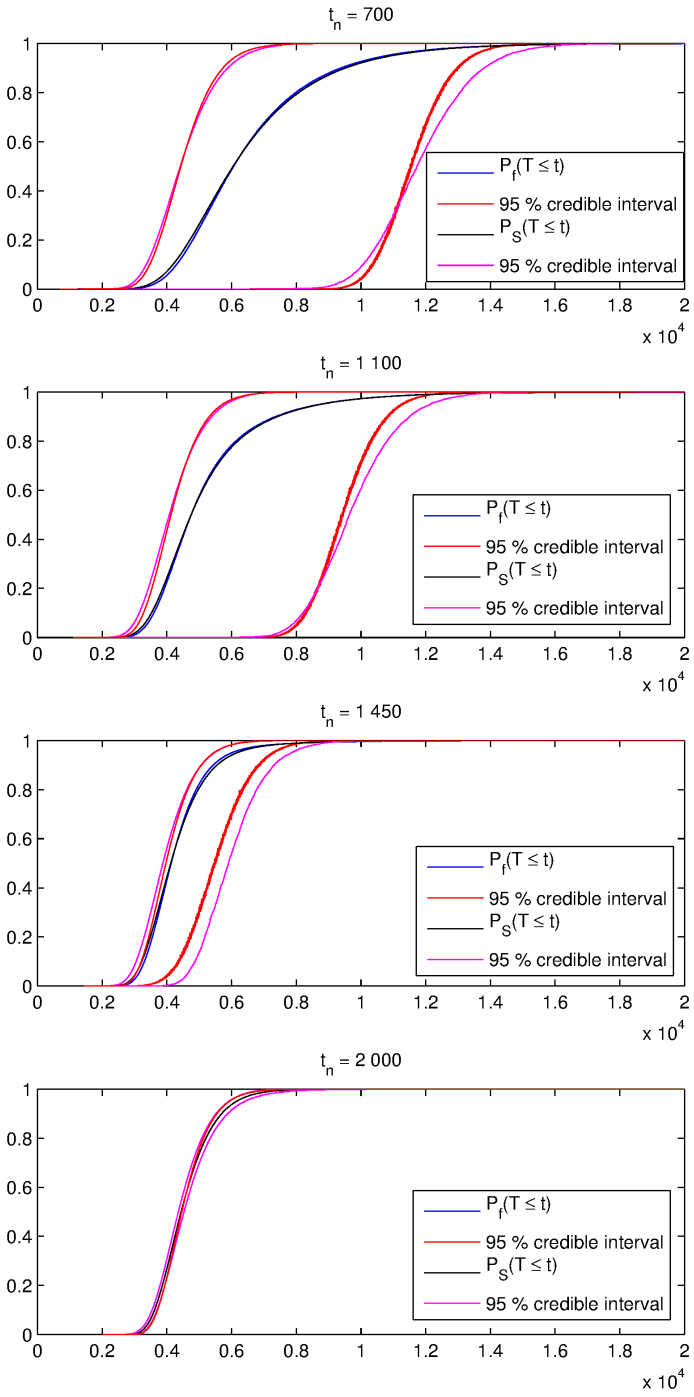

Figure 5.2.8: Comparison of the simulation and formulaic approaches, $\widehat{P}_S(T \leq t)$ and $\widehat{P}_f(T \leq t)$ shown with 95 % credible intervals with τ, ν and σ as known parameters, for three different observation lengths t_n. Here, τ_1 and τ_2 are unknown parameters, ν, σ and α are known parameters. Top panel: $t_n = 700$. Second panel: $t_n = 1\ 100$. Third panel: $t_n = 1\ 450$. Bottom panel: $t_n = 2\ 000$. The x-axis is time and y-axis is probability.

Table 5.17: Percentiles of the estimated CDF by the formulaic solution, $P_f(T \leq t)$ with ν, σ and α are known parameters.

Percentile	2.5	5	10	25	50	75	90	95	97.5
$t_n = 700$									
Upper 95 %	3 074	3 250	3 473	3890	4 434	5 071	5 733	6 174	6 586
Estimated CDF	3 682	3 950	4 302	5 015	6 059	7 531	9 375	10 762	12 149
Lower 2.5 %	9 789	10 049	10 330	10 850	11 519	12 239	12 948	13 433	13 870
$t_n = 1\ 100$									
Upper 95 %	2 935	3 086	3 277	3 644	4 127	4 705	5 315	5 725	6 110
Estimated CDF	3 180	3 363	3 599	4 067	4 753	5 782	7 368	8 719	10 116
Lower 2.5 %	7 746	7 959	8 278	8 791	9 395	10 095	10 808	11 238	11 680
$t_n = 1\ 450$									
Upper 95 %	2 911	3 042	3 210	3 534	3 970	4 499	5 066	5 451	5 814
Estimated CDF	2 985	3 125	3 306	3 657	4 133	4 732	5 433	6 007	6 721
Lower 2.5 %	3 811	4 038	4 339	4 857	5 474	6 138	6 820	7 256	7 649
$t_n = 2\ 000$									
Upper 95 %	3 419	3 547	3 712	4 032	4 462	4 985	5 547	5 928	6 289
Estimated CDF	3 419	3 547	3 712	4 032	4 462	4 985	5 547	5 928	6 289
Lower 2.5 %	3 419	3 547	3 712	4 032	4 462	4 985	5 547	5 928	6 289
True distribution									
$t_n = 2\ 000$	3 305	3 438	3 612	3 956	4 433	5 029	5 684	6 136	6 567

Table 5.18: Percentiles of the estimated CDF by the simulation solution, $P_S(T \leq t)$ with ν, σ and α are known parameters.

Percentile	2.5	5	10	25	50	75	90	95	97.5
$t_n = 700$									
Upper 95 %	2 945	3 138	3 370	3 818	4 419	5 136	5 891	6 387	6 820
Estimated CDF	3 549	3 827	4 192	4 939	6 050	7 605	9 519	10 934	12 338
Lower 2.5 %	9 319	9 679	10 075	10 804	11 733	12 786	13 811	14 478	15 117
$t_n = 1\ 100$									
Upper 95 %	2 775	2 946	3 149	3 547	4 080	4 721	5 395	5 834	6 212
Estimated CDF	3 084	3 276	3 524	4 021	4 752	5 830	7 420	8 768	10 175
Lower 2.5 %	7 553	7 857	8 206	8 858	9 637	10 563	11 508	12 148	12 723
$t_n = 1\ 450$									
Upper 95 %	2 723	2 872	3 055	3 405	3 880	4 458	5 072	5 475	5 826
Estimated CDF	2 885	3 033	3 225	3 602	4 121	4 779	5 554	6 166	6 876
Lower 2.5 %	4 471	4 645	4 869	5 305	5 906	6 601	7 330	7 859	8 348
$t_n = 2\ 000$									
Upper 95 %	3 220	3 368	3 548	3 894	4 360	4 926	5 528	5 926	6 262
Estimated CDF	3 311	3 445	3 619	3 964	4 443	5 039	5 695	6 148	6 582
Lower 2.5 %	3 405	3 534	3 700	4 043	4 532	5 166	5 887	6 416	6 935
True distribution									
$t_n = 2\ 000$	3 305	3 438	3 612	3 956	4 433	5 029	5 684	6 136	6 567

5.2.2 τ_1, τ_2 and ν unknown, σ, α known

The priors on τ_1 and $\rho_2 = \tau_2 - \tau_1$ are continued, both geometric, with parameters $\lambda_1 = 1/2000$ and $\lambda_2 = 1/1000$ respectively. Now the drift parameter ν is also considered unknown. We put a gamma prior on ν such that $\nu \sim \mathcal{G}am(4, 0.0006)$, as in the situation with one change point, where it was necessary to balance out an objective prior with the numerical difficulties that followed for large ν values.

This section will follow the structure of the previous. Firstly, the MCMC diagnostics of the Markov chains for the unknown parameters will be reviewed for all t_n, and next the resulting estimated hitting time CDF by the time transformed formulaic approach, $\widehat{P}_f(T \leq t)$ and the simulation approach $\widehat{P}_S(T \leq t)$ will be examined.

Observation until $t_n = 700$

Figure 5.2.9 shows the prior and posterior distributions for τ_1, $\rho_2 = \tau_2 - \tau_1$, and ν, $\pi(\tau_1)$, $\pi(\tau_1|x_1, \ldots, x_n)$, $\pi(\tau_2 - \tau_1)$, $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n)$, $\pi(\nu)$, $\pi(\nu|x_1, \ldots, x_n)$ as well as the posterior distribution of $\pi(\tau_2)$. Table 5.19 gives the estimates of the mean and median as well as the 2.5th and 97.5th percentile and 95 % HPD region for τ_1, τ_2 and ν.

Further diagnostic plots are shown in appendix A.2.2: the trace plot of τ_1, τ_2 and ν is shown in figure A.2.17, the histogram of the three variables and $\tau_2 - \tau_1$ is shown in figure A.2.18, figure A.2.19 shows the three pairwise contour plots of the joint posterior density of the three parameters, namely $\pi(\tau_1, \tau_2|x_1, \ldots, x_n)$, $\pi(\tau_1, \nu|x_1, \ldots, x_n)$ and $\pi(\tau_2, \nu|x_1, \ldots, x_n)$, and finally, figure A.2.20 shows the ACF as function of lag for τ_1, τ_2 and ν. The ACF is decaying very quickly, and the acceptance rate is clearly high, in fact it is 0.794, as seen from the trace plots. The Rubin-Gelman statistic was calculated for five chains with run length $2 \cdot 10^6$ and burn-in 10^4 to $\widehat{R}^3 = 1$. Everything is thus pointing to convergence of the chain. Thus the analysis is continued with the assumption that the Markov chain is a representable sample from the posterior distribution of τ_1, τ_2 and ν, $\pi(\tau_1, \tau_2, \nu|x_1, \ldots, x_n; \alpha, \sigma).0$

We see that, as the situation with one change point, when the change point has not occurred, nearly all proposals of ν are accepted, as is natural, since none of the temperature increment realizations x_i are drawn with expectation ν. This also increases the probability mass of $\tau_1 < t_n$ compared to figure 5.2.2. This is seen in table 5.19 as well, compared to table 5.13, where ν was known. However, the effect on τ_2 is not as noticeable. The estimated mean and median as well as interval estimates of ν in table 5.19 is mostly reflecting the prior distribution as discussed in the situation with two change points and ν known.

Observation until $t_n = 1\ 100$

Figure 5.2.10 shows the prior and posterior distributions for τ_1, $\rho_2 = \tau_2 - \tau_1$, and ν, $\pi(\tau_1)$, $\pi(\tau_1|x_1, \ldots, x_n)$, $\pi(\tau_2 - \tau_1)$, $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n)$, $\pi(\nu)$, $\pi(\nu|x_1, \ldots, x_n)$ as well as the posterior distribution of $\pi(\tau_2)$. Table 5.20 gives the estimates of the mean and median as well as the 2.5th and 97.5th percentile and 95 % HPD region for τ_1, τ_2 and ν.

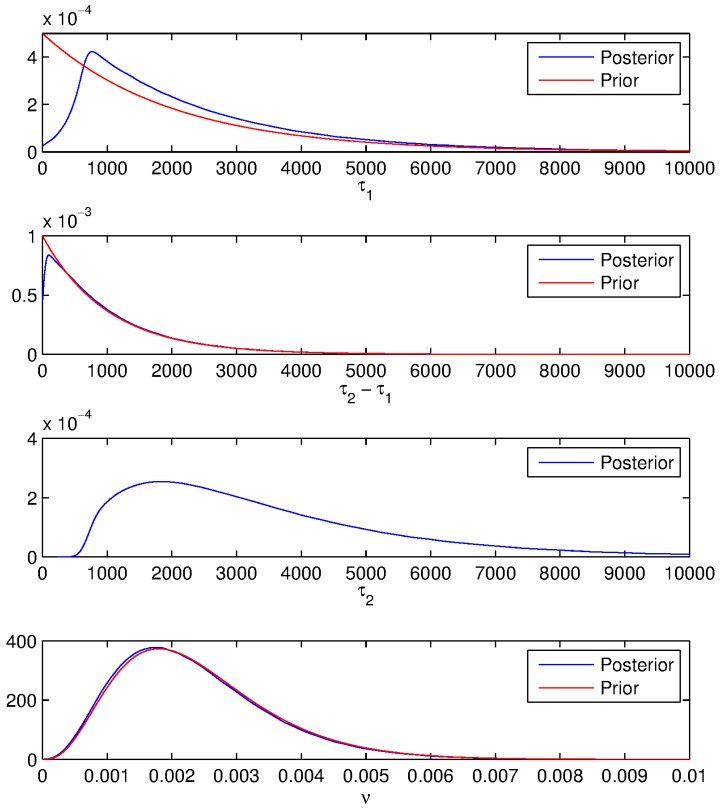

Figure 5.2.9: Prior distribution of τ_1, $\pi(\tau_1)$, $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1)$, and ν, $\pi(\nu)$, and posterior distributions of τ_1,$\pi(\tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \sigma)$, ν, $\pi(\nu|x_1, \ldots, x_n; \alpha, \sigma)$ given the observed temperature $W(t)$ up until $t_n = 700$.

Further diagnostic plots are shown in appendix A.2.2: the trace plot of τ_1, τ_2 and ν is shown in figure A.2.21, the histogram of the three variables and $\tau_2 - \tau_1$ is shown in figure A.2.22, figure A.2.23 shows the three pairwise contour plots of the joint posterior density of the three parameters, namely $\pi(\tau_1, \tau_2|x_1, \ldots, x_n)$, $\pi(\tau_1, \nu|x_1, \ldots, x_n)$ and $\pi(\tau_2, \nu|x_1, \ldots, x_n)$, and figure A.2.24 shows the ACF as function of lag for τ_1, τ_2 and ν. The ACF, especially of ν, is now decaying at a slower rate than for $t_n = 700$, however, it is not worrisome. The acceptance rate was 0.508. The Rubin-Gelman statistic was calculated for five chains with run length $2 \cdot 10^6$ and burn-in 10^4 to $\widehat{R}^3 = 1$. Thus, it is believed that the Markov chain has converged to the posterior distribution $\pi(\tau_1, \tau_2, \nu|x_1, \ldots, x_n)$.

Looking at figure 5.2.10, we see that the posterior distribution of τ_1, $\pi(\tau_1|x_1, \ldots, x_n)$ now has a large peak around $\tau_1 = 800$, which is pushing the posterior distribution of τ_2, $\pi(\tau_2|x_1, \ldots, x_n)$, towards larger values of τ_2, compared to figure 5.2.9. It is also evident that ν is pushed further towards larger values of ν. These changes are also detectable when comparing table 5.20 to table 5.19, where we see that the interval estimates of τ_1 and τ_2 have narrowed, whereas the interval estimates of ν has widened, however the estimated mean and median are now closer to the true value. Comparing table 5.20 to table 5.14, where the drift parameter ν was known, we see that the interval estimates of τ_1 and τ_2 are slightly wider with ν unknown, as is natural.

Observation until $t_n = 1\ 450$

Figure 5.2.11 shows the prior and posterior distributions for τ_1, $\rho_2 = \tau_2 - \tau_1$, and ν, $\pi(\tau_1)$, $\pi(\tau_1|x_1, \ldots, x_n)$, $\pi(\tau_2 - \tau_1)$, $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n)$, $\pi(\nu)$, $\pi(\nu|x_1, \ldots, x_n)$ as well as the posterior distribution of $\pi(\tau_2)$. Table 5.21 gives the estimates of the mean and median as well as the 2.5th and 97.5th percentile and 95 % HPD region for τ_1, τ_2 and ν.

Further diagnostic plots are shown in appendix A.2.2: the trace plot of τ_1, τ_2 and ν is shown in figure A.2.25, the histogram of the three variables and $\tau_2 - \tau_1$ is shown in figure A.2.26, figure A.2.27 shows the three pairwise contour plots of the joint posterior density of the three parameters, namely $\pi(\tau_1, \tau_2|x_1, \ldots, x_n)$, $\pi(\tau_1, \nu|x_1, \ldots, x_n)$ and $\pi(\tau_2, \nu|x_1, \ldots, x_n)$, and figure A.2.28 shows the ACF as function of lag for τ_1, τ_2 and ν. The trace plots shows that the acceptance rate has dropped markedly, to 0.130, and this is reflecting in the ACF, especially of ν. The Rubin-Gelman statistic was calculated for five chains with run length $2 \cdot 10^6$ and burn-in 10^4 to $\widehat{R}^3 = 1$. The long run length is assumed to wash out the autocorrelation effects, and allow the chain to traverse the entire range of the parameters.

The posterior distribution now has a sharp peak around $\tau_2 = 1\ 400$, as was the case when ν was known, and the tail is still considerable, as seen by the interval estimates in table 5.15. The interval estimates of τ_1 and τ_2 have narrowed, especially from above, and the estimated means and medians are closer. This points to the likelihood of large values for either change point being small. Note that the intervals of ν are moving further towards larger values of ν, and the estimate of ν is now very close to the true value. However, the realized mean of the temperature increments x_i in the time interval $[\tau_1, \tau_2) = [800, 1400)$ is $\bar{x}_i = 0.006$, thus it is underestimated. This may be the reason that the interval estimate is moving towards larger values of ν. However, remember that

Table 5.19: Statistics of the posterior distribution of τ_1, τ_2 and ν, $\pi(\tau_1, \tau_2, \text{nu}|x_1, \ldots, x_n; \alpha, \sigma)$, as estimated by MCMC methods. Here, α and σ were considered known, and the endpoint of observation was $t_n = 700$.

	τ_1	
Point estimates	Mean	2 453
	Median	1 849
Interval estimates	2.5 and 97.5 percentile	[370, 7 843]
	95 % HPD	[1, 6 464]
	τ_2	
Point estimates	Mean	3 474
	Median	2 923
Interval estimates	2.5 and 97.5 percentile	[875, 9 218]
	95 % HPD	[686, 7 833]
	ν	
Point estimates	Mean	0.0024
	Median	0.0022
Interval estimates	2.5 and 97.5 percentile	[0.0006, 0.0052]
	95 % HPD	[0.0004, 0.0047]

Table 5.20: Statistics of the posterior distribution of τ_1, τ_2 and ν, $\pi(\tau_1, \tau_2, \text{nu}|x_1, \ldots, x_n; \alpha, \sigma)$, as estimated by MCMC methods. Here, α and σ were considered known, and the endpoint of observation was $t_n = 1\ 100$.

	τ_1	
Point estimates	Mean	1 715
	Median	885
Interval estimates	2.5 and 97.5 percentile	[279, 6 779]
	95 % HPD	[1, 5 394]
	τ_2	
Point estimates	Mean	2 948
	Median	2 338
Interval estimates	2.5 and 97.5 percentile	[1 136, 8 176]
	95 % HPD	[1 070, 6 821]
	ν	
Point estimates	Mean	0.0027
	Median	0.0025
Interval estimates	2.5 and 97.5 percentile	[0.0007, 0.0057]
	95 % HPD	[0.0005, 0.0052]

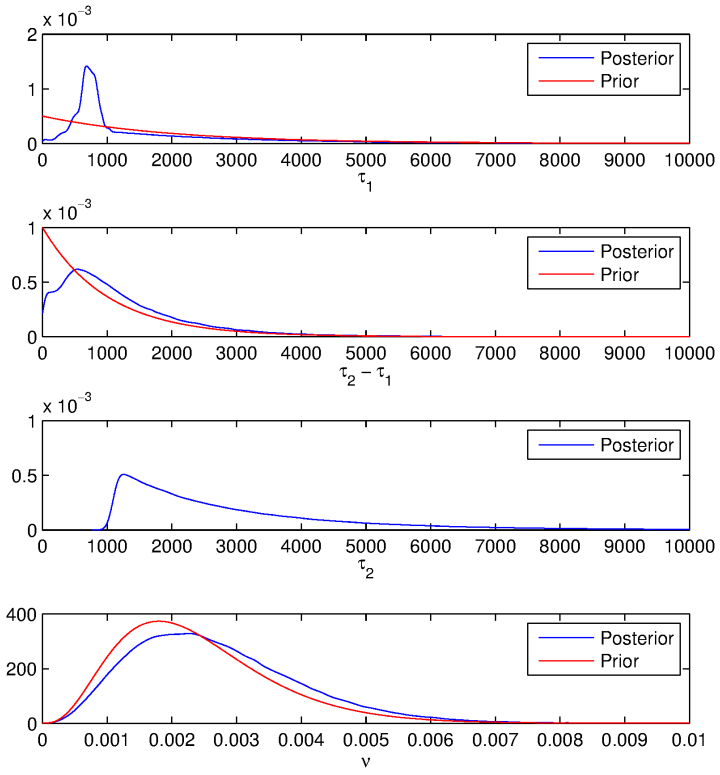

Figure 5.2.10: Prior distribution of τ_1, $\pi(\tau_1)$, $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1)$, and ν, $\pi(\nu)$, and posterior distributions of τ_1,$\pi(\tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \sigma)$, ν, $\pi(\nu|x_1, \ldots, x_n; \alpha, \sigma)$ given the observed temperature $W(t)$ up until $t_n = 1\ 100$.

when we estimate ν we are implicitly estimating $\alpha\nu$ as well. This may be the cause of the underestimation, as there are few temperature increments x_i drawn with expectation $\alpha\nu$.

Observation until t_n = 2 000

Figure 5.2.12 shows the prior and posterior distributions for τ_1, $\rho_2 = \tau_2 - \tau_1$, and ν, $\pi(\tau_1)$, $\pi(\tau_1|x_1, \ldots, x_n)$, $\pi(\tau_2 - \tau_1)$, $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n)$, $\pi(\nu)$, $\pi(\nu|x_1, \ldots, x_n)$ as well as the posterior distribution of $\pi(\tau_2)$. Table 5.22 gives the estimates of the mean and median as well as the 2.5th and 97.5th percentile and 95 % HPD region for τ_1, τ_2 and ν.

As before, further diagnostic plots are shown in appendix A.2.2: the trace plot of τ_1, τ_2 and ν is shown in figure A.2.29, the histogram of the three variables and $\tau_2 - \tau_1$ is shown in figure A.2.30, figure A.2.31 shows the three pairwise contour plots of the joint posterior density of the three parameters, namely $\pi(\tau_1, \tau_2|x_1, \ldots, x_n)$, $\pi(\tau_1, \nu|x_1, \ldots, x_n)$ and $\pi(\tau_2, \nu|x_1, \ldots, x_n)$, and figure A.2.32 shows the ACF as function of lag for τ_1, τ_2 and ν. The trace plots shows that the acceptance rate has dropped markedly, and this is reflecting in the ACF, quite severely, and it is now 0.005. The Rubin-Gelman statistic was calculated for five chains with run length $2 \cdot 10^6$ and burn-in 10^4 to $\widehat{R}^3 = 1$. The long run length is thus assumed to wash out the autocorrelation effects, and allow the chain to traverse the entire range of the parameters.

As when ν was known, the interval estimates of τ_1 and τ_2 have narrowed distinctly. The interval estimates of ν has also narrowed, mainly from above. It is believed, as in the case when ν was known, that it is the change in the variance parameter to $\alpha\sigma$ that causes the narrowness in the posterior distribution $\pi(\tau_2|x_1, \ldots, x_n)$ of τ_2 compared to the posterior distribution $\pi(\tau_1|x_1, \ldots, x_n)$ of τ_1.

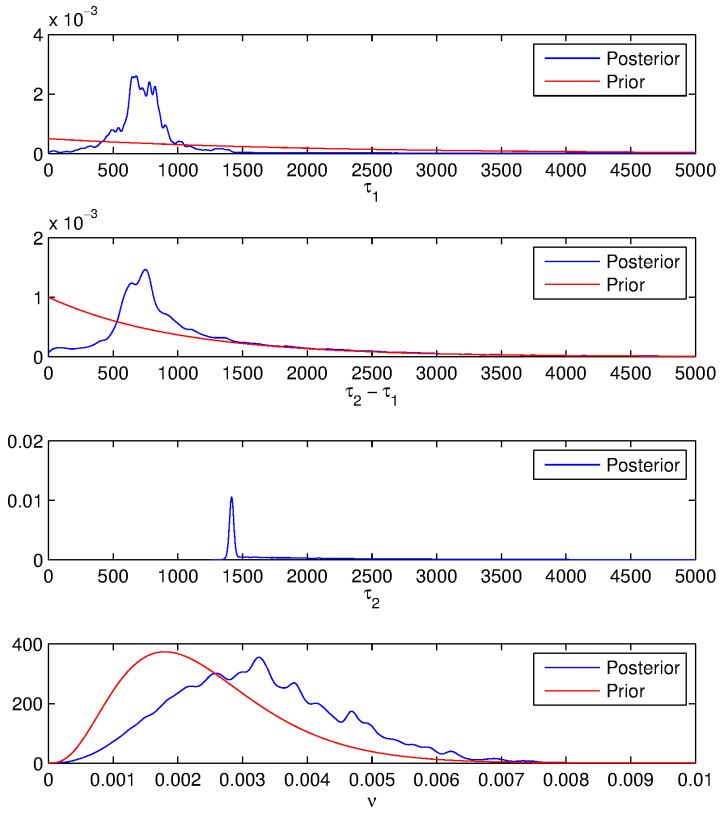

Figure 5.2.11: Prior distribution of τ_1, $\pi(\tau_1)$, $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1)$, and ν, $\pi(\nu)$, and posterior distributions of τ_1,$\pi(\tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \sigma)$, ν, $\pi(\nu|x_1, \ldots, x_n; \alpha, \sigma)$ given the observed temperature $W(t)$ up until $t_n = 1\ 450$.

Table 5.21: Statistics of the posterior distribution of τ_1, τ_2 and ν, $\pi(\tau_1, \tau_2, \text{nu}|x_1, \ldots, x_n; \alpha, \sigma)$, as estimated by MCMC methods. Here, α and σ were considered known, and the endpoint of observation was $t_n = 1\ 450$

τ_1		
Point estimates	Mean	935
	Median	737
Interval estimates	2.5 and 97.5 percentile	[281, 3 713]
	95 % HPD	[1, 2 318]
τ_2		
Point estimates	Mean	2 142
	Median	1 572
Interval estimates	2.5 and 97.5 percentile	[1 398, 5 626]
	95 % HPD	[1 535, 4 572]
ν		
Point estimates	Mean	0.0032
	Median	0.0032
Interval estimates	2.5 and 97.5 percentile	[0.001, 0.0061]
	95 % HPD	[0.001, 0.0059]

Table 5.22: Statistics of the posterior distribution of τ_1, τ_2 and ν, $\pi(\tau_1, \tau_2, \text{nu}|x_1, \ldots, x_n; \alpha, \sigma)$, as estimated by MCMC methods. Here, α and σ were considered known, and the endpoint of observation was $t_n = 2\ 000$

τ_1		
Point estimates	Mean	714
	Median	714
Interval estimates	2.5 and 97.5 percentile	[212, 1 287]
	95 % HPD	[193, 1 264]
τ_2		
Point estimates	Mean	1 415
	Median	1 416
Interval estimates	2.5 and 97.5 percentile	[1 388, 1 443]
	95 % HPD	[1 389, 1 443]
ν		
Point estimates	Mean	0.0027
	Median	0.0025
Interval estimates	2.5 and 97.5 percentile	[0.0009, 0.0047]
	95 % HPD	[0.0008, 0.0045]

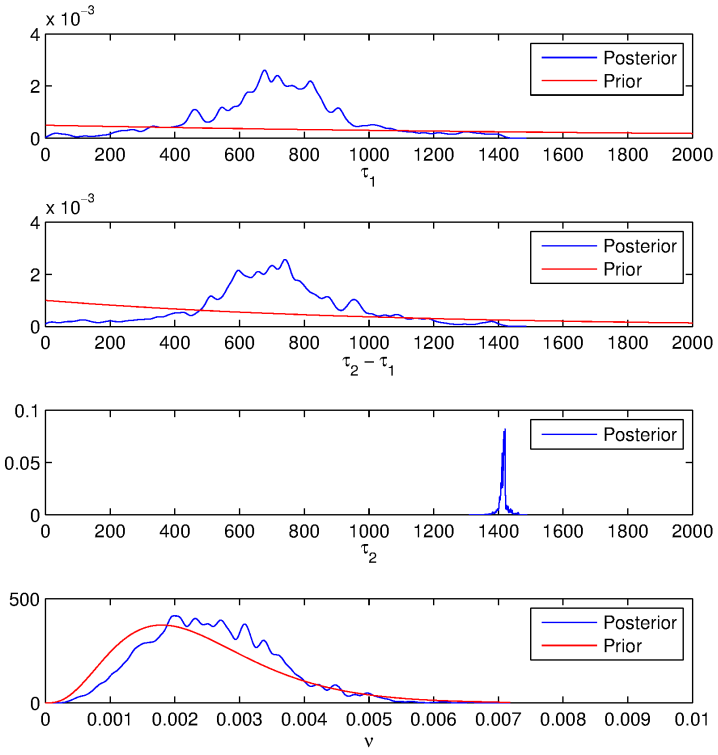

Figure 5.2.12: Prior distribution of τ_1, $\pi(\tau_1)$, $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1)$, and ν, $\pi(\nu)$, and posterior distributions of τ_1,$\pi(\tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \sigma)$, ν, $\pi(\nu|x_1, \ldots, x_n; \alpha, \sigma)$ given the observed temperature $W(t)$ up until $t_n = 2\ 000$.

Hitting time CDF estimates

In the formulaic approach, $\widehat{P}_f(T \leq t)$, there were $m = 500\ 000$ draws from the posterior distribution in the MC summation. In the simulation approach, $\widehat{P}_S(T \leq t)$, there were 10 000 draws from the posterior distribution and 100 simulations of Wiener processes for each draw from the posterior distribution. As before, the difference in number of draws from the posterior distribution stems from computational limitations. The estimated hitting time CDF by the formulaic approach, $\widehat{P}_f(T \leq t)$, is shown in figure 5.2.13 with the 95 % credible intervals and the true distribution. Selected percentiles of $\widehat{P}_f(T \leq t)$ and the corresponding credible intervals are given in table 5.23. The estimated hitting time CDF by the simulation approach, $\widehat{P}_S(T \leq t)$, is shown in figure 5.2.14 with the 95 % credible intervals and the true distribution. Table 5.24 displays selected percentiles of $\widehat{P}_S(T \leq t)$ and the corresponding credible intervals. The two approaches are compared in figure 5.2.15.

The general behavior of the estimates $\widehat{P}_f(T \leq t)$ and $\widehat{P}_S(T \leq t)$ are much like the estimates when ν was known. The credible intervals narrow as more information is considered, at $t_n = 1\ 100$, we can see the effect of the first change point, by elevating and steepening $\widehat{P}_f(T \leq t)$ and $\widehat{P}_S(T \leq t)$ for smaller t. Also the effect of $t_n = 1\ 450$ is clear, and further steepens $\widehat{P}_f(T \leq t)$ and $\widehat{P}_S(T \leq t)$ for smaller t.

However, the width of the credible intervals for each is remarkably larger, and the shapes are also changed. This is of course caused by ν being a random variable to be estimated. This naturally changes the estimates themselves as well, $\widehat{P}_f(T \leq t)$ and $\widehat{P}_S(T \leq t)$, such that the estimated curves become less steep. It seems that it was more difficult to estimate ν that it was for one change point, because the effect seems more severe. Perhaps this is linked to the fact that $\nu_2 = \alpha\nu_1$, such that in reality, we are estimating both ν and $\alpha\nu$.

The two estimation approaches render similar results, as seen in figure 5.2.15, and it is not possible say with certainty whether $\widehat{P}_f(T \leq t)$ is steeper than $\widehat{P}_S(T \leq t)$, as it was with ν unknown. However, it seems that for all t_n, the credible intervals for $\widehat{P}_f(T \leq t)$ are steeper than those for $\widehat{P}_S(T \leq t)$, but it is difficult to say if this is caused by the smaller number of draws from the posterior distribution in the simulation approach or whether it is the same effect that caused steepness in the case where ν was known, that is now masked by the added variability caused by ν being unknown.

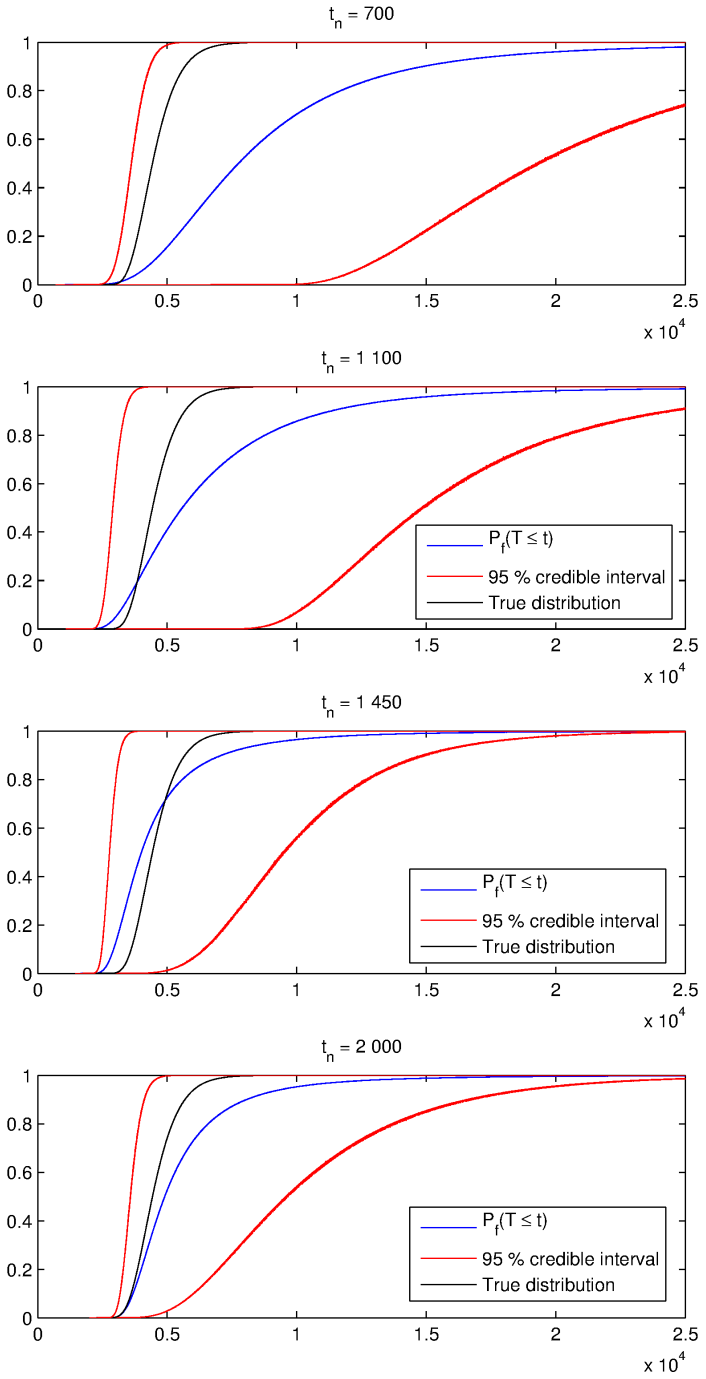

Figure 5.2.13: Formulaic approach to estimating CDF, $\widehat{P}_f(T \leq t)$ shown with 95 % credible intervals for different observation lengths, t_n. Shown with true distribution. Here, τ_1, τ_2 and ν are unknown parameters, σ and α are known parameters. Top panel: $t_n = 700$. Second panel: $t_n = 1\ 100$. Third panel: $t_n = 1\ 450$. Bottom panel: $t_n = 2\ 000$. The x-axis is time and y-axis is probability. Note that the same legend applies to all panels and has been omitted in the top panel for visibility.

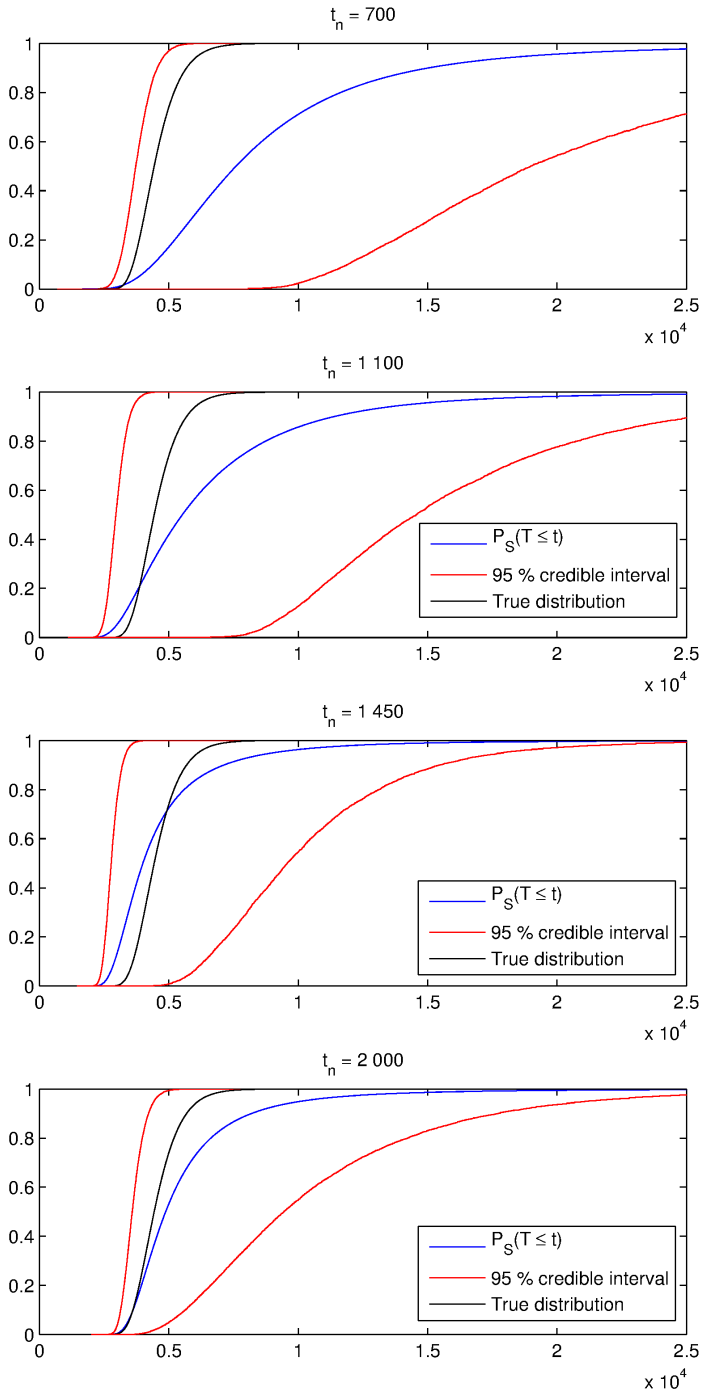

Figure 5.2.14: Simulation approach to estimating CDF, $\widehat{P}_S(T \leq t)$ shown with 95 % credible intervals for different observation lengths, t_n. Shown with true distribution. Here, τ_1, τ_2 and ν are unknown parameters, σ and α are known parameters. Top panel: $t_n = 700$. Second panel: $t_n = 1\ 100$. Third panel: $t_n = 1\ 450$. Bottom panel: $t_n = 2\ 000$. The x-axis is time and y-axis is probability. Note that the same legend applies to all panels and has been omitted in the top panel for visibility.

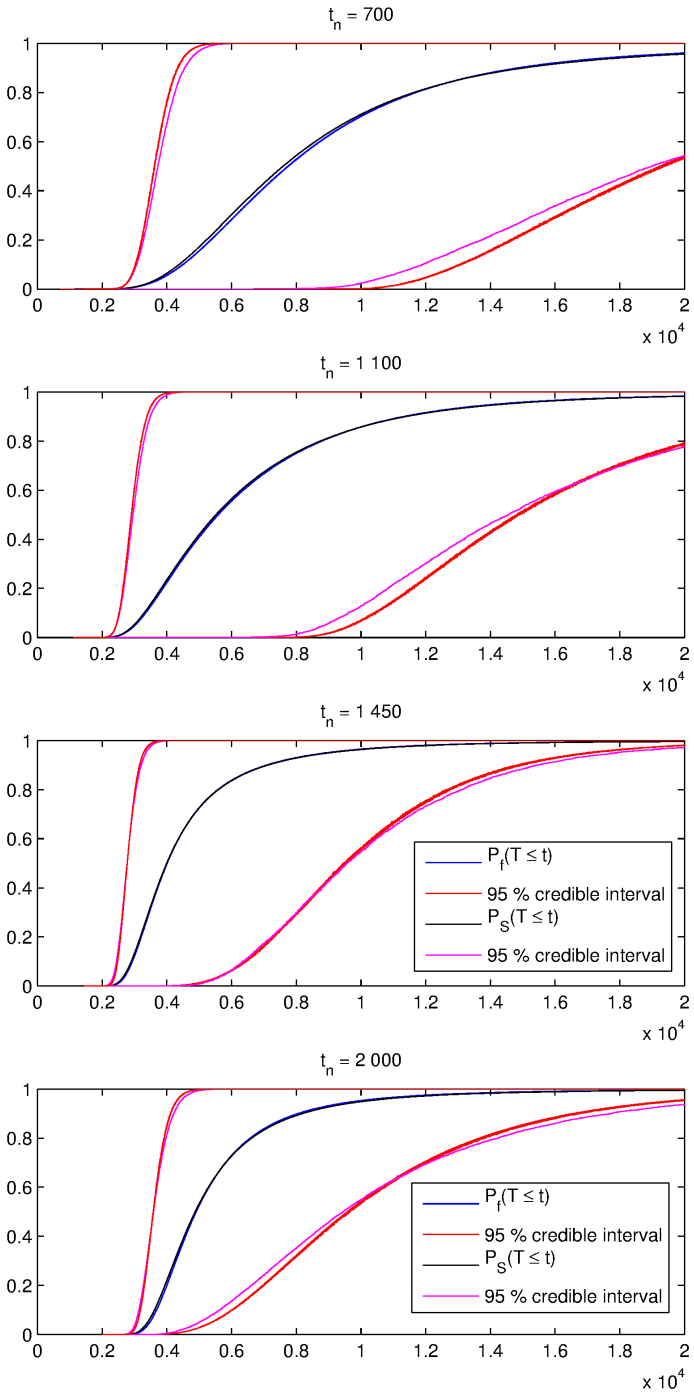

Figure 5.2.15: Comparison of the simulation and formulaic approaches, $\widehat{P}_S(T \leq t)$ and $\widehat{P}_f(T \leq t)$ for different observation lengths t_n. Here, τ_1, τ_2 and ν are unknown parameters, σ and α are known parameters. Top panel: $t_n = 700$. Second panel: $t_n = 1\ 100$. Third panel: $t_n = 1\ 450$. Bottom panel: $t_n = 2\ 000$. The x-axis is time and y-axis is probability. Note that the same legend applies to all panels and has been omitted in the top panels for visibility.

Table 5.23: Percentiles of the estimated CDF by the formulaic solution, $\widehat{P}_f(T \leq t)$ with ν unknown, α and σ known.

Percentile	2.5	5	10	25	50	75	90	95	97.5
$t_n = 700$									
Upper 95 %	2 765	2 883	3 031	3 294	3 619	3 976	4 337	4 554	4 767
Estimated CDF	3 514	3 930	4 494	5 750	7 751	10 725	14 851	18 522	23 022
Lower 2.5 %	11 386	12 067	13 043	15 324	19 229	25 117	$> 30\ 000$		
$t_n = 1\ 100$									
Upper 95 %	2 355	2 432	2 524	2 692	2 899	3 142	3 384	3 542	3 690
Estimated CDF	2 788	3 916	3 344	4 128	5 582	7 985	11 321	14 122	17 442
Lower 2.5 %	9 164	9 695	10 402	12 065	14 828	18 893	24 282	28 269	>30 000
$t_n = 1\ 450$									
Upper 95 %	2 355	2 410	2 47	2 605	2 776	2 971	3 157	3 280	3 398
Estimated CDF	2 620	2 752	2 933	3 333	4 005	5 164	7 084	8 920	11 086
Lower 2.5 %	5 325	5 779	6 389	7 640	9 462	11 979	14 835	16 986	19 178
$t_n = 2\ 000$									
Upper 95 %	3 015	3 090	3 184	3 355	3 578	3 840	4 144	4 291	4 457
Estimated CDF	3 320	3 473	3 682	4 141	4 904	6 151	8 051	9 814	11 972
Lower 2.5 %	4 928	5 380	6 014	7 396	9 591	12 701	16 536	19 407	22 278
True distribution									
$t_n = 2\ 000$	3 305	3 438	3 612	3 956	4 433	5 029	5 684	6 136	6 567

Table 5.24: Percentiles of the estimated CDF by the simulation solution, $\widehat{P}_S(T \leq t)$ with ν unknown, α and σ known.

Percentile	2.5	5	10	25	50	75	90	95	97.5
$t_n = 700$									
Upper 95 %	2 778	2 904	3 071	3 361	3 722	4 148	4 558	4 831	5 072
Estimated CDF	3 442	3 835	4 384	5 607	7 598	10 641	15 031	19 092	24 176
Lower 2.5 %	10 058	10 776	11 932	14 582	18 995	26 455	$> 30\,000$		
$t_n = 1\,100$									
Upper 95 %	2 354	2 436	2 537	2 719	2 956	3 224	3 496	3 677	3 848
Estimated CDF	2 753	2 976	3 299	4 076	5 519	7 938	11 387	14 366	17 864
Lower 2.5 %	8 368	8 903	9 622	11 412	14 573	19 276	25 511	$> 30\,000$	
$t_n = 1\,450$									
Upper 95 %	2 303	2 370	2 445	2 590	2 774	2 985	3 209	3 355	3 485
Estimated CDF	2 580	2 714	2 900	3 310	3 997	5 171	7 123	9 030	11 305
Lower 2.5 %	5 445	5 856	6 363	7 650	9 573	12 283	15 471	17 938	20 637
$t_n = 2\,000$									
Upper 95 %	2 957	3 039	3 137	3 333	3 588	3 891	4 218	4 432	4 627
Estimated CDF	3 252	3 405	3 617	4 087	4 876	6 174	8 190	10 088	12 435
Lower 2.5 %	4 585	5 021	5 650	7 079	9 453	13 064	17 722	21 232	24 744
True distribution									
$t_n = 2\,000$	3 305	3 438	3 612	3 956	4 433	5 029	5 684	6 136	6 567

Chapter 6

Concluding remarks

In this thesis, the aim was to find a method to predict the hitting time of a specified threshold for a piecewise Wiener process with change points. The motivating example was the failure development in a wind turbine bearing, as described in Valland et al. (2012). The work was partly based on the theory developed in Lindqvist and Slimacek (2013), and their modeling of the failure development as a hidden Markov model was adopted. The theory developed in Lindqvist and Slimacek (2013) was expanded to find the hitting time CDF, which was the method chosen to predict the hitting time. This allowed for the user to find the level of risk to take by choosing a credible level α such that $\widehat{P}(T \leq t) \leq \alpha$.

The hitting time CDF was developed in detail in the case where the failure development had one and two change points. For one change point, the approaches to find the hitting time CDF were a simulation approach and a formulaic approach based on the inverse Gaussian CDF, as seen in section 4.2.2 and 4.2.1 respectively. The model for one change point was expanded to include the drift parameter ν and the variance parameter σ as unknown parameters. For two change points, the simulation approach was extended, as seen in section 4.3.3. The formulaic approach was extended first straightforwardly in section 4.3.1, leading to a method with complicated and computer intensive expressions. Section 4.3.2 extended the formulaic approach via an elegant time transformation following Doksum and Høyland (1992). The model for two change points was expanded to include the drift parameter ν_1 as an unknown parameter. The methods used Bayesian inference in the sense that the approaches conditioned on the change points, and MC summation was employed over the Bayesian posterior distribution of the change point(s). Generalizations on how to extend the simulation approach and the formulaic time transformed approach were given for m change points.

To assess the uncertainty in the estimated hitting time CDFs, it was noted that all approaches used MC summation with m draws, however m varied for the different approaches. Thus, in each approach, m estimates of the hitting time were available for each t, and the 1 - α credible interval defined by α and $1-\alpha$ percentile of the m estimates were chosen to measure the uncertainty in the estimate.

Finally, numerical examples were provided for one and two change points. For one

change point, examples were given with τ unknown, τ and ν unknown and finally τ, ν and σ unknown. The methods employed were the simulation and the formulaic approach. The examples showed that both the formulaic and the simulation approach gave good results, although the simulation approach proved more computer intensive. The results were significantly more accurate for ν and σ known. The parameters ν and σ have a bigger impact on the shape of the hitting time CDF than τ. It was also seen that when there were many measurements after the change point, the Markov chain in the MCMC procedure was moving slowly.

For two change points, examples were given with τ_1 and τ_2 unknown, and with τ_1, τ_2 and ν_1 unknown. The simulation and the formulaic time transformed approach were employed to estimate the hitting time CDF. The simulation approach gave good results also for two change point, and again proved to be the more computer intensive of the two methods, as expected. For the formulaic time transformed approach, there was an unexplained steepness in the hitting time CDF estimate which slightly exceeded the steepness of the true distribution (as calculated the last observation point) which manifested when there was an abundance of measurements after the second change points. This was perhaps the reason that the uncertainty estimates for the distribution were different from the uncertainty estimates for the simulation approach. The formulaic time transformed approach gave good results despite of this. However, it was not possible to exclude the possibility that the differences in the uncertainty estimates were due to the differences in number of draws from the posterior distribution by the two methods. It was seen that, as for one change point, when there were many measurements after the change point, the Markov chain in the MCMC procedure was moving slowly.

6.1 Further work

In both the one change point setting and two change point setting, it was clear that a more efficient MCMC algorithm would be beneficial. In the situation where there were many observations after the change point(s), in both the case of one and two change points, the chain was moving slowly as noted above. Thus, if one were worried about the Markov chain getting stuck, one could for example implement a random walk routine instead of the independence sampler, taking care not to violate the restrictions that $\tau_2 > \tau_1 > 0$ and $\nu > 0$. However, we make the note that being in a situation where an error has occurred a long time ago, and wanting to make predictions a long time in the future, is perhaps not to realistic. It is more plausible that if one finds that the failure development started a long time ago, one wishes to do maintenance sooner rather than later. Thus, this shortcoming of the estimation procedure shown may not be the most critical for realistic applications.

The unexplained steepness in the formulaic approach must be assessed. By employing the simulation approach with a larger number of measurements, it may be checked whether the steepness in the uncertainty estimates is due to the number of draws from the posterior distribution. If not, the formulaic time transformation approach must be evaluated.

It can be noted that when one wishes to calculate the hitting time CDF as a function of the time t for many t, the calculations for each t is independent of each other, for the two approaches used with both one and two change points. Thus, the computations may be parallelized, to speed up the computations.

It was seen in the one and two change point setting that when the drift parameter ν was unknown, the hitting time CDF was greatly affected. Thus, the importance of assessment of the model parameters is evident.

This thesis has also implicitly considered the number of change points in the process known. Another approach could be to condition on the number of change points, if the failure development process is not known to such a degree as the motivating example.

In the TCI framework, other health indicators were suggested. If these were to be taken into account, and the Wiener process was appropriate for the indicators, one could model a multi-dimensional Wiener process. The health indicators would be correlated. It would be necessary find a measure of a inverse Gaussian distribution corresponding to multidimensional measurements, or to measure the distance from each health indicators hitting time distribution in multi-dimensional space.

Another suggestion for model extension is to the include other failure modes, that is, competing risks. The failure that has been considered in this model is a 'soft' failure from mechanical wear. One could add the possibility of a shock, which will have a greater effect further along in the failure development.

In this thesis, it was suggested to choose an estimator for T according to the risk one was willing to take, such that $\widehat{T}_\alpha$ is the t for which $P(T \leq t) = \alpha$. Another approach to include risk could be to use loss functions to estimate the parameters. In this setting it may be reasonable to penalize more for overestimation than underestimation.

The time transformation that was seen in section 4.3.2 and its following restrictions on ν_2 and σ_2 may not be reasonable for all types of failure development. Several other models are proposed in Doksum and Høyland (1992), including continuously varying stress and generally changing stress, which give rise to a flexible group of models. Common for all the models they propose is that the time transformation $\xi(t)$ must be a nonnegative strictly increasing and continuous function.

The Wiener process was the mathematical model in this thesis. For problems where the deterioration is cumulative, such as corrosion or crack growth, the gamma process may be more suitable. An extensive study of the gamma process can be found in van Noortwijk (2009). A frequentistic approach to finding change points is described in Fouladirad et al. (2008) and is exemplified through the gamma process.

In order to effectively use the distributions found in this thesis, a maintenance optimization strategy should be found, combining the hitting time CDFs of the bearing temperature in different wind turbines.

Chapter 7

Bibliography

Aalen, O. O. and Gjessing, H. K. (2001). Understanding the shape of the hazard rate: A process point of view. *Statistical Science*, 16(1):1 – 22.

Brooks, S. and Gelman, A. (1998). General methods for monitoring convergence by iterative simulations. *Journal of Computational and Graphical Statistics*, 7:434 – 455.

Doksum, K. and Høyland, A. (1992). Models for variable-stress accelerate life testing experiments based on Wiener processes and the inverse Gaussian distribution. *Technometrics*, 34(1):74 – 82.

Fouladirad, M., Grall, A., and Dieulle, L. (2008). On the use on on-line detection for maintenance of gradually deteriorating systems. *Reliability Engineering and System Safety*, 93:1814 – 1820.

Givens, G. H. and Hoeting, J. A. (2013). *Computational Statistics*. John Wiley & Sons, Inc., 2nd edition.

Hawkins, D. M. (2001). Fitting multiple change-point models to data. *Computational Statistics & Data Analysis*, 37:323 – 341.

Horrocks, J. and Thompson, M. E. (2004). Modeling event times with multiple outcomes using the Wiener process with drift. *Lifetime Data Analysis*, 10:29 – 49.

Lindqvist, B. and Slimacek, V. (2013). Failure prediction in monitored systems. *Proceedings to ESREL*.

Link, W. A. and Eaton, M. J. (2012). On thinning of chains in MCMC. *Methods of Ecology and Evolution*, 3:112 – 115.

Rausand, M. and Høyland, A. (2004). *System Reliability Theory: Models, Statistical Methods and Applications*. John Wiley & Sons, Inc., 2nd edition.

Ross, S. M. (2007). *Introduction to Probability Models*. Academic Press, 9th edition.

Shiryaev, A. N. (1963). On optimum methods in quickest detection problems. *Theory and probability and its applications*, 8(1):22 – 46.

Valland, A., Brurok, T., Slimacek, V., and Lindqvist, B. (2012). Probabilistic modeling of wind farm using TCI framework. Manuscript draft.

van Noortwijk, J. M. (2009). A survey of the application of gamma processes in maintenance. *Reliability Engineering and System Safety*, 94:2 – 21.

Whitmore, G. A. (1986). First-passage-time models for duration data: regression structures and competing risks. *The Statistician*, 35:207 – 219.

Whitmore, G. A., Crowder, M. J., and Lawless, J. F. (1998). Failure inference from a marker process based on a bivariate Wiener model. *Lifetime Data Analysis*, 4:229 – 251.

Whitmore, G. A. and Schenkelberg, F. (1997). Modelling accelerated degradation data using Wiener diffusion with a time scale transformation. *Lifetime Data Analysis*, 3:27 – 45.

Appendix A

Markov Chain Monte Carlo diagnostic plots

A.1 One change point

A.1.1 τ unknown, ν and σ known

Observation until $t_n = 1\ 400$

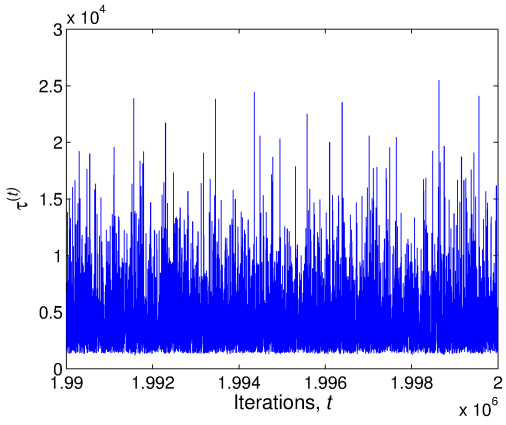

Figure A.1.1: Trace plot of the last 10 000 iterations of the Markov chain from the posterior distribution of τ with ν and σ considered known, $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, and the endpoint of observations was $t_n = 1\ 400$. Note that few iterations in the Markov chain are below $\tau^{(t)} = 1400$. This is also reflected in the histogram in figure A.1.2 and the 2.5th percentile in table 5.1.

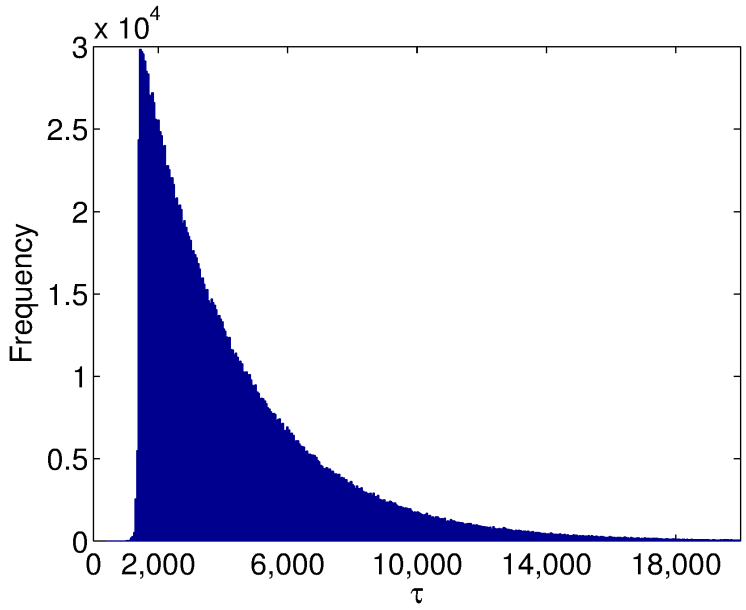

Figure A.1.2: Histogram of the MCMC samples from $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, the posterior distribution of τ with ν and σ considered known, and the endpoint of observation was $t_n = 1\ 400$.

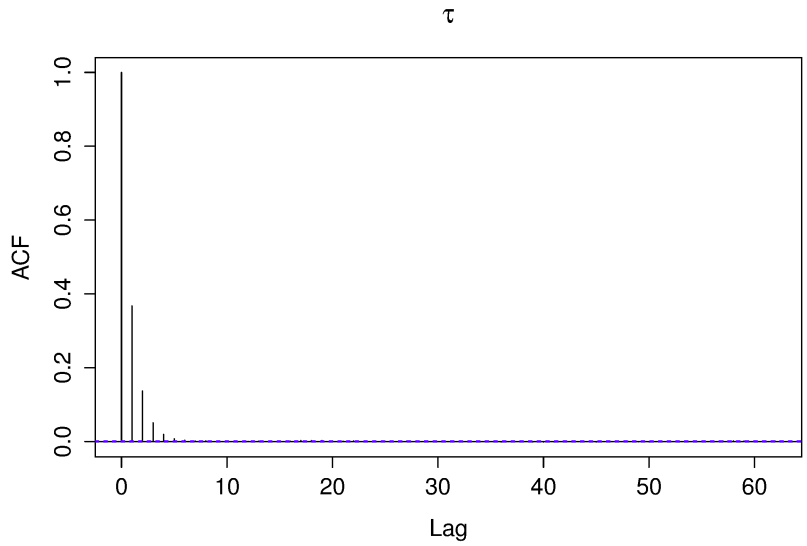

Figure A.1.3: Autocorrelation of the MCMC sample from $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, the posterior distribution of τ, as function of the lag. Here, ν and σ are considered known, and the endpoint of observation was $t_n = 1\ 400$.

Observation until $t_n = 1\ 700$

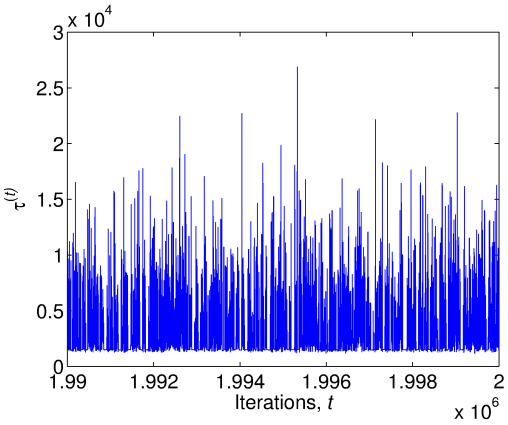

Figure A.1.4: Trace plot of the last 10 000 iterations of the Markov chain from the posterior distribution of τ with ν and σ considered known, $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, and the endpoint of observation was $t_n = 1\ 700$.

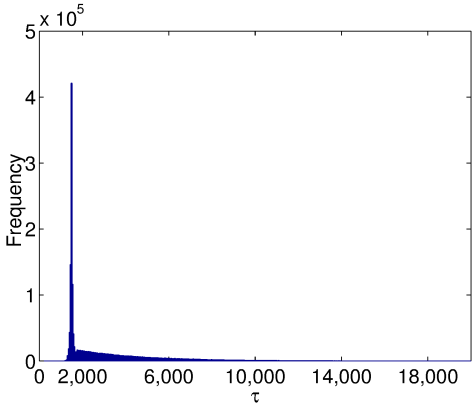

Figure A.1.5: Histogram of the MCMC samples from $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, the posterior distribution of τ with ν and σ considered known, and the endpoint of observation was $t_n = 1\ 700$.

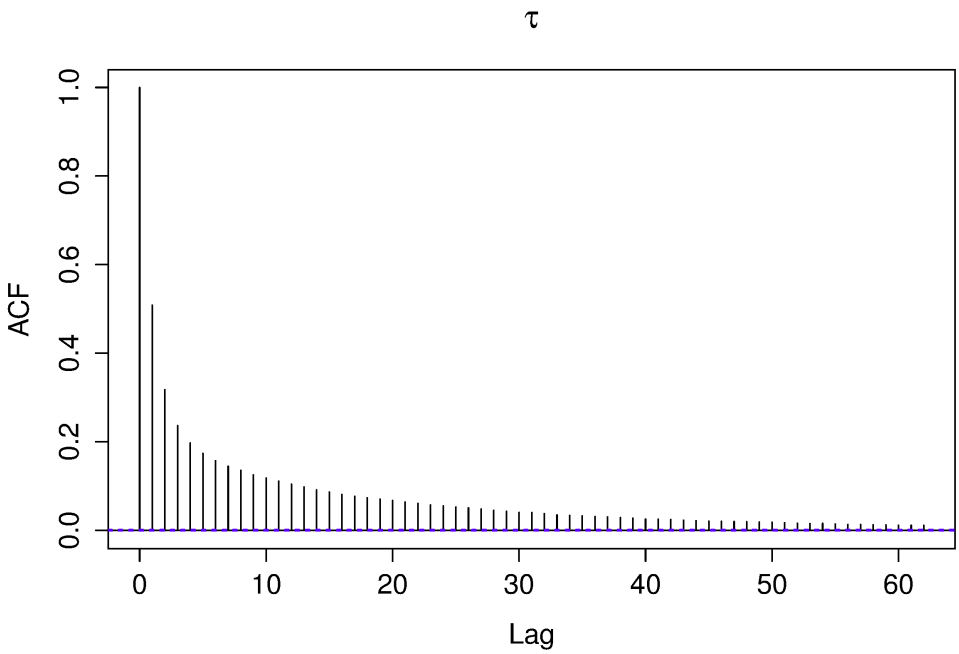

Figure A.1.6: Autocorrelation of the MCMC samples from $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, the posterior distribution of τ, as function of the lag. Here, ν and σ are considered known, and the endpoint of observation was $t_n = 1\ 700$.

Observation until t_n = 3 000

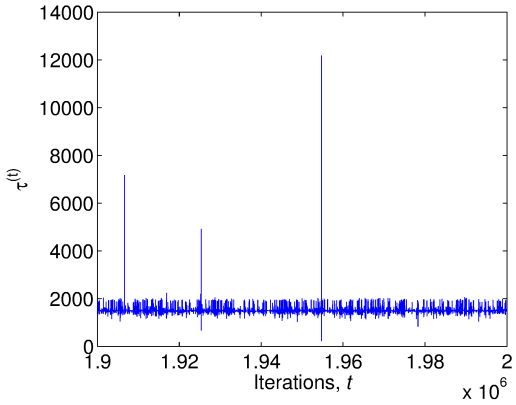

Figure A.1.7: Trace plot of the last 100 000 iterations of the Markov chain from the posterior distribution of τ with ν and σ considered known, $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, and the endpoint of observations was $t_n = 3\ 000$.

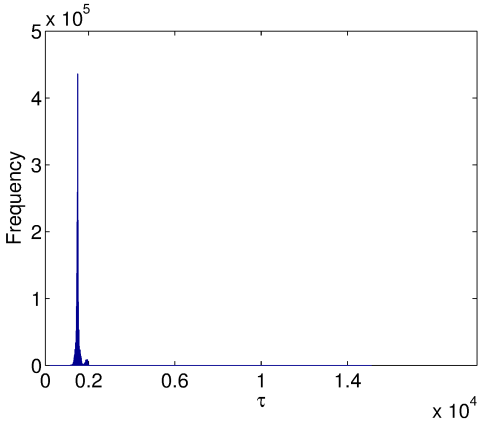

Figure A.1.8: Histogram of the MCMC samples from $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, the posterior distribution of τ with ν and σ considered known, and the endpoint of observation was $t_n = 3\ 000$.

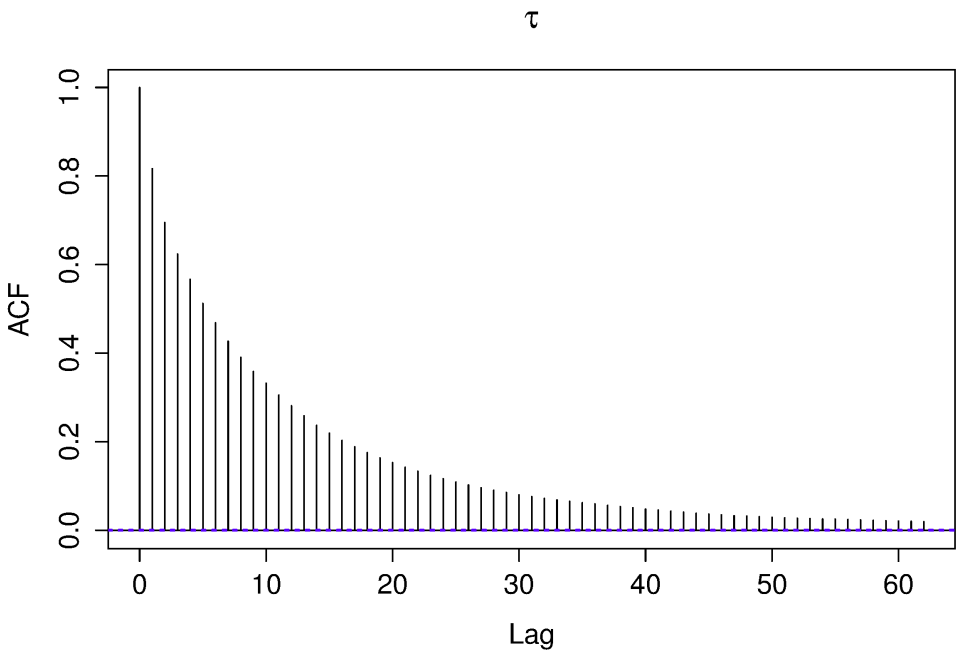

Figure A.1.9: Autocorrelation of the MCMC samples from $\pi(\tau|x_1, \ldots, x_n; \nu, \sigma)$, the posterior distribution of τ, as function of the lag. Here, ν and σ are considered known, and the endpoint of observation was $t_n = 3\ 000$.

A.1.2 τ and ν unknown, σ known

Observation until $t_n = 1\ 400$

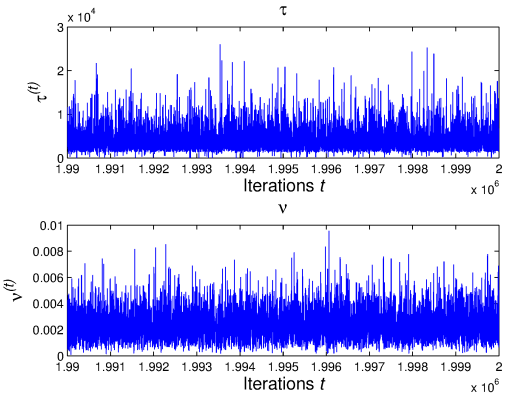

Figure A.1.10: Marginal trace plot of the last 100 000 iterations of the Markov chain from the posterior distribution of τ, $\pi(\tau|x_1, \ldots, x_n; \sigma)$, and ν, $\pi(\nu|x_1, \ldots, x_n; \sigma)$. Here, σ was considered known, and the endpoint of observation was $t_n = 1\ 400$.

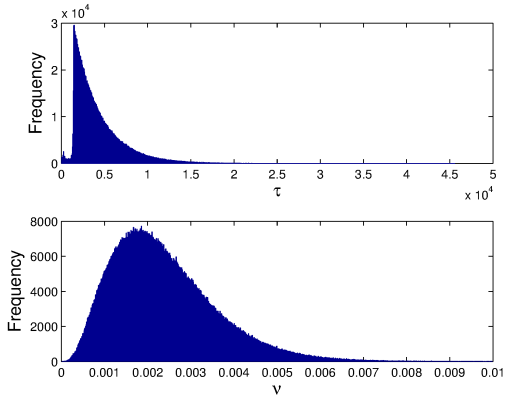

Figure A.1.11: Histograms of the MCMC sample from the marginal posterior distributions of τ, $\pi(\tau|x_1, \ldots, x_n; \sigma)$, and ν, $\pi(\nu|x_1, \ldots, x_n; \sigma)$. Here, σ was considered known, and the endpoint of observation was $t_n = 1\ 400$.

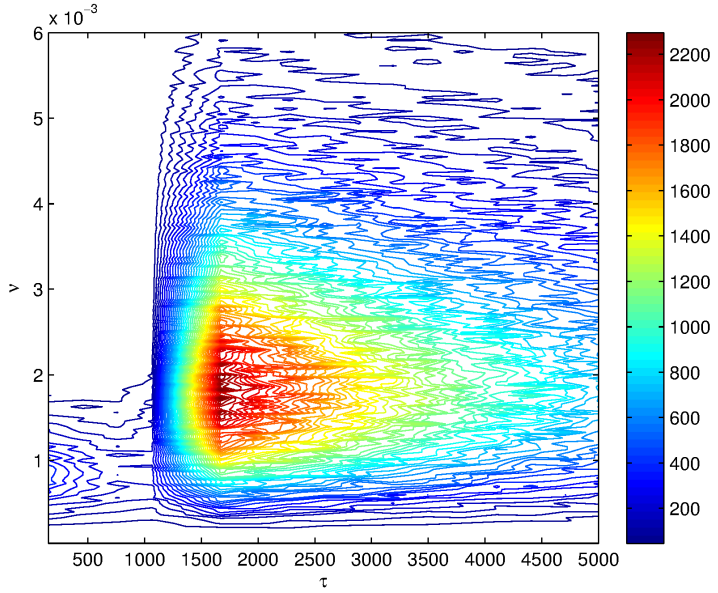

Figure A.1.12: Contour plot of the joint posterior distribution of τ and ν, $\pi(\tau, \nu | x_1, \ldots, x_n; \sigma)$, obtained by MCMC methods with σ considered known, and the endpoint of observation was $t_n = 1\ 400$.

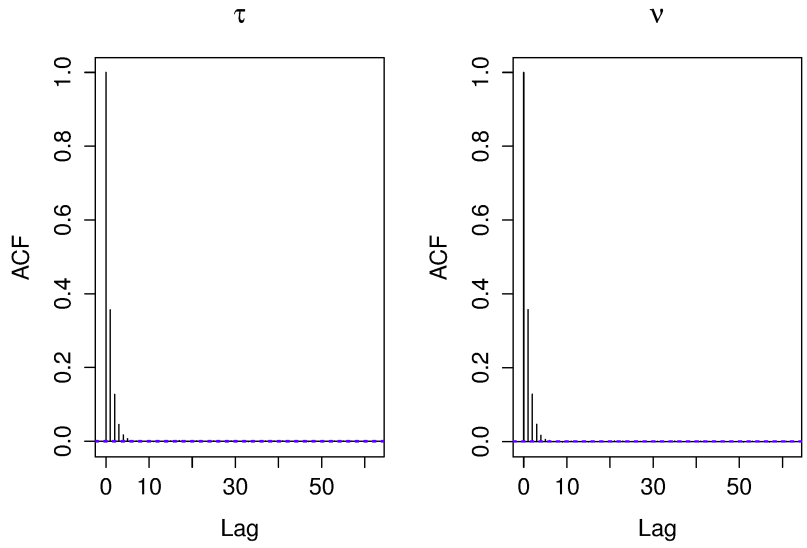

Figure A.1.13: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ, $\pi(\tau | x_1, \ldots, x_n; \sigma)$, and ν, $\pi(\nu | x_1, \ldots, x_n; \sigma)$, as function of the lag. Here, σ is considered known, and the endpoint of observation was $t_n = 1\ 400$.

Observation until $t_n = 1\ 700$

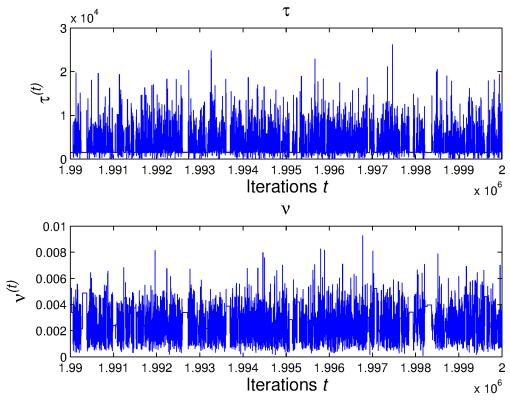

Figure A.1.14: Marginal trace plot of the last 100 000 iterations of the Markov chain from the posterior distribution of τ, $\pi(\tau|x_1, \ldots, x_n; \sigma)$, and ν, $\pi(\nu|x_1, \ldots, x_n; \sigma)$. Here, σ was considered known, and the endpoint of observation was $t_n = 1\ 700$.

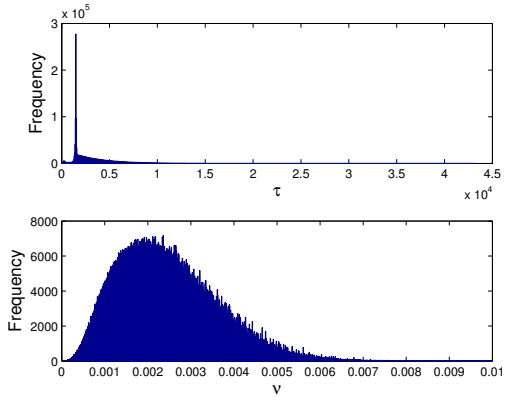

Figure A.1.15: Histograms of the MCMC sample from the marginal posterior distributions of τ, $\pi(\tau|x_1, \ldots, x_n; \sigma)$, and ν, $\pi(\nu|x_1, \ldots, x_n; \sigma)$. Here, σ was considered known, and the endpoint of observation was $t_n = 1\ 700$.

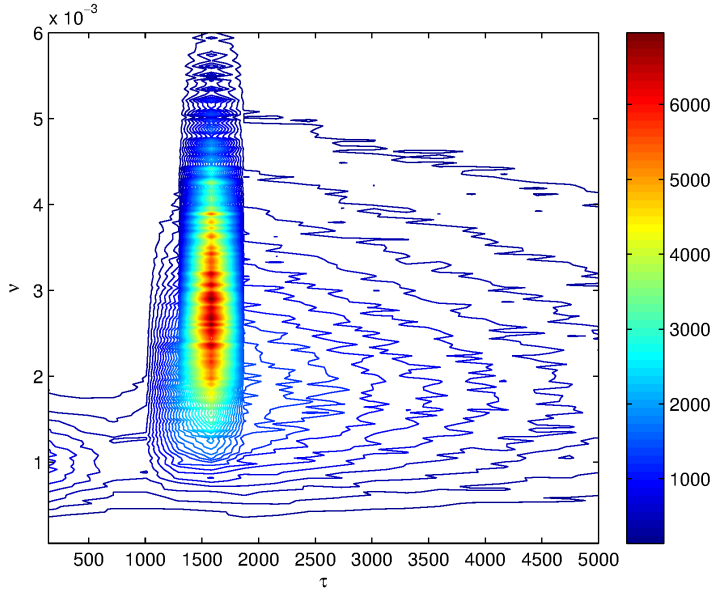

Figure A.1.16: Contour plot of the joint posterior distribution of τ and ν, $\pi(\tau, \nu | x_1, \ldots, x_n; \sigma)$, obtained by MCMC methods with σ considered known, and the endpoint of observation was $t_n = 1\ 700$.

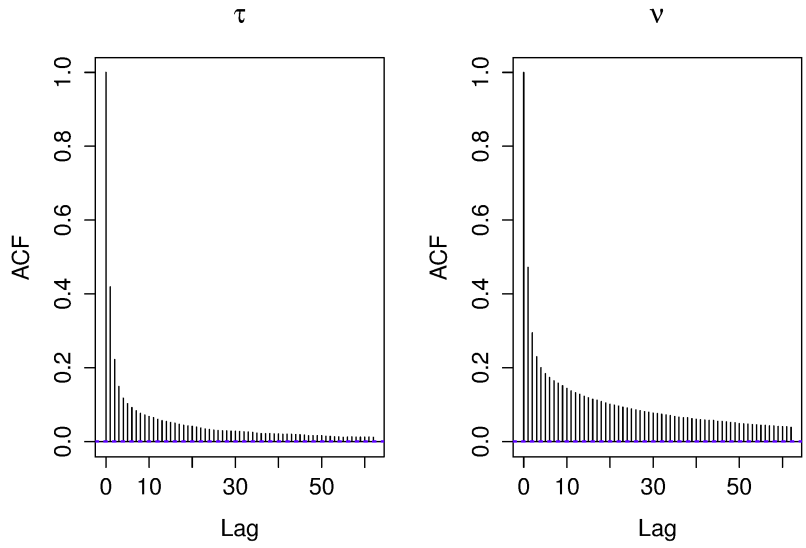

Figure A.1.17: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ, $\pi(\tau | x_1, \ldots, x_n; \sigma)$, and ν, $\pi(\nu | x_1, \ldots, x_n; \sigma)$, as function of the lag. Here, σ is considered known, and the endpoint of observation was $t_n = 1\ 700$.

Observation until $t_n = 3\ 000$

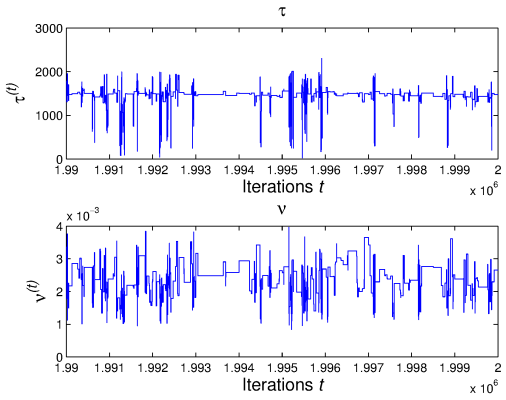

Figure A.1.18: Marginal trace plot of the last 100 000 iterations of the Markov chain from the posterior distribution of τ, $\pi(\tau|x_1, \ldots, x_n; \sigma)$, and ν, $\pi(\nu|x_1, \ldots, x_n; \sigma)$. Here, σ was considered known, and the endpoint of observation was $t_n = 3\ 000$.

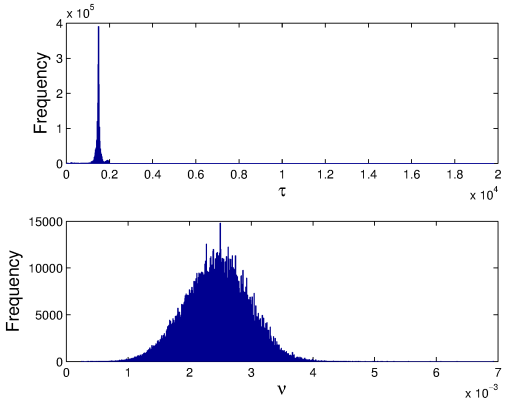

Figure A.1.19: Histograms of the MCMC sample from the marginal posterior distributions of τ, $\pi(\tau|x_1, \ldots, x_n; \sigma)$, and ν, $\pi(\nu|x_1, \ldots, x_n; \sigma)$. Here, σ was considered known, and the endpoint of observation was $t_n = 3\ 000$.

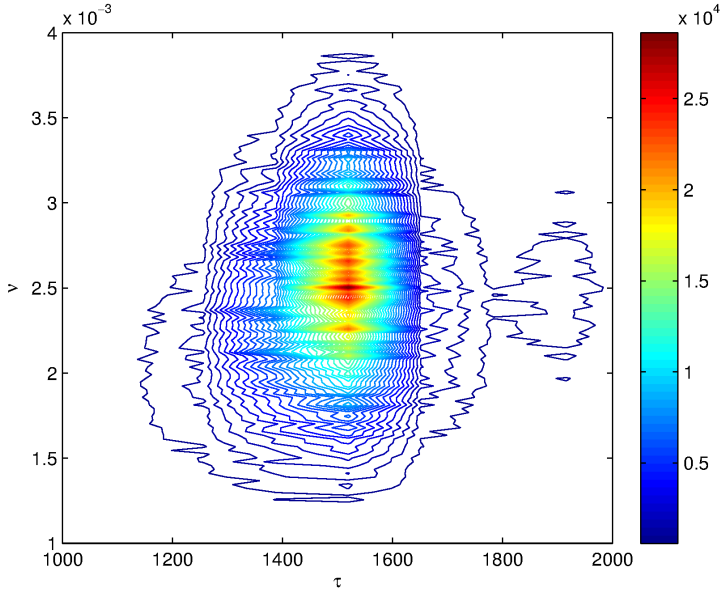

Figure A.1.20: Contour plot of the joint posterior distribution of τ and ν, $\pi(\tau, \nu | x_1, \ldots, x_n; \sigma)$, obtained by MCMC methods with σ considered known, and the endpoint of observation was $t_n = 3\ 000$.

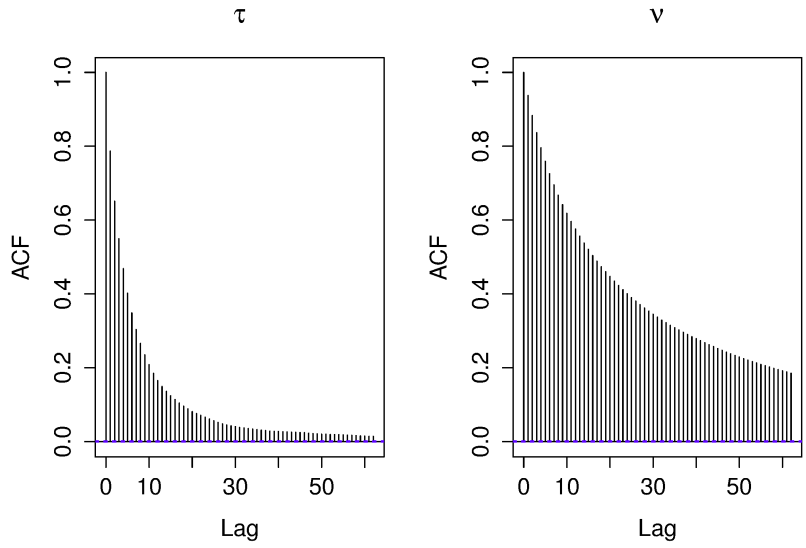

Figure A.1.21: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ, $\pi(\tau | x_1, \ldots, x_n; \sigma)$, and ν, $\pi(\nu | x_1, \ldots, x_n; \sigma)$, as function of the lag. Here, σ is considered known, and the endpoint of observation was $t_n = 3\ 000$.

A.1.3 τ, ν and σ unknown

Observation until t_n = 1 400

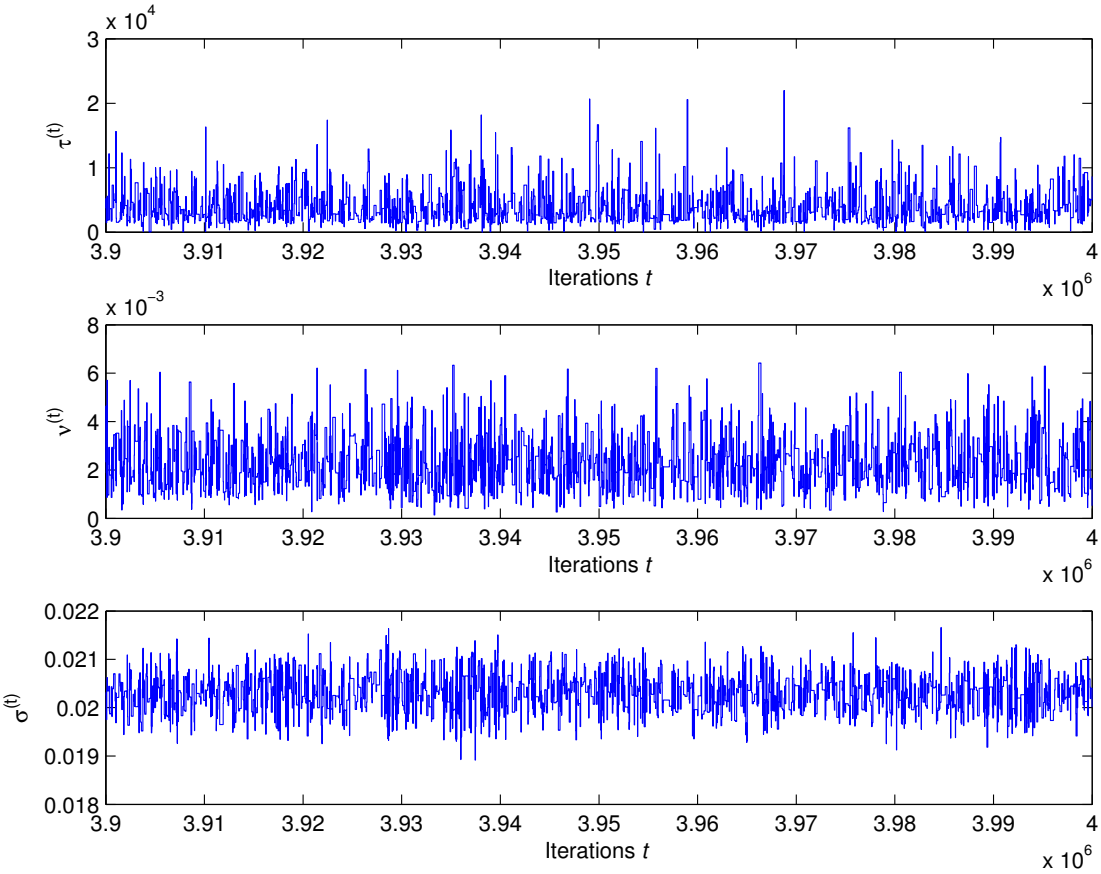

Figure A.1.22: Trace plot of the last 100 000 iterations of the Markov chain from the marginal posterior distributions of τ, $\pi(\tau|x_1,\ldots,x_n)$, ν, $\pi(\nu|x_1,\ldots,x_n)$, and σ, $\pi(\sigma|x_1,\ldots,x_n)$, and the endpoint of observation was $t_n = 1\ 400$.

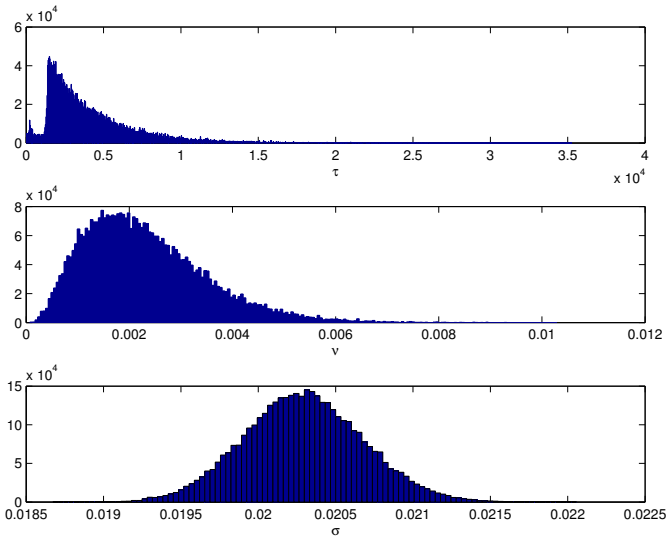

Figure A.1.23: Histograms of the MCMC sample from the marginal posterior distributions of τ, $\pi(\tau|x_1,\ldots,x_n)$, ν, $\pi(\nu|x_1,\ldots,x_n)$, and σ, $\pi(\sigma|x_1,\ldots,x_n)$, and the endpoint of observation was $t_n = 1\ 400$.

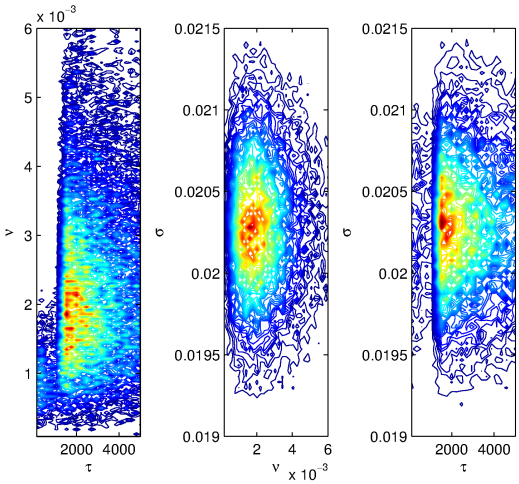

Figure A.1.24: Contour plot of the pairwise joint posterior distributions of τ, ν and σ, $\pi(\tau,\nu|x_1,\ldots,x_n)$, $\pi(\nu,\sigma|x_1,\ldots,x_n)$ and $\pi(\tau,\sigma|x_1,\ldots,x_n)$, obtained by MCMC methods, and the endpoint of observation was $t_n = 1\ 400$.

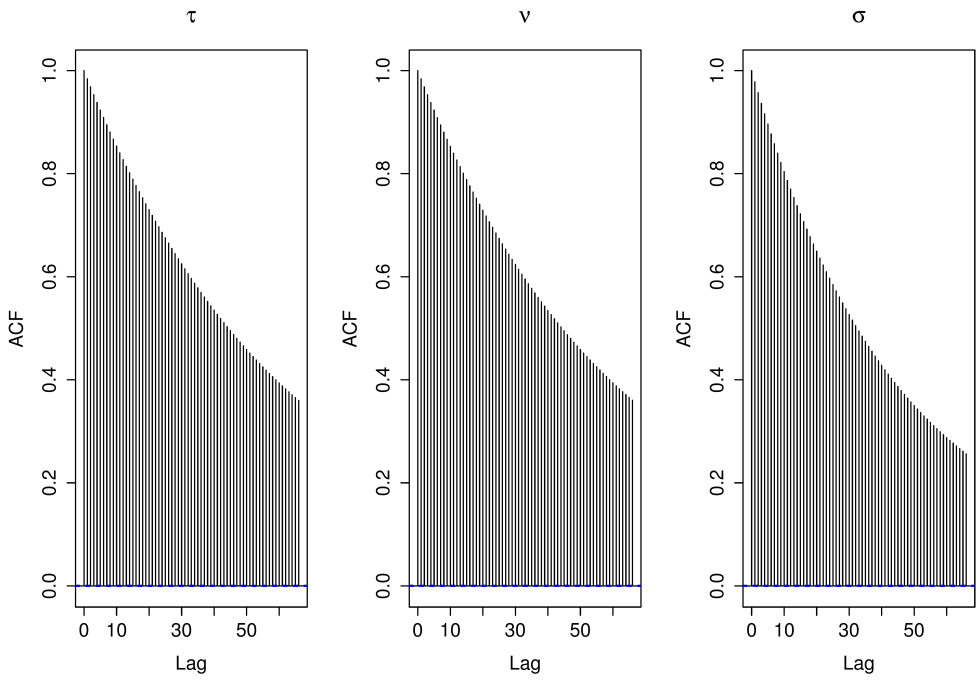

Figure A.1.25: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ, $\pi(\tau|x_1,\ldots,x_n)$, ν, $\pi(\nu|x_1,\ldots,x_n)$ and σ, $\pi(\sigma|x_1,\ldots,x_n)$, as function of the lag. The endpoint of observation was $t_n = 1\ 400$.

Observation until t_n = 1 700

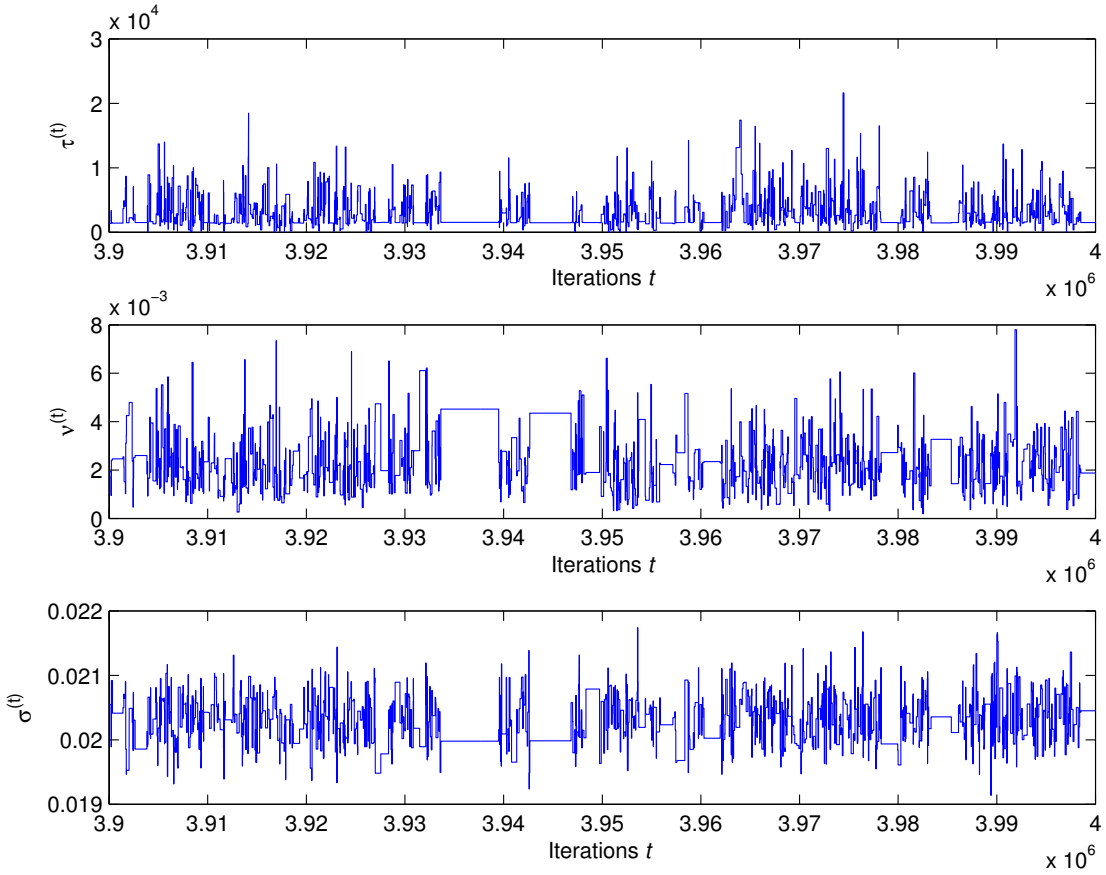

Figure A.1.26: Trace plot of the last 100 000 iterations of the Markov chain from the marginal posterior distributions of τ, $\pi(\tau|x_1,\ldots,x_n)$, ν, $\pi(\nu|x_1,\ldots,x_n)$, and σ, $\pi(\sigma|x_1,\ldots,x_n)$, and the endpoint of observation was $t_n = 1\ 700$.

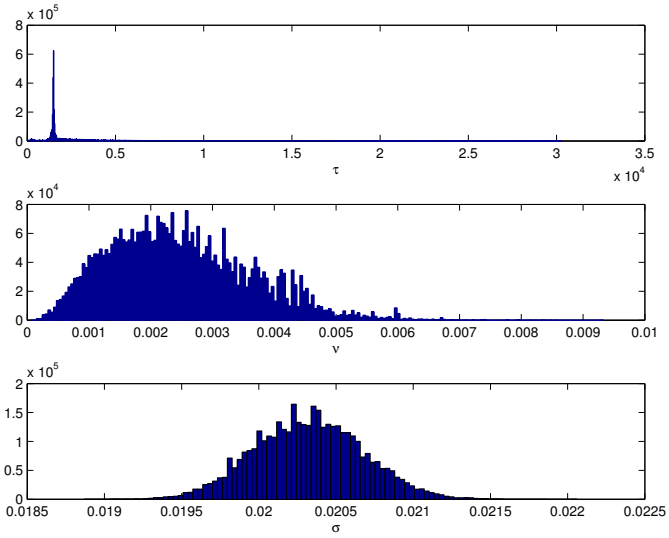

Figure A.1.27: Histograms of the MCMC sample from the marginal posterior distributions of τ, $\pi(\tau|x_1,\ldots,x_n)$, ν, $\pi(\nu|x_1,\ldots,x_n)$, and σ, $\pi(\sigma|x_1,\ldots,x_n)$, and the endpoint of observation was $t_n = 1\ 700$.

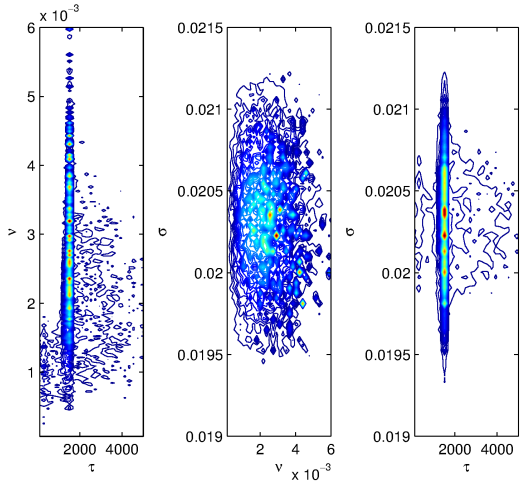

Figure A.1.28: Contour plot of the pairwise joint posterior distributions of τ, ν and σ, $\pi(\tau,\nu|x_1,\ldots,x_n)$, $\pi(\nu,\sigma|x_1,\ldots,x_n)$ and $\pi(\tau,\sigma|x_1,\ldots,x_n)$, obtained by MCMC methods, and the endpoint of observation was $t_n = 1\ 700$.

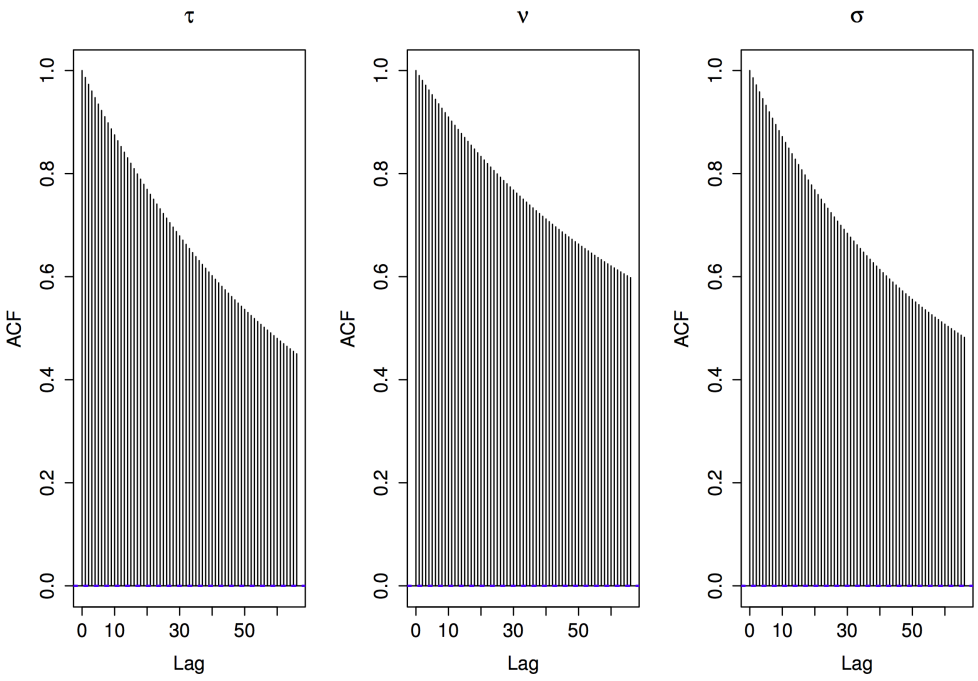

Figure A.1.29: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ, $\pi(\tau|x_1,\ldots,x_n)$, ν, $\pi(\nu|x_1,\ldots,x_n)$ and σ, $\pi(\sigma|x_1,\ldots,x_n)$, as function of the lag. The endpoint of observations was $t_n = 1\ 700$.

Observation until t_n = 3 000

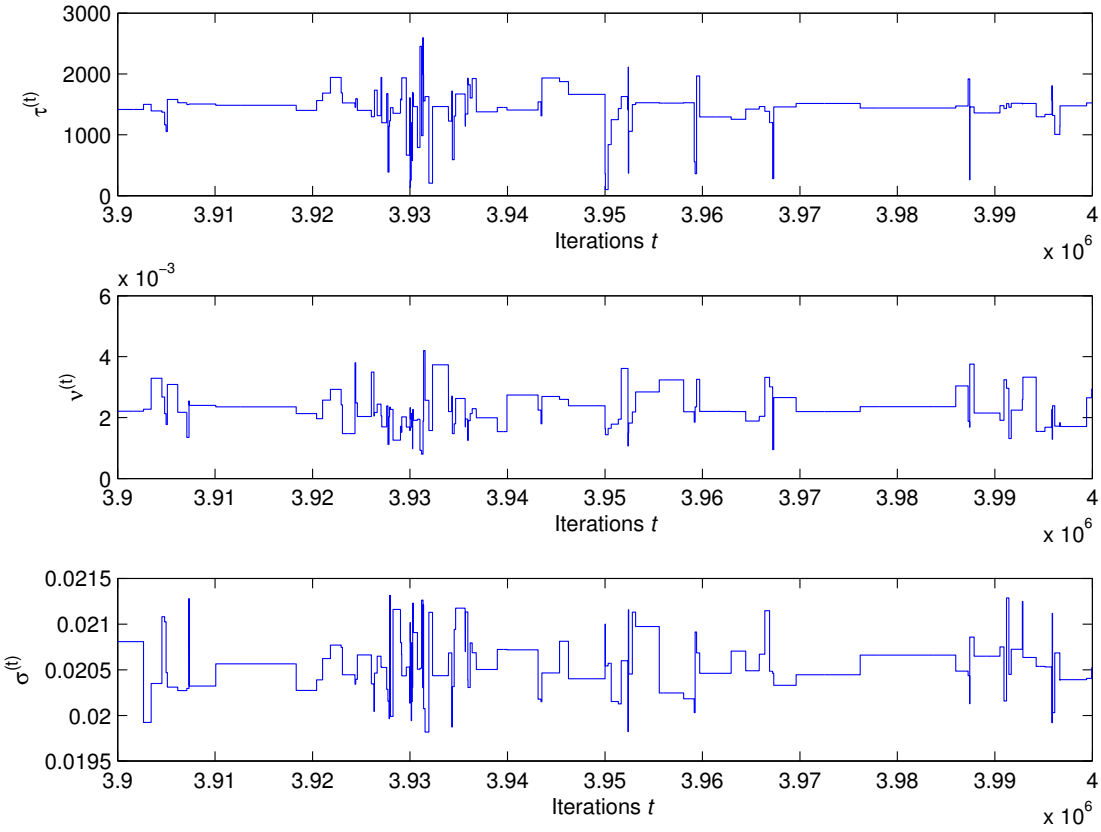

Figure A.1.30: Trace plot of the last 100 000 iterations of the Markov chain from the marginal posterior distributions of τ, $\pi(\tau|x_1, \ldots, x_n)$, ν, $\pi(\nu|x_1, \ldots, x_n)$, and σ, $\pi(\sigma|x_1, \ldots, x_n)$, and the endpoint of observation was $t_n = 3\ 000$.

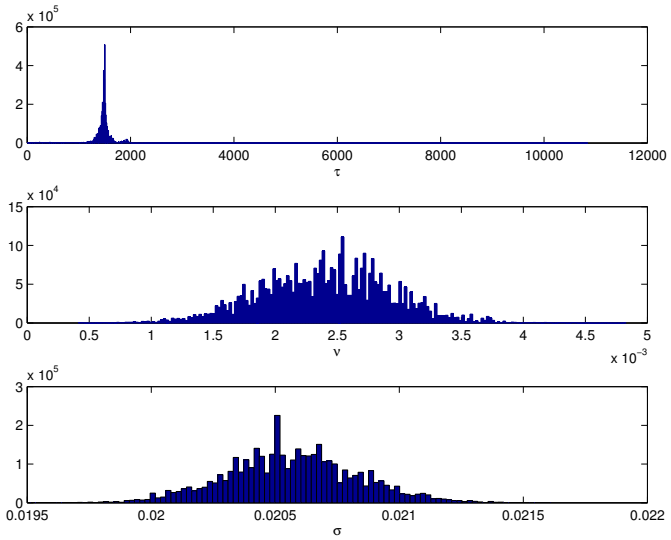

Figure A.1.31: Histograms of the MCMC sample from the marginal posterior distributions of τ, $\pi(\tau|x_1,\ldots,x_n)$, ν, $\pi(\nu|x_1,\ldots,x_n)$, and σ, $\pi(\sigma|x_1,\ldots,x_n)$, and the endpoint of observation was $t_n = 3\ 000$.

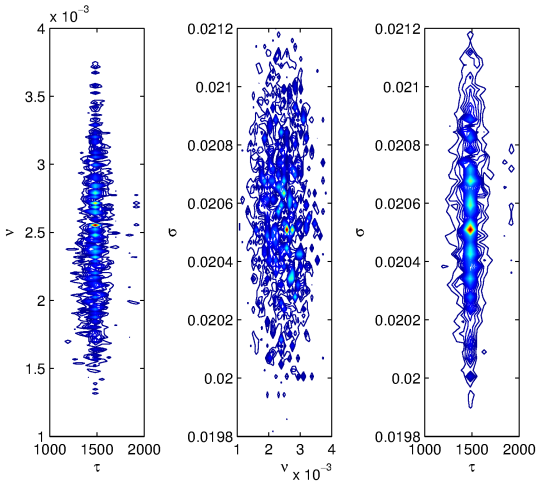

Figure A.1.32: Contour plot of the pairwise joint posterior distributions of τ, ν and σ, $\pi(\tau,\nu|x_1,\ldots,x_n)$, $\pi(\nu,\sigma|x_1,\ldots,x_n)$ and $\pi(\tau,\sigma|x_1,\ldots,x_n)$, obtained by MCMC methods, and the endpoint of observation was $t_n = 3\ 000$.

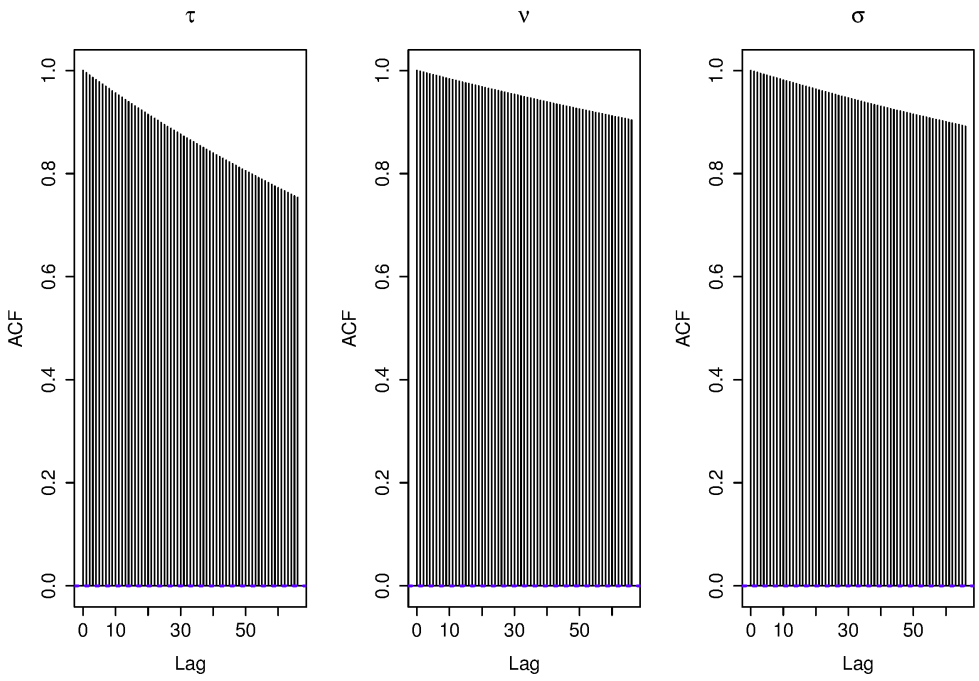

Figure A.1.33: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ, $\pi(\tau|x_1,\ldots,x_n)$, ν, $\pi(\nu|x_1,\ldots,x_n)$ and σ, $\pi(\sigma|x_1,\ldots,x_n)$, as function of the lag. The endpoint of observation was $t_n = 3\ 000$.

A.2 Two change points

A.2.1 τ_1 and τ_2 unknown, ν, σ and α known

Observation until $t_n = 700$

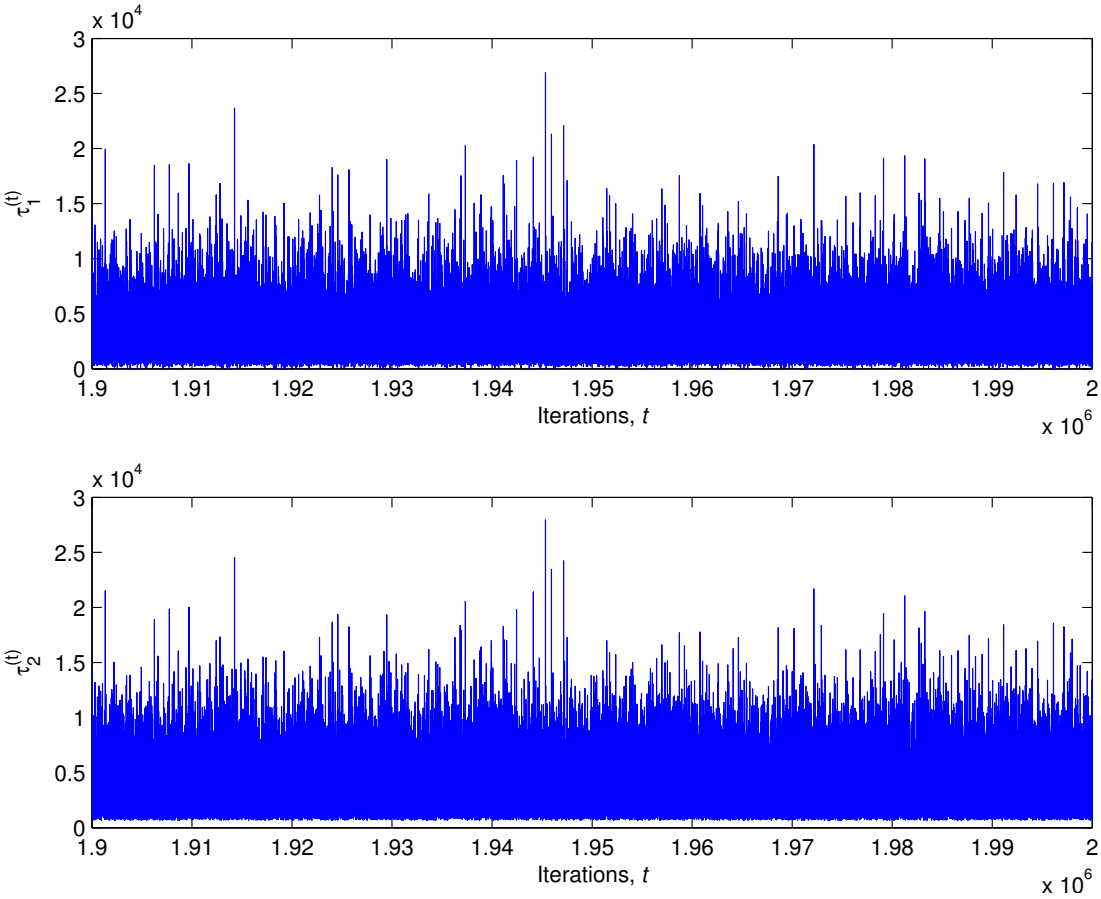

Figure A.2.1: Trace plot of the last 100 000 iterations of the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1, \ldots, x_n; \alpha, \nu, \sigma)$ and τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$. Here, ν, σ and α are considered known, and the endpoint of observation was $t_n = 700$.

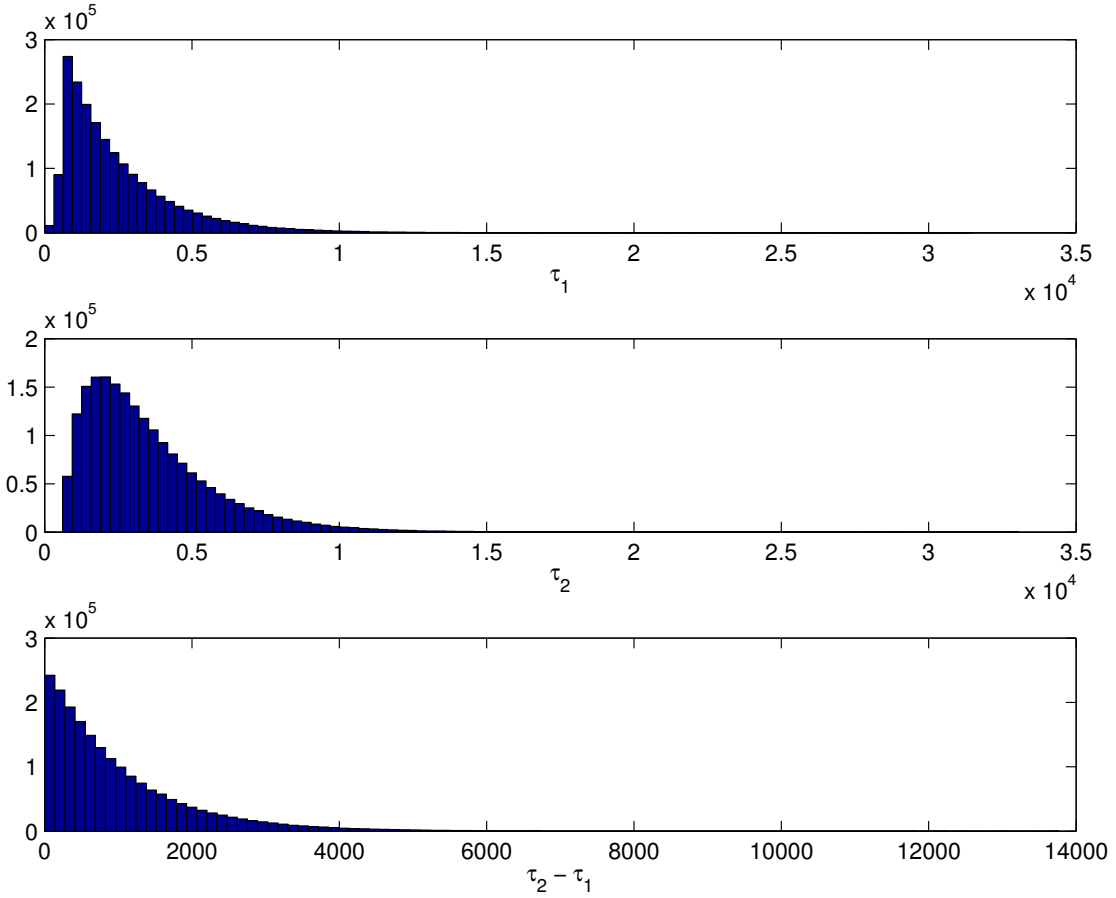

Figure A.2.2: Histograms of the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\nu,\sigma)$, τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\nu,\sigma)$ and $\tau_2-\tau_1$, $\pi(\tau_2-\tau_1|x_1,\ldots,x_n;\alpha,\nu,\sigma)$. Here, ν, σ and α are considered known, and the endpoint of observation was $t_n = 700$.

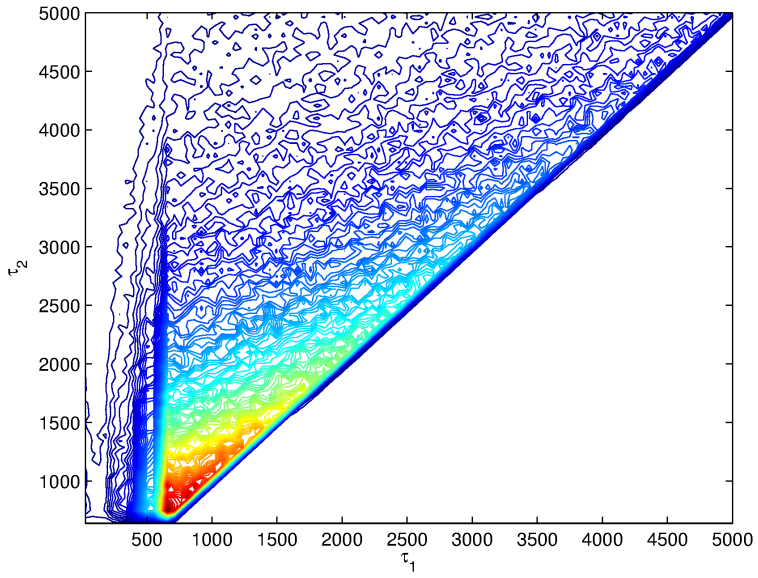

Figure A.2.3: Contour plot of the joint posterior distribution of τ_1 and τ_2, $\pi(\tau_1, \tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, as obtained by MCMC methods. Here, ν, σ and α were considered known, and the endpoint of observation was $t_n = 700$.

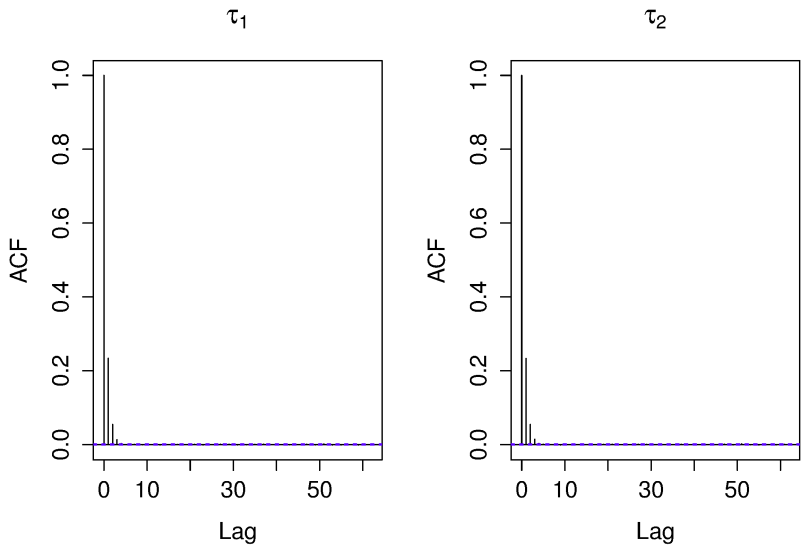

Figure A.2.4: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, and τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, as function of the lag. Here, ν and σ are considered known, and the endpoint of observation was $t_n = 700$.

Observation until $t_n = 1\ 100$

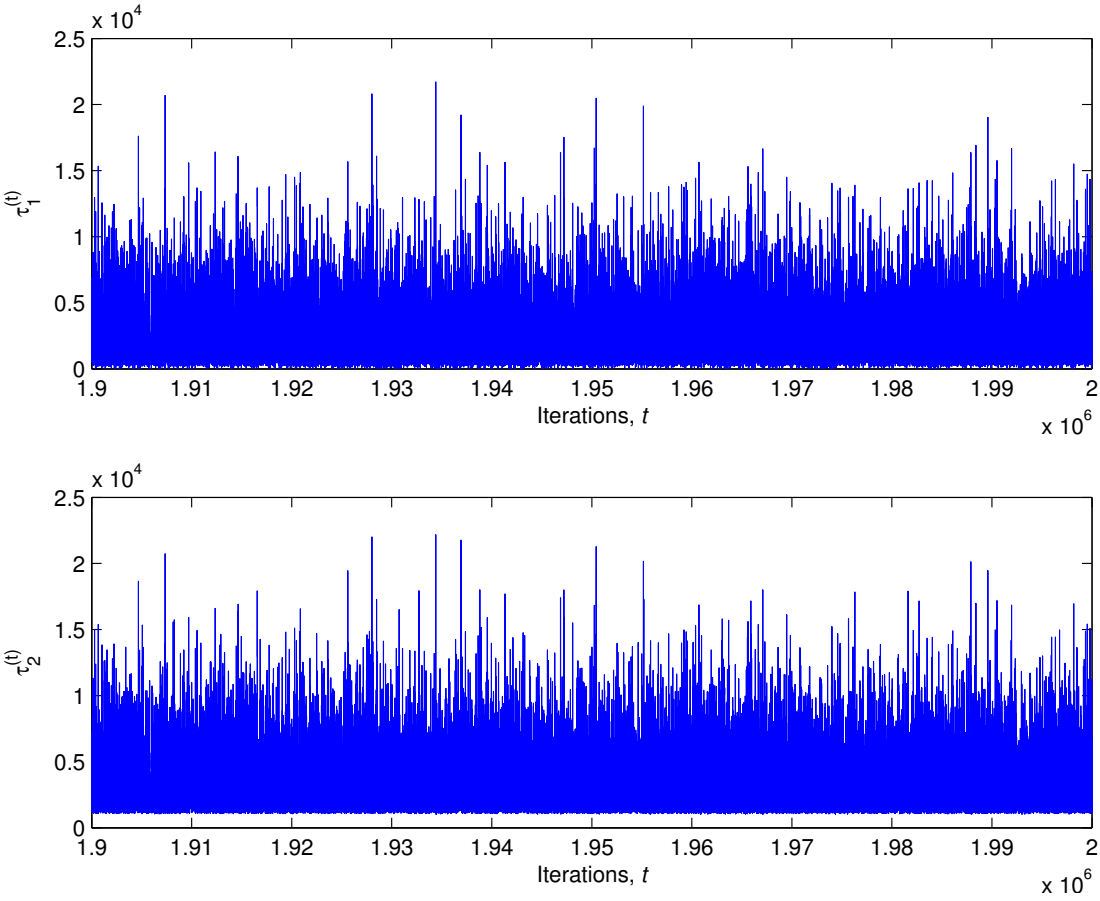

Figure A.2.5: Trace plot of the last 100 000 iterations of the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1, \ldots, x_n; \alpha, \nu, \sigma)$ and τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$. Here, ν, σ and α are considered known, and the endpoint of observation was $t_n = 1\ 100$.

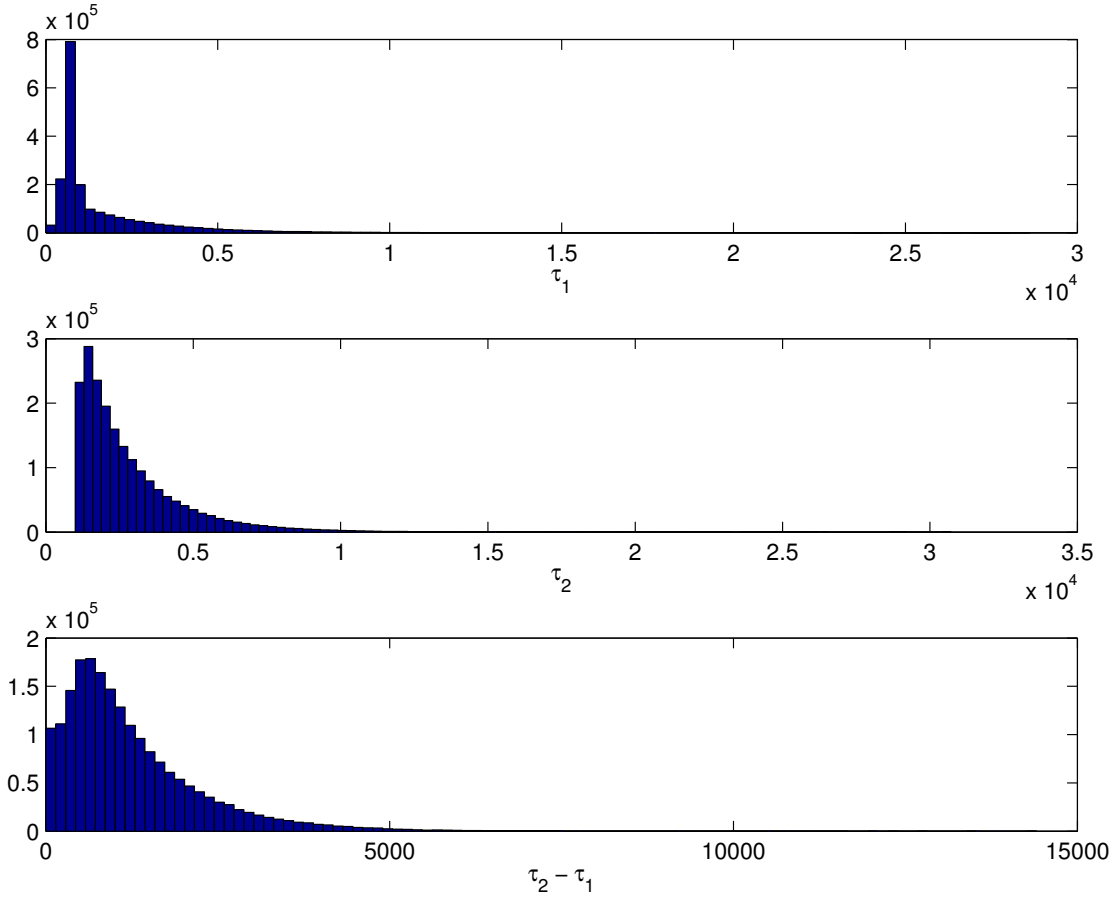

Figure A.2.6: Histograms of the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\nu,\sigma)$, τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\nu,\sigma)$ and $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1|x_1,\ldots,x_n;\alpha,\nu,\sigma)$. Here, ν, σ and α are considered known, and the endpoint of observation was $t_n = 1\ 100$.

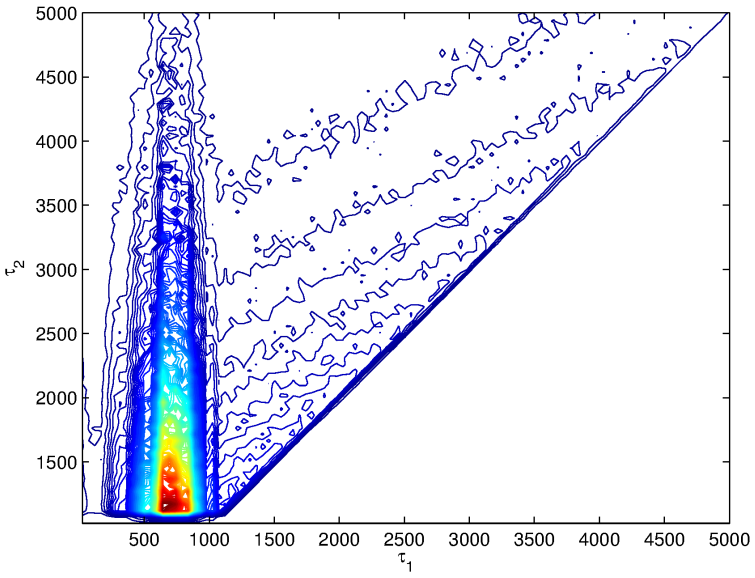

Figure A.2.7: Contour plot of the joint posterior distribution of τ_1 and τ_2, $\pi(\tau_1, \tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, as obtained by MCMC methods. Here, ν, σ and α were considered known, and the endpoint of observation was $t_n = 1\ 100$.

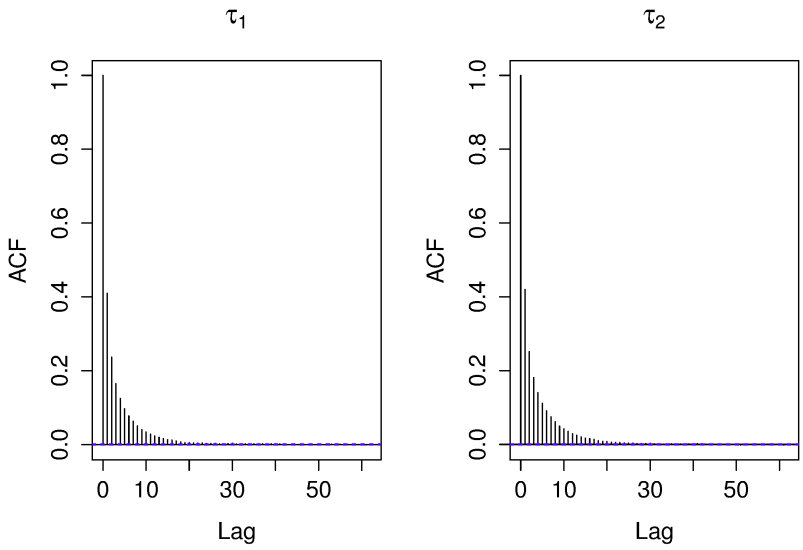

Figure A.2.8: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, and τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \nu, \sigma)$, as function of the lag. Here, ν and σ are considered known, and the endpoint of observation was $t_n = 1\ 100$.

Observation until t_n = 1 450

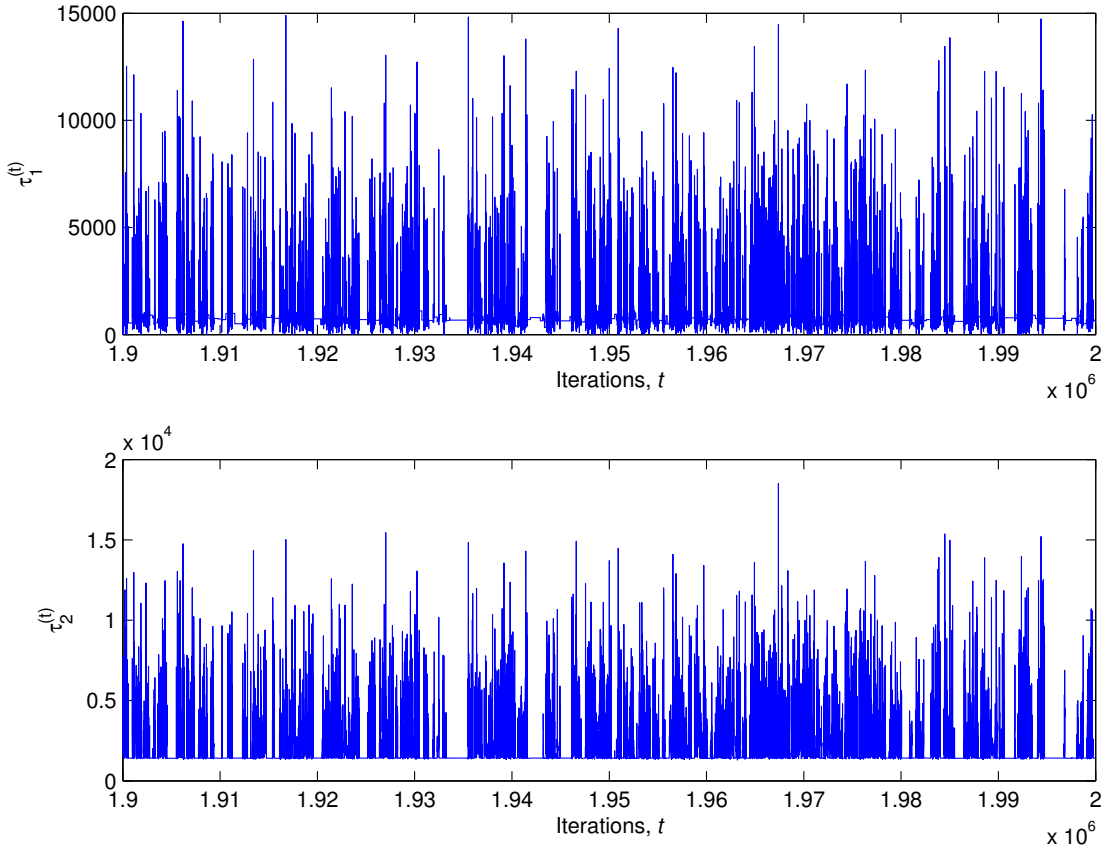

Figure A.2.9: Trace plot of the last 100 000 iterations of the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\nu,\sigma)$ and τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\nu,\sigma)$. Here, ν, σ and α are considered known, and the endpoint of observation was $t_n = 1\ 450$.

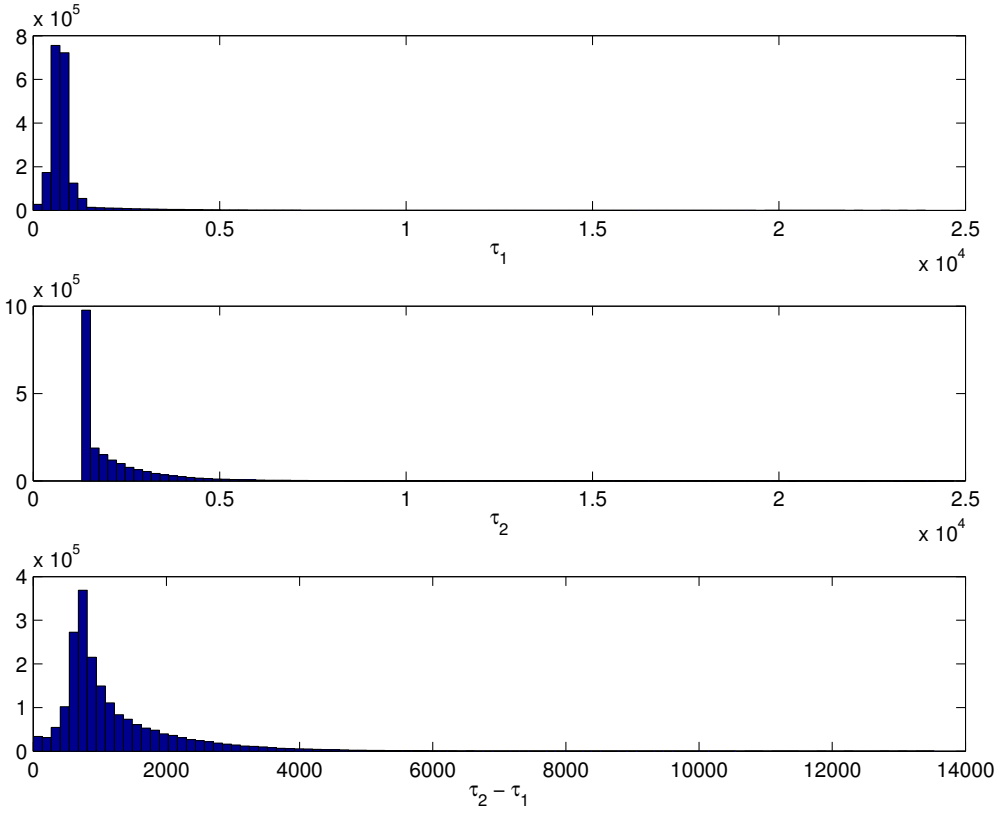

Figure A.2.10: Histograms of the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\nu,\sigma)$, τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\nu,\sigma)$ and $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1|x_1,\ldots,x_n;\alpha,\nu,\sigma)$. Here, ν, σ and α are considered known, and the endpoint of observation was $t_n = 1\ 450$.

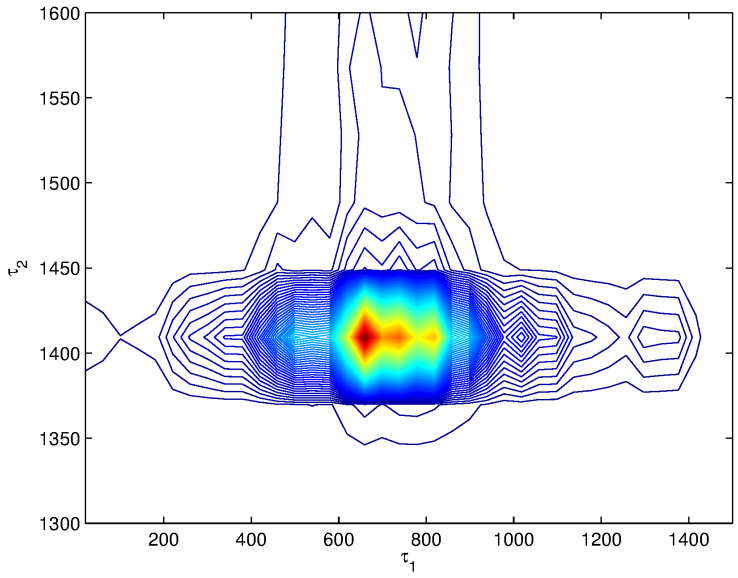

Figure A.2.11: Contour plot of the joint posterior distribution of τ_1 and τ_2, $\pi(\tau_1, \tau_2 | x_1, \ldots, x_n; \alpha, \nu, \sigma)$, as obtained by MCMC methods. Here, ν, σ and α were considered known, and the endpoint of observation was $t_n = 1\ 450$.

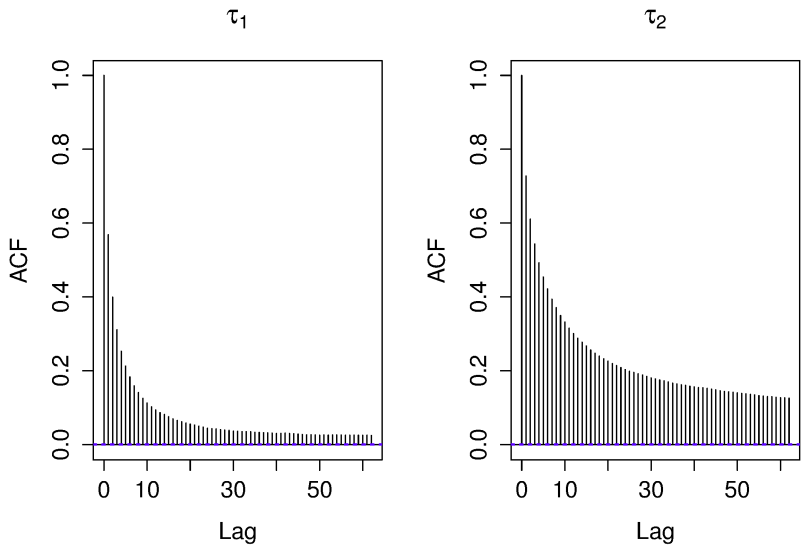

Figure A.2.12: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1 | x_1, \ldots, x_n; \alpha, \nu, \sigma)$, and τ_2, $\pi(\tau_2 | x_1, \ldots, x_n; \alpha, \nu, \sigma)$, as function of the lag. Here, ν and σ are considered known, and the endpoint of observation was $t_n = 1\ 450$.

Observation until t_n = 2 000

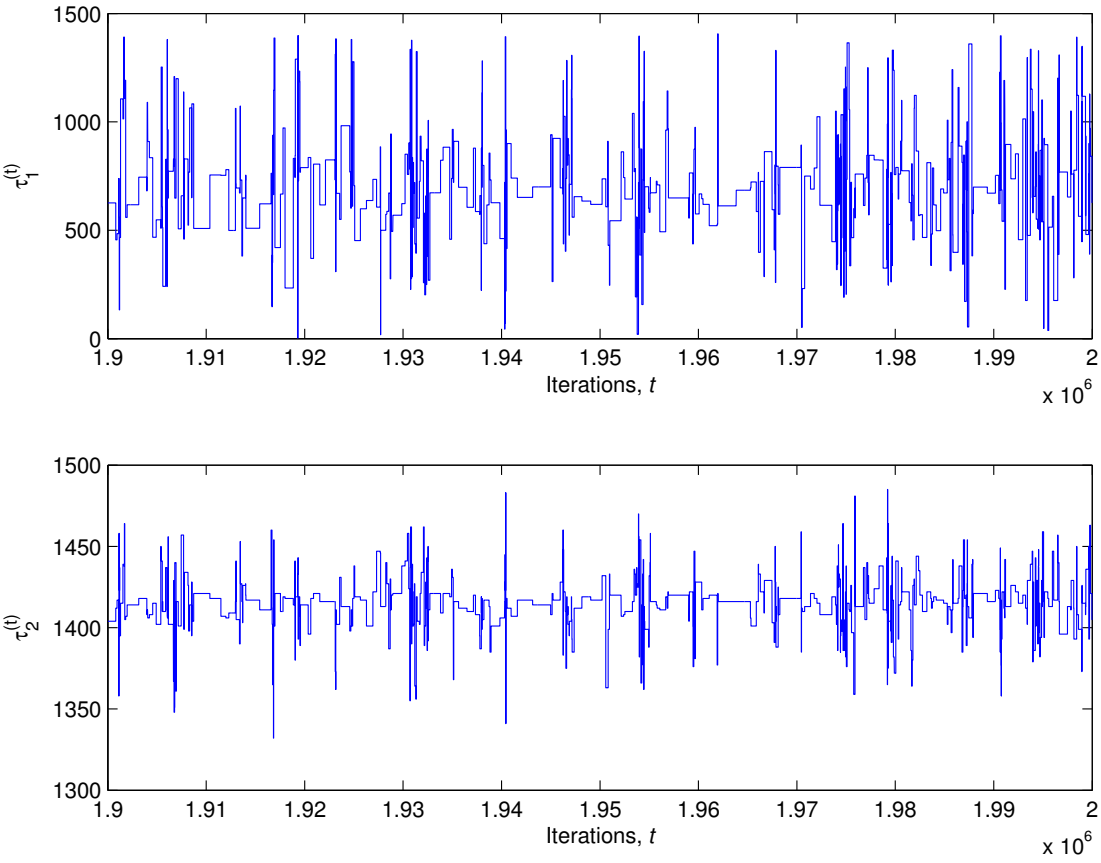

Figure A.2.13: Trace plot of the last 100 000 iterations of the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\nu,\sigma)$ and τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\nu,\sigma)$. Here, ν, σ and α are considered known, and the endpoint of observation was $t_n = 2\ 000$.

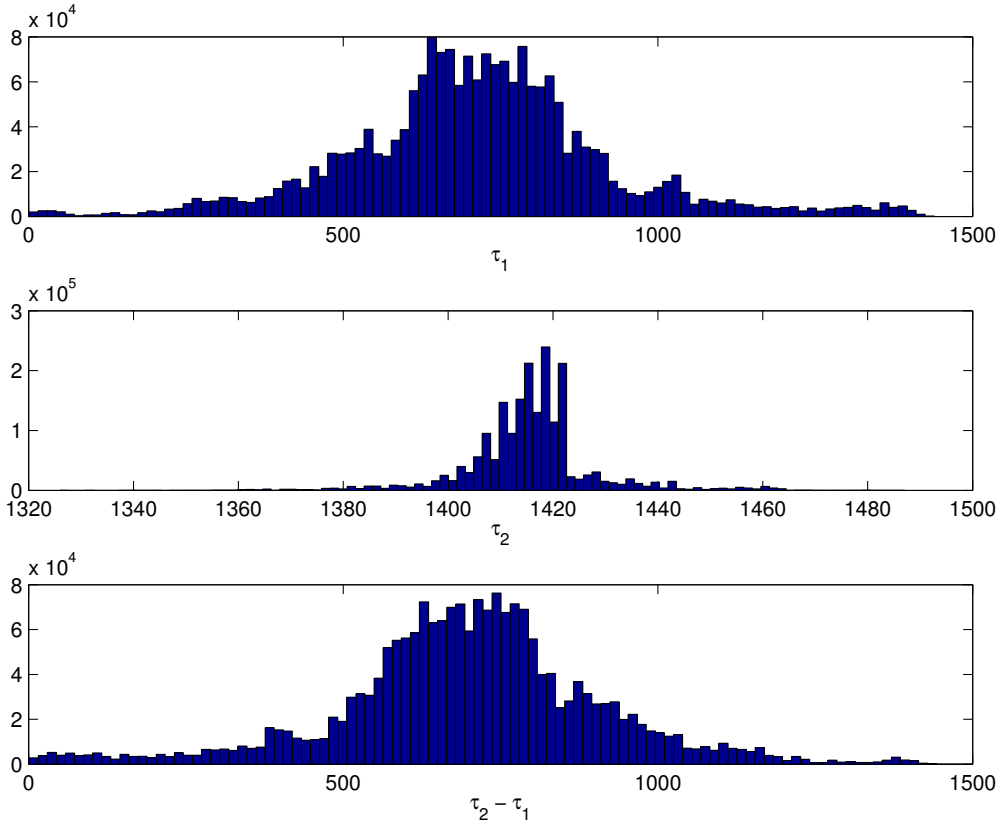

Figure A.2.14: Histograms of the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\nu,\sigma)$, τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\nu,\sigma)$ and $\tau_2 - \tau_1$, $\pi(\tau_2 - \tau_1|x_1,\ldots,x_n;\alpha,\nu,\sigma)$. Here, ν, σ and α are considered known, and the endpoint of observation was $t_n = 2\ 000$.

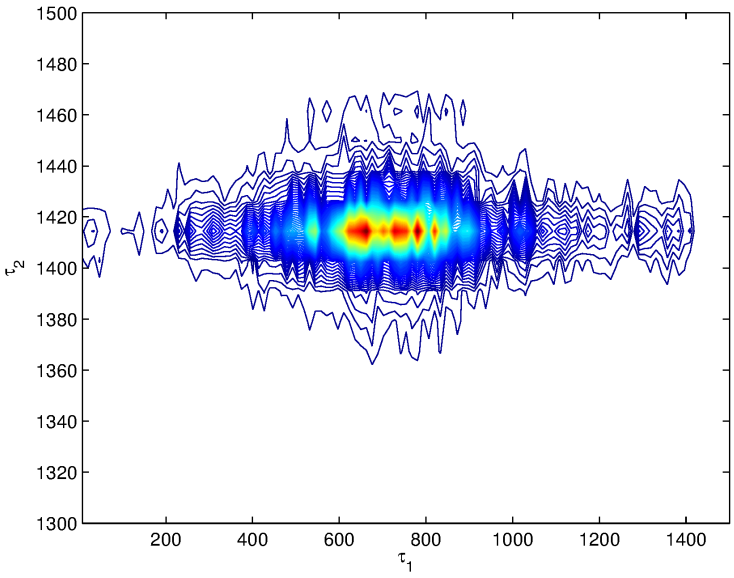

Figure A.2.15: Contour plot of the joint posterior distribution of τ_1 and τ_2, $\pi(\tau_1, \tau_2 | x_1, \ldots, x_n; \alpha, \nu, \sigma)$, as obtained by MCMC methods. Here, ν, σ and α were considered known, and the endpoint of observation was $t_n = 2\ 000$.

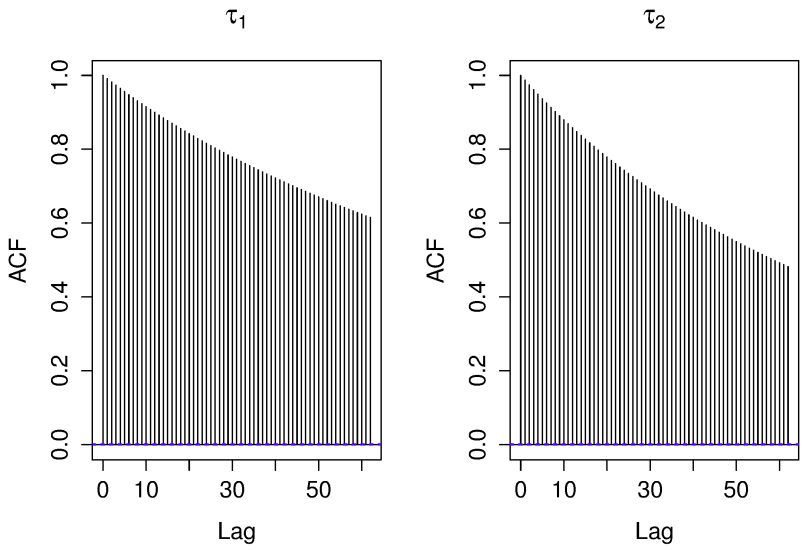

Figure A.2.16: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1 | x_1, \ldots, x_n; \alpha, \nu, \sigma)$, and τ_2, $\pi(\tau_2 | x_1, \ldots, x_n; \alpha, \nu, \sigma)$, as function of the lag. Here, ν and σ are considered known, and the endpoint of observation was $t_n = 2\ 000$.

A.2.2 τ_1, τ_2 and ν unknown, σ and α known

Observation until $t_n = 700$

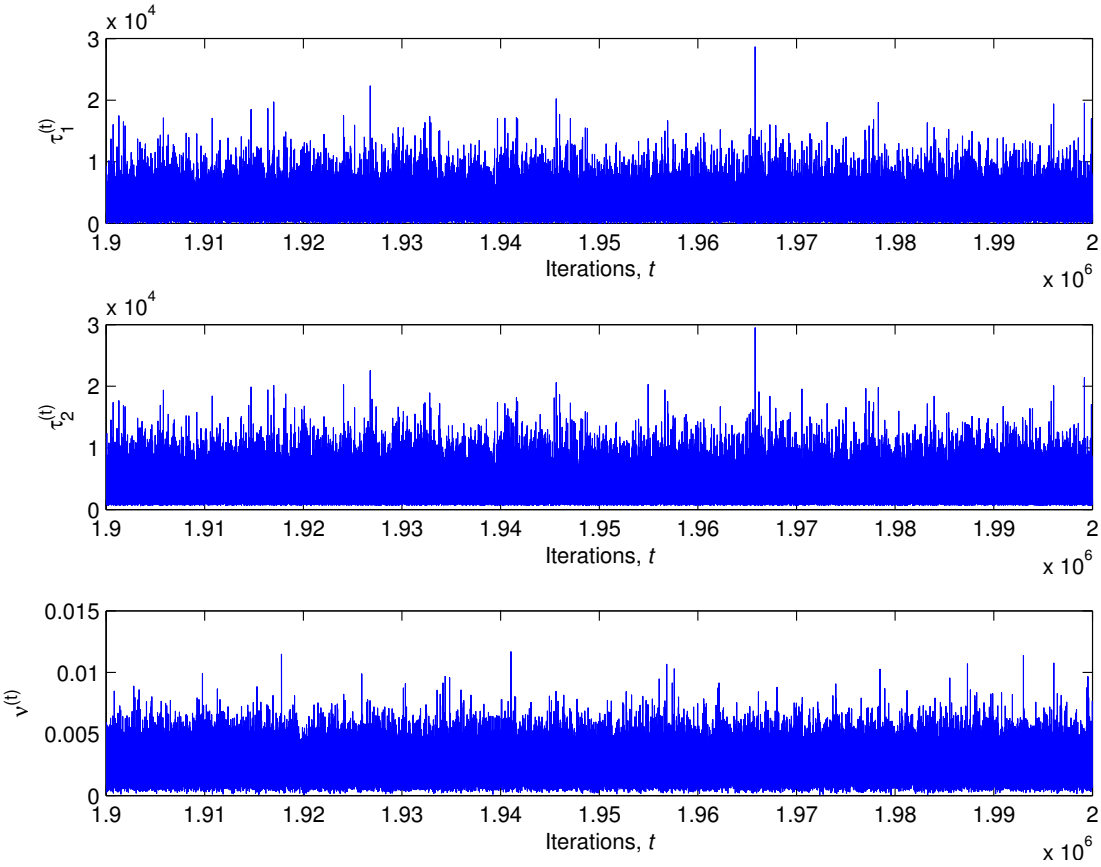

Figure A.2.17: Trace plot of the last 100 000 iterations of the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \sigma)$ and ν, $\pi(\nu|x_1, \ldots, x_n; \alpha, \sigma)$. Here, σ and α were considered known, and the endpoint of observations was $t_n = 700$.

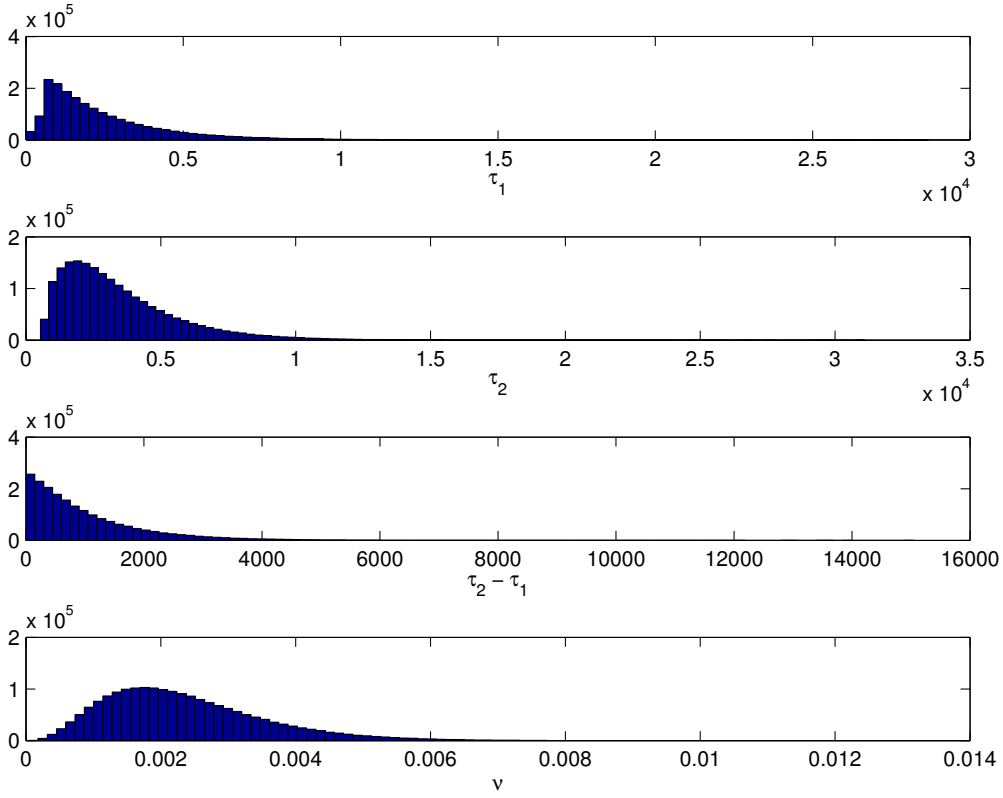

Figure A.2.18: Histograms of the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\sigma)$, τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\sigma)$, $\tau_2-\tau_1$, $\pi(\tau_2-\tau_1|x_1,\ldots,x_n;\alpha,\sigma)$ and ν, $\pi(\nu|x_1,\ldots,x_n;\alpha,\sigma)$. Here, σ and α were considered known, and the endpoint of observation was $t_n = 700$.

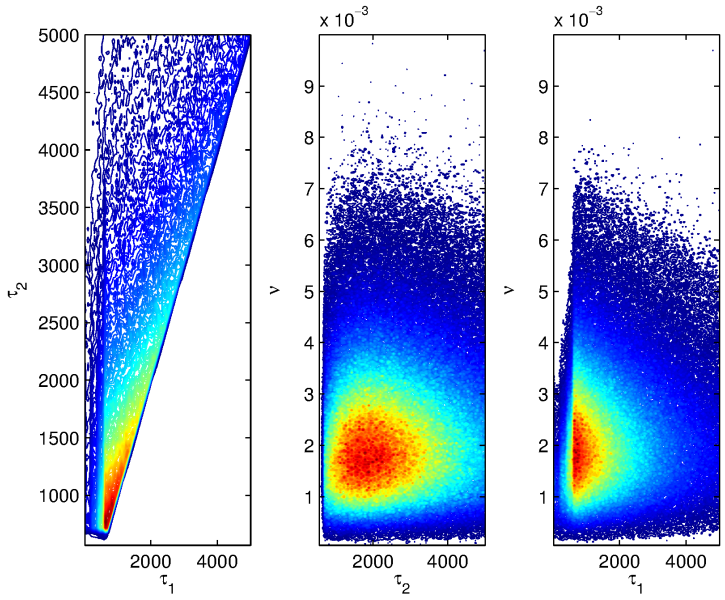

Figure A.2.19: Contour plot of the pairwise joint posterior distributions of τ_1, τ_2 and ν, $\pi(\tau_1, \tau_2|x_1, \ldots, x_n; \alpha, \sigma)$, $\pi(\tau_2, \nu|x_1, \ldots, x_n; \alpha, \sigma)$ and $\pi(\tau_1, \nu|x_1, \ldots, x_n; \alpha, \sigma)$. Here, σ and α were considered known, and the endpoint of observations was $t_n = 700$.

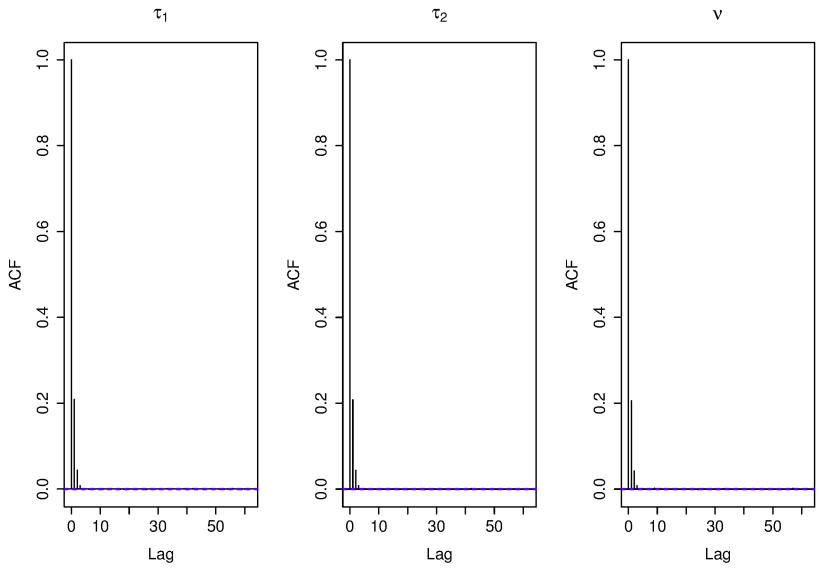

Figure A.2.20: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \sigma)$, and ν, $\pi(\nu|x_1, \ldots, x_n; \alpha, \sigma)$as function of the lag. Here, ν and σ are considered known, and the endpoint of observation was $t_n = 700$.

Observation until t_n = 1 100

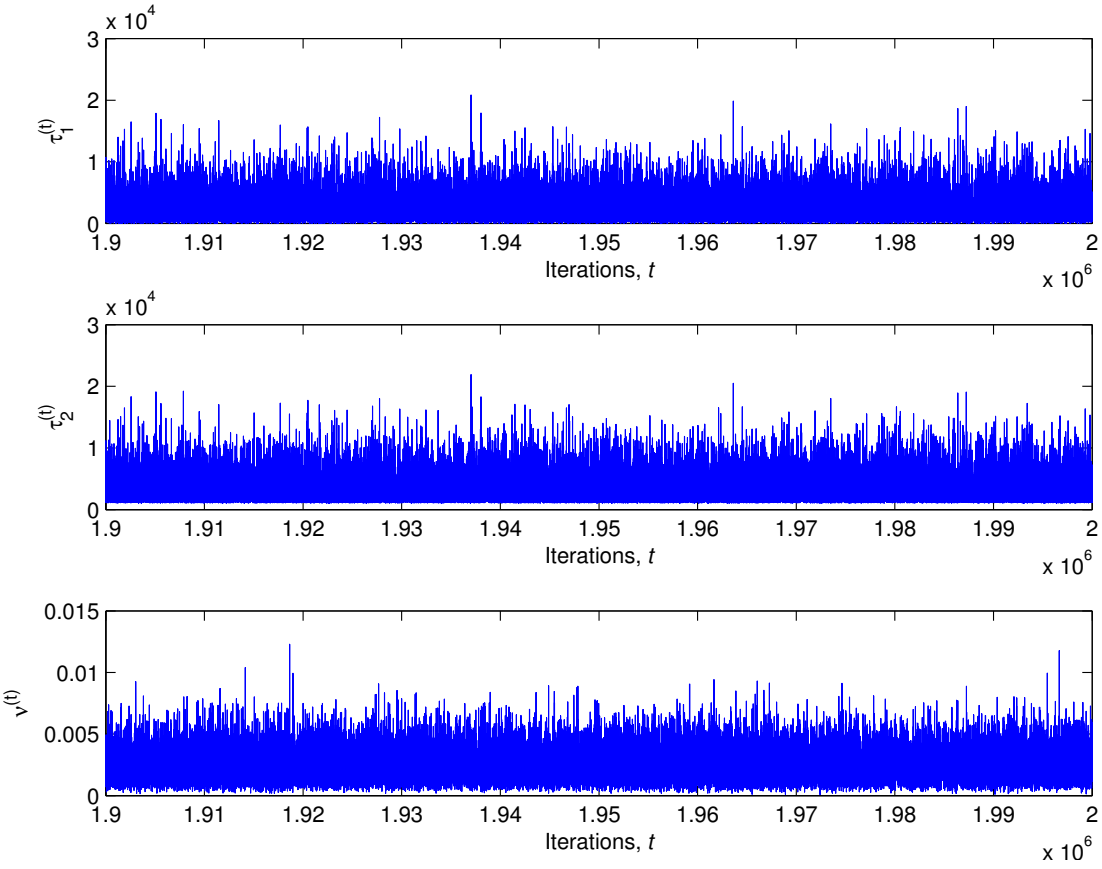

Figure A.2.21: Trace plot of the last 100 000 iterations of the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\sigma)$, τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\sigma)$ and ν, $\pi(\nu|x_1,\ldots,x_n;\alpha,\sigma)$. Here, σ and α were considered known, and the endpoint of observations was $t_n = 1\ 100$.

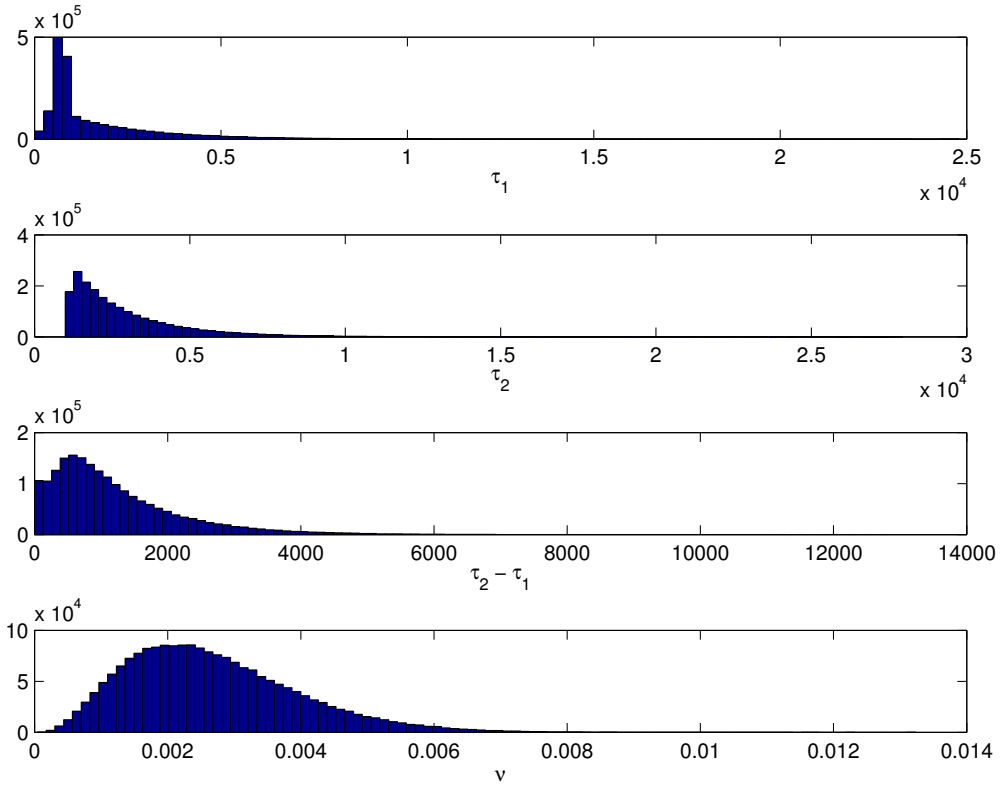

Figure A.2.22: Histograms of the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\sigma)$, τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\sigma)$, $\tau_2-\tau_1$, $\pi(\tau_2-\tau_1|x_1,\ldots,x_n;\alpha,\sigma)$ and ν, $\pi(\nu|x_1,\ldots,x_n;\alpha,\sigma)$. Here, σ and α were considered known, and the endpoint of observation was $t_n = 1\ 100$.

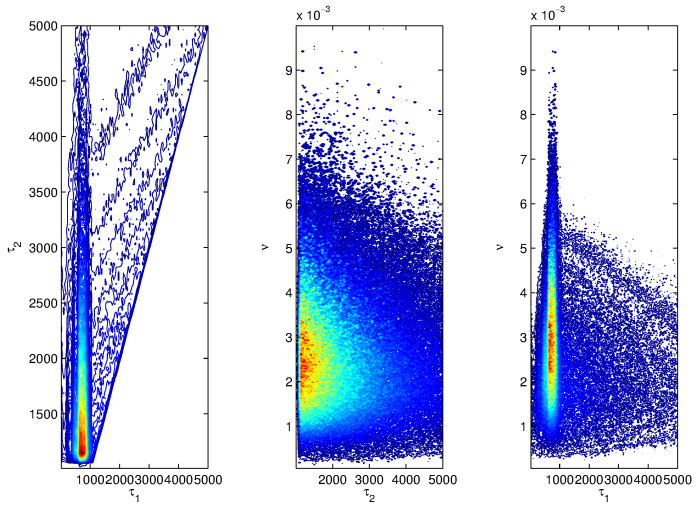

Figure A.2.23: Contour plot of the pairwise joint posterior distributions of τ_1, τ_2 and ν, $\pi(\tau_1, \tau_2|x_1, \ldots, x_n; \alpha, \sigma)$, $\pi(\tau_2, \nu|x_1, \ldots, x_n; \alpha, \sigma)$ and $\pi(\tau_1, \nu|x_1, \ldots, x_n; \alpha, \sigma)$. Here, σ and α were considered known, and the endpoint of observations was $t_n = 1\ 100$.

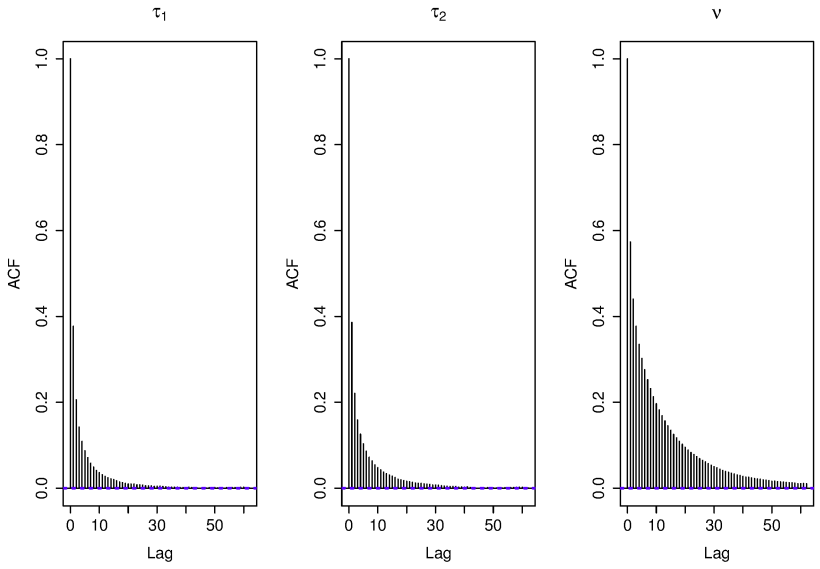

Figure A.2.24: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \sigma)$, and ν, $\pi(\nu|x_1, \ldots, x_n; \alpha, \sigma)$as function of the lag. Here, ν and σ are considered known, and the endpoint of observation was $t_n = 1\ 100$.

Observation until t_n = 1 450

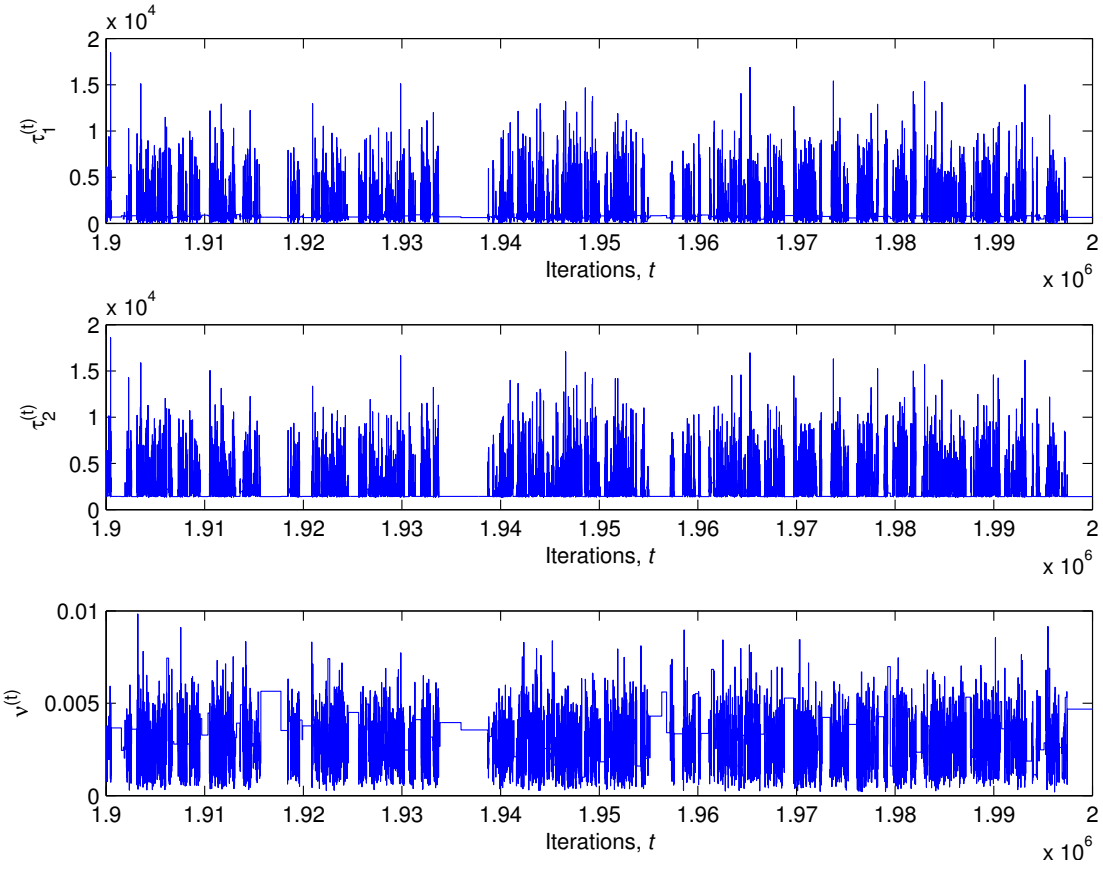

Figure A.2.25: Trace plot of the last 100 000 iterations of the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\sigma)$, τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\sigma)$ and ν, $\pi(\nu|x_1,\ldots,x_n;\alpha,\sigma)$. Here, σ and α were considered known, and the endpoint of observations was $t_n = 1\ 450$.

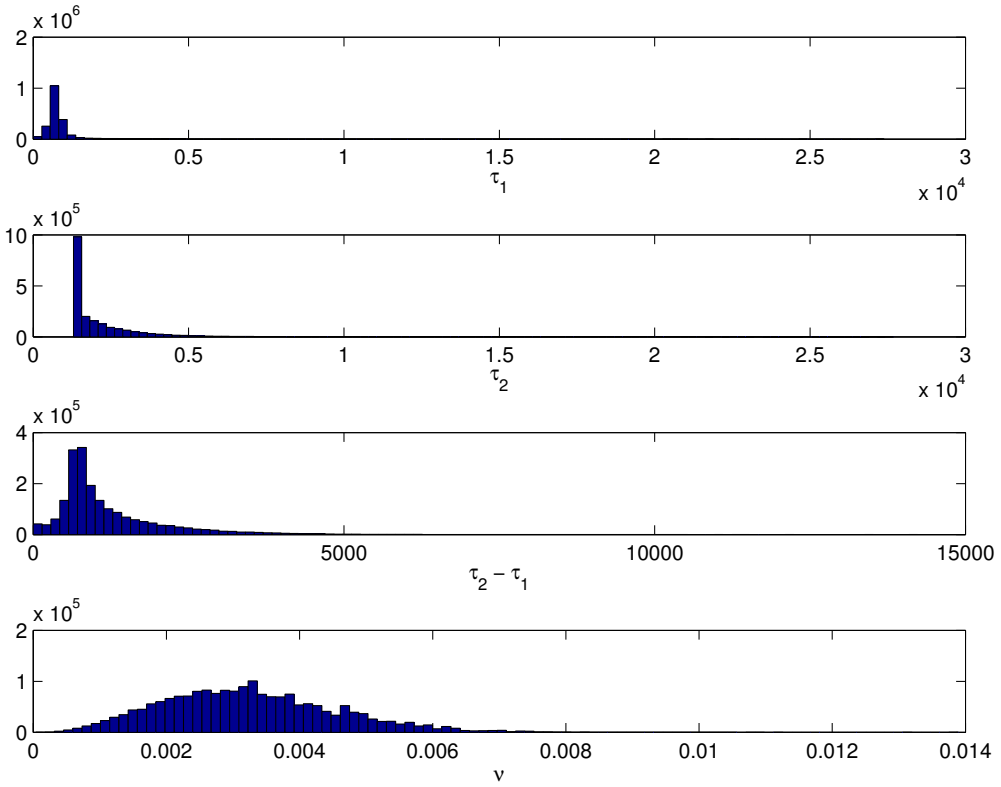

Figure A.2.26: Histograms of the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\sigma)$, τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\sigma)$, $\tau_2-\tau_1$, $\pi(\tau_2-\tau_1|x_1,\ldots,x_n;\alpha,\sigma)$ and ν, $\pi(\nu|x_1,\ldots,x_n;\alpha,\sigma)$. Here, σ and α were considered known, and the endpoint of observation was $t_n = 1\ 450$.

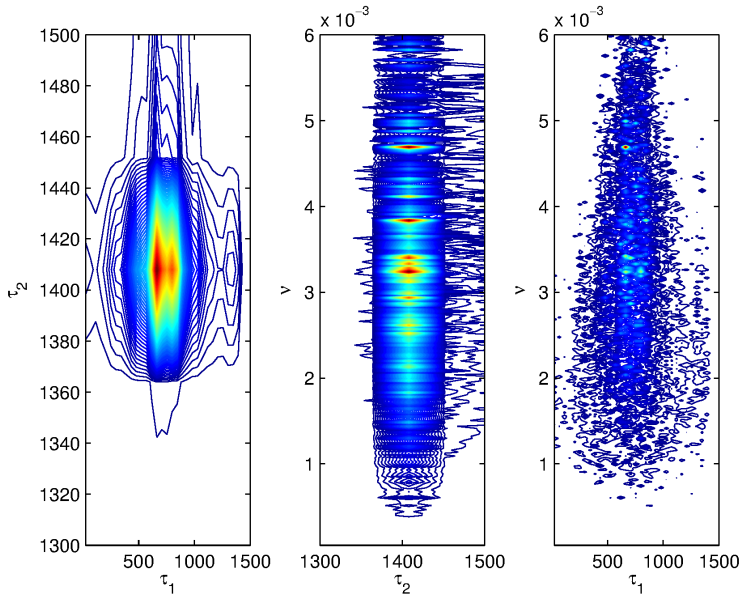

Figure A.2.27: Contour plot of the pairwise joint posterior distributions of τ_1, τ_2 and ν, $\pi(\tau_1, \tau_2|x_1, \ldots, x_n; \alpha, \sigma)$, $\pi(\tau_2, \nu|x_1, \ldots, x_n; \alpha, \sigma)$ and $\pi(\tau_1, \nu|x_1, \ldots, x_n; \alpha, \sigma)$. Here, σ and α were considered known, and the endpoint of observations was $t_n = 1\ 450$.

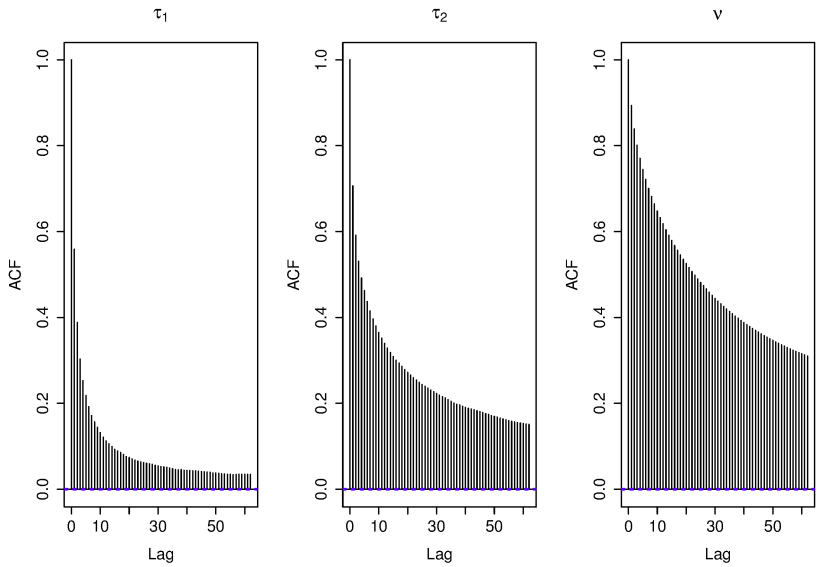

Figure A.2.28: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \sigma)$, and ν, $\pi(\nu|x_1, \ldots, x_n; \alpha, \sigma)$as function of the lag. Here, ν and σ are considered known, and the endpoint of observation was $t_n = 1\ 450$.

Observation until $t_n = 2\ 000$

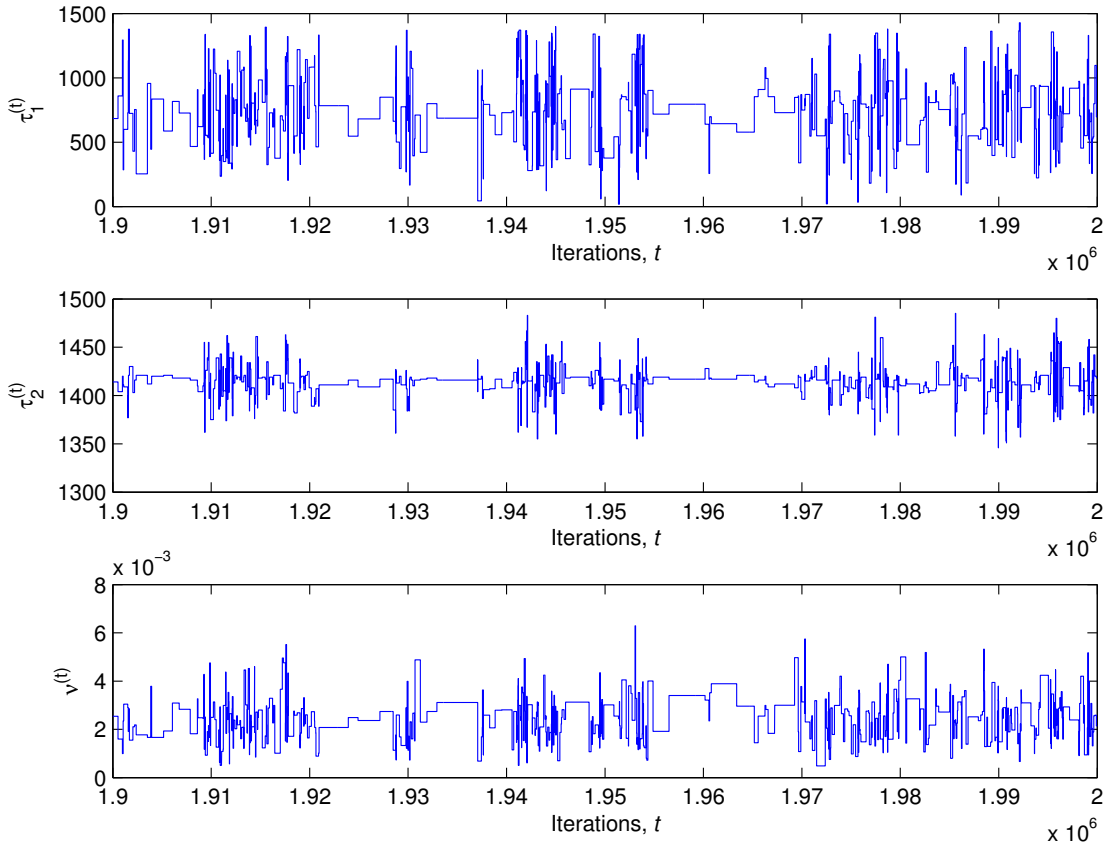

Figure A.2.29: Trace plot of the last 100 000 iterations of the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\sigma)$, τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\sigma)$ and ν, $\pi(\nu|x_1,\ldots,x_n;\alpha,\sigma)$. Here, σ and α were considered known, and the endpoint of observations was $t_n = 2\ 000$.

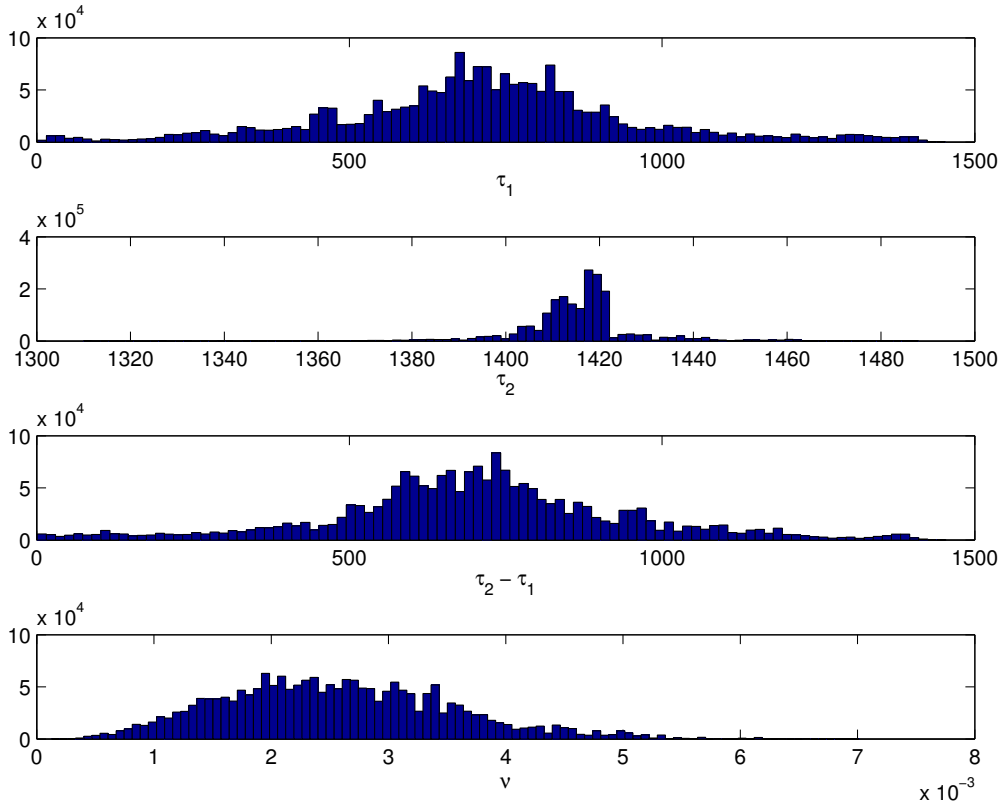

Figure A.2.30: Histograms of the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1,\ldots,x_n;\alpha,\sigma)$, τ_2, $\pi(\tau_2|x_1,\ldots,x_n;\alpha,\sigma)$, $\tau_2-\tau_1$, $\pi(\tau_2-\tau_1|x_1,\ldots,x_n;\alpha,\sigma)$ and ν, $\pi(\nu|x_1,\ldots,x_n;\alpha,\sigma)$. Here, σ and α were considered known, and the endpoint of observation was $t_n = 2\ 000$.

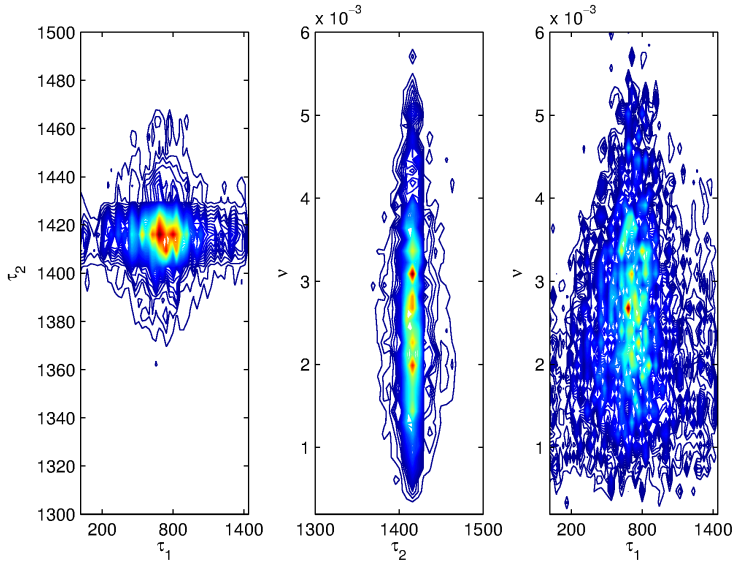

Figure A.2.31: Contour plot of the pairwise joint posterior distributions of τ_1, τ_2 and ν, $\pi(\tau_1, \tau_2|x_1, \ldots, x_n; \alpha, \sigma)$, $\pi(\tau_2, \nu|x_1, \ldots, x_n; \alpha, \sigma)$ and $\pi(\tau_1, \nu|x_1, \ldots, x_n; \alpha, \sigma)$. Here, σ and α were considered known, and the endpoint of observations was $t_n = 2\ 000$.

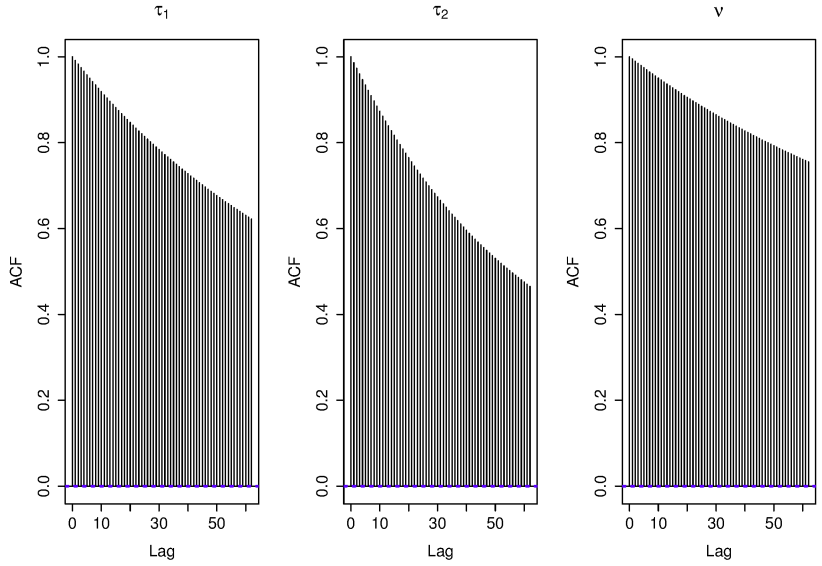

Figure A.2.32: Autocorrelation of the iterations in the Markov chain from the marginal posterior distributions of τ_1, $\pi(\tau_1|x_1, \ldots, x_n; \alpha, \sigma)$, τ_2, $\pi(\tau_2|x_1, \ldots, x_n; \alpha, \sigma)$, and ν, $\pi(\nu|x_1, \ldots, x_n; \alpha, \sigma)$as function of the lag. Here, ν and σ are considered known, and the endpoint of observation was $t_n = 2\ 000$.

Appendix B

Matlab code

B.1 One change point

B.1.1 Simulation

```
% Function sampling Wiener process
% One change point
% Different variance parameter
function [w] = WP(t, n, nu, sigma, tau)
% Input parameters
%  t      end time point
%  n      number of sampling points
%  nu     drift parameter vector [0, nu_1]
%  sigma  variance parameter vector [sigma_0, sigma_1]
%  tau    change point (time)

% Resolution of the sampling
dt = t/n;

% Sampling parameters
mean = nu*dt;
sd = sigma*sqrt(dt);

tau_index = floor(tau/dt);

% Sampling normal distributed increments
wpbefore = mean(1) + sd(1).*randn(tau_index,1);
wpafter = mean(2) + sd(2).*randn(n - tau_index - 1,1);

w = [0; wpbefore; wpafter];

% Cumulative sum of increments
w = cumsum(w);
end
```

MATLAB code B.1: Code to simulate a piecewise Wiener process with one predetermined change point.

B.1.2 MCMC algorithm

```
% Metropolis-Hastings algorithm
% Proposal distribution: prior distribution

function tau = mhTau(w, t, n, sigma, tau_lambda, nu, runs)
% Input parameters
%  t          end time point
%  n          number of sampling points
%  nu         drift parameter
%  sigma      variance parameter
%  w          vector containing observations
% tau_lambda prior parameter for tau
% runs       number of iterations for the mh
dt = t/n;
x = diff(w);

% Parameter matrix [theta, nu]
tau= zeros(runs,1);

% Initial tau
tau_0 = random('geo', tau_lambda) + 1;

tau(1) = tau_0;
tau_star_mat = random('geo', tau_lambda, [runs,1]) + 1;


accept = 0;
for i = 2 : runs
    % Proposal
    tau_star = tau_star_mat(i);
    % Current value
    tau_curr = tau(i-1,1);

    % alpha calculated on log scale
    logalpha = loglikelihood(tau_star, n, dt, nu, sigma, w) - ...
        loglikelihood(tau_curr, n, dt, nu, sigma, w);

    u = log(rand(1));
    if u < min(0, logalpha)
        tau(i) = tau_star;
        accept = accept + 1;
    else
        tau(i) = tau_curr;
    end
end
acceptancerate = accept/runs

end
```

MATLAB code B.2: Code to simulate from the posterior distribution of τ, by the Metropolis-Hastings algorithm.

```
function theta = mh(w, t, n, sigma, tau_lambda, nu_lambda, nu_nu, runs)
% Metropolis-Hastings
% nu unknown, sigma known
% nu sampled from normal distribution
% Metropolis-Hastings algorithm
% Proposal distribution: prior distribution
% tau_lambda: prior parameter for tau
% nu_lambda, nu_nu: prior parameters for nu
dt = t/n;
x = diff(w);

% Parameter matrix [theta, nu]
theta = zeros(runs,2);

% Initial tau
tau_0 = random('geo', tau_lambda) + 1;
% Initial nu
nu_0 = random('gam', nu_lambda, nu_nu);

theta(1,:) = [tau_0, nu_0];
tau_star_mat = random('geo', tau_lambda, [runs,1]) + 1;
nu_star_mat = random('gam', nu_lambda, nu_nu, [runs,1]);

accept = 0;
for i = 2 : runs
    % Proposal
    tau_star = tau_star_mat(i);
    nu_star = nu_star_mat(i);
    % Current value
    tau_curr = theta(i-1,1);
    nu_curr = theta(i-1, 2);
     % Log M-H ratio
    logalpha = loglikelihood(tau_star, n, dt, nu_star, sigma, w) - ...
        loglikelihood(tau_curr, n, dt, nu_curr, sigma, w);
    % Log random number
    u = log(rand(1));
    if u < min(0, logalpha)
        theta(i,1) = tau_star; theta(i,2) = nu_star;
        accept = accept + 1;
    else
        theta(i,1) = tau_curr; theta(i,2) = nu_curr;
    end
end
acceptancerate = accept/runs

end
```

MATLAB code B.3: Code to simulate from the joint posterior distribution of τ and ν, by the Metropolis-Hastings algorithm.

```
function [llh] = loglikelihood(tau, n, dt, nu, sigma, w)
% Input parameters
%  t         end time point
%  n         number of sampling points
%  nu        drift parameter vector
%  sigma     variance parameter
%  w         vector containing observations

% Extracting increments
w = diff(w);
t = n*dt;
n = n - 1;
    if tau <= n
        llh =  - 1/(2*dt*sigma^2)*( - 2*dt*nu*sum(w(tau/dt:n)) + ...
            (n-tau+1)*(dt*nu)^2);
    else
        llh = 1;
    end
end
```

MATLAB code B.4: Code to calculate log likelihood in `mhTau.m` and `mh.m`.

```
function theta = mhnusigma(w, t, n, tau_lambda, nu_lambda, nu_nu, ...
    sigma_lambda, sigma_nu, runs)
% Metropolis-Hastings
% nu unknown, sigma known
% Metropolis-Hastings algorithm
% Proposal distribution: prior distribution
% tau_lambda: prior parameter for tau
% nu_lambda, nu_nu: prior parameters for nu
% sigma_lambda, sigma_nu: prior parameters for sigma
dt = t/n;
x = diff(w);

% Parameter matrix [theta, nu]
theta = zeros(runs,3);

% Initial tau
tau_0 = random('geo', tau_lambda) + 1;
% Initial nu
nu_0 = random('gam', nu_lambda, nu_nu);
% Inital sigma
sigma_0 = random('gam', sigma_lambda, sigma_nu);

theta(1,:) = [tau_0, nu_0, sigma_0];
tau_star_mat = random('geo', tau_lambda, [runs,1]) + 1;
nu_star_mat = random('gam', nu_lambda, nu_nu, [runs,1]);
sigma_star_mat = random('gam', sigma_lambda, sigma_nu, [runs,1]);

accept = 0;
for i = 2 : runs
    % Proposal
    tau_star = tau_star_mat(i);
    nu_star = nu_star_mat(i);
    sigma_star = sigma_star_mat(i);
    % Current value
    tau_curr = theta(i - 1, 1);
    nu_curr = theta(i - 1, 2);
    sigma_curr = theta(i - 1, 3);
    % log M-H ratio
    logalpha = loglikelihoodnusigma(tau_star, nu_star, sigma_star, ...
        n, dt, w) - ...
        loglikelihoodnusigma(tau_curr, nu_curr, sigma_curr, n, dt, w);
    % log random number
    u = log(rand(1));
    if u < min(0, logalpha)
        theta(i, 1) = tau_star;
        theta(i, 2) = nu_star;
        theta(i, 3) = sigma_star;
        accept = accept + 1;
    else
        theta(i, 1) = tau_curr;
        theta(i, 2) = nu_curr;
        theta(i, 3) = sigma_curr;
    end
```

```
end
acceptancerate = accept/runs

end
```

MATLAB code B.5: Code to simulate from the joint posterior distribution of τ, ν and σ, by the Metropolis-Hastings algorithm.

```
% Log likelihood nu, sigma, tau unknown
function L = loglikelihoodnusigma(tau, nu, sigma, n, dt, w)
% Input parameters
%  t          end time point
%  n          number of sampling points
%  nu         drift parameter vector
%  sigma      variance parameter
%  w          vector containing observations
x = diff(w); n = n - 1;
if tau <= n
    L = - n*log(sigma) - 1/(2*dt*sigma^2)*(sum(x.^2) - ...
        2*dt*nu*sum(x(tau/dt : n)) +(n - tau/dt + 1)*(dt*nu)^2);
else
    L = - n*log(sigma) - 1/(2*dt*sigma^2)*(sum(x.^2));
end
```

MATLAB code B.6: Code to calculate log likelihood in `mhnusigma.m`.

B.1.3 Hitting time CDF estimates

```
function [cumdist, upper, lower, T_vec] = formulaicCDF(t_n, n, w, nu, ...
    sigma, burnin, m, endTau, a, tau_mc)
% burnin : the burnin length for the Markov chain
% tau_mc : the tau parameter Markov chain
% m :       number of draws from the posterior distribution
% a:        threshold temperature parameter
% endTau:   the upper t-value for which to calculate P(T<= t)
dt = t_n/n;
T_vec = [t_n + dt : dt : endTau];

cumdist = zeros(size(T_vec));
upper = zeros(size(T_vec));
lower = zeros(size(T_vec));

for i = 1 : length(T_vec)
    if mod(i,1000)==1
        disp(T_vec(i))
    end
    % Simulate posterior values
    % tau_sample = randsample(1:dt:endTau, m, true, post_vec);
    idx_sample = randsample(1:length(tau_mc(burnin:end)), m, true);
    tau_sample = tau_mc(idx_sample);
    idx_crossed = [tau_sample < T_vec(i)];
    tau_sample = tau_sample(idx_crossed);

    % Calculate relevant a-value vector  using the expected value of w(t)
    % at w(tau) if tau > t_n which is w(t_n)
    a_vec = a - w(n);

    % Calculate relevant t-vector
    t_vec = T_vec(i) - t_n*[(tau_sample)<=t_n] - ...
        tau_sample.*[tau_sample>t_n];

    % Calculate CDF based on posterior values
    cumdist_vec = normcdf((nu.*t_vec - a_vec)./(sigma.*sqrt(t_vec))) + ...
        exp(2.*a_vec.*nu./sigma^2).*normcdf((-a_vec - ...
        nu.*t_vec)./(sigma.*sqrt(t_vec)));

    % return cdf as mean
    cumdist(i) = 1/m*sum(cumdist_vec);

    % Find 95 % conf bound
    upper(i) = prctile(cumdist_vec, 97.5);
    lower(i) = prctile(cumdist_vec, 2.5);
end

end
```

MATLAB code B.7: Code to estimate the hitting time CDF by the formulaic approach, with ν and σ known.

```
function [cumdist, upper, lower, T_vec] = simulationCDF(w, t_n, ...
    n, sigma, nu, a, m_simu, tau_mc, m_tau, endTau, burnin)
% a: threshold temperature parameter
% m_simu: number of simulations for each draw from the posterior
% distribution
% tau_mc: the Markov chain from the posterior distribution
% m_tau: number of tau draws from the posterior distribution
% endTau: the upper limit for t for which to calculate P(T <= t)
% burnin: burnin length to apply to Markov chain from posterior
% distribution
s0 = t_n; dt = t_n/n;


% Simulated tau-values from the posterior distribution
idx_sample = randsample(burnin:length(tau_mc), m_tau, true);

% Allocating space for threshold times
threshold_timedt = zeros(m_simu,m_tau);

for i = 1 : m_tau
    % Take a tau sample
    tau = tau_mc(idx_sample(i));
    for j = 1 : m_simu
    if(tau < s0)
        % Tau has already happened
        crossed = 0; % False
        % Expected time to hit threshold
        Exp_T = tau + (a - w(round(tau/dt)))/nu;
        Exp_left = Exp_T - s0;
        endtime = s0;
        endtemp = w(n);

        while (crossed==0)
            % Sample for twice as long
            dw = dt*nu + ...
                sqrt(dt)*sigma.*randn(round(max(s0/2, 2*Exp_left)),1);
            % At least one value larger
            if(sum(find(endtemp+cumsum(dw)>=a))>0)
                threshold_timedt(j,i) = endtime + ...
                    dt*find(endtemp+cumsum(dw)>=a,1);
                crossed = 1;
            else
                endtime = endtime + dt*length(dw);
                endtemp = endtemp+sum(dw);
            end

        end

    else % Tau has not happened
        % Must sample with mean 0 until tau
        % Then mean mu
        crossed = 0; % false
        % Sampling until tau
```

```
        dw = sqrt(dt)*sigma.*randn(tau-n, 1);
        endtime = tau;
        endtemp = w(n) + sum(dw);
        Exp_T = tau + (a - endtemp)/nu;

        if(sum(find(endtemp+cumsum(dw)>=a))>0) % Threshold crossed
            threshold_timedt(j,i) = endtime + ...
                dt*find(endtemp+cumsum(dw)>=a,1);
            crossed = 1;
        end

        while (crossed == 0)
            % Sample for twice as long
            dw = dt*nu + sqrt(dt)*sigma.*randn(round(Exp_T),1);
            if(sum(find(endtemp+cumsum(dw)>=a))>0)
                threshold_timedt(j,i) = endtime + ...
                    dt*find(endtemp+cumsum(dw)>=a,1);
                crossed = 1;
            else
                endtime = endtime + dt*length(dw);
                endtemp = endtemp + sum(dw);
            end
        end
    end
    end
end
threshold_timedt = sort(threshold_timedt,1);
T_vec = t_n + dt : dt : min(max(max(threshold_timedt)), endTau);
% Uncertainty and F estimate
cumdist = zeros(size(T_vec));
upper = zeros(size(T_vec));
lower = zeros(size(T_vec));

for i = 1 : length(T_vec);
    if mod(T_vec(i), 1000) == 1
        disp(T_vec(i))
    end
    cumdist_vec = sum(threshold_timedt<= T_vec(i));
    cumdist(i) = mean(cumdist_vec);
    upper(i) = prctile(cumdist_vec, 97.5);
    lower(i) = prctile(cumdist_vec, 2.5);

end
cumdist = cumdist/m_simu; upper = upper/m_simu; lower = lower/m_simu;
end
```

MATLAB code B.8: Code to estimate the hitting time CDF by the simulation approach, with ν and σ known.

```
function [cumdist, T_vec, theta_mc, upper, lower] = formulaicCDFnu(...
    t_n, n, w, sigma, theta_mc, endTau, a, burnin, m)
% theta_mc: Markov chain from posterior distribution of tau and nu
% endTau:   upper value of t for which to calculate P(T <= t)
% a:        temperature threshold parameter
% burnin:   burnin value for the Markov chain
% m         number of draws in the posterior distribution for each t
dt = t_n/n;
T_vec = [t_n + dt : dt : endTau];

cumdist = zeros(size(T_vec));
upper = zeros(size(T_vec));
lower = zeros(size(T_vec));

for i = 1 : length(T_vec)

    % Simulate posterior values
    idx_sample = randsample(burnin:length(theta_mc), m, true);
    tau_sample = theta_mc(idx_sample,1);
    nu_sample = theta_mc(idx_sample,2);
    idx_crossed = [tau_sample < T_vec(i)];
    tau_sample = tau_sample(idx_crossed);
    nu_sample = nu_sample(idx_crossed);

    % Calculate relevant a-value vector
    a_vec = a - w(n);

    % Calculate relevant t-vector
    t_vec = T_vec(i) - t_n*[(tau_sample)<=t_n] - ...
        tau_sample.*[tau_sample>t_n];

    % Calculate CDF based on posterior values
    cumdist_vec = normcdf((nu_sample.*t_vec -...
        a_vec)./(sigma.*sqrt(t_vec))) + ...
        exp(2.*a_vec.*nu_sample./sigma^2).*normcdf((-a_vec -...
        nu_sample.*t_vec)./(sigma.*sqrt(t_vec)));

    % return cdf as mean
    cumdist(i) = 1/m*sum(cumdist_vec);

    upper(i) = prctile(cumdist_vec, 97.5);
    lower(i) = prctile(cumdist_vec, 2.5);

end

end
```

MATLAB code B.9: Code to estimate the hitting time CDF by the formulaic approach, with ν unknown and σ known.

```
function [cumdist, upper, lower, T_vec] = simulationCDFnu(w, ...
    t_n, n, sigma, ...
     a, m_simu, theta_mc, m_tau, endTau, burnin)
% theta_mc: Markov chain from posterior distribution of tau and nu
% endTau:   upper value of t for which to calculate P(T <= t)
% a:        temperature threshold parameter
% burnin:   burnin value for the Markov chain
% m_tau     number of draws in the posterior distribution for each t
% m_simu    number of Wiener processes to simulate for each tau
s0 = t_n; dt = t_n/n;
% Simulated tau-values from the posterior distribution
idx_sample = randsample(burnin:length(theta_mc), m_tau, true);

% Allocating space for threshold times
threshold_timedt = zeros(m_simu,m_tau);

for i = 1 : m_tau
    % Take a tau sample
    tau = theta_mc(idx_sample(i), 1);
    nu = theta_mc(idx_sample(i), 2);
    for j = 1 : m_simu
    if(tau < s0)
        % Tau has already happened
        crossed = 0; % False
        % Expected time to hit threshold
        Exp_T = tau + (a - w(round(tau/dt)))/nu;
        Exp_left = Exp_T - s0;
        endtime = s0;
        endtemp = w(n);
        while (crossed==0)
            % Sample for twice as long
            dw = dt*nu + ...
                sqrt(dt)*sigma.*randn(round(max(s0/2, 2*Exp_left)),1);
            % At least one value larger
            if(sum(find(endtemp+cumsum(dw)>=a))>0)
                threshold_timedt(j,i) = endtime +...
                    dt*find(endtemp+cumsum(dw)>=a,1);
                crossed = 1;
            else
                endtime = endtime + dt*length(dw);
                endtemp = endtemp+sum(dw);
            end
        end
    else % Tau has not happened
        % Must sample with mean 0 until tau
        % Then mean mu
        crossed = 0; % false
        % Sampling until tau
        dw = sqrt(dt)*sigma.*randn(tau-n, 1);
        endtime = tau;
        endtemp = w(n) + sum(dw);
        Exp_T = tau + (a - endtemp)/nu;
```

```
54          if(sum(find(endtemp+cumsum(dw)>=a))>0) % Threshold crossed
                threshold_timedt(j,i) = endtime + ...
56                  dt*find(endtemp+cumsum(dw)>=a,1);
                crossed = 1;
58          end
            while (crossed == 0)
60              % Sample for twice as long
                dw = dt*nu + sqrt(dt)*sigma.*randn(round(Exp_T),1);
62              if(sum(find(endtemp+cumsum(dw)>=a))>0)
                    threshold_timedt(j,i) = endtime +...
64                      dt*find(endtemp+cumsum(dw)>=a,1);
                    crossed = 1;
66              else
                    endtime = endtime + dt*length(dw);
68                  endtemp = endtemp + sum(dw);
                end
70          end
        end
72      end
end
74 threshold_timedt = sort(threshold_timedt,1);
T_vec = t_n + dt : dt : min(max(max(threshold_timedt)), endTau);
76 % Uncertainty and F estimate
cumdist = zeros(size(T_vec));
78 upper = zeros(size(T_vec));
lower = zeros(size(T_vec));
80
for i = 1 : length(T_vec);
82
    cumdist_vec = sum(threshold_timedt<= T_vec(i));
84  cumdist(i) = mean(cumdist_vec);
    upper(i) = prctile(cumdist_vec, 97.5);
86  lower(i) = prctile(cumdist_vec, 2.5);
end
88 cumdist = cumdist/m_simu; upper = upper/m_simu; lower = lower/m_simu;
end
```

MATLAB code B.10: Code to estimate the hitting time CDF by the simulation approach, with ν unknown and σ known.

```
function [cumdist, T_vec, upper, lower,theta_mc] = formulaicCDFnusigma(...
    t_n, n, w, endTau, a, m, burnin, theta_mc)
% theta_mc: Markov chain from posterior distribution of tau and nu
% endTau:   upper value of t for which to calculate P(T <= t)
% a:        temperature threshold parameter
% burnin:   burnin value for the Markov chain
% m         number of draws in the posterior distribution for each t
dt = t_n/n;
T_vec = [t_n + dt : dt : endTau];

cumdist = zeros(size(T_vec));
upper = zeros(size(T_vec));
lower = zeros(size(T_vec));

for i = 1 : length(T_vec)

    % Simulate posterior values
    idx_sample = randsample(burnin:runs, m, true);
    tau_sample = theta_mc(idx_sample,1);
    nu_sample = theta_mc(idx_sample,2);
    sigma_sample = theta_mc(idx_sample, 3);
    idx_crossed = [tau_sample < T_vec(i)];
    tau_sample = tau_sample(idx_crossed);
    nu_sample = nu_sample(idx_crossed);
    sigma_sample = sigma_sample(idx_crossed);

    % Calculate relevant a-value vector
    a_vec = a - w(n);

    % Calculate relevant t-vector
    t_vec = T_vec(i) - t_n*[(tau_sample)<=t_n] - ...
        tau_sample.*[tau_sample>t_n];

    % Calculate CDF based on posterior values
    cumdist_vec = normcdf((nu_sample.*t_vec - ...
        a_vec)./(sigma_sample.*sqrt(t_vec))) + ...
        exp(2.*a_vec.*nu_sample./sigma_sample.^2).*normcdf((-a_vec - ...
        nu_sample.*t_vec)./(sigma_sample.*sqrt(t_vec)));

    % return cdf as mean
    cumdist(i) = 1/m*sum(cumdist_vec);
    upper(i) = prctile(cumdist_vec, 97.5);
    lower(i) = prctile(cumdist_vec, 2.5);

end
end
```

MATLAB code B.11: Code to estimate the hitting time CDF by the formulaic approach, with ν and σ unknown.

```
function [cumdist, upper, lower, T_vec] = simulationCDFnusigma(w, s0, n,
    ...
    theta_mc, m_tau, a, burnin, m_simu, endTau)
% theta_mc: Markov chain from posterior distribution of tau and nu
% endTau:   upper value of t for which to calculate P(T <= t)
% a:        temperature threshold parameter
% burnin:   burnin value for the Markov chain
% m_tau     number of draws in the posterior distribution for each t
% m_simu    number of Wiener processes to simulate for each tau
t_n = s0; dt = t_n/n;
% Simulated tau-values from the posterior distribution
runs = length(theta_mc);
idx_sample = randsample(burnin:runs, m_tau, true);

% Allocating space for threshold times
threshold_timedt = zeros( m_simu, m_tau);

for i = 1 : m_tau
    % Take a tau sample
    tau = theta_mc(idx_sample(i),1);
    % Take a nu sample
    nu = theta_mc(idx_sample(i),2);
    % Take a sigma sample
    sigma = theta_mc(idx_sample(i),3);
    for j = 1 : m_simu
    if(tau < s0)
        % Tau has already happened
        crossed = 0; % False
        % Expected time to hit threshold
        Exp_T = tau + (a - w(round(tau/dt)))/nu;
        Exp_left = Exp_T - s0;
        endtime = s0;
        endtemp = w(n);
        while (crossed==0)
            % Sample for twice as long
            dw = dt*nu + ...
                sqrt(dt)*sigma.*randn(round(max(s0/2, 2*Exp_left)),1);
            % At least one value larger
            if(sum(find(endtemp+cumsum(dw)>=a))>0)
                threshold_timedt(j,i) = endtime +...
                    dt*find(endtemp+cumsum(dw)>=a,1);
                crossed = 1;
            else
                endtime = endtime + dt*length(dw);
                endtemp = endtemp+sum(dw);
            end
            %keyboard;
        end
    else % Tau has not happened
        % Must sample with mean 0 until tau
        % Then mean mu
        crossed = 0; % false
        % Sampling until tau
```

```
        dw = sqrt(dt)*sigma.*randn(tau-n, 1);
        endtime = tau;
        endtemp = w(n) + sum(dw);
        Exp_T = tau + (a - endtemp)/nu;
        if(sum(find(endtemp+cumsum(dw)>=a))>0) % Threshold crossed
            threshold_timedt(j,i) = endtime + ...
                dt*find(endtemp+cumsum(dw)>=a, 1);
            crossed = 1;
        end
        while (crossed == 0)
            % Sample for twice as long
            dw = dt*nu + sqrt(dt)*sigma.*randn(round(Exp_T),1);
            if(sum(find(endtemp+cumsum(dw)>=a))>0)
                threshold_timedt(j,i) = endtime + ...
                    dt*find(endtemp+cumsum(dw)>=a,1);
                crossed = 1;
            else
                endtime = endtime + dt*length(dw);
                endtemp = endtemp + sum(dw);
            end
        end
    end
    end
end

threshold_timedt = sort(threshold_timedt,1);
T_vec = t_n + dt : dt : min(max(max(threshold_timedt)), endTau);
% Uncertainty and F estimate
cumdist = zeros(size(T_vec));
upper = zeros(size(T_vec));
lower = zeros(size(T_vec));

for i = 1 : length(T_vec);
    cumdist_vec = sum(threshold_timedt<= T_vec(i));
    cumdist(i) = mean(cumdist_vec);
    upper(i) = prctile(cumdist_vec, 97.5);
    lower(i) = prctile(cumdist_vec, 2.5);
end
cumdist = cumdist/m_simu; upper = upper/m_simu; lower = lower/m_simu;

end
```

MATLAB code B.12: Code to estimate the hitting time CDF by the simulation approach, with ν and σ unknown.

B.2 Two change points

B.2.1 Simulation

```
% Function sampling Wiener process
% Two change point
% Different variance parameter
function [w] = WP2(t, n, nu, sigma, tau)
% Input parameters
%  t      end time point
%  n      number of sampling points
%  nu     drift parameter vector [0, nu_1, nu_2]
%  sigma  variance parameter vector [sigma_0, sigma_1, sigma_2]
%  tau    change points [tau_1, tau_2] (time)

% Resolution of the sampling
dt = t/n;

% Sampling parameters
mean = nu*dt;
sd = sigma*sqrt(dt);

tau_index = floor(tau/dt);

% Sampling normal distributed increments
wp0 = mean(1) + sd(1).*randn(tau_index(1)-1,1);
wp1 = mean(2) + sd(2).*randn(tau_index(2) - tau_index(1),1);
wp2 = mean(3) + sd(3).*randn(n - tau_index(2),1);
%keyboard
w = [0; wp0; wp1; wp2];

% Cumulative sum of increments
w = cumsum(w);
end
```

MATLAB code B.13: Code to simulate a piecewise Wiener process with two predetermined change points.

B.2.2 MCMC algorithm

```
% Metropolis-Hastings algorithm
% Proposal distribution: prior distribution
function tau = mh(w, t, n, nu, sigma, alpha, lambda1, lambda2, runs)
dt = t/n;
tau = zeros(runs,2);

% Initial tau
[tau1_0, tau2_0] = priordraw(lambda1, lambda2,dt);


tau(1,:) = [tau1_0, tau2_0];

accept = 0;
for i = 2 : runs
    % Proposal
    [tau1_star, tau2_star] = priordraw(lambda1, lambda2,dt);
    % Current tau values
    tau1_curr = tau(i-1,1);
    tau2_curr = tau(i-1, 2);
    %log of Likelihood ratio
    logalpha = loglikelihood(tau1_star, tau2_star, w,...
        nu, sigma, alpha, n, dt) - ...
         loglikelihood(tau1_curr, tau2_curr, w, nu, sigma, alpha, n, dt);
    u = log(rand(1));
    if u < min(0, logalpha)
        tau(i,1) = tau1_star; tau(i,2) = tau2_star;
        accept = accept + 1;
    else
        tau(i,1) = tau1_curr; tau(i,2) = tau2_curr;
    end
end
acceptancerate = accept/runs

end
```

MATLAB code B.14: Code to simulate from the joint posterior distribution of τ_1 and τ_2, by the Metropolis-Hastings algorithm.

```
function logL = loglikelihood(tau1, tau2, w, nu, sigma, alpha, n, dt)
% Function to calculate log likelihood
% Input parameters
%  dt       resolution of samples
%  n        number of sampling points
%  nu       drift parameter vector
%  sigma    variance parameter
%  w        vector containing observations

nu1 = nu(2);
nu2 = nu(3);

x = diff(w);

if tau2 < n*dt
    logL = -(n-tau2/dt)*log(alpha) -...
        1/(2*dt*sigma^2)*(sum(x(1:ceil((tau1-dt)/dt)).^2) + ...
        sum((x(ceil(tau1/dt) : ceil((tau2-dt)/dt))-dt*nu1).^2) + ...
        sum((x(ceil(tau2/dt):n-1)-dt*nu2).^2)/alpha^2);
elseif tau1 < n*dt
    logL = -1/(2*dt*sigma^2)*(sum(x(1:max(1,ceil((tau1-dt)/dt))).^2) + ...
        sum((x(ceil(tau1/dt):n-1)-dt*nu1).^2));
else
    logL = -1/(2*dt*sigma^2)*(sum(x.^2));
end

end
```

MATLAB code B.15: Code to Calculate the log likelihood in `mh2CP.m` and `mhnu2CP.m`

```
function tau = mhnu(w, t, n, sigma, alpha, nu_lambda, ...
    nu_nu, lambda1, lambda2, runs)
dt = t/n;
tau = zeros(runs,3);

% nu draws
nudraws = random('gam', nu_lambda, nu_nu, runs, 1);
rho1draws = geornd(lambda1, runs, 1);
rho2draws = geornd(lambda2, runs, 1);

tau1draws = rho1draws + 1; tau2draws = tau1draws + rho2draws + 1;

tau(1,:) = [tau1draws(1), tau2draws(1), nudraws(1)];

accept = 0;

for i = 2 : runs
    tau1_star = tau1draws(i);
    tau2_star = tau2draws(i);
    nu_star = nudraws(i);
    nu_star_vec = [0, nu_star, alpha*nu_star];

    tau1_curr = tau(i - 1, 1);
    tau2_curr = tau(i - 1, 2);
    nu_curr = tau(i - 1, 3);
    nu_curr_vec = [0, nu_curr, alpha*nu_curr];

    logalpha = loglikelihood(tau1_star, tau2_star, w,...
        nu_star_vec, sigma, alpha, n, dt) - ...
         loglikelihood(tau1_curr, tau2_curr, w, ...
         nu_curr_vec, sigma, alpha, n, dt);

    u = log(rand(1));

    if u < min(0, logalpha)
        tau(i,1) = tau1_star; tau(i,2) = tau2_star; tau(i,3) = nu_star;
        accept = accept + 1;
    else
        tau(i,1) = tau1_curr; tau(i,2) = tau2_curr; tau(i,3) = nu_curr;
    end
end
acceptancerate = accept/runs

end
```

MATLAB code B.16: Code to simulate from the joint posterior distribution of τ_1, τ_2 and ν_1 by the Metropolis-Hastings algorithm.

B.2.3 Hitting time CDF estimates

```
function [cumdist, T_vec, upper, lower, post_mc] = formulaicCDF2cp(w, ...
    t_n, n, nu, sigma, ...
    alpha, m, endTau, burnin,a, post_mc)
% theta_mc: Markov chain from joint posterior distribution of tau1 and tau2
% endTau:   upper value of t for which to calculate P(T <= t)
% a:        temperature threshold parameter
% burnin:   burnin value for the Markov chain
% m         number of draws in the posterior distribution for each t
nu1 = nu(2);
dt = t_n/n;
T_vec = [t_n + dt : dt : endTau];
cumdist = zeros(size(T_vec);
upper = zeros(size(T_vec));
lower = zeros(size(T_vec));
for i = 1 : length(T_vec)

    % Sample posterior tau-values
    tau_idx_sample = randsample(burnin:length(post_mc), m, true);
    tau1_sample = post_mc(tau_idx_sample,1);
    tau2_sample = post_mc(tau_idx_sample,2);

    % Is tau_1 crossed?
    idx_crossed = [tau1_sample < T_vec(i)];

    tau1_sample = tau1_sample(idx_crossed);
    tau2_sample = tau2_sample(idx_crossed);

    % Calculate relevant a-value vector uses expected temp
    a_vec = a - w(n);

    % Calculate relevant t-vector
    t_vec = [alpha*(T_vec(i) - t_n)].*[tau2_sample<t_n] + ...
        [tau2_sample + alpha*(T_vec(i) -tau2_sample) -...
        t_n].*[tau1_sample<t_n].*[tau2_sample > t_n].*...
        [tau2_sample<= T_vec(i)] + ...
        [tau2_sample + alpha*(T_vec(i) - tau2_sample) - ...
        tau1_sample].*[tau1_sample>=t_n].*[tau2_sample > t_n].*...
        [tau2_sample <= T_vec(i)] + ...
        [T_vec(i) - tau1_sample].*[tau1_sample>=t_n].*...
        [tau2_sample >= T_vec(i)] + ...
        [T_vec(i) - t_n].*[tau1_sample<t_n].*[tau2_sample>=T_vec(i)];


    % Calculate distribution based on posterior values
    cumdist_vec = 1 - (normcdf((a_vec - nu1.*t_vec)./...
        (sigma.*sqrt(t_vec))) - ...
        exp(2*a_vec.*nu1./sigma^2).*normcdf((-a_vec - ...'
        nu1.*t_vec)./(sigma.*sqrt(t_vec))));

   % Normalize and store in vector
    cumdist(i) = 1/m*sum(cumdist_vec);
```

```
52      upper(i) = prctile(cumdist_vec, 2.5);
        lower(i) = prctile(cumdist_vec, 97.5);
54
    end
56  end
```

MATLAB code B.17: Code to estimate the hitting time CDF by the formulaic approach, with ν and σ known.

```
function [cumdist, upper, lower, T_vec] = simulationCDF2cp(w, ...
    t_n, n, nu, sigma, alpha, ...
    post_mc, m_tau, burnin, a, m_simu, endTau)
% theta_mc: Markov chain from joint posterior distribution of tau1 and tau2
% endTau:   upper value of t for which to calculate P(T <= t)
% a:        temperature threshold parameter
% burnin:   burnin value for the Markov chain
% m_tau     number of draws in the posterior distribution
% m_simu    number of Wiener processes to simulate for each tau1 tau2 draw
nu1 = nu(2); nu2 = nu(3);
dt = t_n/n;
% Sampling tau1, tau2
tau_idx_sample = randsample(burnin:length(post_mc), m_tau, true);
tau1_sample = post_mc(tau_idx_sample,1);
tau2_sample = post_mc(tau_idx_sample,2);

% Allocating space
threshold_time = zeros(m_simu,m_tau);

for i = 1 : m_tau
    % Take one of the tau-samples
    tau1 = tau1_sample(i); tau2 = tau2_sample(i);
    for j = 1 : m_simu
    % Both change points passed
    if tau2 < t_n
        crossed = 0;
        endtime = t_n;
        endtemp = w(n);
        while crossed == 0
            % Sample for t_n time longer
            dw = dt*nu2 + sqrt(dt)*alpha*sigma*randn(t_n/dt,1);
            % At least one value larger
            if(sum((endtemp + cumsum(dw))>=a)>0)
                crossed = 1;
                threshold_time(j,i) = endtime + ...
                    dt*find((endtemp + cumsum(dw))>=a,1);
            else
                endtime = endtime + t_n;
                endtemp = endtemp + sum(dw);
            end
        end
    % Only the first change point is passed
    elseif tau1 < t_n
        crossed = 0;
        endtime = t_n;
        endtemp = w(n);
        % Sample until tau2
        dw = dt*nu1 + sqrt(dt)*sigma*randn((tau2-t_n)/dt,1);

        % Check for crossing
        if(sum((endtemp + cumsum(dw))>=a)>0)
            crossed = 1;
            threshold_time(j,i) = endtime + ...
```

```
54                  dt*find((endtemp + cumsum(dw))>=a,1);
        else
56              endtime = tau2;
                endtemp = endtemp + sum(dw);
58          end

60          while crossed == 0
                dw = dt*nu2 + sqrt(dt)*sigma*alpha*randn(t_n/dt, 1);
62              % At least one value larger
                if(sum((endtemp + cumsum(dw))>=a)>0)
64                  crossed = 1;
                    threshold_time(j,i) = endtime + ...
66                      dt*find((endtemp + cumsum(dw))>=a,1);
                else
68                  endtime = endtime + t_n;
                    endtemp = endtemp + sum(dw);
70              end
            end
72      % Neither of the change points passed
        else
74          crossed = 0;
            endtime = t_n;
76          endtemp = w(n);

78          % Sample until tau1
            dw = sqrt(dt)*sigma*randn((tau1-t_n)/dt, 1);
80          % Check for crossing
            if (sum((endtemp + cumsum(dw))>=a)>0)
82              crossed = 1;
                threshold_time(j,i) = endtime + ...
84                  dt*find((endtemp + cumsum(dw))>=a,1);
            else
86              endtime = tau1;
                endtemp = endtemp + sum(dw);
88              % Sample until tau2
                dw = dt*nu1 + sqrt(dt)*sigma*randn((tau2 - tau1)/dt,1);
90              % Check for crossing
                if (sum((endtemp + cumsum(dw))>=a)>0)
92                  crossed = 1;
                    threshold_time(j,i) = endtime + ...
94                      dt*find((endtemp + cumsum(dw))>=a,1);
                else
96                  endtime = tau2;
                    endtemp = endtemp + sum(dw);
98              end
            end
100         while crossed == 0
                dw = dt*nu2 + sqrt(dt)*sigma*alpha*randn(t_n/dt, 1);
102             % At least one value larger
                if(sum((endtemp + cumsum(dw))>=a)>0)
104                 crossed = 1;
                    threshold_time(j,i) = endtime + ...
106                     dt*find((endtemp + cumsum(dw))>=a,1);
```

```
            else
108             endtime = endtime + t_n;
                endtemp = endtemp + sum(dw);
110         end
        end
112 end
    end
114 end

116 threshold_time = sort(threshold_time,1);
T_vec = t_n + dt : dt : min(max(max(threshold_time)), endTau);
118 % Uncertainty and F estimate
cumdist = zeros(size(T_vec));
120 upper = zeros(size(T_vec));
lower = zeros(size(T_vec));
122
for i = 1 : length(T_vec);
124     cumdist_vec = sum(threshold_time<= T_vec(i));
    cumdist(i) = mean(cumdist_vec);
126     upper(i) = prctile(cumdist_vec, 97.5);
    lower(i) = prctile(cumdist_vec, 2.5);
128 end
cumdist = cumdist/m_simu; upper = upper/m_simu; lower = lower/m_simu;
130
end
```

MATLAB code B.18: Code to estimate the hitting time CDF by the simulation approach, with ν and σ known.

```
function [cumdist, T_vec, upper, lower, post_mc] = formulaicCDF2cpnu(w,...
    t_n, n, sigma, alpha, ...
    a, m, burnin, endTau, post_mc)
% theta_mc: Markov chain from posterior distribution of tau1, tau2 and nu
% endTau:   upper value of t for which to calculate P(T <= t)
% a:        temperature threshold parameter
% burnin:   burnin value for the Markov chain
% m         number of draws in the posterior distribution for each t

dt = t_n/n;
T_vec = [t_n+dt : dt : endTau];

% Allocate space
cumdist = zeros(size(T_vec));
upper = zeros(size(T_vec));
lower = zeros(size(T_vec));

for i = 1 : length(T_vec)
    % Sample posterior tau-values
    tau_idx_sample = randsample(burnin:length(post_mc), m, true);
    tau1_sample = post_mc(tau_idx_sample, 1);
    tau2_sample = post_mc(tau_idx_sample, 2);
    nu_sample = post_mc(tau_idx_sample, 3);
    % Is tau_1 crossed?
    idx_crossed = [tau1_sample < T_vec(i)];

    tau1_sample = tau1_sample(idx_crossed);
    tau2_sample = tau2_sample(idx_crossed);
    nu_sample = nu_sample(idx_crossed);
    % Calculate relevant a-value vector uses expected temp
    a_vec = a - w(n);


    % Calculate relevant t-vector
    t_vec = [alpha*(T_vec(i) - t_n)].*[tau2_sample<t_n] + ...
        [tau2_sample + alpha*(T_vec(i) -tau2_sample) - ...
        t_n].*[tau1_sample<t_n].*[tau2_sample > t_n].*...
        [tau2_sample<= T_vec(i)] + ...
        [tau2_sample + alpha*(T_vec(i) - tau2_sample) - ...
        tau1_sample].*[tau1_sample>=t_n].*[tau2_sample > t_n].*...
        [tau2_sample <= T_vec(i)] + ...
        [T_vec(i) - tau1_sample].*[tau1_sample>=t_n].*...
        [tau2_sample >= T_vec(i)] + ...
        [T_vec(i) - t_n].*[tau1_sample<t_n].*[tau2_sample>=T_vec(i)];

    % Calculate distribution based on posterior values
    cumdist_vec = 1 - (normcdf((a_vec -...
        nu_sample.*t_vec)./(sigma.*sqrt(t_vec))) - ...
        exp(2*a_vec.*nu_sample./sigma^2).*normcdf((-a_vec -...
        nu_sample.*t_vec)./(sigma.*sqrt(t_vec))));

   % Normalize and store in vector
    cumdist(i) = 1/m*sum(cumdist_vec);
```

```
    upper(i) = prctile(cumdist_vec, 2.5);
    lower(i) = prctile(cumdist_vec, 97.5);
end
end
```

MATLAB code B.19: Code to estimate the hitting time CDF by the formulaic approach, with ν unknown and σ known.

```
function [cumdist, upper, lower, T_vec] = simulationCDF2cpnu(w,...
    t_n, n, sigma, alpha, post_mc, ...
    m_tau, burnin, a, m_simu, endTau)
% theta_mc: Markov chain from joint posterior distribution of tau1, tau2
%  and nu
% endTau:   upper value of t for which to calculate P(T <= t)
% a:        temperature threshold parameter
% burnin:   burnin value for the Markov chain
% m_tau     number of draws in the posterior distribution
% m_simu    number of Wiener processes to simulate for each tau1 tau2 draw
dt = t_n/n;

% Sampling tau1, tau2
tau_idx_sample = randsample(burnin:length(post_mc), m_tau, true);
tau1_sample = post_mc(tau_idx_sample,1);
tau2_sample = post_mc(tau_idx_sample,2);
nu_sample = post_mc(tau_idx_sample, 3);

% Allocating space
threshold_time = zeros(m_simu,m_tau);

for i = 1 : m_tau
    % Take one of the tau-samples
    tau1 = tau1_sample(i); tau2 = tau2_sample(i); nu1 = nu_sample(i);
    nu2 = alpha*nu1;
    for j = 1 : m_simu
    % Both change points passed
    if tau2 < t_n
        crossed = 0;
        endtime = t_n;
        endtemp = w(n);

        while crossed == 0
            % Sample for t_n time longer
            dw = dt*nu2 + sqrt(dt)*alpha*sigma*randn(t_n/dt,1);
             % At least one value larger
            if(sum((endtemp + cumsum(dw))>a)>0)
                crossed = 1;
                threshold_time(j,i) = endtime + ...
                    dt*find((endtemp + cumsum(dw))>=a,1);

            else
                endtime = endtime + t_n;
                endtemp = endtemp + sum(dw);

            end
        end
    % Only the first change point is passed
    elseif tau1 < t_n
        crossed = 0;
        endtime = t_n;
        endtemp = w(n);

```

```
        % Sample until tau2
55      dw = dt*nu1 + sqrt(dt)*sigma*randn((tau2-t_n)/dt,1);

57      % Check for crossing
        if(sum((endtemp + cumsum(dw))>=a)>0)
59          crossed = 1;
            threshold_time(j,i) = endtime + ...
61              dt*find((endtemp + cumsum(dw))>=a,1);

63      else
            endtime = tau2;
65          endtemp = endtemp + sum(dw);
        end
67
        while crossed == 0
69          dw = dt*nu2 + sqrt(dt)*sigma*alpha*randn(t_n/dt, 1);
            % At least one value larger
71          if(sum((endtemp + cumsum(dw))>=a)>0)
                crossed = 1;
73              threshold_time(j,i) = endtime +...
                    dt*find((endtemp + cumsum(dw))>=a,1);
75
            else
77              endtime = endtime + t_n;
                endtemp = endtemp + sum(dw);
79          end
        end
81
    % Neither of the change points passed
83  else
        crossed = 0;
85      endtime = t_n;
        endtemp = w(n);
87
        % Sample until tau1
89      dw = sqrt(dt)*sigma*randn((tau1-t_n)/dt, 1);
        % Check for crossing
91      if (sum((endtemp + cumsum(dw))>=a)>0)
            crossed = 1;
93          threshold_time(j,i) = endtime + ...
                dt*find((endtemp + cumsum(dw))>=a,1);
95
        else
97          endtime = tau1;
            endtemp = endtemp + sum(dw);
99
            % Sample until tau2
101         dw = dt*nu1 + sqrt(dt)*sigma*randn((tau2 - tau1)/dt,1);
            % Check for crossing
103         if (sum((endtemp + cumsum(dw))>a)>0)
                crossed = 1;
105             threshold_time(j,i) = endtime + ...
                    dt*find((endtemp + cumsum(dw))>=a,1);
```

```
107
                else
109                 endtime = tau2;
                    endtemp = endtemp + sum(dw);
111             end
            end
113
            while crossed == 0
115             dw = dt*nu2 + sqrt(dt)*sigma*alpha*randn(t_n/dt, 1);
                % At least one value larger
117             if(sum((endtemp + cumsum(dw))>=a)>0)
                    crossed = 1;
119                 threshold_time(j,i) = endtime +...
                        dt*find((endtemp + cumsum(dw))>=a,1);
121
                else
123                 endtime = endtime + t_n;
                    endtemp = endtemp + sum(dw);
125             end
            end
127


129     end
        end
131
    end
133
    threshold_time = sort(threshold_time,1);
135 T_vec = t_n + dt : dt : min(max(max(threshold_time)), endTau);
    % Uncertainty and F estimate
137 cumdist = zeros(size(T_vec));
    upper = zeros(size(T_vec));
139 lower = zeros(size(T_vec));

141 for i = 1 : length(T_vec);

143     cumdist_vec = sum(threshold_time<= T_vec(i));
        cumdist(i) = mean(cumdist_vec);
145     upper(i) = prctile(cumdist_vec, 97.5);
        lower(i) = prctile(cumdist_vec, 2.5);
147
    end
149 cumdist = cumdist/m_simu; upper = upper/m_simu; lower = lower/m_simu;

151 end
```

MATLAB code B.20: Code to estimate the hitting time CDF by the simulation approach, with ν unknown and σ known.