Generalized Bent-Cable Methodology for Changepoint Data: A Bayesian Approach

A Thesis Submitted to the

College of Graduate and Postdoctoral Studies

in Partial Fulfillment of the Requirements

for the degree of Master of Science

in the Department of Mathematics and Statistics

University of Saskatchewan

Saskatoon

By

Setu Chandra Kar

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Mathematics & Statistics
Room 142 McLean Hall
106 Wiggins Road
University of Saskatchewan
Saskatoon SK Canada
S7N 5E6
Email: math@math.usask.ca

College of Graduate & Postdoctoral Studies
Room 116 Thorvaldson Building
110 Science Place
Saskatoon SK Canada
S7N 5C9
Email: grad.studies@usask.ca

i

ABSTRACT

The choice of the model framework in a regression setting depends on the nature of the data. The focus of this study is on changepoint data, exhibiting three phases: incoming and outgoing, both of which are linear, joined by a curved transition. These types of data can arise in many applications, including medical, health and environmental sciences. Piecewise linear models have been extensively utilized to characterize such changepoint trajectories in different scientific fields. However, although appealing due to its simple structure, a piecewise linear model is not realistic in many applications where data exhibit a gradual change over time.

The most important aspect of characterizing a changepoint trajectory involves identifying the transition zone accurately. It is not only because the location of the transition zone is of particular interest in many areas of study, but also because it plays an important role in adequately describing the incoming and the outgoing phases of a changepoint trajectory. Note that once the transition is detected, the incoming and the outgoing phases can be modeled using linear functions. Overall, it is desirable to formulate a model in such a way that it can capture all the three phases satisfactorily, while being parsimonious with greatly interpretable regression coefficients. Since data may exhibit an either gradual or abrupt transition, it is also important for the transition model to be flexible. Bent-cable regression is an appealing statistical tool to characterize such trajectories, quantifying the nature of the transition between the two linear phases by modeling the transition as a quadratic phase with unknown width. We demonstrate that a quadratic function may not be appropriate to adequately describe many changepoint data. In practice, the quadratic function of the bent-cable model may lead to a wider or narrower interval than what could possibly be necessary to adequately describe a transition phase. We propose a generalization of the bent-cable model by relaxing the assumption of the quadratic bend. Specifically, an additional transition parameter is included in the bent-cable model to provide sufficient flexibility so that inference about the transition zone (i.e., shape and width of the bend) can be data driven, rather than pre-assumed as a specific type.

We discuss the properties of the generalized model, and then propose a Bayesian approach for statistical inference. The generalized model is then demonstrated with applications to three data

ii

sets taken from environmental science and economics. We also consider a comparison among the quadratic bent-cable, generalized bent-cable and piecewise linear models in terms of goodness of fit in analyzing both real-world and simulated data. Moreover, we supplement the motivation for our generalized bent-cable methodology via extensive simulations – we simulate changepoint data under some realistic assumptions, and then fit the quadratic bent-cable, generalized bent-cable and piecewise linear models to each of the simulated data sets to compare the performance of these models with respect to the overall quality of fit. A sensitivity analysis is also performed to investigate the sensitivity of Bayesian inference to prior specifications. This study suggests that the proposed generalization of the bent-cable model can be valuable in adequately describing changepoint data that exhibit either an abrupt or gradual transition over time.

ACKNOWLEDGEMENTS

This dissertation will be dedicated to my thesis supervisor Dr. Shahedul A. Khan

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Data showing a trend that characterizes a change are commonly observed in medical, health and environmental sciences. An example is the atmospheric concentrations of chloroflurocarbons (CFCs) monitored from different stations across the globe (NOAA/ESRL Global Monitoring Division 2016; see also Chapter 2). The atmospheric concentrations of CFCs exhibit a similarly shaped trajectory, where it rises initially in a linear fashion, then goes through a curved transition phase, followed by a linear decreasing trend. The piecewise linear model (Muggeo 2003) is a natural candidate to describe a continuous trend with an abrupt change. An issue with such modeling is that the abrupt change can be artificial, with possibly a more natural smoothness reflecting the trajectory change. In practice, the piecewise linear is not realistic in applications where data exhibit a gradual change over time. Therefore, it is desirable to formulate a model that is flexible enough to handle both gradual and abrupt transitions. One such methodology is the bent-cable regression (Chiu et al. 2006), which provides a regression framework to analyze changepoint data that exhibit a transition between two approximately linear phases. It quantifies the nature of the transition by modeling the transition as a quadratic phase with unknown width.

In this study, we demonstrate that the assumption of a quadratic bend of the bent-cable model is arguably a restrictive assumption in modeling the transition zone for changepoint data. Specifically, this assumption may lead to the starting and the end points of the transition period to be either underestimated or overestimated. We propose a generalization of the bent-cable model to overcome this problem. It provides sufficient flexibility so that inference about the transition zone can be data driven, rather than a specific type.

1.1 Background of the Study

The bent-cable methodology (Chiu et al. 2006) provides a regression framework to analyze change-point data that exhibit a transition between two approximately linear phases. The model is parsimonious, and appealing due to its simple structure, great flexibility and interpretability. Mathematically, the model can be expressed as (see also Figure 1.1)

$$y_i = \beta_0 + \beta_1 t_i + \beta_2 q(t_i; \tau, \gamma) + \epsilon_i, \tag{1.1}$$

where

$$q(t_i; \tau, \gamma) = \frac{(t_i - \tau + \gamma)^2}{4\gamma} I\{\tau - \gamma < t_i \leq \tau + \gamma\} + (t_i - \tau)I\{t_i > \tau + \gamma\}; \tag{1.2}$$

y_i is the response at time t_i ($i = 1, 2, \ldots, n$); I(A) is an indicator function that equals 1 if A is true and 0 otherwise; β_0 and β_1 are the intercept and slope of the linear incoming phase, respectively; $\beta_1 + \beta_2$ is the slope of the linear outgoing phase; τ and γ are the transition parameters, characterizing the center and half-width of the bend, respectively; and ϵ_i is the random error component. Under this formulation, the transition begins at time $\tau_1 = \tau - \gamma$ and ends at $\tau_2 = \tau + \gamma$, and the critical time point (CTP) at which the slope of the bent-cable changes signs is $\tau - \gamma - \frac{2\beta_1 \gamma}{\beta_2}$ (Chiu and Lockhart 2010).

Figure 1.1: A graphical description of the bent-cable function $f(t; \beta_0, \beta_1, \beta_2, \tau, \gamma) = \beta_0 + \beta_1 t + \beta_2 q(t; \tau, \gamma)$.

The bent-cable model assumes a quadratic bend (Equation (1.2)) to characterize the transition zone. In practice, the quadratic function of the bent-cable model may lead to an interval $[\tau_1, \tau_2]$ which is either wider or narrower than what could possibly be necessary to adequately describe the transition zone (see Chapter 2 for detail). This may also lead to biased estimates of the linear parameters β_0, β_1 and β_2. In this study, we propose a generalization of the bent-cable model by relaxing the assumption of the quadratic bend. We demonstrate that the proposed model can be valuable in adequately describing different types of changepoint trajectories.

1.2 Objectives of the Study

In light of the apparent three phases of a changepoint trajectory (linear incoming and outgoing, joined by a curved bend), there are two main objectives of changepoint data analyses (Muggeo 2003): (a) quantification of the transition $[\tau_1, \tau_2]$, and (b) estimation of the incoming and outgoing slopes β_1 and $\beta_1 + \beta_2$. Note that the estimates of the linear parameters β_0, β_1 and β_2 largely depend on the estimated transition, which as a whole determine the quality of the overall fit of a model. Therefore, it is important to describe a changepoint trajectory using a model that provides an accurate estimate of the transition. Although the assumption of a quadratic bend of the bent-cable model could be reasonable to characterize the transition phase for many changepoint data, it can also lead to unsatisfactory fits to many other changepoint trajectories (see Chapter 2 for detail). In this study, we propose a generalization of the bent-cable model by relaxing the assumption of the quadratic bend. Specifically, an additional transition parameter is included in the bent-cable model to provide sufficient flexibility so that inference about the transition zone (i.e., shape and width of the bend) can be data driven, rather than pre-assumed as a specific type. Although the primary intention is to model the transition phase more accurately, the proposed model is also expected to provide more precise estimates of the linear parameters β_0, β_1 and β_2. We develop a Bayesian framework for statistical inference. Extensive simulations are conducted to evaluate the performance of the proposed model in comparison with the bent-cable and piecewise linear models. Henceforth, we will use the term *quadratic* bent-cable to refer to the model Equations defined by (1.1) and (1.2) and *generalized* bent-cable to refer to the proposed generalization of the quadratic bent-cable model.

Apart from the methodological development, we also apply our method to three data sets taken from environmental science and economics, and discuss our findings in the context of scientific inquiry. The first example involves characterizing the atmospheric concentrations of CFC measurements over time, and the second example considers modeling the trends in housing values in the United States during two different time periods (January of 2001 to December of 2008 and June of 2006 to July of 2016). The objectives are to address some questions of broad interests, including (i) How long did it take for the CFC/housing value trend to show an obvious change? (ii) What were the rates of increase/decrease before and after the change? (iii) What was the time point at which the trend went from increasing to decreasing, or vice versa?

1.3   Changepoint Models: A Review from Literature

Prior to 1950's, the exploratory data analyses were extensively used to roughly identify the change point of a trajectory (Page 1954). Since this approach was not reliable, regression method was subsequently proposed (Quandt 1958), which is now the most widely used approach for changepoint data analyses. The general framework of a regression setting can be expressed as

$$y_i = f(t_i) + \epsilon_i, \tag{1.3}$$

where $f(\cdot)$ is a suitably chosen function that describes the trend over time and ϵ_i is the random error term. Depending on the form of $f(\cdot)$, changepoint models can be broadly classified to fall into one of three families: piecewise linear (also known as broken-stick or segmented) models, smooth changepoint models, and polynomial and spline regression.

1.3.1   Piecewise Linear Models

Piecewise linear models have been extensively utilized to characterize changepoint trajectories in different scientific fields (e.g. Ghosh and Vaida 2007, Slate and Turnbull 2000, Bellera et al. 2008, Toms and Lesperance 2003, Ruch and Claridge 1992, Wu et al. 2001). The general form of $f(\cdot)$

4

for a piecewise linear model is

$$f(t_i; \beta_0, \beta_1, \beta_2, \tau) = \beta_0 + \beta_1 t_i + \beta_2 (t_i - \tau) \mathrm{I}(t_i > \tau), \tag{1.4}$$

where β_0 and β_1 are, respectively, the intercept and slope of the incoming phase, $\beta_1 + \beta_2$ is the slope of the outgoing phase, and τ is the CTP (see Figure 1.2).
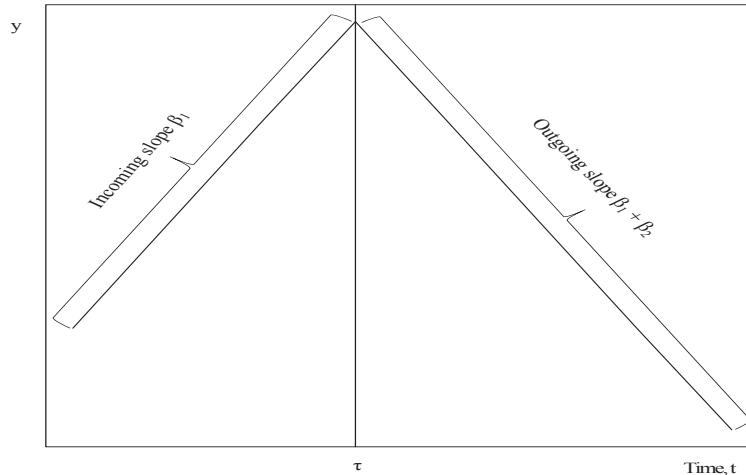
Figure 1.2: A graphical description of the piecewise linear model.

The piecewise linear function (1.4) is also popular for graphic analyses, including curve fitting, interpolation and extrapolation (e.g. Bian and Menz 1998, Dai et al. 2007, Magnani and Boyd 2009, Misener and Floudas 2010, Jimenez-Fernandez et al. 2014). Although appealing due to its simple structure, a major limitation of the piecewise linear model is that the first derivative of the function is discontinuous at the breakpoint (i.e. not continuously differentiable). This leads to considerable challenges in asymptotic theory for frequentist approach (Chiu et al. 2006, Chiu and Lockhart 2010, Kelly et al. 2004, Muggeo 2003, Tishler and Zang 1981) – either the frequentist approach is used under some non-standard regularity conditions (Muggeo 2003) or a Bayesian framework is considered for statistical inference so that the concern about unsatisfactory performance of asymptotics is irrelevant (Kiuchi et al. 1995, Bellera et al. 2008). Another shortcoming is that a piecewise linear model is not realistic in many applications where data exhibit a gradual change over time. To overcome these problems, smooth changepoint models have been proposed

(e.g. Bacon and Watts 1971, Chiu et al. 2006, van den Hout et al. 2010), utilizing a curved bend between the two linear phases to model the transition zone.

1.3.2 Smooth Changepoint Models

Bacon and Watts (1971) proposed a class of smooth changepoint models by considering a different parameterization of the piecewise linear model. In terms of the average slope $\alpha_1 = (\beta_1 + \beta_2)/2$ and the average difference in slope $\alpha_2 = (\beta_2 - \beta_1)/2$, the piecewise linear model was written as

$$y_i = \alpha_0 + \alpha_1(t_i - \tau) + \alpha_2(t_i - \tau)\text{sgn}(t_i - \tau) + \epsilon_i, \qquad (1.5)$$

where τ is the unknown join point, and $\text{sgn}(\cdot)$ is a sign function defined by

$$\text{sgn}(d) = \begin{cases} -1, & \text{if d} < 0 \\ 0, & \text{if d} = 0 \\ +1, & \text{if d} > 0 \end{cases} \qquad (1.6)$$

A smooth changepoint model was then formulated by replacing the sign function by a transition function $\text{trn}(\frac{t_i - \tau}{\gamma})$, where $\gamma > 0$ is a transition parameter. The model can be written as

$$y_i = \alpha_0 + \alpha_1(t_i - \tau) + \alpha_2(t_i - \tau)\,\text{trn}\!\left(\frac{t_i - \tau}{\gamma}\right) + \epsilon_i, \qquad (1.7)$$

where the transition function should satisfy the following conditions (Bacon and Watts 1971):

(a) $\lim_{d \to \infty} \text{trn}(\frac{|d|}{\gamma}) = 1$, so that $\text{trn}(\frac{d}{\gamma})$ behaves like $\text{sgn}(d)$ for large d;

(b) $\text{trn}(0) = 0$, so that $\text{trn}(\frac{d}{\gamma}) = \text{sgn}(d)$ for large $d = 0$;

(c) $\lim_{\gamma \to 0} \text{trn}(\frac{d}{\gamma}) = \text{sgn}(d)$, so that $\text{trn}(\frac{d}{\gamma})$ behaves like $\text{sgn}(d)$ for small γ; and

(d) $\lim_{d \to \infty} d\,\text{trn}(\frac{d}{\gamma}) = d$, so that $d\,\text{trn}(\frac{d}{\gamma})$ behaves like $d\,\text{sgn}(d) = |d|$ for large d.

Under these conditions, large values of γ lead to a gradual transition, whereas γ close to zero results in an abrupt transition. These conditions also ensure that the Bacon-Watts model is a con-

tinuously differentiable model. Examples of Bacon-Watts' transition function include cumulative distribution function of any symmetric probability density function and hyperbolic tangent. One difficulty of the Bacon-Watts model is that interpretation of α_1 and α_2 is not straightforward, as these parameters are linked to the shape of the transition (van den Hout et al. 2010). Consequently, inference about the slopes of the linear incoming and outgoing phases is not possible.

As described in Section 1.1, Chiu et al. (2006) introduced another class of smooth change-point models, referred to as the bent-cable regression methodology. A quadratic function was considered to model the transition phase. As opposed to the Bacon-Watts model, an appealing feature of the bent-cable model is that inference about the slopes of the linear incoming and outgoing phases is straightforward. Although the bent-cable model is continuously differentiable, the second derivative of the likelihood function does not exist everywhere. Therefore, the asymptotics were developed under some non-standard regularity conditions (Chiu et al. 2006). Khan et al. (2009), Khan et al. (2012) and Khan et al. (2013) subsequently proposed Bayesian method of inference, which avoids the substantial complexity of asymptotics. As indicated in Section 1.2, our work is motivated by Chiu et al.'s bent-cable regression methodology, and considers a generalization of the bent-cable model to describe the transition phase more accurately.

1.3.3   Polynomial and Spline Regression

A polynomial regression model (Fan and Gijbels 1996) is sometimes used to characterize the overall trend of a changepoint trajectory. However, although a pth order polynomial $f(t_i; \beta_0, \beta_1, \ldots, \beta_p) = \beta_0 + \beta_1 t_i + \beta_2 t_i^2 + \ldots + \beta_p t_i^p$ might be appealing to characterize the overall convexity of the trend, it would not be expected to fit the observed data all that well, in light of the apparent three phases: incoming and outgoing, joined by the curved transition. That is, as opposed to a changepoint model (see the characterization given in Figure 1.1), a polynomial model does not have the ability to reveal useful information regarding the rates of change in the incoming and outgoing phases and the transition. Other flexible modeling approaches such as penalized spline regression (Ruppert et al. 2003) can also handle abrupt and/or gradual changes, though the added flexibility in the shape of the fitted model can come at a cost of interpretability.

1.4 Organization of the Thesis

This is a manuscript-based thesis. We present our manuscript (Khan and Kar 2017) in Chapter 2, which includes (a) a motivating example to demonstrate the importance of this work, (b) the proposed model and its properties, (c) a Bayesian framework for statistical inference, (d) applications of our methodology to three real-world data sets, and (e) a simulation study to evaluate the performance of our methodology. We conclude in Chapter 3 by summarizing our findings.

GENERALIZED BENT-CABLE METHODOLOGY FOR CHANGEPOINT DATA

The bent-cable methodology (Chiu et al. 2006) provides a regression framework to analyze changepoint data that exhibit a transition between two approximately linear phases. It quantifies the nature of the transition between the two linear phases by modeling the transition as a quadratic phase with unknown, positive width. The model is appealing due to its simple structure, great flexibility and interpretability, and the model can be expressed as (see also Figure 2.1)

$$y_i = \beta_0 + \beta_1 t_i + \beta_2 \, q(t_i; \tau, \gamma) + \epsilon_i, \tag{2.1}$$

where

$$q(t_i; \tau, \gamma) = \frac{(t_i - \tau + \gamma)^2}{4\gamma} \mathrm{I}\{\tau - \gamma < t_i \le \tau + \gamma\} + (t_i - \tau)\mathrm{I}\{t_i > \tau + \gamma\}; \tag{2.2}$$

y_i is the response at time t_i ($i = 1, 2, \ldots, n$); I{A} is an indicator function that equals 1 if A is true and 0 otherwise; β_0 and β_1 are the intercept and slope of the linear incoming phase, respectively; $\beta_1 + \beta_2$ is the slope of the linear outgoing phase; τ and γ are the transition parameters, characterizing the center and half-width of the bend, respectively; and ϵ_i is the random error component. Under this formulation, the transition begins at time $\tau_1 = \tau - \gamma$ and ends at $\tau_2 = \tau + \gamma$, and the critical time point (CTP) at which the slope of the bent-cable changes signs is $\tau - \gamma - 2\beta_1\gamma/\beta_2$ (Chiu and Lockhart 2010). The model is identifiable, though the frequentist estimation method and asymptotics of bent-cable regression are highly complicated (Chiu et al. 2006). Since the second derivative of the likelihood function does not exist everywhere, the asymptotics were developed under non-standard regularity conditions (Chiu et al. 2006). Khan et al. (2013; 2009; 2012) subsequently proposed the
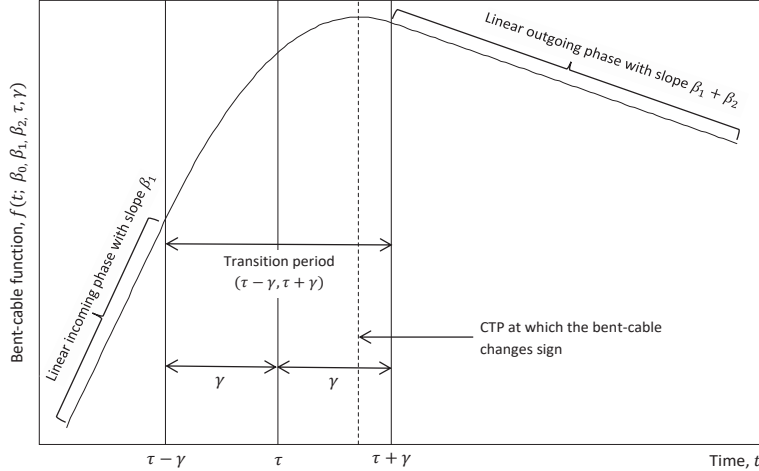
Figure 2.1: A graphical description of the bent-cable function $f(t; \beta_0, \beta_1, \beta_2, \tau, \gamma) = \beta_0 + \beta_1 t + \beta_2 \, q(t; \tau, \gamma)$ (for completeness of this chapter, we re-introduce this figure from chapter 1).

Bayesian method of inference, which avoids the substantial complexity of asymptotics.

The bent-cable model assumes a quadratic bend (Equation (2.2)) to characterize the transition zone. Although the assumption of the quadratic bend could be reasonable to characterize many changepoint data (e.g. the housing data described in Section 2.3.2), it may also lead to unsatisfactory fits to many other changepoint trajectories. To illustrate, we reconsider here the monthly mean chlorofluorocarbon-11 (CFC-11) data described in Khan et al. (2009): "CFCs are nontoxic, nonflammable chemicals containing atoms of carbon, chlorine, and fluorine. CFCs were extensively used in air conditioning/cooling units and as aerosol propellants prior to the 1980s. While CFCs are safe to use in most applications and are inert in the lower atmosphere, they do undergo significant reaction in the upper atmosphere. Chlorine inside the CFCs is one of the most important free-radical catalysts to destroy ozone. Because of this, CFCs were banned globally by the 1987 Montréal Protocol on Substances That Deplete the Ozone Layer. Since this protocol came into effect, the atmospheric concentration of CFCs has either leveled off or decreased." CFCs are monitored from different stations all over the globe by the Global Monitoring Division of the National Oceanic and Atmospheric Administration (NOAA/ESRL Global Monitoring Division 2016). We fit the bent-cable model to the monthly mean CFC-11 data from Barrow, Alaska, with a study period ranging from January of 1988 to September of 2010. The bent-cable fit (Figure 2.2a) suggests
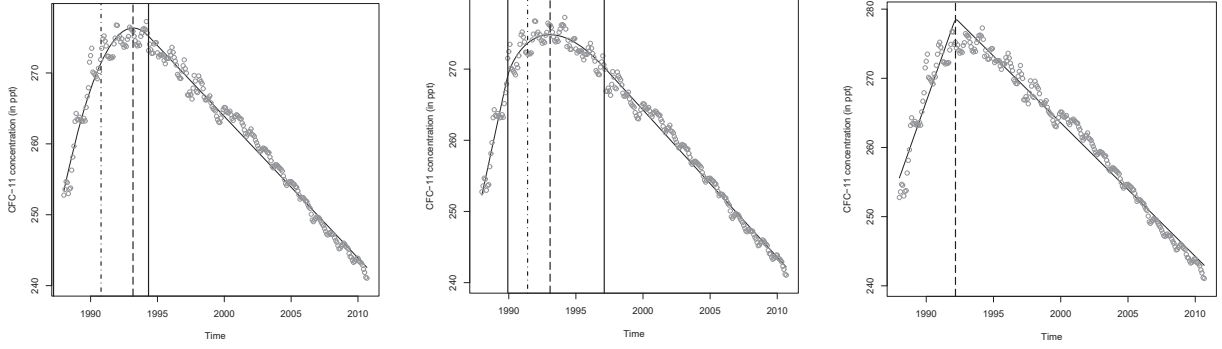
(a) Quadratic bent-cable fit: $\hat{\beta}_0 = 254.00$, $\hat{\beta}_1 = 0.85$, $\hat{\beta}_2 = -1.02$, $\hat{\gamma} = 42.64$, $\hat{\tau} = 33.37$ (Oct. 1990), $\hat{\tau}_1 = -9.28$ (Mar. 1987), $\hat{\tau}_2 = 75.99$ (May 1994), $\widehat{CTP} = 62.06$ (Mar. 1993), and DIC = 940.01.

(b) Generalized bent-cable fit: $\hat{\beta}_0 = 252.30$, $\hat{\beta}_1 = 0.73$, $\hat{\beta}_2 = -0.90$, $\hat{\kappa} = 1.26$, $\hat{\gamma} = 68.55$, $\hat{\tau} = 40.82$ (May 1991), $\hat{\tau}_1 = 23.14$ (Dec. 1989), $\hat{\tau}_2 = 109.36$ (Feb. 1997), $\widehat{CTP} = 60.94$ (Jan. 1993), and DIC = 887.46.

(c) Piecewise linear model fit: $\hat{\beta}_0 = 255.55$, $\hat{\beta}_1 = 0.46$, $\hat{\beta}_2 = -0.62$, $\widehat{CTP} = 50.15$ (Mar. 1992), and DIC = 1263.50.

Figure 2.2: Observed data and the corresponding fitted curves (solid) for Barrow CFC-11 data, with a study period ranging from January of 1988 to September of 2010 (273 months with $t_1 = 0, t_2 = 1, \ldots, t_{273} = 272$). Estimated transitions are marked by solid vertical lines, estimated τ by dot-dashed vertical lines, and estimated critical time points by dashed vertical lines.

that a gradual transition is in progress from the beginning of the study period. However, careful examination of the trajectory indicates apparently three phases within the study period: roughly linear incoming and outgoing phases at the ends of the profile, with a continuous transition between phases (Figure 2.2b). Model selection procedure also provides more support to the fit displayed in Figure 2.2b than the bent-cable fit shown in Figure 2.2a (see Section 2.3 for detail). Therefore, it is desirable to formulate a model that is flexible enough to characterize the transition phase more accurately.

In practice, the quadratic function of the bent-cable model may lead to a wider or narrower interval than what could possibly be necessary to adequately describe a transition phase (see Section 2.1 for a discussion). In this study, we propose a generalization of the bent-cable model by relaxing the assumption of the quadratic bend. Specifically, an additional transition parameter, κ, is included in the bent-cable model to provide sufficient flexibility so that inference about the transition zone (i.e. shape and width of the bend) can be data driven, rather than pre-assumed as a specific type. The fit of the proposed model to the CFC-11 data is displayed in Figure 2.2b,

11

the detail of which is presented in Section 2.3. As mentioned in Chapter 1, we will use the term *quadratic* bent-cable to refer to the model Equations defined by (2.1) and (2.2), and *generalized* bent-cable to refer to the proposed generalization of the quadratic bent-cable model.

The piecewise linear (broken-stick) model (Muggeo 2003) has been extensively utilized to describe a continuous trend exhibiting an abrupt change over time (e.g. Bellera et al. 2008, Hall et al. 2003). As an extremely sharp bend reduces the bent-cable to a piecewise linear model, the former encompasses the latter as a limiting case. In fact, $\gamma = 0$ reduces the bent-cable to a piecewise linear model for which

$$q(t_i; \tau, \gamma = 0) = (t_i - \tau)\mathrm{I}\{t_i > \tau\}. \tag{2.3}$$

When $\gamma = 0$, any sign change of the slope occurs at the point τ, the CTP for an abrupt transition. In practice, the abruptness of change imposed by the piecewise linear model is unrealistic in many applications where data exhibit a gradual transition over time. To illustrate, we fit the piecewise linear model to the CFC-11 data as displayed in Figure 2.2c. A comparison of the fits of Figure 2.2 using a model selection criterion suggests that the generalized bent-cable model would be preferred over the quadratic bent-cable and piecewise linear models in terms of the overall fit (see Section 2.3 for detail). Figure 2.2 also suggests that the estimates of the linear parameters β_0, β_1 and β_2 largely depend on the estimated transition, which as a whole determine the quality of the overall fit of a model. Note that the research interest usually lies not only in quantifying the transition, but also in the slopes (β_1 and $\beta_1 + \beta_2$) of the incoming and outgoing phases (Muggeo 2003). Therefore, it is important to describe a changepoint trajectory using a model that provides a superior fit overall. We present more comparisons of these models in Sections 2.3 and 2.4, using real-world and simulated data.

In Section 2.1, we introduce the generalized model and discuss some of its properties. The model is then formulated under a Bayesian framework for statistical inference (Section 2.2). We then apply our method to three data sets taken from environmental science and economics, and discuss our findings in the context of scientific inquiry (Section 2.3). In Section 2.4, simulations demonstrate (a) the flexibility of the generalized model, and (b) the importance of such flexibility

in adequately characterizing a changepoint trajectory. We conclude in Section 2.5 by summarizing our findings.

2.1 The Generalized Bent-Cable Model

For the types of data under consideration, consider a changepoint trajectory that comprises two linear segments (incoming and outgoing) joined by a curved bend. Let the transition spans $[\tau_1, \tau_2]$, where $\tau_1 = \tau - \gamma_1$, $\tau_2 = \tau + \gamma_2$, $\tau_1 < \tau < \tau_2$ and γ_1, γ_2 are both positive. To characterize the trajectory, we define a continuously differentiable function as follows:

$$f(t; \boldsymbol{\theta}) = \begin{cases} \beta_0 + \beta_1 t & \text{if } t \leq \tau_1 \\ \beta_0 + \beta_1 t + \beta_2 \frac{(\tau_2 - \tau)(t - \tau_1)^\kappa}{(\tau_2 - \tau_1)^\kappa} & \text{if } \tau_1 \leq t \leq \tau_2 \\ \beta_0 + \beta_1 t + \beta_2 (t - \tau) & \text{if } t \geq \tau_2 \end{cases} \tag{2.4}$$

where $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \tau, \gamma, \kappa)'$ and $\kappa > 1$. As written, the function is continuous both at $t = \tau_1$ and $t = \tau_2$. Note that the function is continuously differentiable if $f'(t; \boldsymbol{\theta})$ exists and is itself a continuous function. The left-hand and the right-hand derivatives of $f(t; \boldsymbol{\theta})$ at $t = \tau_1$ are both β_1, and therefore $f'(t; \boldsymbol{\theta})$ is continuous at τ_1. At $t = \tau_2$, the left-hand and the right-hand derivatives of $f(t; \boldsymbol{\theta})$ are $\beta_1 + \beta_2 \frac{\kappa(\tau_2 - \tau)}{(\tau_2 - \tau_1)}$ and $\beta_1 + \beta_2$, respectively. In order for $f'(t; \boldsymbol{\theta})$ to be continuous at $t = \tau_2$, we must have

$$\beta_1 + \beta_2 \frac{\kappa(\tau_2 - \tau)}{(\tau_2 - \tau_1)} = \beta_1 + \beta_2,$$

which leads to the solution $\gamma_1 = (\kappa - 1)\gamma_2$. Taking $\gamma_2 = \gamma$, we have $\tau_1 = \tau - (\kappa - 1)\gamma$ and $\tau_2 = \tau + \gamma$. Model (2.4) can then be written as

$$f(t; \boldsymbol{\theta}) = \begin{cases} \beta_0 + \beta_1 t & \text{if } t \leq \tau - (\kappa - 1)\gamma \\ \beta_0 + \beta_1 t + \beta_2 \frac{\gamma[t - \tau + (\kappa - 1)\gamma]^\kappa}{(\kappa\gamma)^\kappa} & \text{if } \tau - (\kappa - 1)\gamma \leq t \leq \tau + \gamma \\ \beta_0 + \beta_1 t + \beta_2 (t - \tau) & \text{if } t \geq \tau + \gamma. \end{cases} \tag{2.5}$$

Note that (2.5) can also be written using β_0, β_1, β_2, τ_1, τ_2 and κ, in which case γ and τ can be obtained from $\gamma = (\tau_2 - \tau_1)/\kappa$ and $\tau = [\tau_1 + (\kappa - 1)\tau_2]/\kappa$, respectively.

13

We can now formulate a changepoint regression model using Equation (2.5). Let t_i be the ith measurement occasion ($i = 1, 2, \ldots, n$) and y_i be the response at time t_i. Note that the definition of time includes a time scale and a time origin. We assume that t is chronological or calendar time, with origin $t_1 = 0$. We then model the response at time t_i by the relationship

$$y_i = f(t_i; \boldsymbol{\theta}) + \epsilon_i, \tag{2.6}$$

where

$$f(t_i; \boldsymbol{\theta}) = \beta_0 + \beta_1 t_i + \beta_2 \, g(t_i; \tau, \gamma, \kappa), \tag{2.7}$$

and

$$g(t_i; \tau, \gamma, \kappa) = \frac{\gamma[t_i - \tau + (\kappa - 1)\gamma]^\kappa}{(\kappa\gamma)^\kappa} I\{\tau - (\kappa - 1)\gamma < t_i \le \tau + \gamma\} + (t_i - \tau) I\{t_i > \tau + \gamma\}. \tag{2.8}$$

Equations (2.6)-(2.8) constitute our generalized bent-cable model. Some properties of the proposed model are presented below.

1. The generalized bent-cable function (2.7) is continuously differentiable when $\kappa > 1$.

2. The generalized bent-cable reduces to the quadratic bent-cable model when $\kappa = 2$, and to a piecewise linear model when $\gamma = 0, \kappa > 1$ or $\gamma > 0, \kappa = 1$. Note that when $\gamma > 0$ and $\kappa = 1$, Equation (2.8) reduces to $g(t_i; \tau, \gamma, \kappa = 1) = (t_i - \tau)I\{\tau < t_i \le \tau + \gamma\} + (t_i - \tau)I\{t_i > \tau + \gamma\} = (t_i - \tau)I\{t_i > \tau\}$. This leads to the piecewise linear function $f(t_i; \boldsymbol{\theta}) = \beta_0 + \beta_1 t_i + \beta_2(t_i - \tau)I(t_i > \tau)$.

3. As opposed to the quadratic model, τ for the generalized model is, in general, not the center of the bend.

4. The width of the bend region is $\kappa\gamma$. Thus, κ not only determines the shape of the bend, but also controls the width of the transition zone. This is a desirable property of the generalized bent-cable model. For the quadratic model, $[\tau_1, \tau_2]$ is determined by the underlying assumption of a quadratic bend for the transition zone. As described earlier of this chapter, this assumption forces τ and γ to be the center and half-width of the bend, respectively. For example, the estimates $\hat{\tau} = 33.37$ and $\hat{\gamma} = 42.64$ lead to $\hat{\tau}_2 = \widehat{\tau + \gamma} \approx 76$ for the fit displayed in Figure 2.2a, which appears a reasonable estimate for the end of the transition

period. However, the assumption of quadratic bend forces the starting point of the transition to be $\hat{\tau}_1 = \widehat{\tau - \gamma} \approx -9.3$ (i.e. nine months before the start of the study, which is March 1987), leading to a wider interval than what could possibly be necessary to adequately describe the transition phase. As a consequence, the quadratic model failed to identify the apparent linear incoming phase exhibited by the data. One way to overcome this problem is to consider an adjusting factor c such that $\tau_1 = \tau - c\gamma$. For the generalized model, k serves $\tau_1 = \tau - (\kappa - 1)\gamma$. This also allows γ to take a value from a wider interval (i.e. a small/large value of γ can be compensated by a large/small value of κ without affecting the width of the transition zone), which may help to overcome computational difficulties in estimating the bent-cable parameters as described by Chiu (2002).

5. The CTP for the generalized bend-cable model is $\tau - (\kappa - 1)\gamma + [-(\frac{\beta_1}{\beta_2})(\kappa\gamma)^{\kappa-1}]^{\frac{1}{\kappa-1}}$, which is defined when the slope of the cable changes signs (i.e. β_1 and $\beta_1 + \beta_2$ are of opposite signs). Note that the concept of a CTP is meaningful only when the slope of the cable changes signs (Chiu and Lockhart 2010).

2.2   Bayesian Inference

We consider a hierarchical modeling framework for Bayesian inference. We assume that ϵ_i's in Equation (2.6) are independent and identically normal with mean 0 and variance σ^2. Our choices of distributions for the random quantities allow us to write the generalized bent-cable model as

$$[y_i|\boldsymbol{\theta}] \sim \mathrm{N}(f(t_i; \boldsymbol{\theta}), \sigma^2),$$

$$[\beta_0|a_{01}, a_{02}] \sim \mathrm{N}(a_{01}, a_{02}), \ [\beta_1|a_{11}, a_{12}] \sim \mathrm{N}(a_{11}, a_{12}), \ [\beta_2|a_{21}, a_{22}] \sim \mathrm{N}(a_{21}, a_{22}),$$

$$[\tau|b_{11}, b_{12}] \sim \mathrm{U}(b_{11}, b_{12}), \ [\gamma|b_{21}, b_{22}] \sim \mathrm{U}(b_{21}, b_{22}),$$

$$[\kappa|c_1, c_2] \sim \mathrm{U}(c_1, c_2), \ \left[\sigma^{-2}|d_1, d_2\right] \sim \mathrm{G}(d_1, d_2),$$

where N, U and G stand for normal, uniform and gamma distributions, respectively, and a's, b's, c's and d's are the hyperparameters which are assumed known. The above specification can be

modified to write the quadratic bent-cable and piecewise linear models. For example, we take $\kappa = 2$ for the quadratic model, and $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \tau)'$ and $f(t_i; \boldsymbol{\theta}) = \beta_0 + \beta_1 t_i + \beta_2 (t_i - \tau) \mathrm{I}\{t_i > \tau\}$ for the piecewise linear model.

For β_0, β_1, β_2 and σ^{-2}, we choose the hyperprior values that lead to fairly vague, minimally informative priors: for normal, we take zero mean and a large variance (e.g. 10^3), and for gamma, we take small values for the hyperparameters (e.g. $d_0 = d_1 = 0.1$). For κ, $c_1 = 1$ and $c_2 = 3$ or 4 could be reasonable to model many real-life data. Some remarks are required to explain what would be a reasonable set of hyperprior values for τ and γ. An unbounded τ may lead to a computational breakdown of the Markov chain Monte Carlo (MCMC) for Bayesian inference. This can be explained from the definition of $g(t_i; \tau, \gamma, \kappa)$ in Equation (2.8). We see that $g(t_i; \tau, \gamma, \kappa)$ is degenerate as $\tau \to \infty$, in which case $f(t_i; \boldsymbol{\theta})$ approaches a straight line with intercept β_0 and slope β_1. Moreover, as $\tau \to 0$, a small value of γ may also make $g(t_i; \tau, \gamma, \kappa)$ degenerate (see Chiu 2002 for detail). Therefore, informative priors could be necessary for τ and γ to overcome computational difficulties. Since $\tau_1 < \tau < \tau_2$ and γ play an important role in modeling the width of the interval, a time-series plot of the data is useful to roughly determine (b_{11}, b_{12}) and (b_{21}, b_{22}) for τ and γ, respectively. We have considered this approach for data analyses in Section 2.3. Note that there is a software package "bentcableAR" (Chiu 2015) in R (R Core Team 2016) to fit the quadratic model using frequentist estimation method, where it is highlighted that an estimation algorithm may fail to converge with initial values that are unrefined guesses of the parameters. The recommendation is to generate initial values for τ and γ using a grid-based procedure, based on the deviance statistic (Chiu et al. 2006). This approach may also be used to roughly determine the hyperprior values for τ and γ for Bayesian inference. Our simulation study (Section 2.4) reveals that the Bayesian estimation is roughly robust if (a) b_{11} is a positive number, (b) $b_{21} = 0$, (c) b_{12} and b_{22} are some finite (could be large) upper bounds for τ and γ, respectively, and (d) the true τ and γ fall in (b_{11}, b_{12}) and (b_{21}, b_{22}), respectively.

We have written our code in WinBUGS (Lunn et al. 2000) to generate MCMC samples for Bayesian inference, which we subsequently analyzed using the "coda" package (Plummer et al. 2006) in R (R Core Team 2016). For simulation (Section 2.4), we generated data in R, and then

used the "R2WinBUGS" package (Sturtz et al. 2005) to invoke `WinBUGS` from R for Bayesian analysis.

2.3   Examples

Three data sets taken from the environmental science and economics are considered to demonstrate the application of the generalized bent-cable model. We also fit the quadratic bent-cable and piecewise linear models to each data set, and use the deviance information criterion (DIC) (Spiegelhalter et al. 2002) to compare the fits. In general, the minimum DIC estimates the model that offers the best short-term predictions. Nevertheless, Spiegelhalter et al. (2002) suggest the following rule of thumb: if $\Delta_M = \text{DIC}_M - \text{DIC}_{\min} \leq 2$, then the minimum DIC model and model M deserve equal consideration.

For each example, we construct two Markov chains each of 3,000,000 iterations to approximate the posterior density (see Appendix Sections A1-A3 for `WinBUGS` and R codes). The Gelman-Rubin statistic R (Gelman and Rubin 1992) is used to determine the burn-in and run length of the chains (values of R substantially above 1 indicate lack of convergence; some authors suggest that $R < 1.1$ is acceptable (Gelman and Shirley 2011)). For the examples presented in this article, the initial 20,000 to 50,000 iterations are discarded as burn-in, depending on the mixing behavior of the chains. The inferences are then based on every l^{th} iteration of the chains (thinning), with l set to a value between 100 and 500, depending on how fast the chain autocorrelation decays. For each example, we obtain $R < 1.05$ for all the parameters and quantities of interest. We also consider trace and density plots to diagnose the mixing of the chains.

Note that any function of the parameter vector $\boldsymbol{\theta} = (\beta_1, \beta_2, \tau, \gamma, \kappa)'$ can be estimated using a Markov chain. For example, the marginal posterior mean for τ_1 can be approximated by $\hat{\tau}_1 = \frac{1}{N} \sum_{j=1}^{N} [\tau^{(j)} - (\kappa^{(j)} - 1)\gamma^{(j)}]$, where $\{\boldsymbol{\theta}^{(j)}, j = 1, 2, \ldots, N\}$ is the MCMC output after discarding the burn-in samples and thinning the chain. Similarly, the MCMC sample mean of $f(t_i; \boldsymbol{\theta})$ at each observed t_i is regarded as the fitted value $\hat{f}(t_i; \boldsymbol{\theta})$, and then the fitted curve is interpolated based on the $\hat{f}(t_i; \boldsymbol{\theta})$ values.

2.3.1 Example 1: Barrow CFC-11 Data

We present here analysis of the CFC-11 data described in Chapter 2. The study period ranges from January of 1988 to September of 2010 ($n = 273$), and hence we take $t_1 = 0, t_2 = 1, \ldots, t_{273} = 272$. We then fit the bent-cable and piecewise linear models to these data. The lack of any trend in the trace plots and no signs of multimodality in the density plots indicate good mixing of the chains (see Appendix Section B.1).

Table 2.1: Barrow CFC-11 (in ppt) data analysis with a study period ranging from January of 1988 to September of 2010 (273 months with $t_1 = 0, t_2 = 1, \ldots, t_{273} = 272$) – posterior summaries of the bent cable and piecewise linear parameters.

Parameter	Generalized Bent-Cable (DIC = 887.46)		Quadratic Bent-Cable (DIC = 940.01)		Piecewise Linear (DIC = 1263.50)	
	Posterior median	95% credible interval	Posterior median	95% credible interval	Posterior median	95% credible interval
κ	1.258	$(1.211, 1.340)$	–	–	–	–
β_0	252.300	$(251.300, 253.200)$	254.000	$(252.700, 263.000)$	255.600	$(254.500, 256.600)$
β_1	0.729	$(0.669, 0.809)$	0.851	$(0.644, 1.228)$	0.462	$(0.418, 0.504)$
β_2	-0.901	$(-0.981, -0.842)$	-1.018	$(-1.394, -0.811)$	-0.623	$(-0.664, -0.580)$
$\beta_1 + \beta_2$	-0.172	$(-0.176, -0.169)$	-0.166	$(-0.170, -0.163)$	-0.161	$(-0.165, -0.157)$
γ	68.550	$(57.940, 78.670)$	42.640	$(33.730, 58.170)$	–	–
τ	40.820	$(38.120, 42.900)$	33.370	$(17.820, 42.080)$	–	–
$[\tau_1, \tau_2]$	$[23.140, 109.355]$	–	$[-9.280, 75.990]$	–	–	–
CTP	60.939	$(58.551, 63.398)$	62.062	$(60.223, 63.927)$	50.090	$(47.750, 52.880)$
σ^2	1.468	$(1.245, 1.748)$	1.799	$(1.524, 2.140)$	2.849	$(2.418, 3.390)$

The fits along with the estimates of the parameters are displayed in Figure 2.2, and numerical results are summarized in Table 2.1. We see that the estimated transitions ($\hat{\tau}$, $\hat{\tau}_1$ and $\hat{\tau}_2$) are quite different for the three competing models. The DIC values for the quadratic, generalized and piecewise linear fits are 940.01, 887.46 and 1263.50, respectively, suggesting that the fit of the generalized model is superior to those of the quadratic and piecewise linear models for the Barrow CFC-11 data. Therefore, we report here the key findings obtained from the generalized bent-cable fit. The posterior medians for $\hat{\tau}_1$ and $\hat{\tau}_2$ are December 1989 ($\hat{\tau}_1 = 23.140$) and February 1997 ($\hat{\tau}_2 = 109.355$), respectively, suggesting a linear incoming phase during January 1988 - December 1989, followed by a transition between December 1989 and February 1997, and a linear outgoing phase thereafter (February 1997 - September 2010). The average increase in CFC-

11 was about 0.729 ppt for a one-month increase in the incoming phase ($\hat{\beta}_1 = 0.729$ with 95% credible interval $(0.669, 0.809)$), and the average decrease was about 0.172 ppt in the outgoing phase ($\widehat{\beta_1 + \beta_2} = -0.172$ with 95% credible interval $(-0.176, -0.169)$). During the transition period, CFC-11 went from increasing to decreasing around January 1993 ($\widehat{CTP} = 60.939$ with 95% credible interval $(58.551, 63.398)$ or (November 1992, April 1993)).

2.3.2 Example 2: Privately-owned Housing Units in the United States

Privately-owned homes completed in a month (in thousands of units) in the United States are given in https://www.quandl.com/ (also available in R package "Quandl" (Raymond et al. 2016)). We consider here time-series data from January of 2001 to December of 2008 (series 1) and from June of 2006 to July of 2016 (series 2) to demonstrate another application of the bent-cable and piecewise linear models. For series 1, $n = 96$ months with $t_1 = 0, t_2 = 1, \ldots, t_{96} = 95$; and for series 2, $n = 122$ months with $t_1 = 0, t_2 = 1, \ldots, t_{122} = 121$. For computational convenience, the monthly data are expressed in $100,000$'s of units. The trace and density plots of the parameters are presented in Appendix Section B.2, which suggest no significant evidence of convergence issues.

The three fits (Figure 2.3) appear very similar for series 1: the estimates of the linear parameters are identical and the estimates of the CTP are very close to each other. A comparison of DIC also suggests that the three models agree closely in terms of the overall fit (DIC = 222.71, 222.76 and 221.38 for the generalized bent-cable, quadratic bent-cable and piecewise linear models, respectively). Posterior summaries of the parameters are given in Table 2.2. We see that privately-owned homes in the United States increased significantly since January 2001 ($\hat{\beta}_1 = 0.08$ with 95% credible interval $(0.07, 0.09)$) before entering into a transition phase around March 2006 ($\hat{\tau}_1 \approx 62$). It took about five months to complete the transition ($\hat{\tau}_1 \approx 62$ or March 2006, and $\hat{\tau}_2 \approx 67$ or August 2006), during which privately-owned homes increased to a maximum around March 2006 ($\widehat{CTP} \approx 63$ or March 2006). After August 2006, privately-owned homes decreased significantly at the rate of 0.36 units (in $100,000$) per month ($\widehat{\beta_1 + \beta_2} \approx -0.36$ with 95% credible interval $(-0.39, -0.33)$).
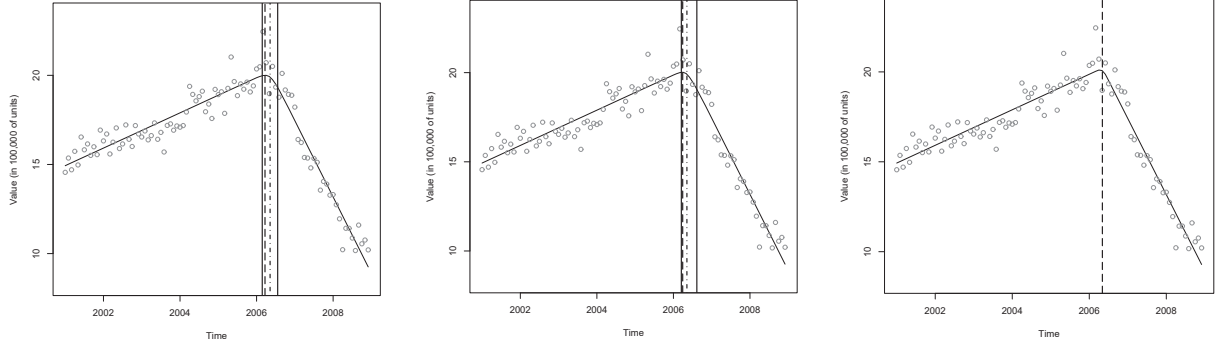
(a) Quadratic bent-cable fit: $\hat{\beta}_0 = 14.93$, $\hat{\beta}_1 = 0.08$, $\hat{\beta}_2 = -0.44$, $\hat{\gamma} = 2.38$, $\hat{\tau} = 64.24$ (May 2006), $\hat{\tau}_1 = 61.78$ (Feb. 2006), $\hat{\tau}_2 = 66.62$ (Jul. 2006), $\widehat{CTP} = 62.64$ (Mar. 2006), and DIC = 222.76.

(b) Generalized bent-cable fit: $\hat{\beta}_0 = 14.93$, $\hat{\beta}_1 = 0.08$, $\hat{\beta}_2 = -0.44$, $\hat{\kappa} = 1.64$, $\hat{\gamma} = 3.05$, $\hat{\tau} = 64.23$ (May 2006), $\hat{\tau}_1 = 62.48$ (Mar. 2006), $\hat{\tau}_2 = 67.34$ (Aug. 2006), $\widehat{CTP} = 62.90$ (Mar. 2006), and DIC = 222.71.

(c) Piecewise linear model fit: $\hat{\beta}_0 = 14.93$, $\hat{\beta}_1 = 0.08$, $\hat{\beta}_2 = -0.44$, $\widehat{CTP} = 64.08$ (May 2006), and DIC = 221.38.

Figure 2.3: Observed data and the corresponding fitted curves (solid) for housing data with a study period ranging from January of 2001 to December of 2008 (series 1 for which $t_1 = 0, t_2 = 1, \ldots, t_{96} = 95$). Estimated transitions are marked by solid vertical lines, estimated τ by dot-dashed vertical lines, and estimated critical time points by dashed vertical lines.

Table 2.2: Housing data analysis with a study period ranging from January of 2001 to December of 2008 (series 1 for which $t_1 = 0, t_2 = 1, \ldots, t_{96} = 95$) – posterior summaries of the bent cable and piecewise linear parameters.

Parameter	Generalized Bent-Cable (DIC = 222.71)		Quadratic Bent-Cable (DIC = 222.76)		Piecewise Linear (DIC = 221.38)	
	Posterior median	95% credible interval	Posterior median	95% credible interval	Posterior median	95% credible interval
κ	1.644	$(1.022, 2.917)$	–	–	–	–
β_0	14.930	$(14.570, 15.300)$	14.930	$(14.560, 15.300)$	14.930	$(14.570, 15.300)$
β_1	0.082	$(0.073, 0.092)$	0.082	$(0.072, 0.092)$	0.082	$(0.072, 0.092)$
β_2	-0.438	$(-0.471, -0.407)$	-0.439	$(-0.472, -0.407)$	-0.435	$(-0.467, -0.405)$
$\beta_1 + \beta_2$	-0.356	$(-0.388, -0.326)$	-0.357	$(-0.389, -0.327)$	-0.353	$(-0.384, -0.323)$
γ	3.049	$(0.165, 12.250)$	2.378	$(0.117, 6.350)$	–	–
τ	64.230	$(62.640, 65.740)$	64.240	$(62.710, 65.830)$	–	–
$[\tau_1, \tau_2]$	$[62.480, 67.344]$	–	$[61.779, 66.618]$	–	–	–
CTP	62.908	$(60.510, 64.987)$	62.643	$(60.243, 64.830)$	64.080	$(62.450, 65.650)$
σ^2	0.559	$(0.422, 0.760)$	0.560	$(0.426, 0.764)$	0.558	$(0.425, 0.756)$

For series 2, a comparison of DIC suggests that the two bent-cable models are virtually indistinguishable in terms of the overall fit, and both are superior to the piecewise linear model (DIC = 220.56, 220.01 and 301.02 for the generalized bent-cable, quadratic bent-cable and piecewise

linear models, respectively; also see Figure 2.4). Posterior medians of the parameters also agree closely for the bent-cable models (Table 2.3). As opposed to series 1, series 2 exhibits a different scenario: there was a significant decrease in privately-owned homes in the United States around June 2006 - July 2007 (incoming phase, for which $\hat{\beta}_1 = -0.40$ with 95% credible interval $(-0.42, -0.36)$), which followed by a transition phase until around March 2012 ($\hat{\tau}_2 \approx 69$ or March 2012), and a significant increase thereafter ($\widehat{\beta_1 + \beta_2} = 0.09$ with 95% credible interval $(0.079, 0.095)$). The number of privately-owned homes reached to a minimum around May 2011 ($\widehat{\text{CTP}} \approx 59$ or May 2011).

(a) Quadratic bent-cable fit: $\hat{\beta}_0 = 20.25$, $\hat{\beta}_1 = -0.40$, $\hat{\beta}_2 = 0.48$, $\hat{\gamma} = 28.18$, $\hat{\tau} = 41.30$ (Nov. 2009), $\hat{\tau}_1 = 13.16$ (Jul. 2007), $\hat{\tau}_2 = 69.50$ (Mar. 2012), $\widehat{\text{CTP}} = 59.38$ (May 2011), and and DIC = 220.01.

(b) Generalized bent-cable fit: $\hat{\beta}_0 = 20.21$, $\hat{\beta}_1 = -0.40$, $\hat{\beta}_2 = 0.48$, $\hat{\kappa} = 2.05$, $\hat{\gamma} = 27.73$, $\hat{\tau} = 41.30$ (Nov. 2009), $\hat{\tau}_1 = 12.26$ (Jun. 2007), $\hat{\tau}_2 = 69.09$ (Mar. 2012), $\widehat{\text{CTP}} = 59.23$ (May 2011), and DIC = 220.56.

(c) Piecewise linear model fit: $\hat{\beta}_0 = 19.25$, $\hat{\beta}_1 = -0.31$, $\hat{\beta}_2 = 0.38$, $\widehat{\text{CTP}} = 46.12$ (Apr. 2010), and DIC = 301.02.
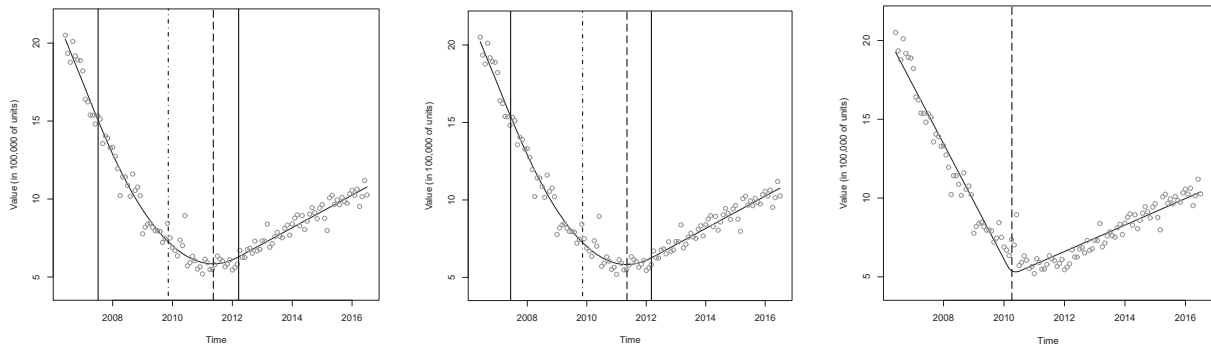
Figure 2.4: Observed data and the corresponding fitted curves (solid) for housing data with a study period ranging from June of 2006 to July of 2016 (series 2 for which $t_1 = 0, t_2 = 1, \ldots, t_{122} = 121$). Estimated transitions are marked by solid vertical lines, estimated τ by dot-dashed vertical lines, and estimated critical time points by dashed vertical lines.

In summary, series 1 analyses demonstrate that the bent-cable and piecewise linear models may perform equally well when data exhibit an abrupt change over time, whereas series 2 analyses reveal that the bent-cable models can produce comparable but significantly superior fits compared to the piecewise linear model (see Section 2.4 for a comparison of these models when data exhibit an extremely sharp bend (i.e. broken-stick) over time). In terms of the trend in States during January 2001 - July 2016, our findings can be summarized as follows: (a) a significant linear

increase in privately-owned homes around January 2001 - March 2006, (b) a transition then occurred around March 2006 - August 2006 when privately-owned homes took a downturn from an increasing trend, (c) a significant decrease thereafter until around July 2007, (c) a transition then occurred again around July 2007 - March 2012 when privately-owned homes took an upturn from a decreasing trend, and (d) a significant increase thereafter until around July 2016.

Table 2.3: Housing data analysis with a study period ranging from June of 2006 to July of 2016 (series 2 for which $t_1 = 0, t_2 = 1, \ldots, t_{122} = 121$) – posterior summaries of the bent cable and piecewise linear parameters.

Parameter	Generalized Bent-Cable (DIC = 220.56)		Quadratic Bent-Cable (DIC = 220.01)		Piecewise Linear (DIC = 301.02)	
	Posterior median	95% credible interval	Posterior median	95% credible interval	Posterior median	95% credible interval
κ	2.054	$(1.742, 2.620)$	–	–	–	–
β_0	20.210	$(19.800, 20.610)$	20.250	$(19.830, 20.630)$	19.250	$(18.680, 19.780)$
β_1	-0.396	$(-0.419, -0.362)$	-0.397	$(-0.420, -0.364)$	-0.306	$(-0.330, -0.278)$
β_2	0.482	$(0.448, 0.509)$	0.484	$(0.450, 0.509)$	0.375	$(0.352, 0.398)$
$\beta_1 + \beta_2$	0.087	$(0.079, 0.095)$	0.087	$(0.079, 0.095)$	0.070	$(0.060, 0.083)$
γ	27.730	$(23.170, 29.890)$	28.180	$(24.490, 29.900)$	–	–
τ	41.300	$(40.060, 43.960)$	41.300	$(40.070, 43.820)$	–	–
$[\tau_1, \tau_2]$	$[12.260, 69.090]$	–	$[13.160, 69.500]$	–	–	–
CTP	59.226	$(57.392, 60.599)$	59.376	$(57.741, 60.652)$	46.120	$(43.590, 49.780)$
σ^2	0.339	$(0.265, 0.444)$	0.335	$(0.262, 0.438)$	0.661	$(0.518, 0.868)$

2.4   Simulation

First, we supplement the motivation for our generalized bent-cable methodology with Scenario 1, where we generate data from the generalized bent-cable model with (a) $\kappa = 2$ (quadratic bend), (b) $\kappa = 1.25$, (c) $\kappa = 3$, and (d) $\gamma = 0$ (piecewise linear or broken stick). We then fit the bent-cable and piecewise linear models to each of the simulated data sets, and compare the fits using coverage probabilities and DIC (see below). Second, to assess the sensitivity of inferences to prior specifications, we present Scenario 2, where the generalized bent-cable fits are compared using different priors for γ (Scenario 2a) and τ (Scenario 2b).

In all the scenarios, we take $n = 273$ and $t_i = i - 1$ for $i = 1, 2, \ldots, n$. Model parameter values were chosen to allow reasonable generalization and are given in Table 2.4. For each simulation,

500 data sets were generated, and 5000 MCMC iterations after burn-in and thinning were used to approximate posterior distributions per set. Posterior summaries were averaged over the 500 sets for each parameters, and the coverage probability of 95% credible intervals (proportion of such intervals out of 500 that capture the truth) was calculated. Letting M = G (generalized bent-cable), Q (quadratic bent-cable) and B (broken stick or piecewise linear), we also defined a DIC-based criterion for model comparison: $p_M = p_{1M} + p_{2M}$, where p_{1M} and p_{2M} are the proportions of model M fits out of 500 for which $DIC_M = DIC_{min}$ and $DIC_M \neq DIC_{min}$ but $DIC_M - DIC_{min} \leq 2$, respectively, and $DIC_{min} = \min\{DIC_G, DIC_Q, DIC_B\}$. Note that p_M is an estimate of approximately how often model M is supported by the data (see Section 2.3 regarding the use of DIC for model comparison).

2.4.1 Results for Scenario 1

Numerical results are summarized in Table 2.4. For Scenario 1a (quadratic bend), the two bent-cable models perform well, and the results are comparable with respect to bias and coverage probability: averages of posterior medians are all close to the true parameter values, and coverage probabilities are all reasonably close to the nominal 0.95. The DIC-based criterion suggests that the quadratic model is supported by the data more often than the generalized model ($p_G = 0.626$ and $p_Q = 0.808$). This finding is not unexpected as the quadratic bent-cable is the true model for Scenario 1a. Nevertheless, the performance of the generalized model in comparison with the true model can be considered quite satisfactory. The performance of the piecewise linear model is poor with respect to both coverage and the DIC-based criterion ($p_{1B} = p_B = 0$), suggesting its inadequacy in characterizing gradual transition over time.

The generalized model performs well with respect to all the parameters when $\kappa \neq 2$ (Scenarios 1b and 1c). We see that the quadratic model can lead to very poor coverage when $\kappa \neq 2$: the coverage probability ranges from 0 to 0.780 when $\kappa = 1.25$, and from 0 to 0.406 when $\kappa = 3$; we also see that the parameters τ_1 and τ_2 are underestimated when $\kappa = 1.25$ (i.e. the transition zone is shifted towards left when $\kappa < 2$), and overestimated when $\kappa = 3$ (i.e. the transition zone is shifted towards right when $\kappa > 2$). Note that Example 1 of Section 2.3 supports the former statement

23

Table 2.4: Simulation results for scenario 1 with $n = 273$: table entries are average of 500 posterior medians of the parameters and coverage of 95% credible intervals; also $p_M = p_{1M} + p_{2M}$, where p_{1M} and p_{2M} are the proportions of model M fits out of 500 for which $DIC_M = DIC_{min}$ and $DIC_M \neq DIC_{min}$ but $DIC_M - DIC_{min} \leq 2$, respectively.

Scenario		True	Generalized Bent-Cable		Quadratic Bent-Cable		Piecewise Linear	
			Mean	Coverage	Mean	Coverage	Mean	Coverage
Scenario 1a	κ	2.00	2.003	0.944	–	–	–	–
(quadratic	β_0	250.00	249.999	0.982	249.997	0.988	253.103	0.000
bend)	β_1	0.70	0.700	0.944	0.700	0.950	0.552	0.000
	β_2	−0.90	−0.900	0.940	−0.900	0.952	−0.740	0.000
	τ_1	30.00	29.869	0.934	30.003	0.946	–	–
	τ_2	120.00	119.980	0.962	120.008	0.960	–	–
	CTP	100.00	99.996	0.956	100.004	0.946	83.543	0.000
	σ^2	0.10	0.101	0.958	0.101	0.958	2.111	0.000
	p_{1M}		0.344	–	0.656	–	0.000	–
	p_M		0.626	–	0.808	–	0.000	–
Scenario 1b	κ	1.25	1.250	0.964	–	–	–	–
	β_0	250.00	250.009	0.980	250.696	0.780	252.795	0.000
	β_1	0.70	0.699	0.938	0.835	0.388	0.504	0.000
	β_2	−0.90	−0.899	0.942	−1.028	0.526	−0.689	0.000
	τ_1	30.00	30.033	0.948	−2.879	0.000	–	–
	τ_2	120.00	120.068	0.944	83.104	0.000	–	–
	CTP	62.94	62.914	0.952	67.005	0.000	54.055	0.000
	σ^2	0.10	0.101	0.944	0.609	0.000	1.990	0.000
	p_{1M}		1.000	–	0.000	–	0.000	–
	p_M		1.000	–	0.000	–	0.000	–
Scenario 1c	κ	3.00	3.054	0.938	–	–	–	–
	β_0	250.00	249.985	0.990	250.330	0.124	252.000	0.000
	β_1	0.70	0.701	0.940	0.681	0.000	0.621	0.000
	β_2	−0.90	−0.901	0.942	−0.883	0.000	−0.812	0.000
	τ_1	30.00	28.381	0.936	58.251	0.000	–	–
	τ_2	120.00	119.904	0.954	125.088	0.000	–	–
	CTP	109.37	109.347	0.950	109.862	0.406	94.948	0.000
	σ^2	0.10	0.102	0.944	0.143	0.012	1.364	0.000
	p_{1M}		1.000	–	0.000	–	0.000	–
	p_M		1.000	–	0.000	–	0.000	–
Scenario 1d	κ	–	1.543	–	–	–	–	–
(broken	β_0	250.00	249.996	0.994	249.994	0.992	250.000	0.992
stick)	β_1	0.70	0.700	0.952	0.700	0.952	0.700	0.952
	β_2	−0.90	−0.900	0.954	−0.900	0.954	−0.900	0.952
	τ_1	–	74.384	–	73.965	–	–	–
	τ_2	–	76.663	–	76.036	–	–	–
	CTP	75.00	75.334	0.992	75.579	0.586	75.006	0.956
	σ^2	0.10	0.102	0.940	0.102	0.946	0.102	0.946
	p_{1M}		0.220	–	0.164	–	0.616	–
	p_M		0.794	–	0.714	–	0.916	–

about under-estimation of the transition parameters (Figure 2.2). The DIC-based criterion also suggests clear superiority of the generalized model over the quadratic model for Scenarios 1b and 1c ($p_{1G} = p_G = 1$ and $p_{1Q} = p_Q = 0$). As indicated above, the piecewise linear model cannot handle 1b and 1c data adequately ($p_{1B} = p_B = 0$).

We see satisfactory performance of the piecewise linear model when data exhibit a broken trend (Scenario 1d in which the piecewise linear is the true model). In terms of bias and coverage probability, the piecewise linear model provides slightly more accurate estimates than those obtained using the generalized model. Note that the quadratic model may lead to under coverage for CTP when data exhibit an abrupt transition (coverage probability for CTP = 0.586). All these results are reflected in the DIC-based criterion: $p_G = 0.794$, $p_Q = 0.714$ and $p_B = 0.916$.

The above results demonstrate the flexibility of the generalized bent-cable model in characterizing a continuous trend exhibiting a gradual change over time. Its performance in describing an abrupt transition is also satisfactory, as demonstrated in Example 2 (Section 2.3) and Scenario 1d.

2.4.2 Results for Scenario 2

To assess the sensitivity of inferences to prior specifications for γ (Scenario 2a) and τ (Scenario 2b), we considered Scenario 1b to generate data for which $\gamma = 72$ and $\tau = 48$, and used N(0, 10000) prior for each of β_0, β_1 and β_2, U(0, 3) prior for κ and G(0.1, 0.1) prior for σ^{-2}. For Scenario 2a, we assumed $\tau \sim$ U(8, 88), and then compared the generalized bent-cable fits for (i) prior$_1$: $\gamma \sim$ U(0, 62), (ii) prior$_2$: $\gamma \sim$ U(0, 67), (iii) prior$_3$: $\gamma \sim$ U(0, 92), and (iv) prior$_4$: $\gamma \sim$ U(0, 372). Note that prior$_1$ and prior$_2$ exclude $\gamma = 72$, whereas prior$_3$ and prior$_4$ include $\gamma = 72$, with prior$_4$ representing a relatively wider interval. Similarly, for Scenario 2b, we assumed $\gamma \sim$ U(0, 92), and then compared the generalized bent-cable fits for (i) prior$_1$: $\tau \sim$ U(5, 40), (ii) prior$_2$: $\tau \sim$ U(5, 45), (iii) prior$_3$: $\tau \sim$ U(8, 88), and (iv) prior$_4$: $\tau \sim$ U(3, 348).

Numerical results are summarized in Table 2.5. We see that prior$_1$ and prior$_2$ lead to very poor coverage and relatively large bias for all the parameters, whereas prior$_3$ and prior$_4$ result in satisfactory performance of our methodology (averages of posterior medians are all close to the true parameter values, and coverage probabilities are all reasonably close to the nominal 0.95).

Table 2.5: Simulation results for scenario 2 (sensitivity analysis to the prior specification for γ and τ), with simulated $\gamma = 72$ and $\tau = 48$, and $n = 273$: table entries are average of 500 posterior medians of the parameters and coverage of 95% credible intervals.

Scenario		True	Prior$_1$ Mean	Coverage	Prior$_2$ Mean	Coverage	Prior$_3$ Mean	Coverage	Prior$_4$ Mean	Coverage
Scenario 2a	κ	1.25	1.308	0.000	1.277	0.000	1.250	0.964	1.250	0.962
Prior$_1$:	β_0	250.00	250.115	0.928	250.054	0.968	250.009	0.980	250.003	0.986
$\gamma \sim U(0, 62)$	β_1	0.70	0.692	0.842	0.697	0.932	0.699	0.938	0.700	0.966
Prior$_2$:	β_2	-0.90	-0.890	0.744	-0.895	0.902	-0.899	0.942	-0.900	0.966
$\gamma \sim U(0, 67)$	τ_1	30.00	28.787	0.202	29.388	0.710	30.033	0.948	29.995	0.964
Prior$_3$:	τ_2	120.00	109.666	0.000	114.623	0.000	120.068	0.944	120.090	0.960
$\gamma \sim U(0, 92)$	CTP	62.94	64.581	0.000	63.861	0.132	62.914	0.952	62.932	0.954
Prior$_4$:	σ^2	0.10	0.127	0.210	0.108	0.866	0.101	0.944	0.101	0.958
$\gamma \sim U(0, 372)$										
Scenario 2b	κ	1.25	1.357	0.000	1.270	0.410	1.250	0.964	1.250	0.972
Prior$_1$:	β_0	250.00	248.229	0.006	249.185	0.000	250.009	0.980	250.004	0.992
$\tau \sim U(5, 40)$	β_1	0.70	0.917	0.000	0.772	0.000	0.699	0.938	0.700	0.950
Prior$_2$:	β_2	-0.90	-1.113	0.000	-0.971	0.000	-0.899	0.942	-0.900	0.952
$\tau \sim U(5, 45)$	τ_1	30.00	16.990	0.000	26.203	0.000	30.033	0.948	30.000	0.930
Prior$_3$:	τ_2	120.00	103.896	0.000	114.571	0.020	120.068	0.944	119.995	0.960
$\tau \sim U(8, 88)$	CTP	62.94	67.307	0.000	64.057	0.252	62.914	0.952	62.939	0.954
Prior$_4$:	σ^2	0.10	0.333	0.000	0.148	0.004	0.101	0.944	0.101	0.948
$\tau \sim U(3, 348)$										

These results suggest that the Bayesian estimation is roughly robust when the true γ and τ fall within the bounds of their respective uniform priors. Note that since both γ and τ are positive, a reasonable choice of the lower bound of γ is 0, and that of τ is a small positive number.

2.5 Conclusion

Bent-cable regression is an appealing statistical tool to model changepoint data due to the model's flexibility while being parsimonious with greatly interpretable regression coefficients. The model assumes a continuous function comprised of three segments connected together: two linear segments to describe the incoming and outgoing phases, joined by a quadratic function to model the transition zone. As demonstrated in this study, the assumption of a quadratic bend is arguably a restrictive assumption in modeling the transition zone for changepoint data. Specifically, this assumption may lead to the starting and the end points of the transition period to be either underestimated or overestimated. In this study, a generalization of the bent-cable model is proposed

to overcome this problem. The generalized model retains the interpretability of the regression coefficients, while improves the accuracy of locating the transition phase at the small cost of estimating one additional parameter. As highlighted in Section 2, it provides sufficient flexibility so that inference about the transition zone can be data driven, rather than pre-assumed as a specific type.

The piecewise linear model has been heavily utilized to describe a continuous trend exhibiting at least one abrupt change over time (Muggeo 2003). An issue with such modeling is that the abrupt change at the point where the stick breaks may be artificial, with possibly a more natural smoothness reflecting the trajectory change. The concept of bent-cable regression allows for a balance between model interpretability and forcing the response to change its trajectory in such an abrupt manner. The examples presented in this article demonstrate that the generalized bent-cable model has the ability to adequately describe changepoint data that exhibit either an abrupt (series 1 housing data) or gradual (CFC-11 and series 2 housing data) transition over time. These examples also demonstrate that the quadratic bent-cable and piecewise linear models can adequately characterize profiles that obviously follow the shape of the quadratic and abrupt bend, respectively. Note that despite the broken stick being the limiting case of the bent-cable, reliable inference for the broken-stick transition may require that the stick trajectory be explicitly characterized by the model; our simulations suggest that the piecewise linear model may provide slightly more accurate estimates than those obtained using the generalized bent-cable model when data follow the shape of the broken stick. Nevertheless, the generalized bent-cable model can be valuable in modeling different types of changepoint data as demonstrated in this article.

CONCLUDING REMARKS AND FUTURE WORK

The focus of this study is to model changepoint data exhibiting three phases, where it rises initially in a linear fashion, then goes through a curved transition phase, followed by a linear decreasing trend. A piecewise linear model is a natural candidate to characterize such trajectories. It consists of two linear components with opposite signs of slopes, and a breakpoint where the two lines intersect. However, the assumption of a broken-stick trend is not realistic in many applications (Chiu et al. 2005). A more general class of models considers a smooth transition (gradual or abrupt) over time, and takes the piecewise linear model as a special case (e.g. Bacon and Watts 1971, Chiu et al. 2006). In this study, we propose a smooth changepoint model by considering a substantial extension of the bent-cable regression model of Chiu et al. (2006). The bent-cable model utilizes a quadratic function to characterize the transition phase. Although a quadratic function can be adequate to describe many changepoint trajectories, it also limits its applications to a broader context. Our model utilizes a more flexible function to describe the transition phase, making it a more general model in characterizing different types of changepoint data. Note that both the quadratic bent-cable and the piecewise linear models are special cases of our generalized bent-cable model. We explore the properties of the generalized model, and develop a Bayesian approach for statistical inference. The proposed methodology is then demonstrated with applications to three examples. This study reveals that the generalized model is more flexible than the quadratic model in characterizing different types of changepoint trajectories; unless a trajectory obviously follows the shape of a quadratic or broken-stick trend, it is highly likely that the generalized model will produce a fit which is superior to those produced by the quadratic bent-cable and piecewise linear models.

3.1 Cautionary Remarks

Although our methodology provides a flexible approach to model changepoint trajectories with greatly interpretable regression coefficients, some caution is required for the following reasons.

1. Unrefined choice of the hyperparameters for γ and τ may lead to poor mixing of a Markov chain, or even a computational breakdown of the MCMC. Therefore, a reasonable set of hyperprior values for γ and τ is necessary for good mixing of the Markov chains and convergence to the stationary distribution; see Sections 2.2 and 2.4 for our recommendation.

2. Despite the quadratic bent-cable and broken-stick being the limiting case of the generalized bent-cable, reliable inference for the broken-stick/quadratic transition may require that the model explicitly acknowledge the stick/quadratic trajectory. Our simulations suggest that (a) the piecewise linear model may provide slightly more accurate estimates than those obtained using the generalized bent-cable model when data follow the shape of the broken-stick, and (b) the quadratic bent-cable model may provide slightly more accurate estimates than those obtained using the generalized bent-cable model when data follow the shape of the quadratic bend. However, it is not straightforward to assess the exact shape of a transition. The generalized bent-cable model can be useful in this regard, as it takes the quadratic bent-cable and the piecewise linear models as special cases. Although the likelihood ratio method of frequentist approach is not justified in Bayesian inference (i.e. testing $H_0{:}\kappa{=}2$ using the likelihood ratio method to check whether the quadratic model fits the data as well as the generalized model), we can use DIC for model comparison: fit all the three models under consideration and choose the one with the smallest DIC value.

3. The bent-cable methodology is intended for data that exhibit only one transition period over time.

3.2 Future Work

There is scope to extend the generalized bent-cable regression to a more general framework, presented as follows.

1. The bent-cable model was utilized to study longitudinal data. Our proposed generalized bent-cable model can also be extended to study longitudinal data. For example, CFC-11 and CFC-12 data might be considered for this purpose.

2. Khan et al. (2012) studied the existing longitudinal bent-cable model to handle spatial effects where they allowed the error terms to be correlated across space in a hierarchical Bayesian framework. It might be the another possibility to extend the generalized bent-cable by considering spatial effects into the model; that would be possible for CFC data and others that comprise spatial data.

3.3 Publication

An article is published in the *Journal of Applied Statistics* (Khan and Kar 2017). We intend to write one more article on the computational aspects and software implementation of the MCMC algorithm, possibly in a journal in computational statistics.

SOFTWARE IMPLEMENTATION

A.1   WinBUGS Model

A.1.1   WinBUGS Model for Generalized Bent-Cable Regression

```
model{
  for (i in 1:n) {
    q0[i] <-
      ((t[i] - tau + gam * (kappa - 1)) / (gam * kappa)) *
      (1 - step(tau - gam * (kappa - 1) - t[i])) *
      (1 - step(t[i] - tau - gam))
    q1[i] <- pow(q0[i], kappa) * gam
    q2[i] <- (t[i] - tau) * step(t[i] - tau - gam)
    q[i] <- q1[i] + q2[i]
    mu[i] <- beta0 + beta1 * t[i] + beta2 * q[i]
    y[i] ~ dnorm(mu[i], inv.s2)
  }
  beta0 ~ dnorm(a01, a02)
  beta1 ~ dnorm(a11, a12)
  beta2 ~ dnorm(a21, a22)
  tau ~ dunif(b11, b12)
  gam ~ dunif(b21, b22)
  kappa ~ dunif(c1, c2)
  inv.s2 ~ dgamma(d1, d2)
  s2 <- 1 / inv.s2
}
```

A.1.2   WinBUGS Model for Quadratic Bent-Cable Regression

```
model{
  for (i in 1:n) {
    q0[i] <-
      ((t[i] - tau + gam * (kappa - 1)) / (gam * kappa)) *
      (1 - step(tau - gam * (kappa - 1) - t[i])) *
      (1 - step(t[i] - tau - gam))
    q1[i] <- pow(q0[i], kappa) * gam
    q2[i] <- (t[i] - tau) * step(t[i] - tau - gam)
```

```
    q[i] <- q1[i] + q2[i]
    mu[i] <- beta0 + beta1 * t[i] + beta2 * q[i]
    y[i] ~ dnorm(mu[i], inv.s2)
  }
  beta0 ~ dnorm(a01, a02)
  beta1 ~ dnorm(a11, a12)
  beta2 ~ dnorm(a21, a22)
  tau ~ dunif(b11, b12)
  gam ~ dunif(b21, b22)
  inv.s2 ~ dgamma(d1, d2)
  s2 <- 1 / inv.s2
}
```

A.1.3 WinBUGS Model for Piecewise Linear Regression

```
model {
  for (i in 1:n) {
    q[i] <- (t[i] - tau) * step(t[i] - tau)
    mu[i] <- beta0 + beta1 * t[i] + beta2 * q[i]
    y[i] ~ dnorm(mu[i], inv.s2)
  }
  beta0 ~ dnorm(a0, a1)
  beta1 ~ dnorm(b0, b1)
  beta2 ~ dnorm(c0, c1)
  tau ~ dunif(d0, d1)
  inv.s2 ~ dgamma(g0, g1)
  s2 <- 1 / inv.s2
}
```

A.2 Barrow CFC-11 Data Analysis

The models given in Sections A.1.1 and A.1.3 can, in general, be used for Bayesian inference of the generalized bent-cable, quadratic bent-cable and piecewise linear models. However, the "data" and "initial values" components of WinBUGS have to be modified, depending on the data to be analyzed. In this section, we present the data and initial values that we used in analyzing the Barrow CFC-11 data.

Data and Initial Values for the Generalized Bent-Cable Model

```
# Barrow CFC-11 Data
list(
  y = c(252.765, 253.583, 254.68, 254.536, 252.981, 253.638, 253.764, 256.299, 258.14, 259.677,
     263.172, 264.297, 263.477, 263.390, 263.763, 263.462, 263.156, 263.221, 263.601, 263.221,
     265.145, 266.688, 267.909, 271.515, 272.458, 273.488, 270.130, 269.972, 269.742, 269.471,
     269.183, 270.650, 270.194, 271.878, 273.478, 274.812, 275.219, 274.504, 273.822, 272.377,
     272.328, 272.061, 272.219, 272.139, 272.276, 274.093, 274.917, 276.758, 276.696, 274.906,
     275.193, 274.379, 274.818, 274.015, 273.628, 273.925, 273.724, 274.747, 275.154, 276.461,
     276.215, 276.171, 275.877, 275.196, 274.542, 273.985, 273.817, 274.201, 273.885, 275.167,
     275.624, 276.181, 276.254, 276.124, 277.264, 275.625, 273.167, 274.053, 273.264, 272.757,
     272.804, 274.074, 274.263, 274.290, 272.839, 272.265, 272.530, 272.520, 272.526, 272.558,
     271.769, 271.299, 272.071, 272.071, 272.071, 272.990, 272.521, 273.614, 272.928, 272.192,
     271.672, 271.115, 269.800, 269.567, 269.261, 269.223, 270.284, 271.019, 271.849, 270.613,
     269.987, 266.912, 266.856, 267.331, 266.935, 266.322, 267.856, 269.644, 268.863, 267.378,
     267.232, 269.583, 269.060, 268.776, 268.490, 267.781, 266.610, 266.206, 266.188, 266.330,
     266.486, 266.769, 265.790, 265.260, 264.970, 264.820, 264.640, 263.890, 263.410, 263.090,
     263.010, 263.640, 264.320, 264.870, 265.140, 265.090, 264.580, 264.370, 264.430, 264.620,
     264.120, 263.580, 263.300, 263.800, 263.880, 264.100, 263.980, 263.760, 263.560, 263.520,
     262.950, 262.610, 261.950, 261.430, 261.070, 261.370, 261.670, 261.770, 262.100, 261.330,
     261.220, 261.030, 260.520, 260.160, 259.720, 259.340, 258.930, 259.020, 259.220, 259.420,
     259.310, 259.250, 258.980, 258.740, 258.780, 258.150, 257.620, 256.830, 256.610, 256.890,
     257.060, 257.040, 256.880, 256.680, 256.550, 256.380, 256.160, 255.470, 254.950, 254.400,
     254.130, 254.380, 254.630, 254.640, 254.680, 254.510, 254.340, 254.250, 253.970, 253.860,
     253.350, 252.810, 252.210, 252.420, 252.420, 252.400, 252.190, 251.930, 251.880, 251.410,
     251.100, 250.890, 250.080, 249.250, 249.030, 249.260, 249.530, 249.720, 249.590, 249.560,
     249.270, 248.950, 248.680, 248.440, 248.050, 247.630, 247.260, 247.140, 247.370, 247.490,
     247.220, 247.250, 247.240, 246.760, 246.380, 246.120, 245.650, 245.190, 245.270, 245.540,
     245.800, 246.010, 245.860, 245.570, 245.410, 245.230, 244.540, 244.090, 244.010, 243.570,
     243.280, 243.410, 243.630, 243.870, 243.860, 243.440, 243.320, 243.170, 242.910, 242.300,
     241.870, 241.190,241.050),
  t = c(0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23,
     24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45,
     46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67,
     68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89,
     90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109,
     110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127,
     128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145,
     146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163,
     164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181,
     182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199,
     200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217,
     218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235,
     236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253,
     254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271,
     272),
  n = 273, a01 = 0, a02 = 0.001, a11 = 0, a12 = 0.001, a21 = 0, a22 = 0.001,
  b11 = 5, b12 = 125, b21 = 0, b22 = 80, c1 = 1, c2 = 3, d1 = 0.1, d2 = 0.1
)
# Initial Values
list(beta0 = 255, beta1 = 0.64, beta2 = -0.81, tau = 50, gam = 50, kappa = 1.5, inv.s2 = 1)
```

Data and Initial Values for the Quadratic Bent-Cable Model

```
# Barrow CFC-11 Data
list(
  y = c(252.765, 253.583, 254.68, 254.536, 252.981, 253.638, 253.764, 256.299, 258.14, 259.677,
    263.172, 264.297, 263.477, 263.390, 263.763, 263.462, 263.156, 263.221, 263.601, 263.221,
    265.145, 266.688, 267.909, 271.515, 272.458, 273.488, 270.130, 269.972, 269.742, 269.471,
    269.183, 270.650, 270.194, 271.878, 273.478, 274.812, 275.219, 274.504, 273.822, 272.377,
    272.328, 272.061, 272.219, 272.139, 272.276, 274.093, 274.917, 276.758, 276.696, 274.906,
    275.193, 274.379, 274.818, 274.015, 273.628, 273.925, 273.724, 274.747, 275.154, 276.461,
    276.215, 276.171, 275.877, 275.196, 274.542, 273.985, 273.817, 274.201, 273.885, 275.167,
    275.624, 276.181, 276.254, 276.124, 277.264, 275.625, 273.167, 274.053, 273.264, 272.757,
    272.804, 274.074, 274.263, 274.290, 272.839, 272.265, 272.530, 272.520, 272.526, 272.558,
    271.769, 271.299, 272.071, 272.071, 272.071, 272.990, 272.521, 273.614, 272.928, 272.192,
    271.672, 271.115, 269.800, 269.567, 269.261, 269.223, 270.284, 271.019, 271.849, 270.613,
    269.987, 266.912, 266.856, 267.331, 266.935, 266.322, 267.856, 269.644, 268.863, 267.378,
    267.232, 269.583, 269.060, 268.776, 268.490, 267.781, 266.610, 266.206, 266.188, 266.330,
    266.486, 266.769, 265.790, 265.260, 264.970, 264.820, 264.640, 263.890, 263.410, 263.090,
    263.010, 263.640, 264.320, 264.870, 265.140, 265.090, 264.580, 264.370, 264.430, 264.620,
    264.120, 263.580, 263.300, 263.800, 263.880, 264.100, 263.980, 263.760, 263.560, 263.520,
    262.950, 262.610, 261.950, 261.430, 261.070, 261.370, 261.670, 261.770, 262.100, 261.330,
    261.220, 261.030, 260.520, 260.160, 259.720, 259.340, 258.930, 259.020, 259.220, 259.420,
    259.310, 259.250, 258.980, 258.740, 258.780, 258.150, 257.620, 256.830, 256.610, 256.890,
    257.060, 257.040, 256.880, 256.680, 256.550, 256.380, 256.160, 255.470, 254.950, 254.400,
    254.130, 254.380, 254.630, 254.640, 254.680, 254.510, 254.340, 254.250, 253.970, 253.860,
    253.350, 252.810, 252.210, 252.420, 252.420, 252.400, 252.190, 251.930, 251.880, 251.410,
    251.100, 250.890, 250.080, 249.250, 249.030, 249.260, 249.530, 249.720, 249.590, 249.560,
    249.270, 248.950, 248.680, 248.440, 248.050, 247.630, 247.260, 247.140, 247.370, 247.490,
    247.220, 247.250, 247.240, 246.760, 246.380, 246.120, 245.650, 245.190, 245.270, 245.540,
    245.800, 246.010, 245.860, 245.570, 245.410, 245.230, 244.540, 244.090, 244.010, 243.570,
    243.280, 243.410, 243.630, 243.870, 243.860, 243.440, 243.320, 243.170, 242.910, 242.300,
    241.870, 241.190,241.050),
  t = c(0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23,
    24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45,
    46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67,
    68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89,
    90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109,
    110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127,
    128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145,
    146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163,
    164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181,
    182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199,
    200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217,
    218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235,
    236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253,
    254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271,
    272),
  n = 273, kappa = 2, a01 = 0, a02 = 0.001, a11 = 0, a12 = 0.001, a21 = 0, a22 = 0.001,
  b11 = 5, b12 = 125, b21 = 0, b22 = 80, d1 = 0.1, d2 = 0.1
)
# Initial Values
list(beta0 = 255, beta1 = 0.64, beta2 = -0.81, tau = 50, gam = 50, inv.s2 = 1)
```

Data and Initial Values for the Piecewise Linear Model

```
# Barrow CFC-11 Data
list(
  y = c(252.765, 253.583, 254.68, 254.536, 252.981, 253.638, 253.764, 256.299, 258.14, 259.677,
    263.172, 264.297, 263.477, 263.390, 263.763, 263.462, 263.156, 263.221, 263.601, 263.221,
    265.145, 266.688, 267.909, 271.515, 272.458, 273.488, 270.130, 269.972, 269.742, 269.471,
    269.183, 270.650, 270.194, 271.878, 273.478, 274.812, 275.219, 274.504, 273.822, 272.377,
    272.328, 272.061, 272.219, 272.139, 272.276, 274.093, 274.917, 276.758, 276.696, 274.906,
    275.193, 274.379, 274.818, 274.015, 273.628, 273.925, 273.724, 274.747, 275.154, 276.461,
    276.215, 276.171, 275.877, 275.196, 274.542, 273.985, 273.817, 274.201, 273.885, 275.167,
    275.624, 276.181, 276.254, 276.124, 277.264, 275.625, 273.167, 274.053, 273.264, 272.757,
    272.804, 274.074, 274.263, 274.290, 272.839, 272.265, 272.530, 272.520, 272.526, 272.558,
    271.769, 271.299, 272.071, 272.071, 272.071, 272.990, 272.521, 273.614, 272.928, 272.192,
    271.672, 271.115, 269.800, 269.567, 269.261, 269.223, 270.284, 271.019, 271.849, 270.613,
    269.987, 266.912, 266.856, 267.331, 266.935, 266.322, 267.856, 269.644, 268.863, 267.378,
    267.232, 269.583, 269.060, 268.776, 268.490, 267.781, 266.610, 266.206, 266.188, 266.330,
    266.486, 266.769, 265.790, 265.260, 264.970, 264.820, 264.640, 263.890, 263.410, 263.090,
    263.010, 263.640, 264.320, 264.870, 265.140, 265.090, 264.580, 264.370, 264.430, 264.620,
    264.120, 263.580, 263.300, 263.800, 263.880, 264.100, 263.980, 263.760, 263.560, 263.520,
    262.950, 262.610, 261.950, 261.430, 261.070, 261.370, 261.670, 261.770, 262.100, 261.330,
    261.220, 261.030, 260.520, 260.160, 259.720, 259.340, 258.930, 259.020, 259.220, 259.420,
    259.310, 259.250, 258.980, 258.740, 258.780, 258.150, 257.620, 256.830, 256.610, 256.890,
    257.060, 257.040, 256.880, 256.680, 256.550, 256.380, 256.160, 255.470, 254.950, 254.400,
    254.130, 254.380, 254.630, 254.640, 254.680, 254.510, 254.340, 254.250, 253.970, 253.860,
    253.350, 252.810, 252.210, 252.420, 252.420, 252.400, 252.190, 251.930, 251.880, 251.410,
    251.100, 250.890, 250.080, 249.250, 249.030, 249.260, 249.530, 249.720, 249.590, 249.560,
    249.270, 248.950, 248.680, 248.440, 248.050, 247.630, 247.260, 247.140, 247.370, 247.490,
    247.220, 247.250, 247.240, 246.760, 246.380, 246.120, 245.650, 245.190, 245.270, 245.540,
    245.800, 246.010, 245.860, 245.570, 245.410, 245.230, 244.540, 244.090, 244.010, 243.570,
    243.280, 243.410, 243.630, 243.870, 243.860, 243.440, 243.320, 243.170, 242.910, 242.300,
    241.870, 241.190,241.050),
  t = c(0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23,
    24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45,
    46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67,
    68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89,
    90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109,
    110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127,
    128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145,
    146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163,
    164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181,
    182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199,
    200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217,
    218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235,
    236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253,
    254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271,
    272),
  n = 273, a0 = 0, a1 = 0.001, b0 = 0, b1 = 0.001, c0 = 0, c1 = 0.001, d0 = 5, d1 = 125, g0 = 0.1, g1 = 0.1
)
# Initial Values
list(beta0 = 255, beta1 = 0.64, beta2 = -0.81, tau = 50, inv.s2 = 1)
```

A.3 Analysis of the MCMC Samples

The MCMC samples generated by the *CODA* output from the "Sample Monitor Tool" in Win-BUGS are saved in external text files, which are subsequently analyzed using the "coda" package (Plummer et al. 2006) in R (R Core Team 2016). In this section, we present R codes to analyze the MCMC samples.

```
library(coda)
##################################################
CTP <- function(b1, b2, tau, gam, kappa = NULL) {
  if (!is.null(kappa)) {
    ctp <- tau - gam * (kappa - 1) + (-((kappa * gam) ^ (kappa - 1)) * (b1 / b2)) ^
      (1 / (kappa - 1))
  } else{
    ctp <- tau - gam - (2 * b1 * gam) / b2
  }
  return(ctp)
}
##################################################
llik <- function(theta, y, t) {
  b0 <- theta["b0"]
  b1 <- theta["b1"]
  b2 <- theta["b2"]
  gam <- theta["gam"]
  tau <- theta["tau"]
  kappa <- as.numeric(ifelse(is.na(theta["kappa"]), 2, theta["kappa"]))
  s2 <- theta["s2"]
  tau1 <- tau - (kappa - 1) * gam
  tau2 <- tau + gam
  q0 <- gam * ((t - tau + (kappa - 1) * gam) / (kappa * gam)) ^ kappa
  q1 <- t - tau
  qq <- c(q0[t > tau1 & t <= tau2], q1[t > tau2])
  qq <- c(rep(0, (length(t) - length(qq))), qq)
  f <- b0 + b1 * t + b2 * qq
  log.lik <- sum(dnorm(y, mean = f, sd = sqrt(s2), log = T))
  return(log.lik)
}
##################################################
llik.b <- function(theta, y, t) {
  b0 <- theta["b0"]
  b1 <- theta["b1"]
  b2 <- theta["b2"]
  tau <- theta["tau"]
  s2 <- theta["s2"]
  q <- (t - tau) * ifelse(t >= tau, 1, 0)
  f <- b0 + b1 * t + b2 * q
  log.lik <- sum(dnorm(y, mean = f, sd = sqrt(s2), log = T))
  return(log.lik)
}
```

```
#####################################################
q.cable <- function(alp, t) {
  gam <- alp["gam"]
  tau <- alp["tau"]
  kappa <- alp["kappa"]
  tau1 <- tau - (kappa - 1) * gam
  tau2 <- tau + gam
  q0 <- gam * ((t - tau + (kappa - 1) * gam) / (kappa * gam)) ^ kappa
  q1 <- t - tau
  qq <- c(q0[t > tau1 & t <= tau2], q1[t > tau2])
  qq <- c(rep(0, (length(t) - length(qq))), qq)
  return(qq)
}
#####################################################
f.cable <- function(theta, t) {
  b <- c(theta["b0"], theta["b1"], theta["b2"])
  kappa <- as.numeric(ifelse(is.na(theta["kappa"]), 2, theta["kappa"]))
  alp <- c(theta["gam"], theta["tau"], kappa)
  names(alp)[3] <- "kappa"
  q <- q.cable(alp, t)
  f <- b[1] + b[2] * t + b[3] * q
  return(f)
}
#####################################################
f.cable.b <- function(theta, t) {
  q <- (t - theta["tau"]) * ifelse(t >= theta["tau"], 1, 0)
  f <- theta["b0"] + theta["b1"] * t + theta["b2"] * q
  return(f)
}
#####################################################
f.cable.post <- function(mcmc.sample1, t) {
  f <- apply(mcmc.sample1, 1, f.cable, t = t)
  f.post <- apply(f, 1, quantile, probs = c(0.025, 0.5, 0.975))
  f.med <- f.post[2, ]
  f.lower <- f.post[1, ]
  f.upper <- f.post[3, ]
  return(list(f.med = f.med, f.lower = f.lower, f.upper = f.upper))
}
#####################################################
f.cable.post.b <- function(mcmc.sample1, t) {
  f <- apply(mcmc.sample1, 1, f.cable.b, t = t)
  f.post <- apply(f, 1, quantile, probs = c(0.025, 0.5, 0.975))
  f.med <- f.post[2, ]
  f.lower <- f.post[1, ]
  f.upper <- f.post[3, ]
  return(list(f.med = f.med, f.lower = f.lower, f.upper = f.upper))
}
#####################################################
DIC <- function(y, t, theta, mcmc.sample1) {
  Dhat <- -2 * llik(theta, y, t)
  Dbar <- mean(-2 * apply(mcmc.sample1, 1, llik, y = y, t = t))
```

```
  pD <- Dbar - Dhat
  DIC.val <- Dbar + pD
  return(DIC.val)
}
###################################################
DIC.b <- function(y, t, theta, mcmc.sample1) {
  Dhat <- -2 * llik.b(theta, y, t)
  Dbar <- mean(-2 * apply(mcmc.sample1, 1, llik.b, y = y, t = t))
  pD <- Dbar - Dhat
  DIC.val <- Dbar + pD
  return(DIC.val)
}
###################################################
# The following function will produce posterior
# summaries.
# model = "G" (generalized bent-cable),
#         "Q" (quadratic bent-cable), or
#         "B" (piecewise linear).
# If time is calender time, give the starting date
# and the end date to produce a time series plot
# for the fit.
###################################################
post.summary <- function(y, t, chain1, chain2, burn, thin, model,
                         dic = FALSE, start.date = NULL, end.date = NULL,
                         fit = TRUE, density = TRUE, trace = TRUE,
                         xlabel = NULL, ylabel = NULL) {
  if (model == "B") {
    sam10 <- cbind(chain1, out.slope = chain1[, 2] + chain1[, 3])
    sam20 <- cbind(chain2, out.slope = chain2[, 2] + chain2[, 3])
  }
  if (model == "G" || model == "Q") {
    if (model == "G") {
      tau1.dat1 <- chain1[, "tau"] - (chain1[, "kappa"] - 1) * chain1[, "gam"]
      tau2.dat1 <- chain1[, "tau"] + chain1[, "gam"]
      ctp.dat1 <- CTP(chain1[, "b1"], chain1[, "b2"], chain1[, "tau"],
                      chain1[, "gam"], chain1[, "kappa"])
      tau1.dat2 <- chain2[, "tau"] - (chain2[, "kappa"] - 1) * chain2[, "gam"]
      tau2.dat2 <- chain2[, "tau"] + chain2[, "gam"]
      ctp.dat2 <- CTP(chain2[, "b1"], chain2[, "b2"], chain2[, "tau"],
                      chain2[, "gam"], chain2[, "kappa"])
    } else{
      tau1.dat1 <- chain1[, "tau"] - chain1[, "gam"]
      tau2.dat1 <- chain1[, "tau"] + chain1[, "gam"]
      ctp.dat1 <- CTP(chain1[, "b1"], chain1[, "b2"], chain1[, "tau"],
                      chain1[, "gam"])
      tau1.dat2 <- chain2[, "tau"] - chain2[, "gam"]
      tau2.dat2 <- chain2[, "tau"] + chain2[, "gam"]
      ctp.dat2 <- CTP(chain2[, "b1"], chain2[, "b2"], chain2[, "tau"],
                      chain2[, "gam"])
    }
    sam1 <- cbind(chain1, out.slope = (chain1[, "b1"] + chain1[, "b2"]),
```

```
      tau1 = tau1.dat1, tau2 = tau2.dat1, ctp = ctp.dat1)
    na.sam1 <- which(is.na(sam1), arr.ind = TRUE)[, 1]
    inf.sam1 <- which(is.infinite(sam1), arr.ind = TRUE)[, 1]
    indx.sam1 <- unique(c(na.sam1, inf.sam1))
    sam2 <- cbind(chain2, out.slope = (chain2[, "b1"] + chain2[, "b2"]),
      tau1 = tau1.dat2, tau2 = tau2.dat2, ctp = ctp.dat2)
    na.sam2 <- which(is.na(sam2), arr.ind = TRUE)[, 1]
    inf.sam2 <- which(is.infinite(sam2), arr.ind = TRUE)[, 1]
    indx.sam2 <- unique(c(na.sam2, inf.sam2))
    indx.sam <- unique(c(indx.sam1, indx.sam2))
    if (length(indx.sam) > 0) {
      sam10 <- sam1[-indx.sam, ]
      sam20 <- sam2[-indx.sam, ]
    } else{
      sam10 <- sam1
      sam20 <- sam2
    }
  }
  sam11 <- sam10[(burn + 1):nrow(sam10), ]
  sam12 <- sam11[(1:thin == thin), ]
  sam21 <- sam20[(burn + 1):nrow(sam20), ]
  sam22 <- sam21[(1:thin == thin), ]
  mcmc.sam <- mcmc.list(mcmc(sam12), mcmc(sam22))
  sum.coda <- summary(mcmc.sam)
  sum.stat <- cbind(mean = sum.coda[1]$statistics[, 1],
    median = sum.coda[2]$quantiles[, 3], SD = sum.coda[1]$statistics[, 2],
    lower.95 = sum.coda[2]$quantiles[, 1],upper.95 = sum.coda[2]$quantiles[, 5]
  )
  if (!is.null(start.date)) {
    if (model == "B") {
      transition.date <- sum.stat[c("tau"), ]
      mean.date <-
        as.Date(ceiling(transition.date[1] / 12 * 365), origin = start.date)
      median.date <-
        as.Date(ceiling(transition.date[2] / 12 * 365), origin = start.date)
      lower.date <-
        as.Date(ceiling(transition.date[4] / 12 * 365), origin = start.date)
      upper.date <-
        as.Date(ceiling(transition.date[5] / 12 * 365), origin = start.date)
      transition.date1 <-
        data.frame(mean.date, median.date, lower.date, upper.date)
      rownames(transition.date1) <- "tau"
    } else{
      transition.date <- sum.stat[c("tau", "tau1", "tau2", "ctp"), ]
      mean.date <-
        as.Date(ceiling(transition.date[, 1] / 12 * 365), origin = start.date)
      median.date <-
        as.Date(ceiling(transition.date[, 2] / 12 * 365), origin = start.date)
      lower.date <-
        as.Date(ceiling(transition.date[, 4] / 12 * 365), origin = start.date)
      upper.date <-
```

```
          as.Date(ceiling(transition.date[, 5] / 12 * 365), origin = start.date)
        transition.date1 <-
          data.frame(mean.date, median.date, lower.date, upper.date)
    }
}
mcmc.sam0 <- mcmc.list(mcmc(sam12[, 1:ncol(chain1)]), mcmc(sam22[, 1:ncol(chain2)]))
R <- gelman.diag(mcmc.sam0)
mcmc.sample1 <- rbind(sam12, sam22)
theta <- sum.stat[, 2]
dic0 <- NA
if (dic) {
  if (model == "B") {
    dic0 <- DIC.b(y, t, theta, mcmc.sample1)
  } else{
    dic0 <- DIC(y, t, theta, mcmc.sample1)
  }
}
if (trace) {
  x11()
  if (model == "G") {
    par(mfrow = c(3, 3))
    plot(mcmc.sam0, trace = T, density = F, auto.layout = F)
  }
  if (model == "Q" || model == "B") {
    par(mfrow = c(2, 3))
    plot(mcmc.sam0, trace = T, density = F, auto.layout = F)
  }
}
if (density) {
  x11()
  if (model == "G") {
    par(mfrow = c(3, 3))
    plot(mcmc.sam0, trace = F, density = T, auto.layout = F)
  }
  if (model == "Q" || model == "B") {
    par(mfrow = c(2, 3))
    plot(mcmc.sam0, trace = F, density = T, auto.layout = F)
  }
}
if (fit) {
  if (model == "B") {
    post.med <- apply(mcmc.sample1, 2, median)
    tau <- post.med["tau"]
    f.post <- f.cable.post.b(mcmc.sample1, t)
    f.med <- f.post$f.med
    f.lower <- f.post$f.lower
    f.upper <- f.post$f.upper
    x11()
    if (!is.null(start.date)) {
      start.y <- as.numeric(substr(start.date, 1, 4))
      start.m <- as.numeric(substr(start.date, 6, 7))
```

```
      end.y <- as.numeric(substr(end.date, 1, 4))
      end.m <- as.numeric(substr(end.date, 6, 7))
      plot(ts(f.med, frequency = 12, start = c(start.y, start.m), end = c(end.y, end.m))
        ylim = c((min(min(f.med), min(y)) - 1), (max(max(f.med), max(y)) + 1)),
        col = 1, lty = 1, ylab = ylabel, xlab = xlabel)
      lines(ts(y, frequency = 12, start = c(start.y, start.m), end = c(end.y, end.m)),
        lty = 0, type = "p", col = "grey50")
      abline(v = (start.y + (start.m - 1) / 12 + tau / 12),
        lty = "longdash", col = 1, lwd = 2)
    } else{
      plot(t, y, lty = 3, col = "grey50", ylim = c(min(y), max(y)),
        ylab = ylabel, xlab = xlabel)
      lines(t, f.med, type = "l")
      abline(v = tau, lty = "dotdash", col = 1, lwd = 2)
    }
  } else{
    post.med <- apply(mcmc.sample1, 2, median)
    tau <- post.med["tau"]
    tau1 <- post.med["tau1"]
    tau2 <- post.med["tau2"]
    ctp <- post.med["ctp"]
    f.post <- f.cable.post(mcmc.sample1, t)
    f.med <- f.post$f.med
    f.lower <- f.post$f.lower
    f.upper <- f.post$f.upper
    x11()
    if (!is.null(start.date)) {
      start.y <- as.numeric(substr(start.date, 1, 4))
      start.m <- as.numeric(substr(start.date, 6, 7))
      end.y <- as.numeric(substr(end.date, 1, 4))
      end.m <- as.numeric(substr(end.date, 6, 7))
      plot(ts(f.med, frequency = 12, start = c(start.y, start.m), end = c(end.y, end.m))
        ylim = c((min(min(f.med), min(y)) - 1), (max(max(f.med), max(y)) + 1)),
        col = 1, lty = 1, ylab = ylabel, xlab = xlabel)
      lines(ts(y, frequency = 12, start = c(start.y, start.m), end = c(end.y, end.m)),
        lty = 0, type = "p", col = "grey50")
      abline(v = (start.y + (start.m - 1) / 12 + tau1 / 12),
        lty = 1, col = 1, lwd = 2)
      abline(v = (start.y + (start.m - 1) / 12 + tau2 / 12),
        lty = 1, col = 1, lwd = 2)
      abline(v = (start.y + (start.m - 1) / 12 + tau / 12),
        lty = "dotdash", col = 1, lwd = 2)
      abline(v = (start.y + (start.m - 1) / 12 + ctp / 12),
        lty = "longdash", col = 1, lwd = 2)
    } else{
      plot(t, y, lty = 3, col = "grey50", ylim = c(min(y), max(y)),
        ylab = ylabel, xlab = xlabel)
      lines(t, f.med, type = "l")
      abline(v = tau1, lty = 1, col = 1, lwd = 2)
      abline(v = tau2, lty = 1, col = 1, lwd = 2)
      abline(v = tau, lty = "dotdash", col = 1, lwd = 2)
```

```
        abline(v = ctp, lty = "longdash", col = 1, lwd = 2)
      }
    }
  }
  if (!is.null(start.date)) {
    return(list(posterior.summary = sum.stat, transition.date = transition.date1,
        gelman.rubin = R, DIC = dic0))
  } else{
    return(list(posterior.summary = sum.stat, gelman.rubin = R, DIC = dic0))
  }
}
```

Below, we present the application of the function "post.summary()" to obtain posterior summaries of the parameters for the CFC-11 data.

```
###################################################
# CFC-11: Generalized Model
###################################################
b0.dat1 <- read.table("c://generalized/b0-1.txt")[, 2]
b1.dat1 <- read.table("c://generalized/b1-1.txt")[, 2]
b2.dat1 <- read.table("c://generalized/b2-1.txt")[, 2]
kappa.dat1 <- read.table("c://generalized/kappa-1.txt")[, 2]
s2.dat1 <- read.table("c://generalized/s2-1.txt")[, 2]
gam.dat1 <- read.table("c://generalized/gam-1.txt")[, 2]
tau.dat1 <- read.table("c://generalized/tau-1.txt")[, 2]
chain1 <- cbind(b0 = b0.dat1, b1 = b1.dat1, b2 = b2.dat1,
  gam = gam.dat1, tau = tau.dat1, kappa = kappa.dat1, s2 = s2.dat1)
b0.dat2 <- read.table("c://generalized/b0-2.txt")[, 2]
b1.dat2 <- read.table("c://generalized/b1-2.txt")[, 2]
b2.dat2 <- read.table("c://generalized/b2-2.txt")[, 2]
kappa.dat2 <- read.table("c://generalized/kappa-2.txt")[, 2]
s2.dat2 <- read.table("c://generalized/s2-2.txt")[, 2]
gam.dat2 <- read.table("c://generalized/gam-2.txt")[, 2]
tau.dat2 <- read.table("c://generalized/tau-2.txt")[, 2]
chain2 <- cbind(b0 = b0.dat2, b1 = b1.dat2, b2 = b2.dat2,
  gam = gam.dat2, tau = tau.dat2, kappa = kappa.dat2, s2 = s2.dat2)
cfc11 <- read.csv("c://CFC-11.csv", header = T)
y <- c(cfc11[, "barrow"])
t <- 0:(length(y) - 1)
burn <- 40000
thin <- 200
start.date <- "1988-01-01"
end.date <- "2010-09-01"
post.summary(y, t, chain1, chain2, burn, thin, model = "G",
  dic = TRUE, start.date = start.date, end.date = end.date,
  fit = TRUE, density = TRUE, trace = TRUE, xlabel = "Time",
  ylabel = "CFC-11 concentration (in ppt)")
```

$posterior.summary

	mean	median	SD	lower.95	upper.95
b0	252.3059764	252.30000	0.492354904	251.30000	253.20000
b1	0.7327021	0.72880	0.043253385	0.66910	0.80860
b2	-0.9051533	-0.90130	0.043144413	-0.98090	-0.84150
gam	68.5643003	68.55000	5.446403719	57.94000	78.67000
tau	40.7067581	40.82000	1.413907161	38.11975	42.90000
kappa	1.2621341	1.25800	0.033564725	1.21100	1.34000
s2	1.4755551	1.46800	0.128179242	1.24500	1.74800
out.slope	-0.1724513	-0.17240	0.001979497	-0.17640	-0.16860
tau1	22.8949471	23.14051	1.910531261	19.00844	25.35122
tau2	109.2710584	109.35500	5.738373598	97.68950	119.74000
ctp	60.9432303	60.93920	1.210756079	58.55082	63.39859

```
$transition.date
      mean.date median.date lower.date upper.date
tau   1991-05-24  1991-05-27 1991-03-06 1991-07-29
tau1  1989-11-28  1989-12-05 1989-08-02 1990-02-11
tau2  1997-02-06  1997-02-09 1996-02-20 1997-12-22
ctp   1993-01-28  1993-01-28 1992-11-16 1993-04-13


$gelman.rubin
Potential scale reduction factors:

      Point est. Upper C.I.
b0             1          1
b1             1          1
b2             1          1
gam            1          1
tau            1          1
kappa          1          1
s2             1          1


Multivariate psrf
1

$DIC
[1] 887.4567




##################################################
# CFC-11: Quadratic Model
##################################################
b0.dat1 <- read.table("c://quadratic/b0-1.txt")[, 2]
b1.dat1 <- read.table("c://quadratic/b1-1.txt")[, 2]
b2.dat1 <- read.table("c://quadratic/b2-1.txt")[, 2]
s2.dat1 <- read.table("c://quadratic/s2-1.txt")[, 2]
gam.dat1 <- read.table("c://quadratic/gam-1.txt")[, 2]
tau.dat1 <- read.table("c://quadratic/tau-1.txt")[, 2]
chain1 <- cbind(b0 = b0.dat1, b1 = b1.dat1, b2 = b2.dat1,
  gam = gam.dat1, tau = tau.dat1, s2 = s2.dat1)
b0.dat2 <- read.table("c://quadratic/b0-2.txt")[, 2]
b1.dat2 <- read.table("c://quadratic/b1-2.txt")[, 2]
b2.dat2 <- read.table("c://quadratic/b2-2.txt")[, 2]
s2.dat2 <- read.table("c://quadratic/s2-2.txt")[, 2]
gam.dat2 <- read.table("c://quadratic/gam-2.txt")[, 2]
tau.dat2 <- read.table("c://quadratic/tau-2.txt")[, 2]
chain2 <- cbind(b0 = b0.dat2, b1 = b1.dat2, b2 = b2.dat2,
  gam = gam.dat2, tau = tau.dat2, s2 = s2.dat2)
cfc11 <- read.csv("c://CFC-11.csv", header = T)
y <- c(cfc11[, "barrow"])
t <- 0:(length(y) - 1)
burn <- 50000
```

```
thin <- 100
start.date <- "1988-01-01"
end.date <- "2010-09-01"
post.summary(y, t, chain1, chain2, burn, thin, model = "Q",
  dic = TRUE, start.date = start.date, end.date = end.date,
  fit = TRUE, density = TRUE, trace = TRUE, xlabel = "Time",
  ylabel = "CFC-11 concentration (in ppt)")

$posterior.summary
                 mean      median            SD   lower.95      upper.95
b0         254.9978186 254.00000  2.686585089  252.70000 263.0000000
b1           0.8716898   0.85120  0.153684854    0.64420   1.2280000
b2          -1.0380954  -1.01800  0.153607030   -1.39400  -0.8107975
gam         43.4385058  42.64000  6.440235096   33.73000  58.1702500
tau         32.5672025  33.37000  6.371176546   17.82000  42.0800000
s2           1.8075426   1.79900  0.157232124    1.52400   2.1400000
out.slope   -0.1664056  -0.16640  0.001669214   -0.16970  -0.1630000
tau1       -10.8713032  -9.28000 12.705561746  -40.31025   7.8900000
tau2        76.0057083  75.99000  1.644912785   72.83000  79.3100000
ctp         62.0621859  62.06188  0.942155157   60.22324  63.9267595


$transition.date
      mean.date median.date lower.date upper.date
tau   1990-09-18  1990-10-13 1989-06-27 1991-07-04
tau1  1987-02-05  1987-03-25 1984-08-23 1988-08-28
tau2  1994-05-01  1994-05-01 1994-01-25 1994-08-10
ctp   1993-03-03  1993-03-03 1993-01-06 1993-04-29


$gelman.rubin
Potential scale reduction factors:

    Point est. Upper C.I.
b0        1.03       1.04
b1        1.00       1.01
b2        1.00       1.01
gam       1.00       1.01
tau       1.00       1.01
s2        1.00       1.00

Multivariate psrf
1

$DIC
[1] 940.0075

##################################################
# CFC-11: Piecewise Linear Model
##################################################
b0.dat1<-read.table("e://piecewise/b0-1.txt")[,2]
b1.dat1<-read.table("e://piecewise/b1-1.txt")[,2]
b2.dat1<-read.table("e://piecewise/b2-1.txt")[,2]
```

```
s2.dat1<-read.table("e://piecewise/s2-1.txt")[,2]
tau.dat1<-read.table("e://piecewise/tau-1.txt")[,2]
chain1<-cbind(b0=b0.dat1,b1=b1.dat1,b2=b2.dat1,tau=tau.dat1,s2=s2.dat1)
b0.dat2<-read.table("e://piecewise/b0-2.txt")[,2]
b1.dat2<-read.table("e://piecewise/b1-2.txt")[,2]
b2.dat2<-read.table("e://piecewise/b2-2.txt")[,2]
s2.dat2<-read.table("e://piecewise/s2-2.txt")[,2]
tau.dat2<-read.table("e://piecewise/tau-2.txt")[,2]
chain2<-cbind(b0=b0.dat2,b1=b1.dat2,b2=b2.dat2,tau=tau.dat2,s2=s2.dat2)
cfc11 <- read.csv("c://CFC-11.csv", header = T)
y<-c(cfc11[,"barrow"])
t<-0:(length(y)-1)
burn<-50000
thin<-100
start.date<-"1988-01-01"
end.date<-"2010-09-01"
post.summary(y, t, chain1, chain2, burn, thin, model = "B",
  dic = TRUE, start.date = start.date, end.date = end.date,
  fit = TRUE, density = TRUE, trace = TRUE, xlabel = "Time",
  ylabel = "CFC-11 concentration (in ppt)")
```

$posterior.summary
```
                 mean    median           SD     lower.95 upper.95
b0        255.5524034 255.6000 0.531403348 254.5000000 256.6000
b1          0.4615851   0.4618 0.022130264   0.4177000   0.5042
b2         -0.6225493  -0.6226 0.021692635  -0.6644025  -0.5797
tau        50.1513824  50.0900 1.313499368  47.7500000  52.8800
s2          2.8638807   2.8490 0.248087657   2.4180000   3.3900
out.slope  -0.1609642  -0.1610 0.001927983  -0.1648000  -0.1572
```

$transition.date
```
     mean.date median.date lower.date upper.date
tau 1992-03-06  1992-03-04 1991-12-24 1992-05-28
```

$gelman.rubin
Potential scale reduction factors:

```
     Point est. Upper C.I.
b0            1          1
b1            1          1
b2            1          1
tau           1          1
s2            1          1
```

Multivariate psrf

1

$DIC
[1] 1263.5

DIAGNOSTIC PLOTS

B.1   Barrow CFC-11 Data Analysis: Trace and Density Plots
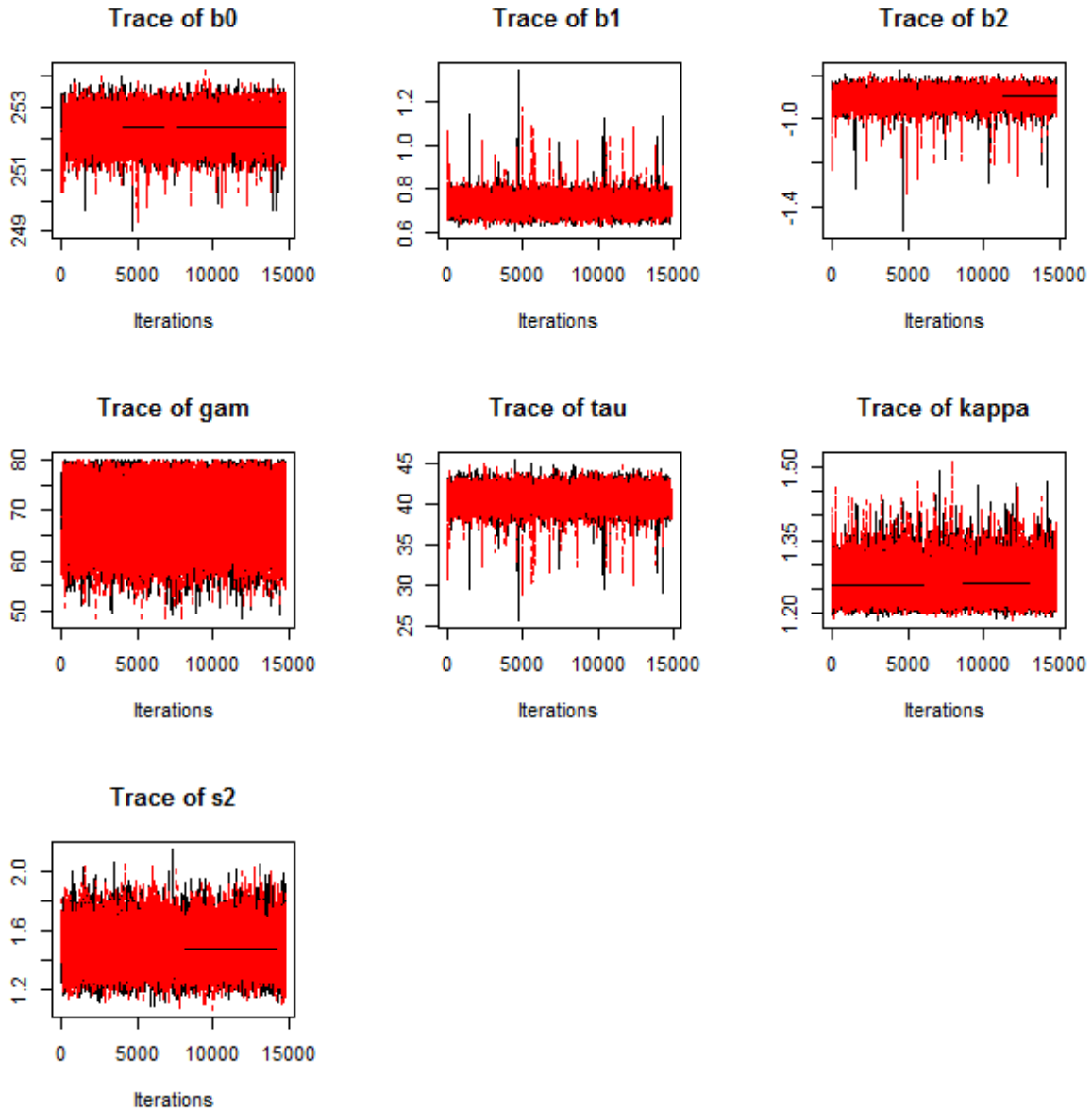
Figure B.1: CFC-11 data analysis using the generalized bent-cable model – trace plots.
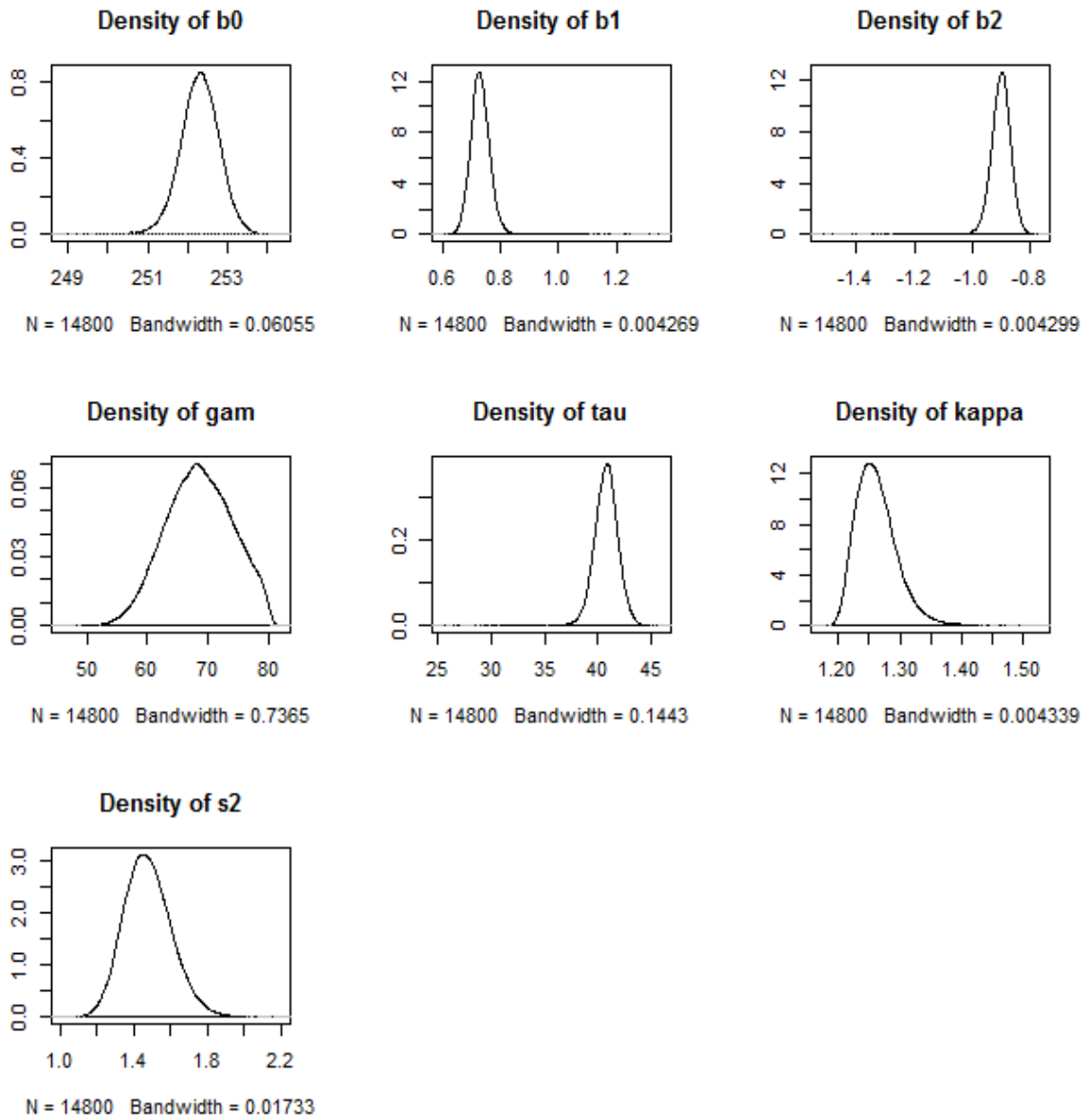
Figure B.2: CFC-11 data analysis using the generalized bent-cable model – density plots.
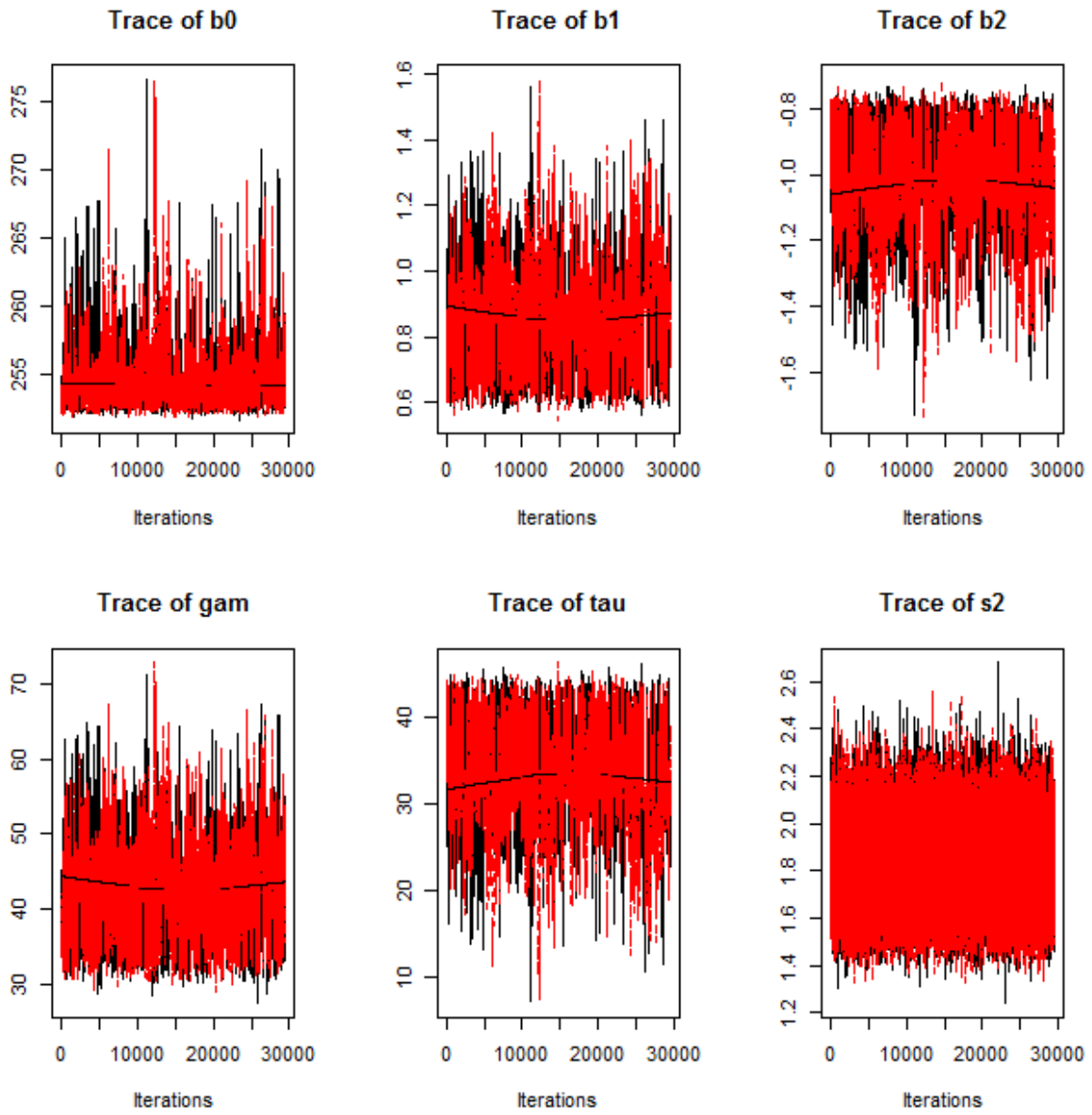
Figure B.3: CFC-11 data analysis using the quadratic bent-cable model – trace plots.
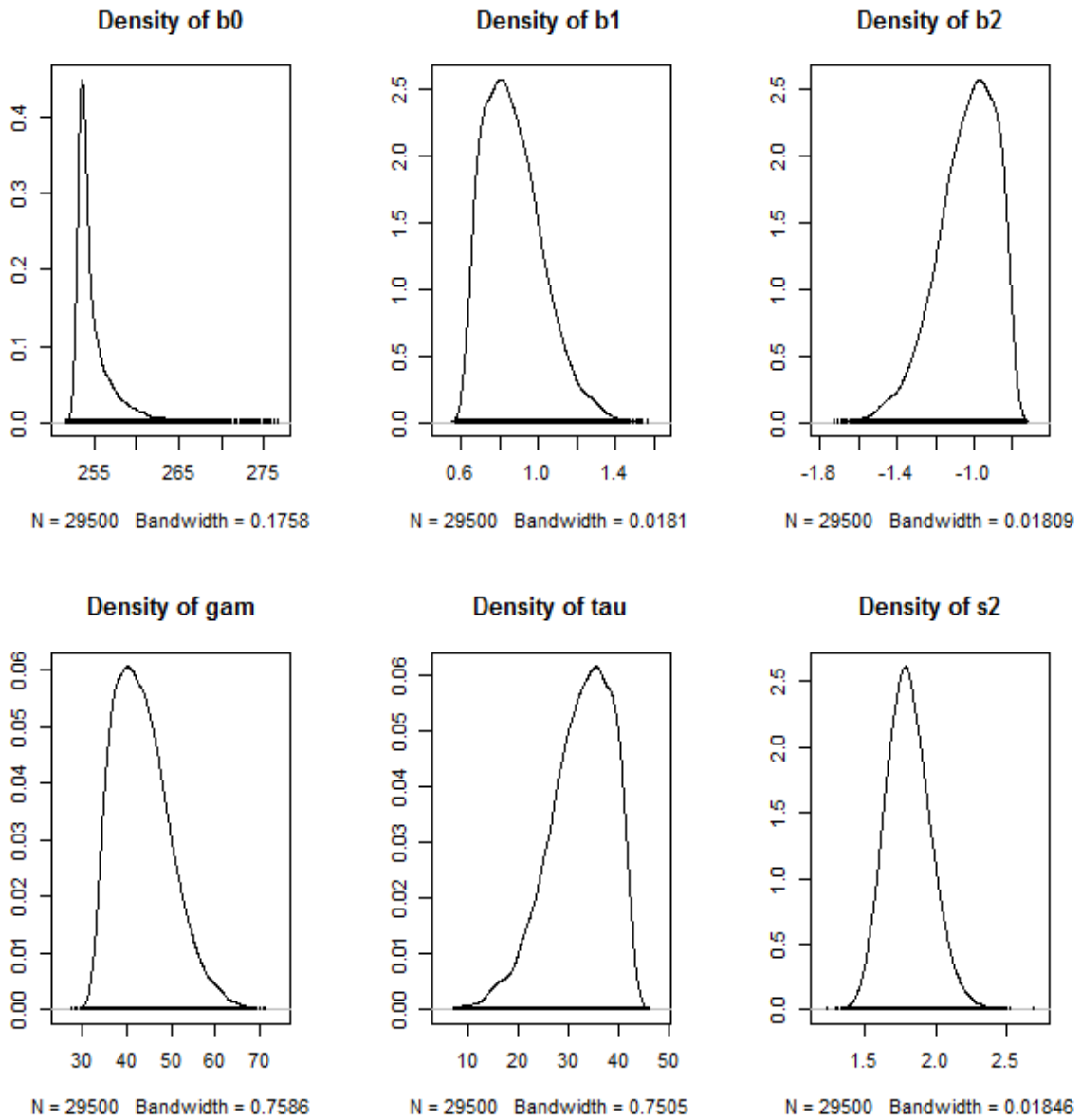
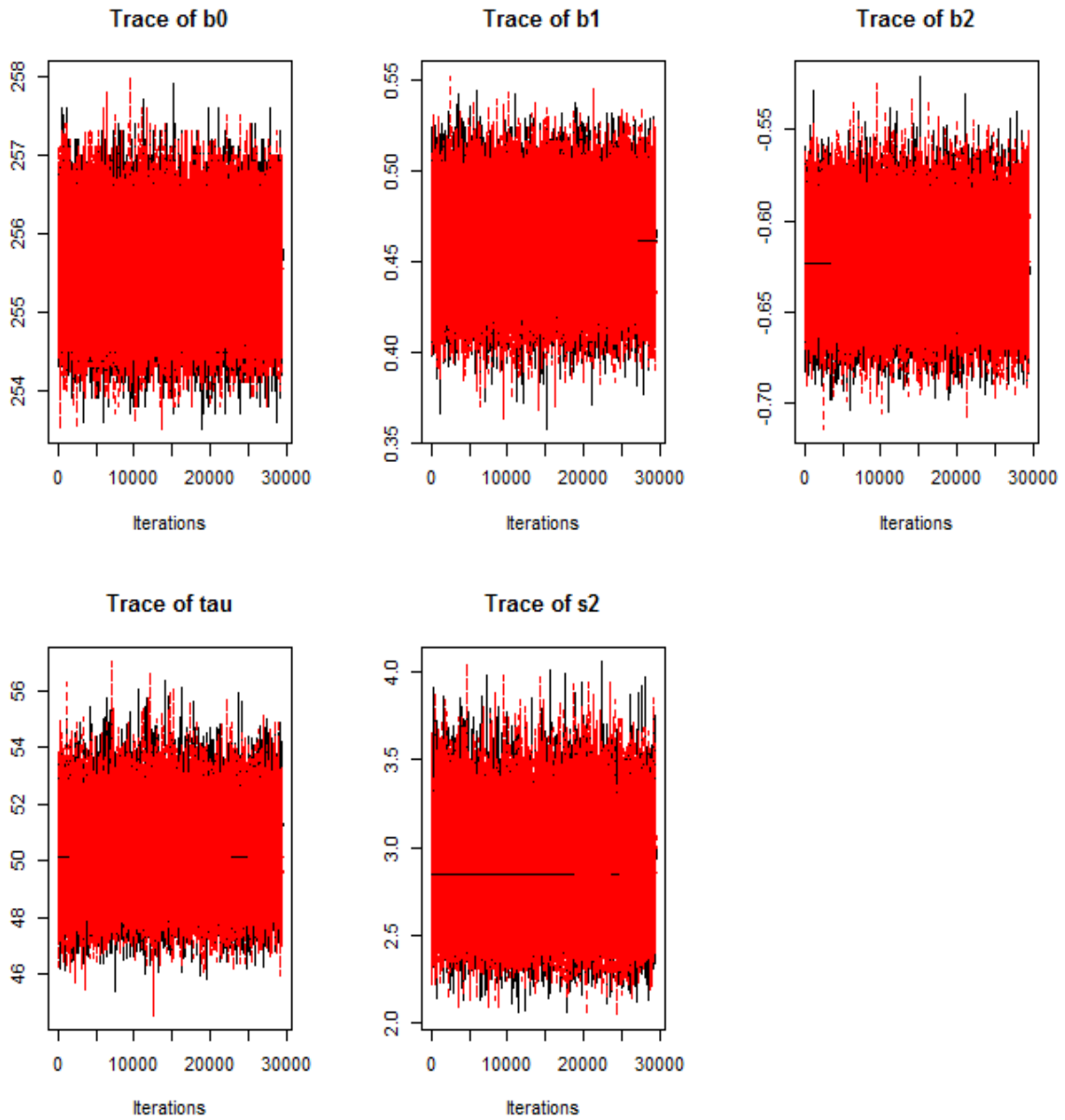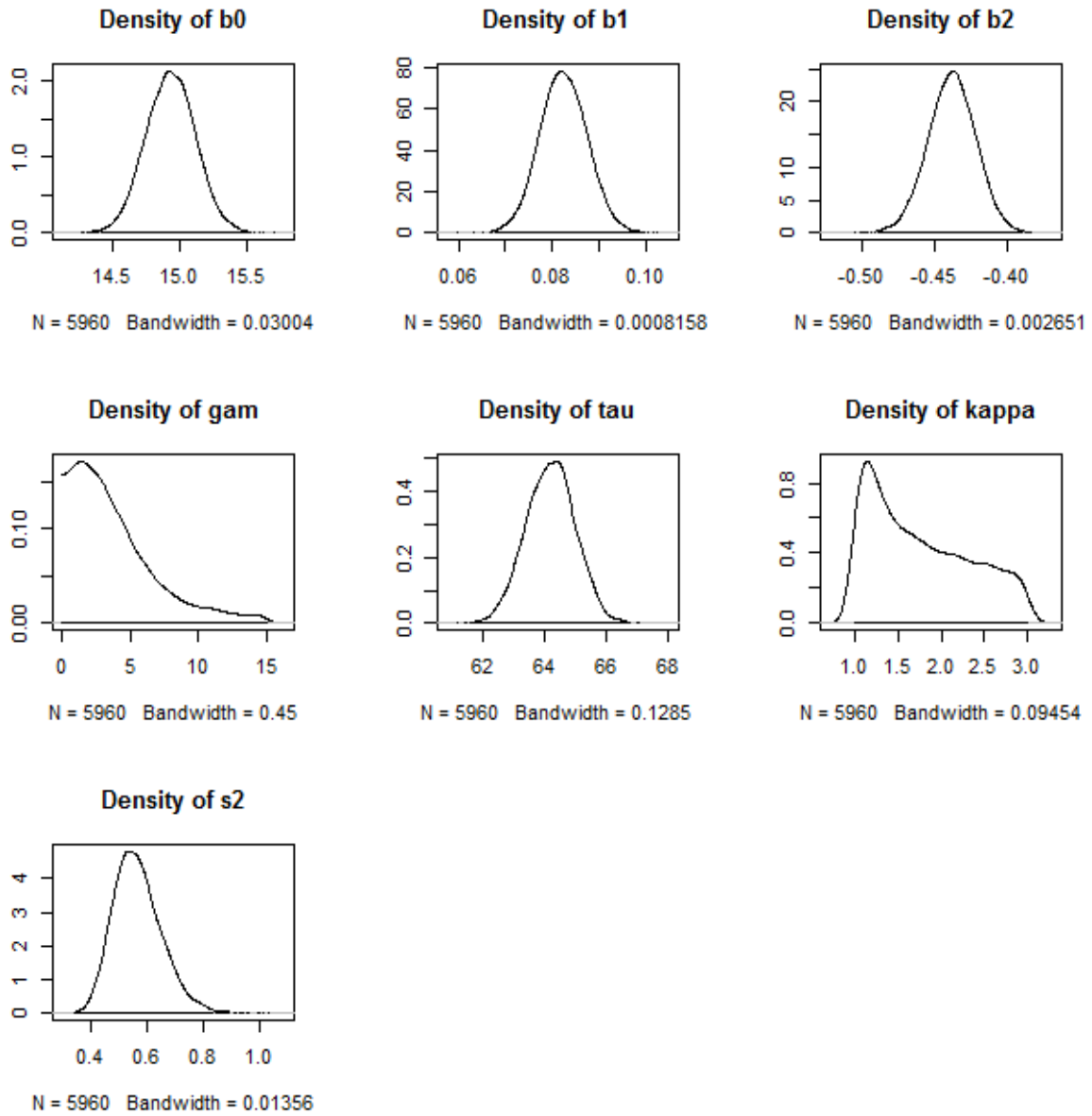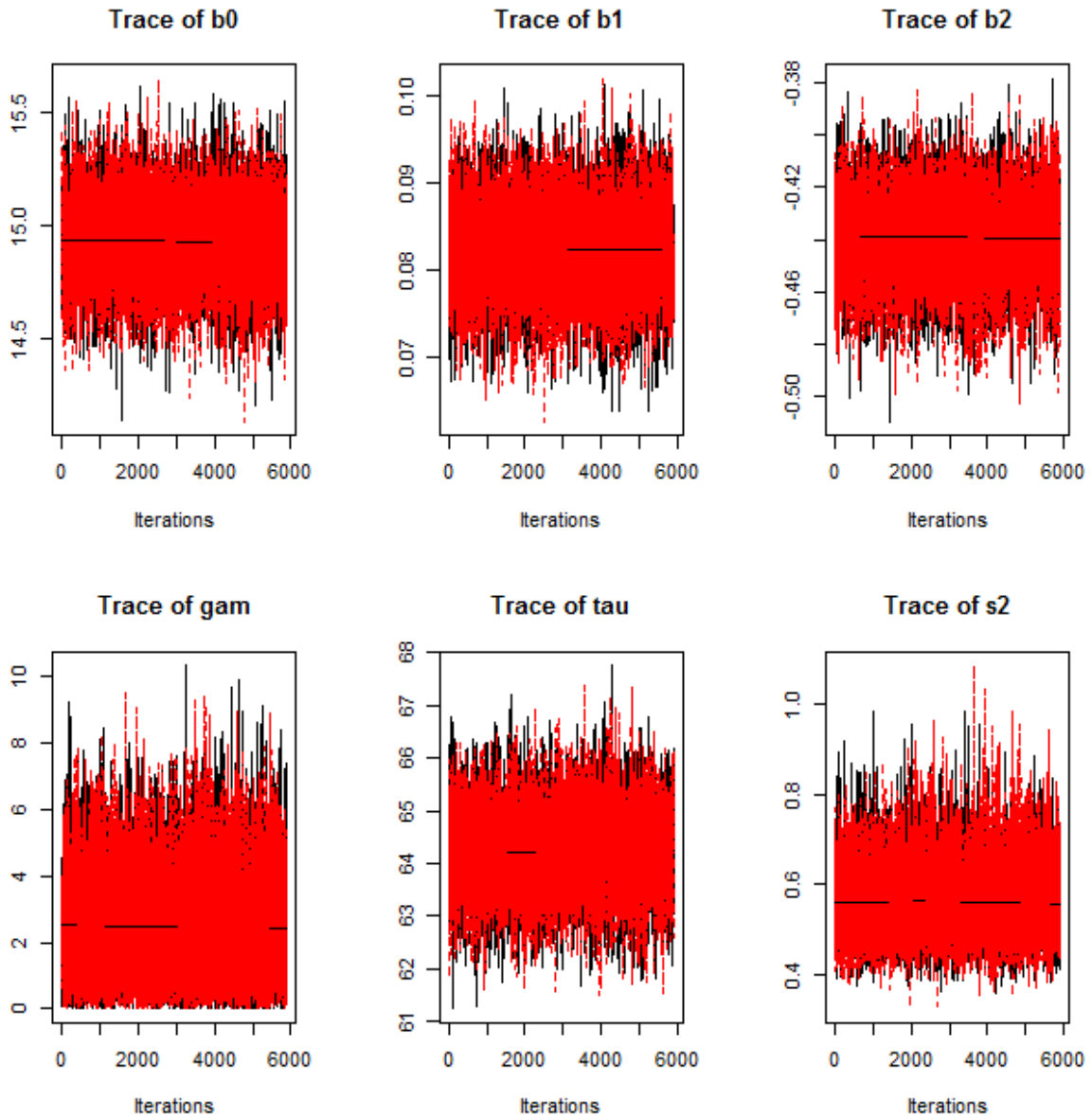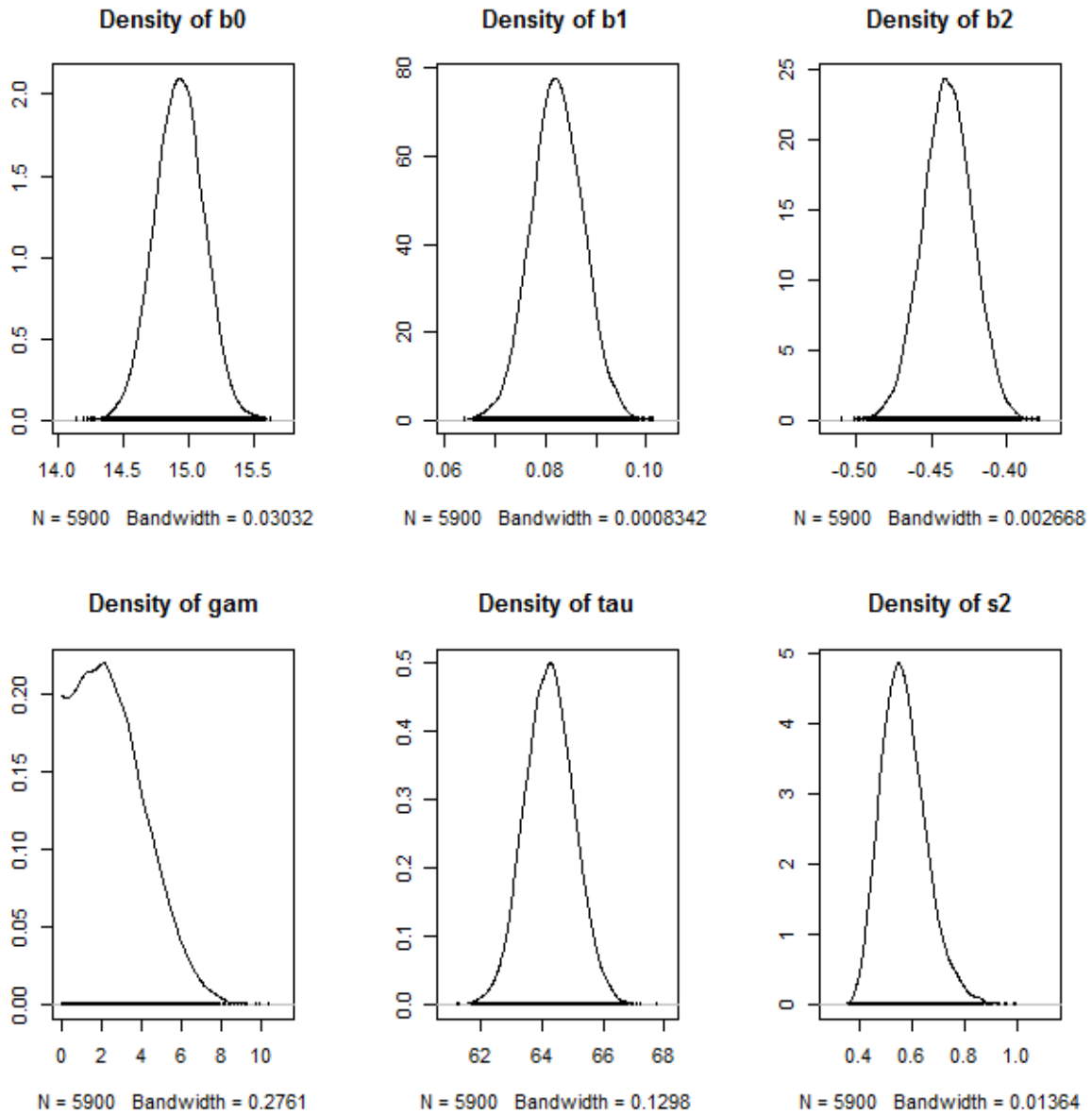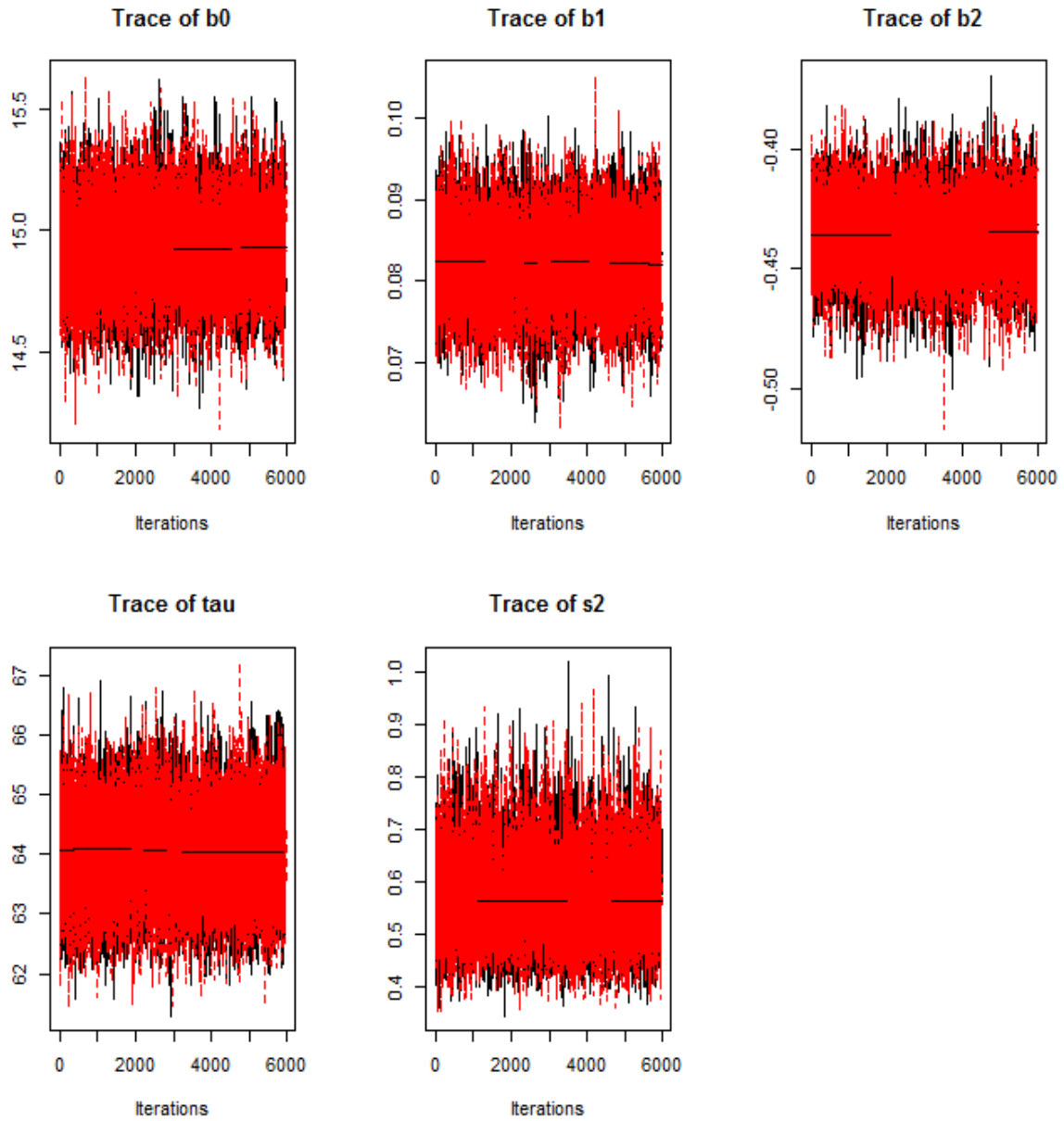**Figure B.4:** CFC-11 data analysis using the quadratic bent-cable model – density plots.

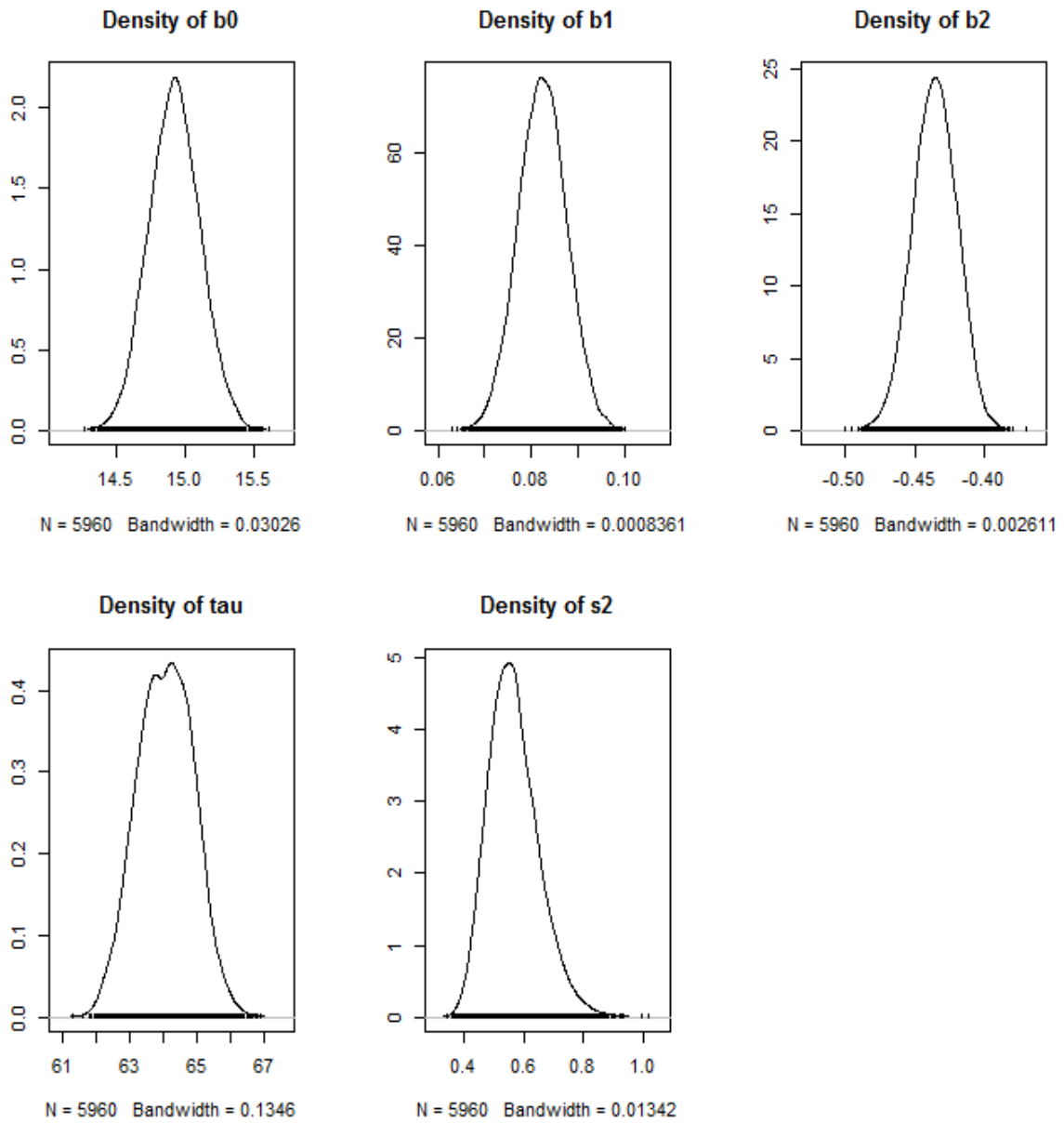**Figure B.5:** CFC-11 data analysis using the piecewise linear model – trace plots.
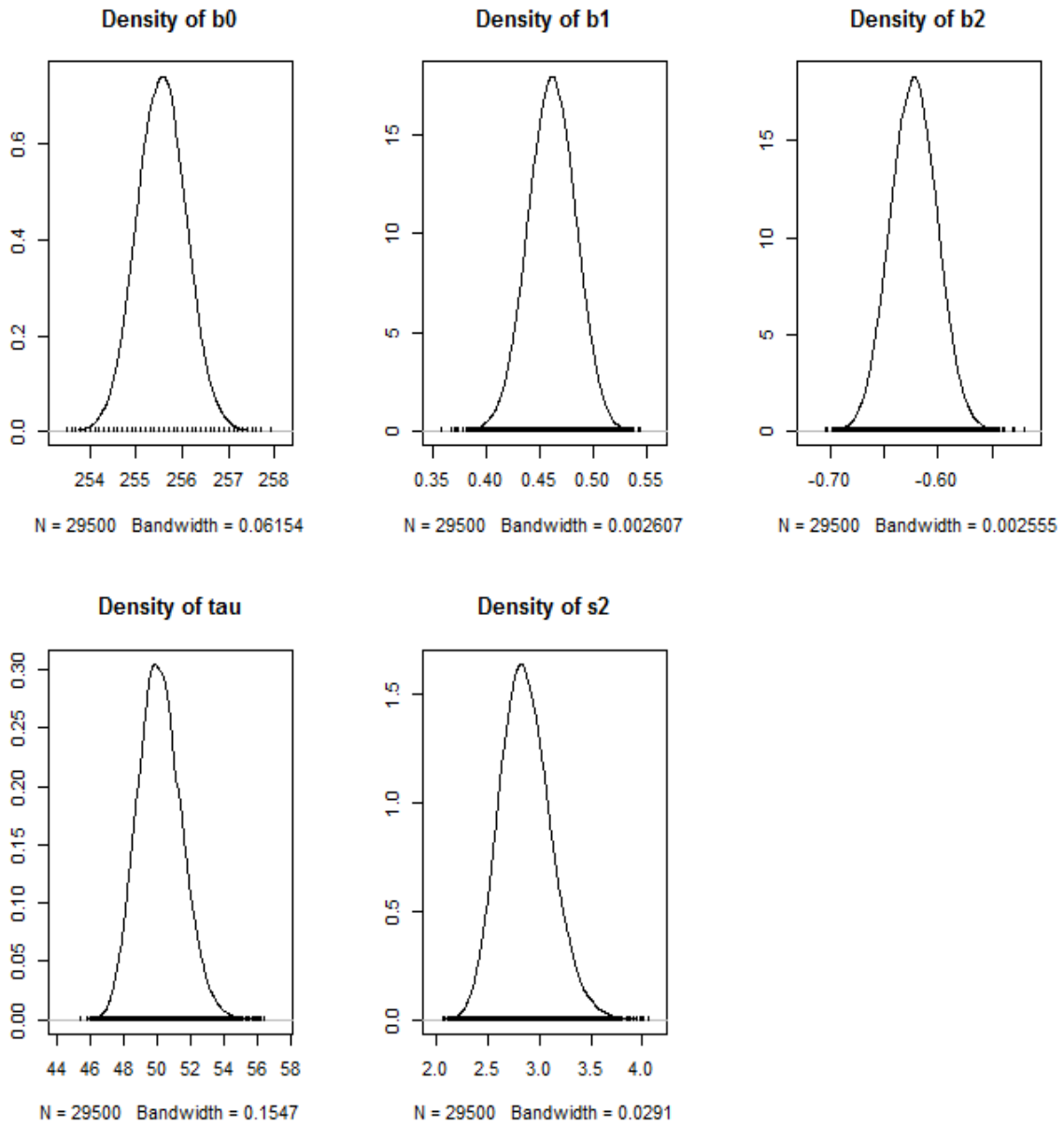
Figure B.6: CFC-11 data analysis using the piecewise linear model – density plots.

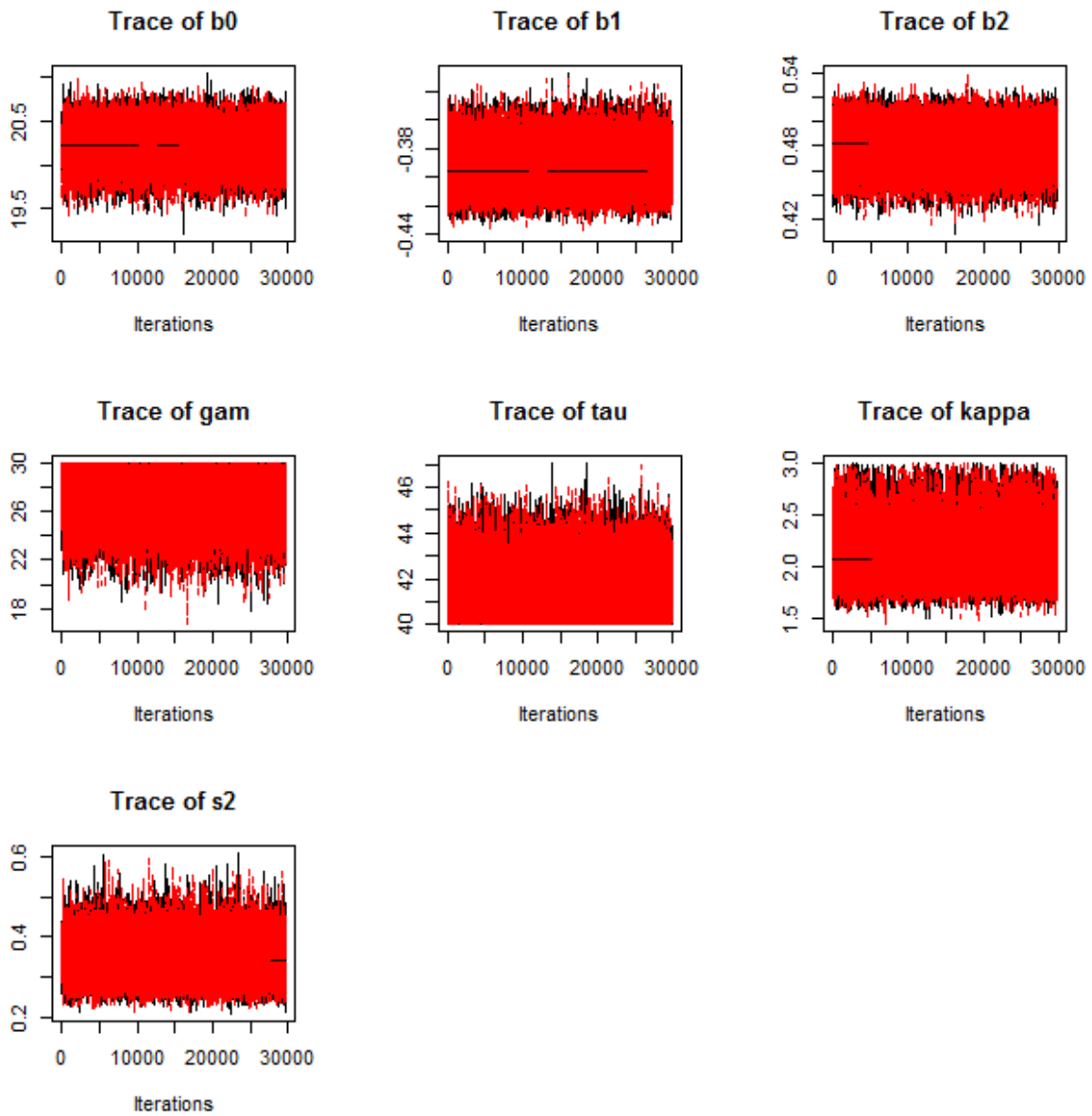# B.2   Housing Data Analysis: Trace and Density Plots
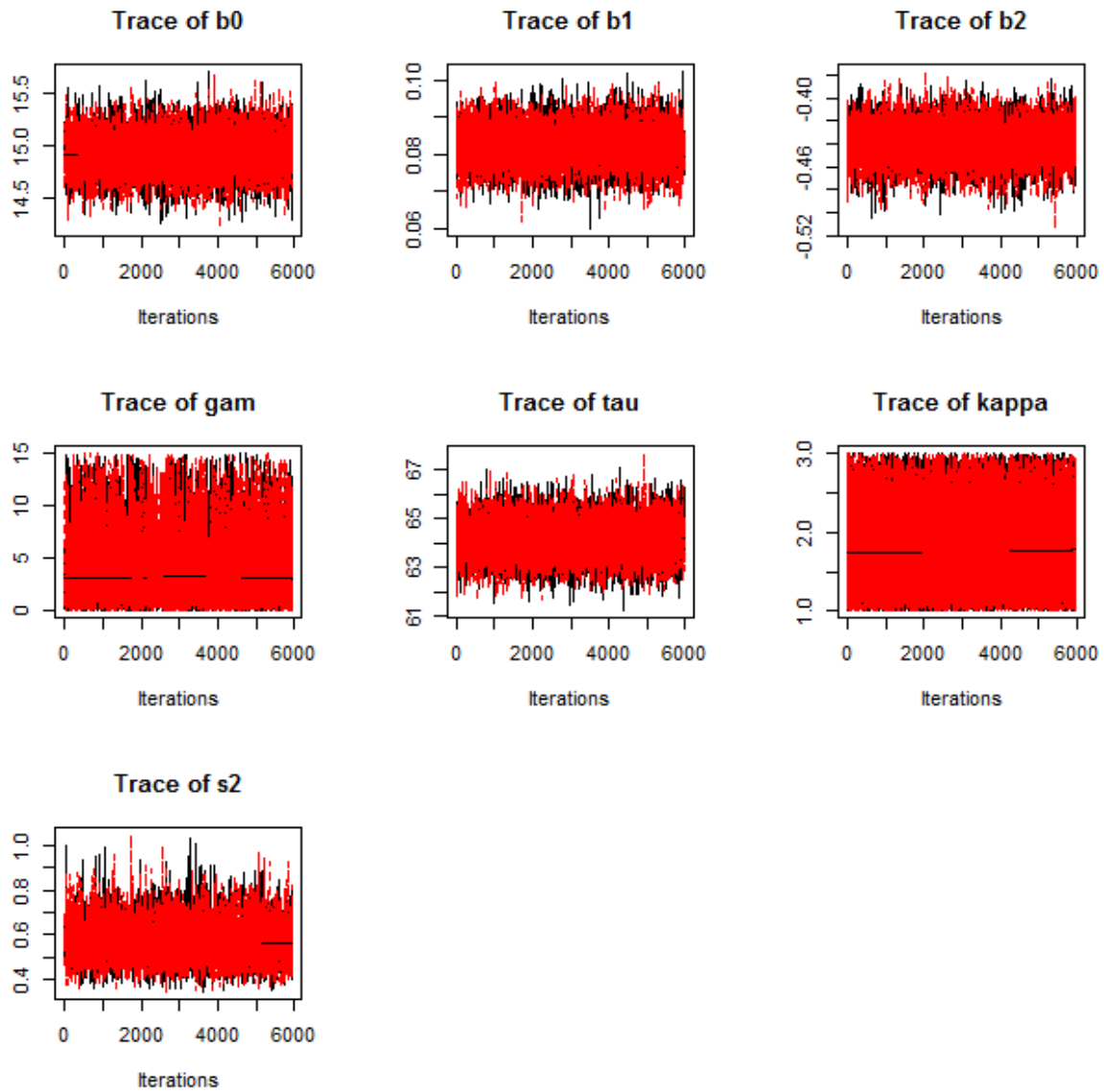
Figure B.7: Housing data (series 1) analysis using the generalized bent-cable model – trace plots.
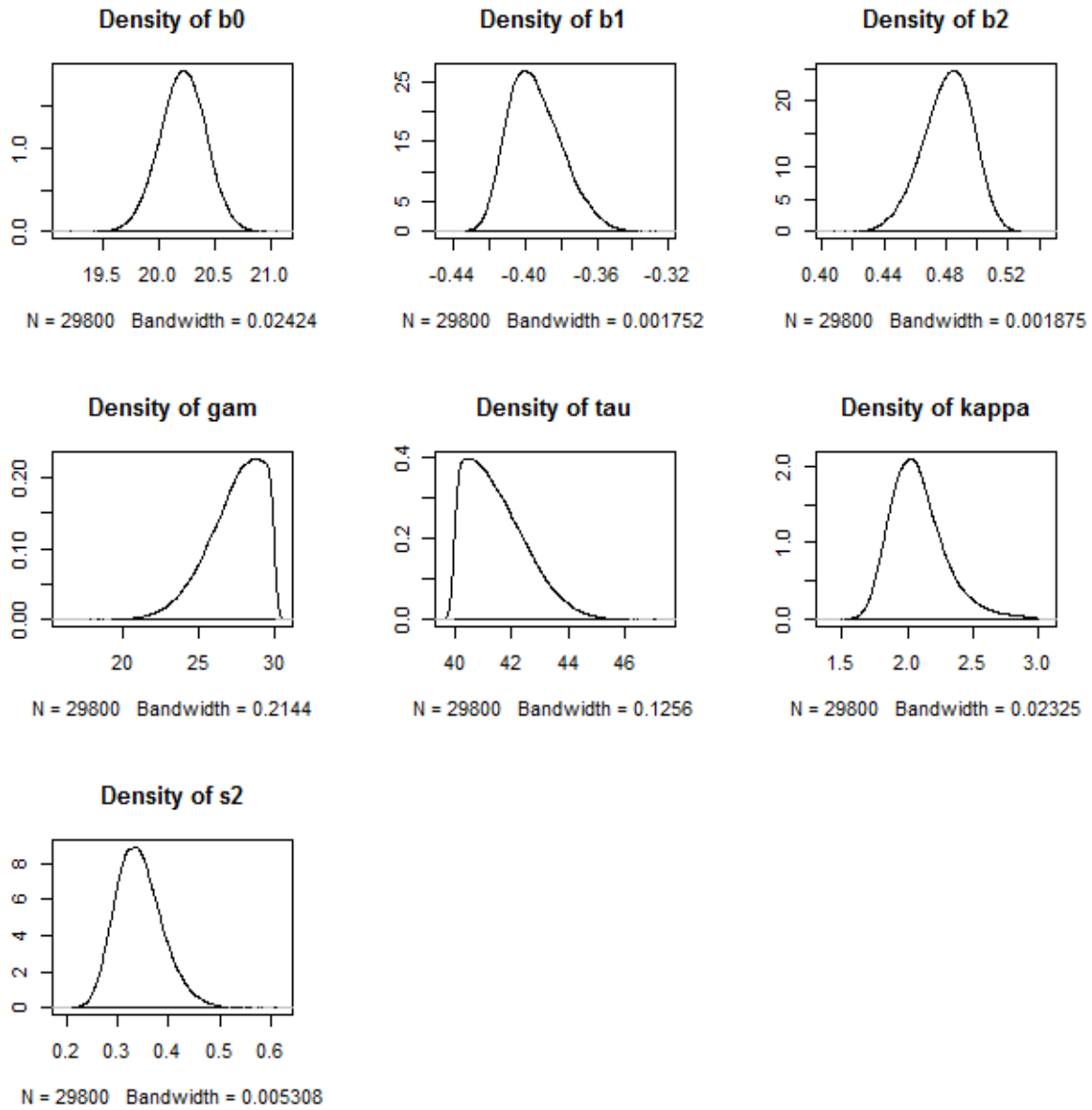
Figure B.8: Housing data (series 1) analysis using the generalized bent-cable model – density plots.

Figure B.9: Housing data (series 1) analysis using the quadratic bent-cable model – trace plots.

Figure B.10: Housing data (series 1) analysis using the quadratic bent-cable model – density plots.

Figure B.11: Housing data (series 1) analysis using the piecewise linear model – trace plots.

Figure B.12: Housing data (series 1) analysis using the piecewise linear model – density plots.

Figure B.13: Housing data (series 2) analysis using the generalized bent-cable model – trace plots.

Figure B.14: Housing data (series 2) analysis using the generalized bent-cable model – density plots.
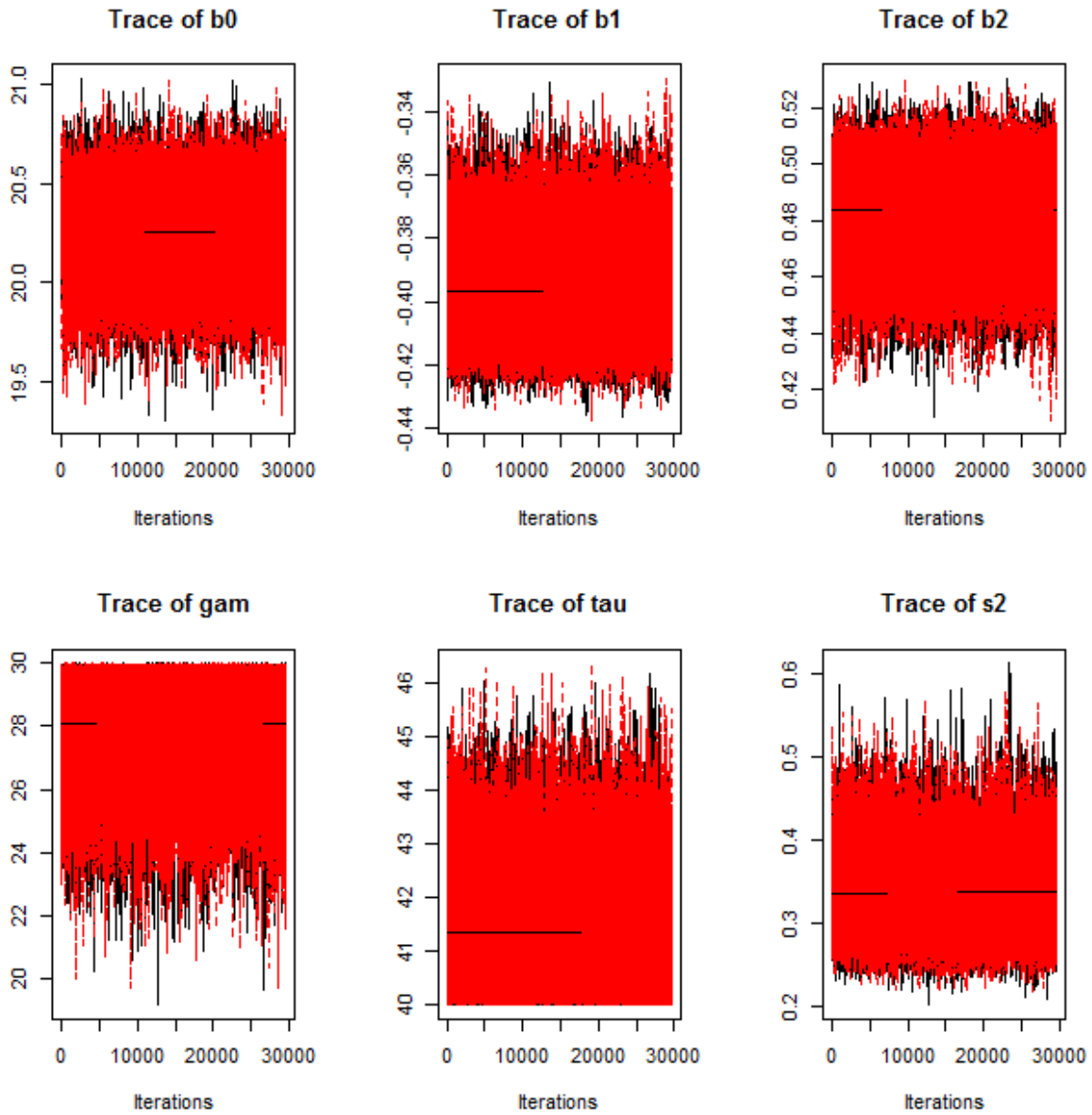
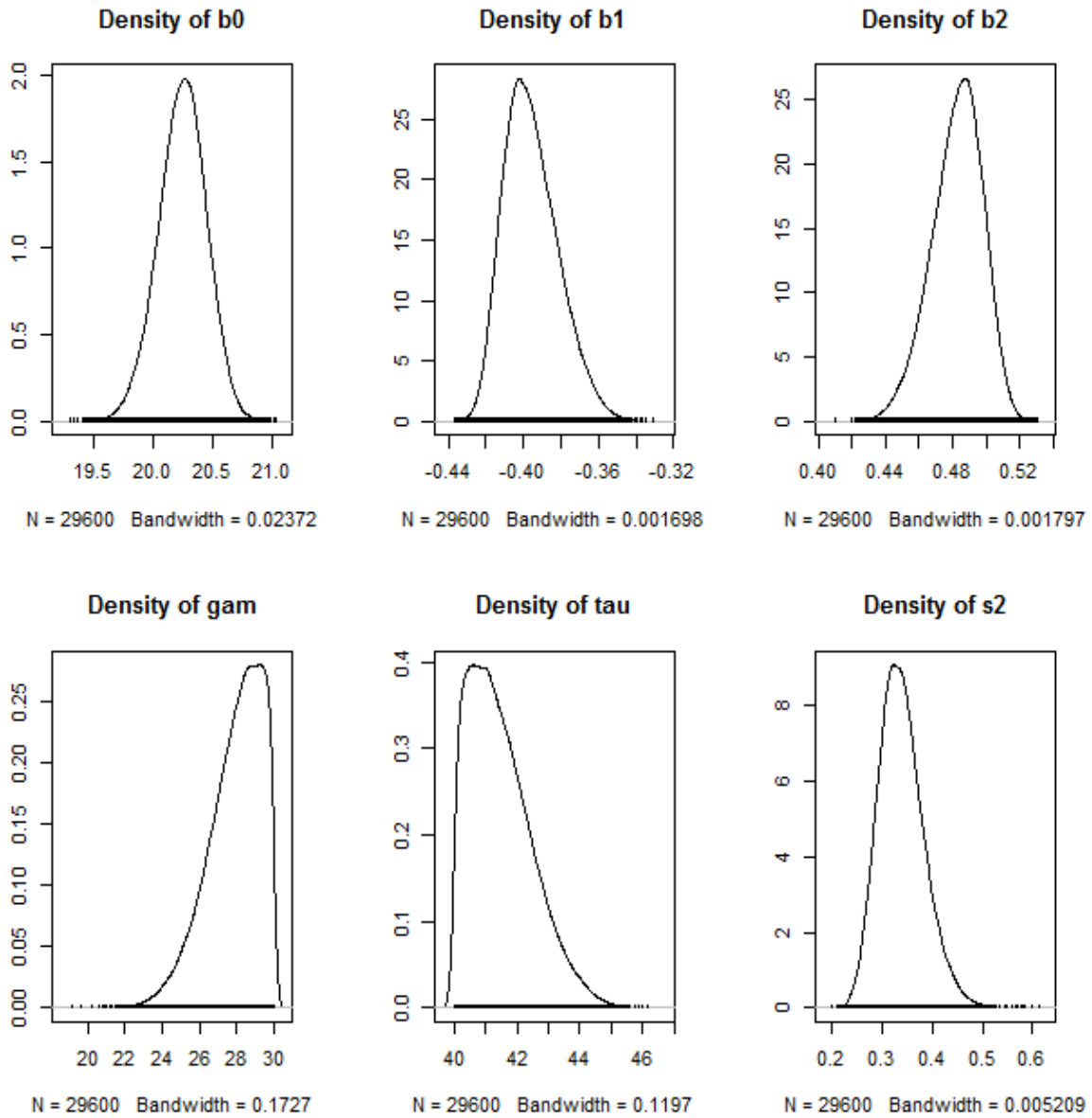**Figure B.15:** Housing data (series 2) analysis using the quadratic bent-cable model – trace plots.

61

Figure B.16: Housing data (series 2) analysis using the quadratic bent-cable model – density plots.
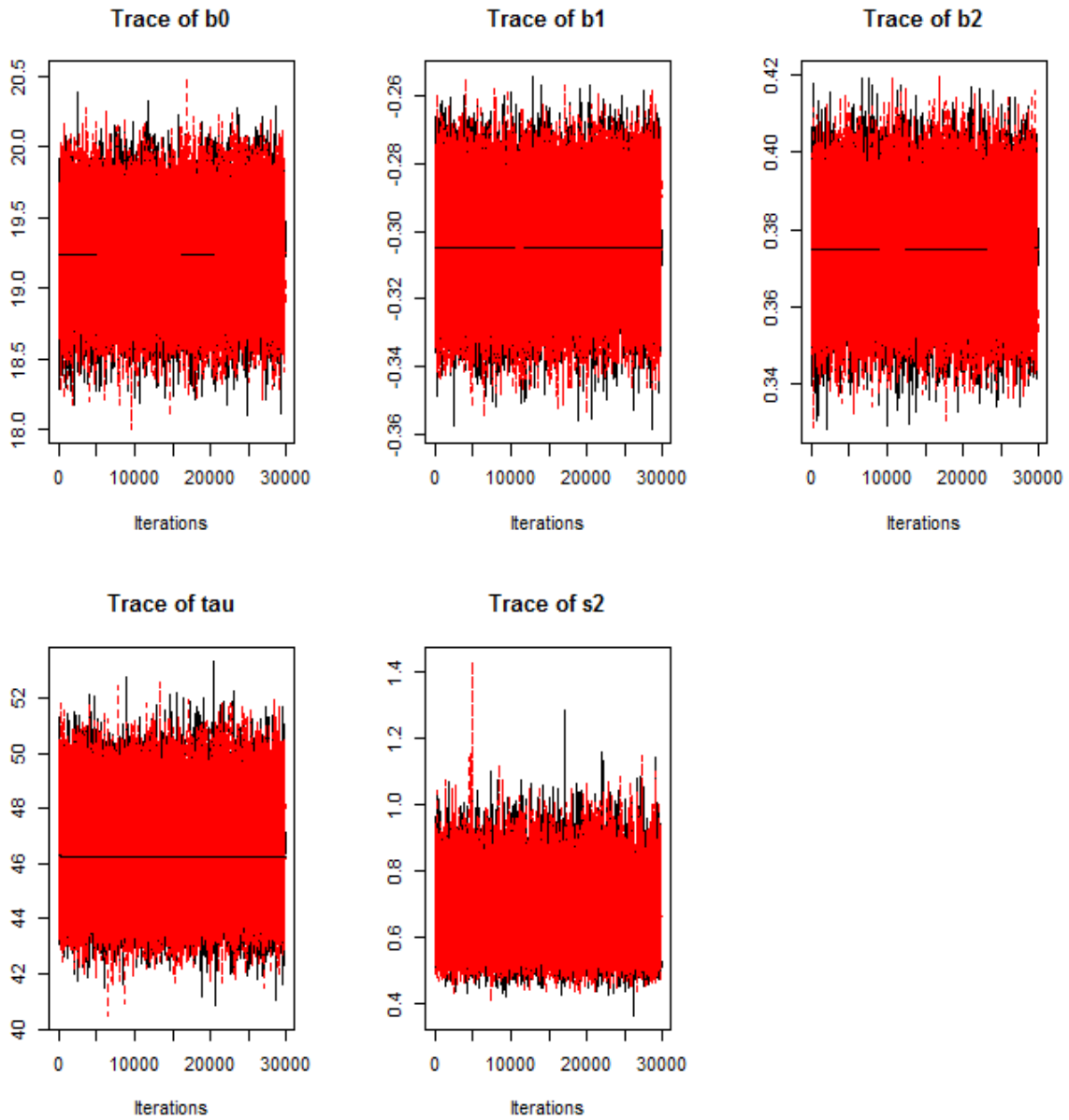
Figure B.17: Housing data (series 2) analysis using the piecewise linear model – trace plots.
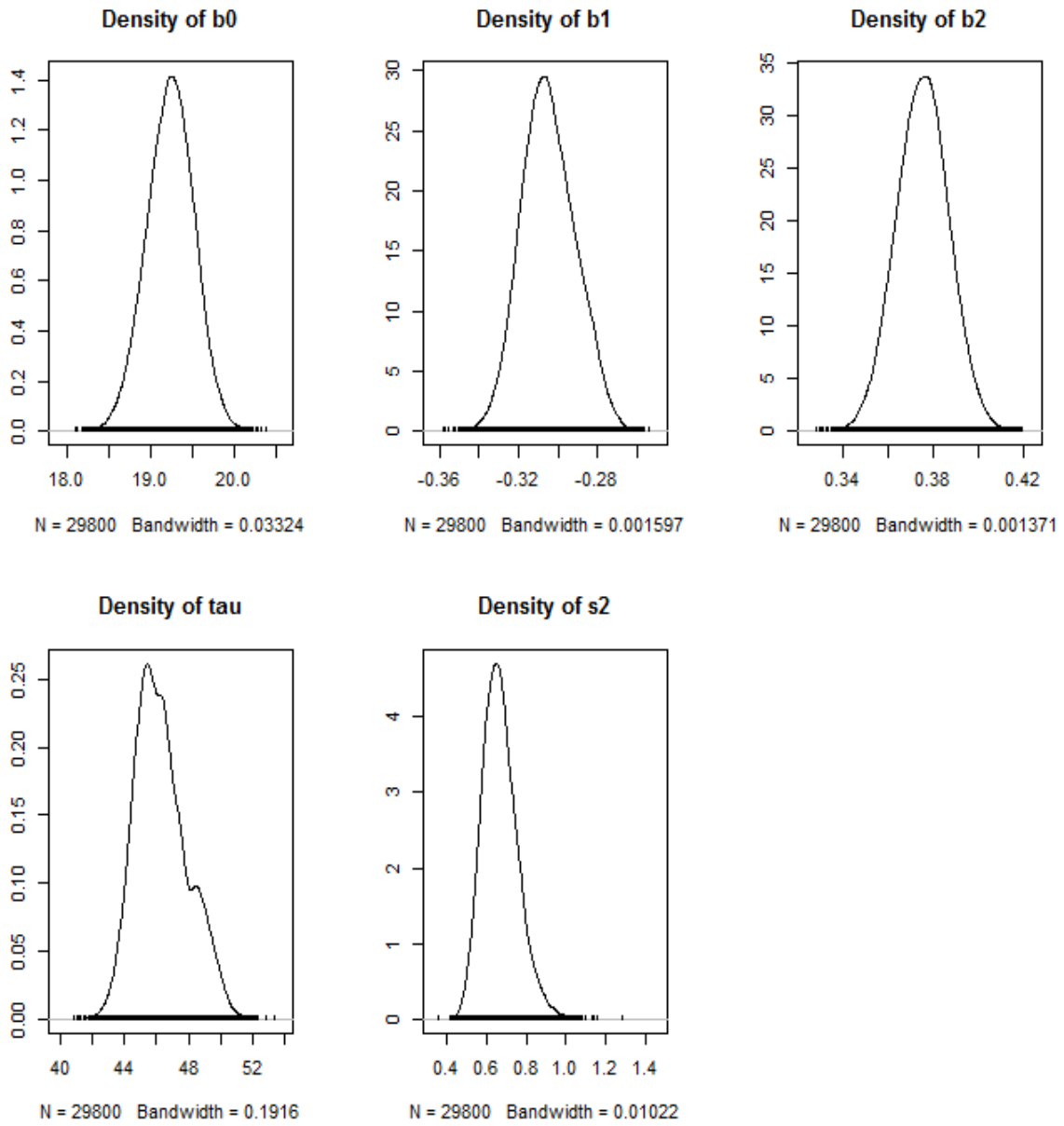
Figure B.18: Housing data (series 2) analysis using the piecewise linear model – density plots.

REFERENCES

Bacon, D. W. and Watts, D. G. (1971). Estimating the transition between two tntersecting straight lines. *Biometrika*, 58(3):525–534. 6, 28

Bellera, C. A., Hanley, J. A., and Joseph, L. (2008). Hierarchical changepoint models for biochemical markers illustrated by tracking postradiotherapy prostate-specific antigen series in men with prostate cancer. *Annals of Epidemiology*, 18:270–282. 4, 5, 12

Bian, S. and Menz, J. (1998). Using piecewise linear interpolations in spherical harmonic analysis. *J. Geodesy*, 72(7):473–481. 5

Chiu, G. (2015). *bentcableAR: Bent-cable regression independent data or autoregressive time series*. R package version 0.3.0. 16

Chiu, G. and Lockhart, R. (2010). Bent-cable regression with autoregressive noise. *The Canadian Journal of Statistics*, 38(3):386–407. 2, 5, 9, 15

Chiu, G., Lockhart, R., and Routledge, R. (2005). Asymptotic theory for bent-cable regression. *Journal of Statistical planning and inference*, 127:143–156. 28

Chiu, G., Lockhart, R., and Routledge, R. (2006). Bent-cable regression theory and applications. *Journal of the American Statistical Association*, 101:542–553. 1, 2, 5, 6, 7, 9, 16, 28

Chiu, G. S. (2002). Bent-cable regression for assessing abruptness of change. *Ph.D diss., Simon Fraser University*. 15, 16

Dai, T. S., Wang, J. Y., and Wei, H. S. (2007). An ingenious, piecewise linear interpolation algorithm for pricing arithmetic average options. In *Lect Notes Comput SC*, pages 131–143. 5

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London. 7

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statist. Sci.*, 7:457–472. 17

Gelman, A. and Shirley, K. (2011). Inference from simulations and monitoring convergence. In Brooks, S., Gelman, A., Jones, G. L., and Meng, X., editors, *Handbook of Markov Chain Monte Carlo*, pages 163–174. Chapman and Hall/CRC, New York. 17

Ghosh, P. and Vaida, F. (2007). Random changepoint modelling of hiv immunologic responses. *Statistics in Medicine*, 26:2074–2087. 4

Hall, C. B., Ying, J., Kuo, L., and Lipton, R. B. (2003). Bayesian and profile likelihood change point methods for modeling cognitive function over time. *Computational Statistics and Data Analysis*, 42(3):91–109. 12

Jimenez-Fernandez, V. M., Cerecedo-Nunez, H. H., Vazquez-Leal, H., Beltran-parrazal, L., and Filobello-Nino, U. A. (2014). A parametric piecewise linear approach to laser projection. *Comput Appl Math*, 33(3):841–858. 5

Kelly, G. E., Lindsey, J. K., and Thin, A. G. (2004). Models for estimating the changepoint in gas exchange data. *Physiological Measurement*, 25:1425–1436. 5

Khan, S. A., Chiu, G., and A.Dubin, J. (2009). Atmospheric concentration of chlorofluorocarbons: Addressing the global concern with the longitudinal bent-cable model. *CHANGE*, 22(3):8–17. 7, 9, 10

Khan, S. A., Chiu, G., and Dubin, J. A. (2013). Therapeutic hypothermia: Quantification of the transition of core body temperature using the flexible bent-cable model for mixture longitudinal data. *Aust. N. Z. J. Stat.*, 55:369–385. 7, 9

Khan, S. A. and Kar, S. C. (2017). Generalized bent-cable methodology for change-point data: a bayesian approach. *Journal of Applied Statistics*, pages 1–14. DOI: 10.1080/02664763.2017.1391754. 8, 30

Khan, S. A., Rana, M., Li, L., and Dubin, J. A. (2012). A statistical investigation to monitor and understand atmospheric cfc decline with the spatial-longitudinal bent-cable model. *Int. J. Stat. Probab.*, 1:56–68. 7, 9, 30

Kiuchi, A. S., Hartigan, J. A., Holford, T. R., Rubinstein, P., and Stevens, C. E. (1995). Change points in the series of t4 counts prior to aids. *Biometrics*, 51:236–248. 5

Lunn, D. J., Thomas, A., Best, N., and Spiegelhalter, D. (2000). Winbugs - a bayesian modeling framework: concepts, structure, and extensibility. *Stat. Comput.*, 10:325–337. 16

Magnani, A. and Boyd, S. P. (2009). Convex piecewise-linear fitting. *Optim Eng*, 10(1):1–17. 5

Misener, R. and Floudas, C. A. (2010). Piecewise-linear approximations of multidimensional functions. *J Optim Theory Appl*, 145(1):120–147. 5

Muggeo, V. M. R. (2003). Estimating regression models with unknown break-points. *Statist. Med.*, 22:3055–3071. 1, 3, 5, 12, 27

NOAA/ESRL Global Monitoring Division (2016). *Chloroflurocarbon-11 data from the NOAA/ESRL halocarbons in situ program*. 1, 10

Page, E. (1954). Continuous inspection schemes. *Biometrica*, 41:100–114. 4

Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). Coda: Convergence diagnosis and output analysis for mcmc. *R News*, 6:7–11. 16, 36

Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the American Statistical Association*, 53:873–880. 4

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. 16, 36

Raymond, M., Gergely, D., and Clement, L. (2016). *Quandl: API wrapper for Quandl.com*. R package version 2.8.0. 19

Ruch, D. and Claridge, D. E. (1992). A four-parameter changepoint model for predicting consumption in commercial buildings. *Journal of Solar Energy Engineering*, 114:77–83. 4

Ruppert, D., Wand, M. P., and Carroll, R. J. (2003). *Semiparametric Regression*. Cambridge University Press, Cambridge. 7

Slate, E. H. and Turnbull, B. W. (2000). Statistical models for longitudinal biomarkers of disease onset. *Statistics in Medicine*, 19:617–637. 4

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Linde, A. V. D. (2002). Bayesian measures of model complexity and fit (with discussion). *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 64:583–616. 17

Sturtz, S., Ligges, U., and Gelman, A. (2005). R2winbugs: A package for running winbugs from r. *J. Stat. Softw.*, 12:1–16. 17

Tishler, A. and Zang, I. (1981). A new maximum likelihood algorithm for piecewise regression. *Journal of the American Statistical Association*, 76:980–987. 5

Toms, J. D. and Lesperance, M. L. (2003). Piecewise regression: a tool for identifying ecological thresholds. *Journal of Ecology*, 84:2034–2041. 4

van den Hout, A., Muniz-Terrera, G., and Matthews, F. E. (2010). Smooth random change point models. *Statistics in Medicine*, 30(6):599–610. 6, 7

Wu, W. B., Woodroofe, M., and Mentz, G. (2001). Isotonic regression: another look at the change point problem. *Biometrika*, 88(3):793–804. 4