47/2010

Raport Badawczy

Research Report

RB/11/2010

Modeling uncertainty structure of greenhouse gases inventories

Z. Nahorski, J. Jarnicka

Instytut Badań Systemowych
Polska Akademia Nauk

Systems Research Institute
Polish Academy of Sciences

POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 3810100

fax: (+48) (22) 3810105

Kierownik Pracowni zgłaszający pracę:
Prof. dr hab. inż. Zbigniew Nahorski

Warszawa 2010

Modeling Uncertainty Structure of Greenhouse Gases Inventories

Zbigniew Nahorski, Jolanta Jarnicka
Systems Research Institute, Polish Academy of Sciences
Newelska 6, 01-477 Warsaw, Poland
email: (Zbigniew.Nahorski, Jolanta.Jarnicka)@ibspan.waw.pl

Contents

Abstract 5

1 Introduction 7

2 Presentation of an idea 9
- 2.1 Basic questions 9
- 2.2 Procedure for the set of revised data for a given country 9
- 2.3 Finding the correlation structure in the data for many countries 12

3 First attempt: revisions treated independently 13
- 3.1 Notation 13
- 3.2 ML estimators of the model parameters 13
- 3.3 Estimator of the model parameters covariance matrix 15

4 Preliminary estimation for several countries 17
- 4.1 CDIAC data analysis for Austria 17
  - 4.1.1 Data 17
  - 4.1.2 Model 1, $f(Y - y_j) = Y - y_j$ 18
  - 4.1.3 Model 2, $f(Y - y_j) = \sqrt{Y - y_j}$ 20
- 4.2 Analysis of the data from the Austrian National Inventory Reports 22
  - 4.2.1 Data 22
  - 4.2.2 Model 1 23
  - 4.2.3 Model 2 24
- 4.3 Analysis of the data from several other countries 26
  - 4.3.1 Belgium 26
  - 4.3.2 Netherlands 28
  - 4.3.3 Denmark 30
  - 4.3.4 Finland 32
  - 4.3.5 UK 34

5 First attempt: All revisions considered together 37
- 5.1 Maximum likelihood estimators 37
- 5.2 Data analysis 38

6 Conclusions 43

Bibliography 45

Abstract

The report addresses the problem of reducing uncertainty in national greenhouse gases inventories by taking into account all possible information from inventories, including all revised values published usually in every few years, provided by national centres for reporting greenhouse gases inventories, as well as independent data gathered by the Carbon Dioxide Information Analysis Center, Oak Ridge, USA. This is done by proposing a parametric model that describes the structure of uncertainties in the inventories. A procedure for estimating parameters is described and preliminary results of fitting the models to data for several countries are given.

Keywords: greenhouse gases inventories, uncertainty, modelling

Chapter 1

Introduction

The problem addressed in this report can informally be described as follows. Estimate uncertainties of national emission inventories taking into account both the data revised in consecutive years[1] and correlations between inventories for different countries. In future, this may constitute a part of further studies where data from other sources will be available in order to constrain the uncertainty of the inventories presented to UN FCC. Only a part of the research planned is discussed here. It is concerned with the use of data revised in different years.

The idea sketched in this report arose during the short stay in IIASA of a group of researchers from SRI PAS in Poland consisting of Zbigniew Nahorski, Joanna Horabik, and Jolanta Jarnicka in the days of 5-7 July 2010. The discussion at that time with Matthias Jonas followed the earlier email correspondence of the Polish group with Matthias Jonas and Gregg Marland. It has been agreed that the group from Poland would focus on a part of the issues raised up during the email discussion, partially presented in the papers by Marland [3] and Marland et al. [4].

The discussion was continued during the Lviv Workshop *3rd International Workshop on Uncertainty in Greenhouse Gas Inventories*, Lviv, 22-24 September, 2010, with additional participation of Gregg Marland and Khrystyna Hamal (Boychuk).

[1] A preliminary methodology for this has already been proposed by Khrystyna Hamal (Boychuk) [2] in her Interim Report.

Chapter 2

Presentation of an idea

2.1 Basic questions

The approach proposed here is based on results presented by Hamal [2] but the problem is attacked from a slightly different point of view.
To solve the problem formulated in the introduction, two questions have to be answered.

(i) how can the data revised in consecutive years be efficiently used in estimating the uncertainty,

(ii) how to find and interpret the correlations between different national inventories.

The interest standing under the former question is obvious. The revisions made in different years use different knowledge, and thus the uncertainties of the different revisions are incomparable. The latter question pertains to the fact, that all inventories are highly correlated due to similar patterns in their evolution.

Let us consider the latter question. We model the data to be composed of the "real" emission, which we call the "deterministic" fraction and a "stochastic" fraction, related to our lack of knowledge and imprecision of observation of the real emission. We assume that the uncertainty is related to the stochastic part of the model. However, there may be possible to explain partly the stochastic fraction by other variables correlated with it. In our case, the other variables will be stochastic fractions of other inventories. Thus, first we should try to extract the "deterministic" fraction of the inventory data and then look for correlation of the residuals, which can be interpreted as realizations of the "stochastic" fraction. Correlation of residuals for different countries can partly explain their stochastic fractions. The unexplained part of the residuals will then be interpreted as "the uncertainty". Therefore, we have to find the "deterministic" fraction.

2.2 Procedure for the set of revised data for a given country

Going to the first question, each revision data, for a given country, forms a realization of a stochastic process. These stochastic processes for a fixed country are different, but related, see Figures 1 and 2, as examples. They form a bunch of stochastic processes.

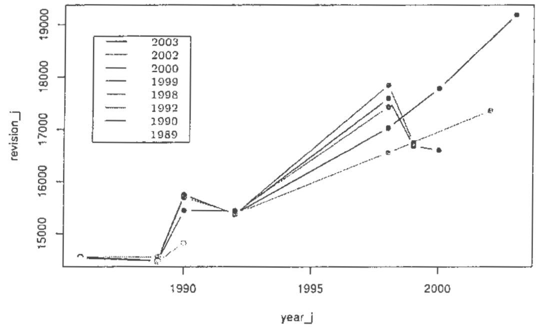

Fig. 1. Graphical illustration of the revisions made in 1989 – 2004; Austria, CDIAC data.

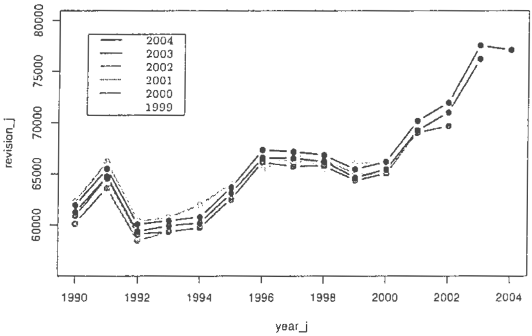

Fig. 2. Graphical illustration of the revisions made in 1999 – 2005; Austria, data from the National Inventory Reports.

The model proposed describes a relationship between different processes related with the revised data for a given country. We assume that the mean values of these processes are linear functions of the differences between the revision years. Similarly, we assume that their standard deviations are also linear functions of the same difference.

To make the above description more precise, let us turn to the mathematical formulation of the model presented. For simplification, we assume that we work with absolute errors, i.e. that the stochastic part is expressed in the weight units. But we can easily interpret the equations below as dealing with the relative errors, expressed in percents (%), by inserting for the inventory data their logarithms. To simplify the description, let us assume that the most recent revised data are for the year 2007. For the present paragraph assume also that we know how to decompose these data into the deterministic and stochastic fraction. Following the notation from Hamal [2], it can be written as

$$E^x_{2007,i} = D^x_{2007,i} + S^x_{2007,i} \tag{2.1}$$

where

E stands for the emission inventory,

D stands for its deterministic fraction,

S stands for the stochastic fraction, and is normally distributed,

i is the country identification number,

x is the year, for which the revised data were recalculated, let's say [1989, 2007].

Now, the data revised in the year $y < 2007$ are modeled as having the same deterministic fraction. Their stochastic fraction is also normally distributed but with different parameters. Thus they follow the same type of decomposition

$$E^x_{y,i} = D^x_{2007,i} + S^x_{y,i} \tag{2.2}$$

with

$$S^x_{y,i} \sim \mathcal{N}(m^x_{y,i}, \sigma^x_{y,i}), \tag{2.3}$$

where the means and the standard deviations are linear functions of the differences in years when the revisions were made

$$m^x_{y,i} = a_i(2007 - y), \qquad \sigma^x_{y,i} = \sigma_{2007,i} + b_i(2007 - y). \tag{2.4}$$

To better explain the indices used we present them in the sketch from Table 1, where Y generalizes the exemplary year 2007. Each revised data for a given year form in it a row.

$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\ldots$	$E^{x-1}_{y,i}$	$E^x_{y,i}$	$E^{x+1}_{y,i}$	$\ldots$	$E^y_{y,i}$	0	$\ldots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\ldots$	$E^{x-1}_{Y,i}$	$E^x_{Y,i}$	$E^{x+1}_{Y,i}$	$\ldots$	$E^y_{Y,i}$	$E^{y+1}_{Y,i}$	$\ldots$	$E^Y_{Y,i}$

Table 1. Indexing the data.

Parameters a_i and b_i above can be estimated from the data together with $\sigma_{2007,i}$. But $\sigma_{2007,i}$ can also be estimated in another way (see the sequel). Parameter a_i describes a shift in the accuracy of the inventory gathering, and b_i – a shift of the precision level, both due to the learning. The above functions can be extended to more complicated nonlinear dependencies, if necessary. Some of them may be linear in parameters, like for example quadratic or other polynomial functions. Then the method outlined below practically does not change.

The above formulation results in a multiinput-multioutput linear regression problem to estimate the parameters a_i and b_i, once the deterministic fraction is extracted from the data. The Bayes estimation can be a competitive approach, taking into account rather limited number of data and possibility to use additional knowledge of *a priori* uncertainties provided by, at least some, countries.

We skip the details of estimation and return to the problem of how to find the deterministic fraction. For this purpose, the smoothing splines can be used, as presented by Nahorski & Jęda [5]. This approach, when applied to the most recently revised data (e.g. for the year 2007), will give not only the estimate of the deterministic fraction, but also an estimate of the variance $\sigma^x_{2007,i}$, which can either be used directly in the formula for above, or as the a priori value in the Bayes estimation.

Finally, the above procedure can be presented in the algorithmic way.

1. For the most recently revised inventory data calculate the smoothing spline and the estimate of the variance of the stochastic fraction.

2. Subtract the spline data from all earlier revisions.

3. Use the data obtained in the linear regression or Bayes estimation to get estimates of a_i and b_i, and hence also $m^x_{y,i}$, and $\sigma^x_{y,i}$.

This gives the full description of the inventory data for all the revisions. Having obtained $m^x_{y,i}$ and $\sigma^x_{y,i}$ it is possible to scale all the data to the $\mathcal{N}(0, \sigma_{2007,i})$ distribution in order to get a set of homogeneous residuals. There will be one or more residuals for a given year, depending on the number of the revised data available for this year.

2.3 Finding the correlation structure in the data for many countries

The sets of homogeneous residuals for each country may be examined for common correlations. One can expect that some group of strongly correlated countries may emerge in the examination. Different methods to cluster the countries into such groups can be used. They may be considered as similar countries. For each cluster a common model, which includes the correlations of data, can be fitted. The models obtained will provide us with the uncertainty estimates. These, as yet slightly vague deliberations will be presented more precisely in future reports.

Chapter 3

First attempt: revisions treated independently

3.1 Notation

From now on, the following notation will be used.

i – a country label,

J – number of revisions made,

y_j – year of revision, $j = 1, \ldots, J$,

Y – the last year, for which the data are available,

$N^{Y}_{y_j,i}$ – number of data revised in the year y_j (for fixed Y and i, simplified to N_j),

x^n – year, for which the data are calculated, $n = 1, \ldots, N_j$,

$E^{x}_{y,i}$ – the data in the year x revised in the year y,

$v^{Y,x^n}_{y_j,i}$ – the difference of $E^{x^n}_{y,i}$ and the spline built for the data from the year Y (for fixed Y and i denoted by v^n_j),

$m^{Y}_{y_j,i}$ – the expected value related to $v^{Y,x^n}_{y_j,i}$, with respect to years $n = 1, \ldots, N_j$ (simplified to m_j for the given Y and i),

$\sigma^{Y}_{y_j,i}$ – the standard deviation related to $v^{Y,x^n}_{y_j,i}$ with respect to years $n = 1, \ldots, N_j$ (simplified to σ_j for the given Y and i).

3.2 ML estimators of the model parameters

The aim is to check that part of the algorithm which consists of construction of the spline functions, as well as to check the basic assumptions of a shift in the mean values, and standard deviations, in dependence of the lag in preparation of the inventories.

Fix i and Y and consider differences v_j^n of $E_{y_j,i}^{x^n}$ and the spline built on the data from the year Y. Assume that

$$v_j^n \sim \mathcal{N}\Big(m_j, \sigma_j\Big), \quad j = 1, \dots, J,$$

where

$$m_j = a_j(Y - y_j) \qquad \sigma_j = \sigma_Y + b_j\, f(Y - y_j), \qquad b_j \neq 0, \tag{3.1}$$

and f is a given function, such that $f(Y - y_j) > -\frac{\sigma_Y}{b_j}$, $j = 1, \dots, J$. Assume also that v_j^n are independent. Our goal is to estimate the unknown parameters a_j, b_j, $j = 1, \dots, J$, and possibly σ_Y.

For a fixed j the likelihood function $L(\mathrm{p})$, where $\mathrm{p} = (a_j, b_j, \sigma_Y)$, has the form

$$L(\mathrm{p}) = \frac{1}{\sqrt{2\pi}^{N_j}(\sigma_j)^{N_j}} \Pi_{n=1}^{N_j} \exp \frac{\left[v_j^n - m_j\right]^2}{2(\sigma_j)^2}$$

so its logarithm is given by

$$\ln L(\mathrm{p}) = -N_j \ln \sqrt{2\pi} - N_j \ln \sigma_j - \frac{1}{2(\sigma_j)^2} \sum_{n=1}^{N_j} \left[v_j^n - m_j\right]^2,$$

where N_j is the number of differences v_j. Taking into account equations (3.1), we obtain the necessary optimality conditions of the form

$$\frac{\partial \ln L(\mathrm{p})}{da_j} = \frac{Y - y_j}{(\sigma_Y + b_j\, f(Y - y_j))^2} \sum_{n=1}^{N_j} \left[v_j^n - a_j(Y - y_j)\right] = 0, \tag{3.2}$$

$$\frac{\partial \ln L(\mathrm{p})}{db_j} = -\frac{N_j\, f(Y - y_j)}{\sigma_Y + b_j\, f(Y - y_j)} + \frac{f(Y - y_j)}{(\sigma_Y + b_j\, f(Y - y_j))^3} \sum_{n=1}^{N_j} \left[v_j^n - a_j(Y - y_j)\right]^2 = 0, \tag{3.3}$$

$$\frac{\partial \ln L(\mathrm{p})}{d\sigma_Y} = -\frac{N_j}{\sigma_Y + b_j\, f(Y - y_j)} + \frac{1}{(\sigma_Y + b_j\, f(Y - y_j))^3} \sum_{n=1}^{N_j} \left[v_j^n - a_j(Y - y_j)\right]^2 = 0. \tag{3.4}$$

From equation (3.2) we have the estimator of a_j

$$\widehat{a}_j = \frac{1}{N_j(Y - y_j)} \sum_{n=1}^{N_j} v_j^n. \tag{3.5}$$

Equations (3.3) and (3.4) are linearly dependent, but assuming that σ_Y is given, we can determine estimator of b_j. We get the following formula

$$\widehat{b}_j = \left(\sqrt{\frac{1}{N_j} \sum_{n=1}^{N_j} \left(v_j^n - \bar{v}_j\right)^2} - \sigma_Y\right) / f(Y - y_j), \tag{3.6}$$

where $\bar{v}_j = \frac{1}{N_j} \sum_{n=1}^{N_j} v_j^n$.

In sections 4.1 and 4.2 we consider functions f of the form

$$f(Y - y_j) = Y - y_j, \quad j = 1, \dots, J, \quad \text{and} \quad f(Y - y_j) = \sqrt{Y - y_j}, \quad j = 1, \dots, J.$$

Given the function f, he sequences $\widehat{a}_j$ and $\widehat{b}_j$, $j = 1, \dots, J$ can be calculated and then depicted in graphs, as functions of y_j, to test visually their constance.

3.3 Estimator of the model parameters covariance matrix

Under mild assumptions the maximum likelihood parameter estimators are asymptotically normal and unbiased with the asymptotic covariance matrix

$$\Gamma = -E\Big[\frac{\partial^2 \ln L(\mathrm{p})}{\partial \hat{\mathrm{p}} \partial \hat{\mathrm{p}}^T}\Big]^{-1} \tag{3.7}$$

where now $\hat{\mathrm{p}}^T = [\hat{a}_j \quad \hat{b}_j]$ is the vector of the maximum likelihood parameter estimators.

To find the estimator of the covariance matrix the Hessian matrix of the second derivatives

$$\mathbf{H} = \frac{\partial^2 \ln L(\mathrm{p})}{\partial \mathrm{p} \partial \mathrm{p}^T} = \begin{bmatrix} \frac{\partial^2 \ln L(\mathrm{p})}{\partial a_j^2} & \frac{\partial^2 \ln L(\mathrm{p})}{\partial b_j \partial a_j} \\ \frac{\partial^2 \ln L(\mathrm{p})}{\partial a_j \partial b_j} & \frac{\partial^2 \ln L(\mathrm{p})}{\partial b_j^2} \end{bmatrix}$$

has to be calculated. Its entries are given by

$$\begin{aligned}
\frac{\partial^2 \ln L(\mathrm{p})}{\partial a_j^2} &= -\frac{(Y-y_j)^2}{(\sigma_Y + b_j f(Y-y_j))^2}, \\
\frac{\partial^2 \ln L(\mathrm{p})}{\partial b_j \partial a_j} =\frac{\partial^2 \ln L(\mathrm{p})}{\partial a_j \partial b_j} &= -\frac{2(Y-y_j) f(Y-y_j)}{(\sigma_Y + b_j f(Y-y_j))^{\frac{3}{2}}} \sum_{n=1}^{N_j} (v_j^n - a_j(Y-y_j)), \\
\frac{\partial^2 \ln L(\mathrm{p})}{\partial b_j^2} &= \frac{f^2(Y-y_j)}{(\sigma_Y + b_j f(Y-y_j))^2} \sum_{n=1}^{N_j} \left(1 - \frac{3(v_j^n - a_j(Y-y_j))^2}{(\sigma_Y + b_j f(Y-y_j))^2}\right).
\end{aligned}$$

Now inserting the maximum likelihood estimators we get

$$\begin{aligned}
\frac{\partial^2 \ln L(\mathrm{p})}{\partial \widehat{a}_j^2} &= -\frac{(Y-y_j)^2}{\frac{1}{N_j}\sum_{n=1}^{N_j}(v_j^n - \bar{v}_j)^2}, \\
\frac{\partial^2 \ln L(\mathrm{p})}{\partial \widehat{b}_j \partial \widehat{a}_j} =\frac{\partial^2 \ln L(\mathrm{p})}{\partial \widehat{a}_j \partial \widehat{b}_j} &= -\frac{2(Y-y_j) f(Y-y_j)}{(\frac{1}{N_j}\sum_{n=1}^{N_j}(v_j^n - \bar{v}_j)^2)^{\frac{3}{2}}} \sum_{n=1}^{N_j} (v_j^n - \bar{v}_j), \\
\frac{\partial^2 \ln L(\mathrm{p})}{\partial \widehat{b}_j^2} &= \frac{f^2(Y-y_j)}{\frac{1}{N_j}\sum_{n=1}^{N_j}(v_j^n - \bar{v}_j)^2} \sum_{n=1}^{N_j} \left(1 - \frac{3(v_j^n - \bar{v}_j)^2}{\frac{1}{N_j}\sum_{n=1}^{N_j}(v_j^n - \bar{v}_j)^2}\right) = \\
&= N_j f^2(Y-y_j) \left(1 - \frac{3}{\frac{1}{N_j}\sum_{n=1}^{N_j}(v_j^n - \bar{v}_j)^2}\right).
\end{aligned}$$

Let us note that the element $\frac{\partial^2 \ln L(\mathrm{p})}{\partial \hat{b}_j^2}$ may either be positive or negative, depending on the data. If it is positive, the Hessian matrix cannot be negative definite. Thus, the likelihood function may possibly have more than one local maximum.

Chapter 4

Preliminary estimation for several countries

4.1 CDIAC data analysis for Austria

4.1.1 Data

We start with the analysis of the CDIAC[1] data presented in Figure 1, for Austria, in 1989 - 2004. For the data for the year $Y = 2004$, we build a smoothing spline, according to the procedure given in [5]. The result obtained, is shown (in red) in Figure 3, together with data from the year Y.

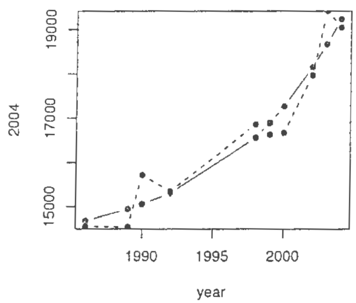

Fig. 3. Smoothing spline; CDIAC data from the year $Y = 2004$, Austria; $\sigma_{2004} = 638.7$.

The data analyzed, come from the years 1986,1989,1990,1992,1998,1999,2000,2002,2003, and 2004. They were recalculated in the years 1989,1990,1992,1998,1999,2000,2002,2003, and

[1]CDIAC – The Carbon Dioxide Information Analysis Center, of the U.S. Department of Energy, located at the Oak Ridge National Laboratory. It provides continuous observations on atmospheric trace gases like e.g. CO_2 emissions. We analyze global CO_2 emissions data from fossil-fuel burning, expressed in million metric tons of carbon.

2004. We calculate differences between the revisions of the data (until 2003) and the smoothing spline (Figure 4).

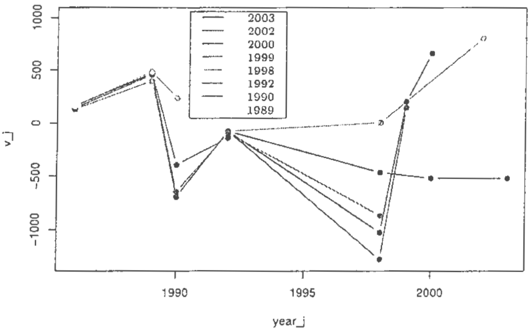

Fig. 4. Differences v_j^n; CDIAC data, $Y = 2004$, Austria

4.1.2 Model 1, $f(Y - y_j) = Y - y_j$

Starting with

$$f(Y - y_j) = Y - y_j, \qquad j = 1, \ldots, J,$$

and hence

$$m_j = a_j(Y - y_j), \qquad \sigma_j = \sigma_Y + b_j\,(Y - y_j), \tag{4.1}$$

we use formulas (3.5) and (3.6) to estimate parameters a_j and b_j (Figure 5), where the estimate of $\sigma_{2004} = 638.7$ was calculated when building the smoothing spline. Having obtained $\widehat{a}_j$ and $\widehat{b}_j$, we get m_j and σ_j (Figure 6). The estimates are gathered in Table 2.

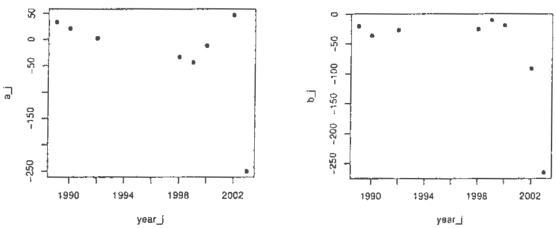

Fig. 5. Estimates of parameters a_j and b_j; model (4.1); Austria.

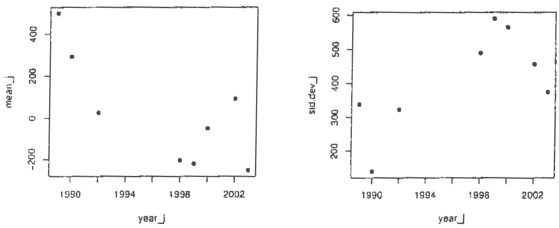

Fig. 6. Calculated values of m_j and σ_j; model (4.1); Austria.

Figure 5 shows that the coefficients $\widehat{a}_j$, until the year 2002, are arranged along consecutive years. It may be noted that $\widehat{a}_{2003}$, stands out from the other values. The same is for the coefficient $\widehat{b}_{2004}$. Except of these coefficients, and perhaps $\widehat{b}_{2002}$, the rest of them have quite similar values in the respective groups. The means of the sequences $\widehat{a}_j$ and $\widehat{b}_j$ are drawn as constant lines in Figure 7. The mean value for the sequence of $\widehat{a}_j$ equals -29.7, and its standard deviation 94.4. For the sequence $\widehat{b}_j$ the mean value is equal to -61.4, and its standard deviation to 85.8.

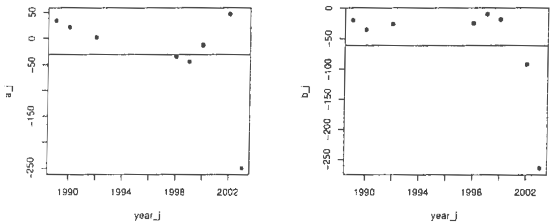

Fig. 7. Mean values for $\widehat{a}_j$ and $\widehat{b}_j$; model (4.1; Austria).

According to (4.1), the values m_j and σ_j should be linear. To check it, straight lines have been fitted to these sequences (Figure 8). For the data analyzed, the ratio R^2 are equal to 61% and 54%.

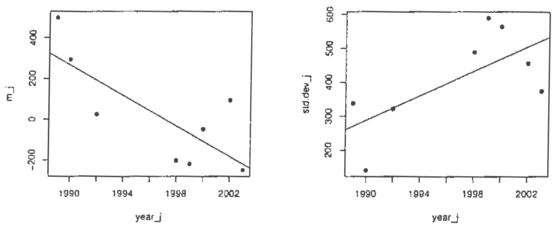

Fig. 8. Linear regression for m_j and σ_j; model (4.1); Austria.

j	1989	1990	1992	1998	1999	2000	2002	2003	mean	std
$\hat{a}_j$	33.18	20.9	2.02	-33.75	-43.74	-12.2	46.34	-250.04	-29.7	94.4
$\hat{b}_j$	-20.07	-35.62	-26.39	-24.92	-9.77	-18.62	-91.11	-264.43	-61.4	85.8
m_j	497.77	292.62	24.2	-202.51	-218.7	-48.81	92.67	-250.04	$R^2 = 0.61$	
σ_j	337.6	140.05	322.1	489.23	589.9	564.26	456.5	374.3	$R^2 = 0.54$	

Table 2. Estimates of parameters in the model (4.1); Austria, CDIAC data.

The question then arises whether the assumption of a linear model for the mean value and standard deviation was justified. To better assess it, the data for other countries will be analyzed, as well as other data sets, e.g. from the National Inventory Reports (Section 4.2). However, before that, we examine another model, with a nonlinear function f.

4.1.3 Model 2, $f(Y - y_j) = \sqrt{Y - y_j}$

Before we analyze other data sets, we consider a modification of the model, i.e.

$$f(Y - y_j) = \sqrt{Y - y_j}, \qquad j = 1, \ldots, J,$$
$$m_j = a_j(Y - y_j), \qquad \sigma_j = \sigma_Y + b_j\sqrt{Y - y_j}. \tag{4.2}$$

Using formulas (3.5) and (3.6) together with (4.2), we obtain the following results (Fig. 9 - 10).

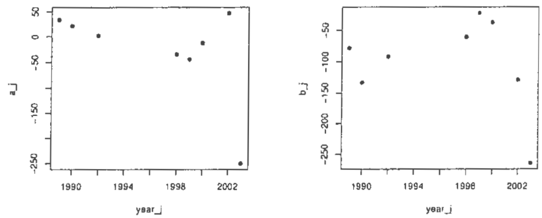

Fig. 9. Estimates of parameters a_j and b_j; model (4.2); Austria.

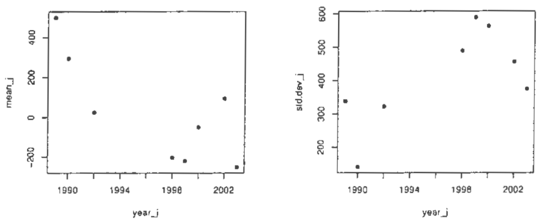

Fig. 10. Calculated values of m_j and σ_j; model (4.2); Austria.

The values estimated are shown in Table 3, and the fit of the model (4.2) is illustrated in Figures 11 and 12. The results for $\hat{a}_j$ and m_j did not change. The main changes occured for $\hat{b}_j$. Now, the mean value is -102,0 and the standard deviation 76,6. The change of model caused decrease of the mean value and standard deviation. Thus, the new sequence of $\hat{b}_j$ is now closer to a constant line, see Figure 11.

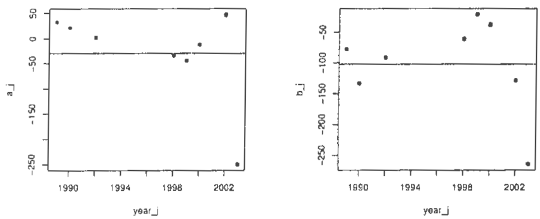

Fig. 11. Mean values for $\hat{a}_j$ and $\hat{b}_j$; model (4.2); Austria.

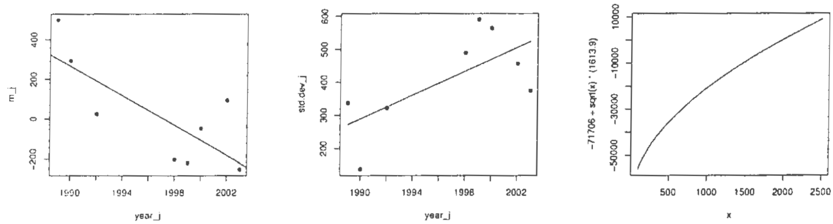

Fig. 12. Regression functions for m_j and σ_j; model (4.2); Austria.

Figure 12 shows the regression function, fitted to sequences of m_j, and σ_j, $j = 1, \ldots, J$. In the case of m_j, we obtain, as in the model (4.1), a linear function. However, to the sequence of σ_j a nonlinear regression function (in blue) was fitted. Due to the fact that locally it gives the impression of a linear one, we also included a graph of the function itself.

j	1989	1990	1992	1998	1999	2000	2002	2003	mean	std
$\hat{a}_j$	33.18	20.9	2.02	-33.75	-43.74	-12.2	46.34	-250.04	-29.7	94.4
$\hat{b}_j$	-77.8	-133.3	-91.4	-61.03	-21.8	-37.24	-128.9	-264.4	-102.0	76.6
m_j	497.77	292.62	24.2	-202.51	-218.7	-48.81	92.67	-250.04	$R^2 = 0.61$	
σ_j	337.6	140.04	322.1	489.2	589.9	564.3	456.5	374.3	$R^2 = 0.46$	

Table 3. Estimates of parameters in the model (4.2); Austria, CDIAC data.

Comparing the results obtained using model (4.2) with those obtained using model (4.1), it can be seen that modification gave a change in the sequence of $\hat{b}_j$, which is now much closer to the constant one. To better examine which of the models allows us to obtain a better fit, we will analyze another data sets (Section 4.2).

4.2 Analysis of the data from the Austrian National Inventory Reports

4.2.1 Data

In this section we analyze data from the National Inventory Reports[2]. For a better comparison of the results obtained, with those from section 4.1, we consider again the data for Austria. The data refer to CO_2 emissions in the years 1990 – 2005, and recalculations (revisions), performed every year, from 1999 to 2005. We start, as before, with building a smoothing spline for the data of year $Y = 2005$ (Figure 13).

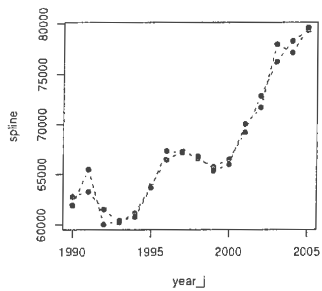

Fig. 13. Smoothing spline; data from National Inventory Reports; $Y = 2005$, Austria; $\sigma_{2005} = 2065.5$.

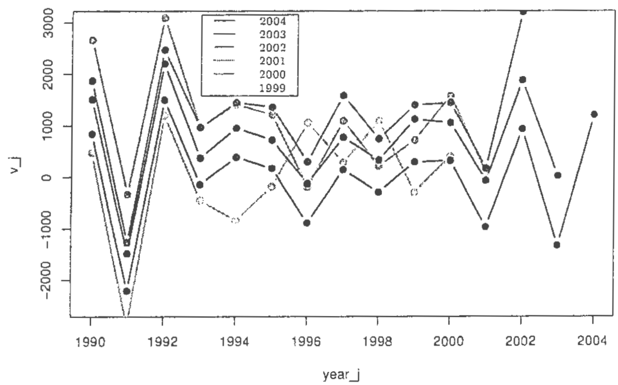

Fig. 14. Differences v_j^n; data from National Inventory Reports; $Y = 2005$, Austria

[2]According to United Nations Framework Convention on Climate Change (UNFCCC), each of the countries which have signed the Convention, is obliged to provide annually data on greenhouse gas inventories.

Then we calculate the differences v_j^n, between the constructed spline, and subsequent revisions of the data, carried out in 1999 - 2004. The results are shown in Figure 14.

4.2.2 Model 1

We start with the model (4.1). Differences v_j^n and formulas (3.5), and (3.6), are used to estimate the parameters a_j and b_j (Figure 15). Estimates $\hat{a}_j$ and $\hat{b}_j$ make it possible to calculate the values m_j and σ_j (Figure 16).

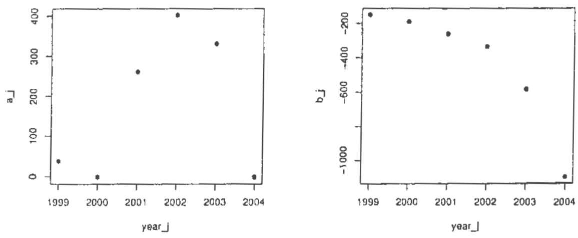

Fig. 15. Estimates of parameters a_j and b_j; model (4.1); Austria.

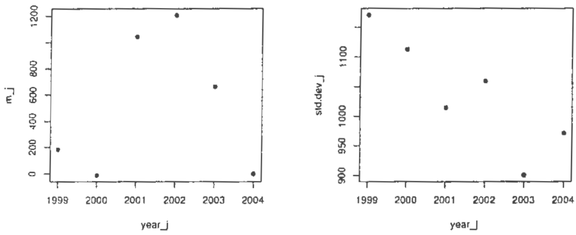

Fig. 16. Calculated values of m_j and σ_j; model (4.1); Austria.

In contrast to the CDIAC data, in this case, the $\hat{b}_j$ sequence is characterized by an evident decreasing linear trend. Coefficients $\hat{a}_j$ seem to be quite dispersed. To investigate the constance of sequences $\hat{a}_j$ and $\hat{b}_j$, as well as the linearity of m_j and σ_j, we calculate, as previously, the mean values and the linear regression, and then test goodness of fit of a linear model. The results are depicted in Figures 17 and 18, and in Table 4.

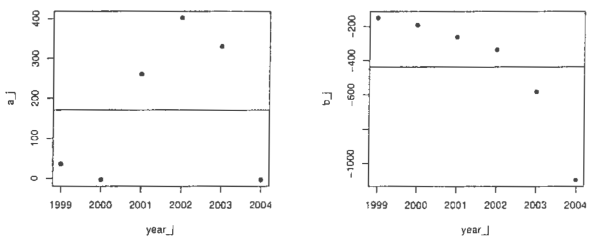

Fig. 17. Mean values for $\hat{a}_j$ and $\hat{b}_j$; model (4.1); Austria.

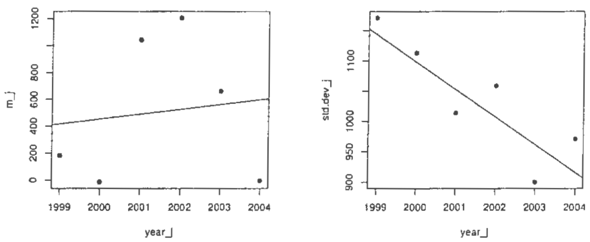

Fig. 18. Linear regression for m_j and σ_j; model (4.1); Austria.

j	1999	2000	2001	2002	2003	2004	mean	std
$\hat{a}_j$	36.25	-2.72	260.81	402.24	330.49	-3.21	170.6	182.0
$\hat{b}_j$	-149.16	-190.50	-262.85	-335.42	-582.58	-1094.44	-435.8	357.1
m_j	181.23	-13.58	1043.25	1206.73	660.98	-3.20	$R^2 = 0.37$	
σ_j	1170.6	1113.01	1014.1	1059.28	900.38	971.1	$R^2 = 0.76$	

Table 4. Estimates of parameters in the model (4.1); Austria; data from the National Inventory Reports.

4.2.3 Model 2

Now, consider the model (4.2). The results of the analysis conducted are shown in Figures 19 - 22, and in Table 5.

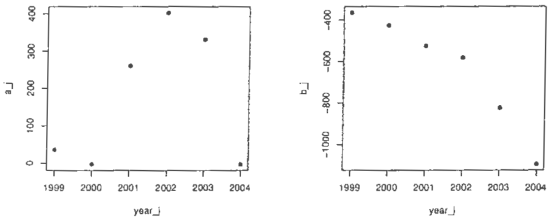

Fig. 19. Estimates of parameters a_j and b_j; model (4.2); Austria.

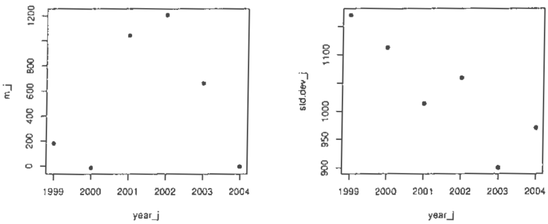

Fig. 20. Calculated values of m_j and σ_j; model (4.2); Austria.

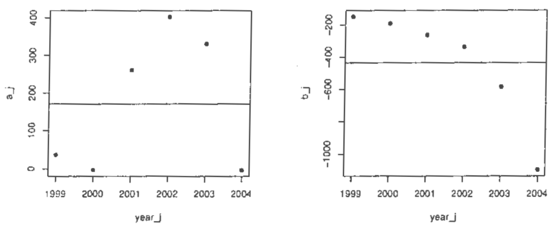

Fig. 21. Mean values for $\hat{a}_j$ and $\hat{b}_j$; model (4.2); Austria.

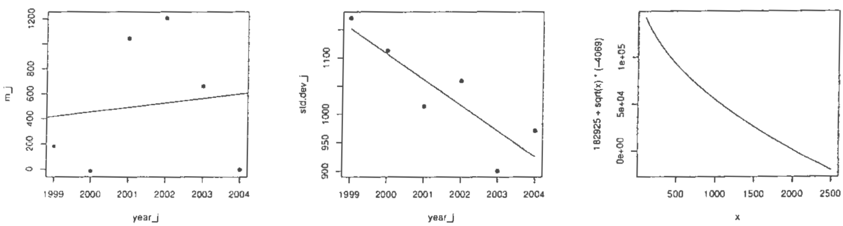

Fig. 22. Regression functions for m_j and σ_j; model (4.2); Austria.

j	1999	2000	2001	2002	2003	2004	mean	std
$\hat{a}_j$	36.25	-2.72	260.81	402.24	330.49	-3.21	170.6	182.0
$\hat{b}_j$	-365.4	-425.9	-525.7	-580.9	-823.9	-1094.4	-636.0	275.0
m_j	181.23	-13.58	1043.25	1206.73	660.98	-3.20	$R^2 = 0.37$	
σ_j	1170.6	1113.1	1014.1	1059.3	900.4	971.1	$R^2 = 0.76$	

Table 5. Estimated parameters in the model (4.2); Austria, data from the National Inventory Reports.

For these data, coming from the National Inventory Reports, model (4.2) again improves constance of the sequences of b_j, $j = 1, \ldots, J$. However, its decreasing character is still evident (Figures 19 and 21). The nonlinear regression function, fitted to σ_j, $j = 1, \ldots, J$, also in this case locally seems to be linear. Its graph (in blue) is shown in Figure 22, on the right.

4.3 Analysis of the data from several other countries

In order to better examine this type of data, the same analysis for a few more EU countries will be conducted. The results obtained are presented below, in form of figures and tables.

4.3.1 Belgium

The linear model for σ_j

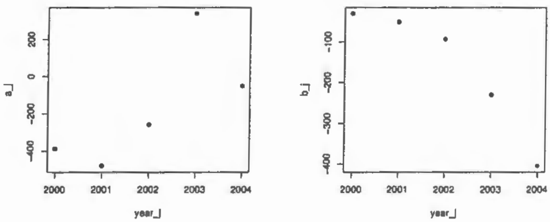

Fig. 23. Estimates of parameters a_j and b_j; model (4.1); Belgium.

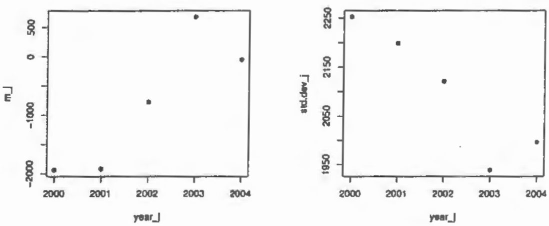

Fig. 24. Calculated values of m_j and σ_j; model (4.1); Belgium.

j	2000	2001	2002	2003	2004	mean	std
$\hat{a}_j$	-388.04	-480.5	-258.48	338.06	-52.45	-168.3	325.5
$\hat{b}_j$	-29.54	-50.41	-93.19	-230.59	-404.29	-161.6	156.6
m_j	-1940.21	-1921.99	-775.45	676.13	-52.45	$R^2 = 0.77$	
σ_j	2252.40	2198.5	2120.55	1938.94	1995.83	$R^2 = 0.85$	

Table 6. Estimates of parameters in the model (4.1); Belgium, data from the National Inventory Reports.

The square root model for σ_j

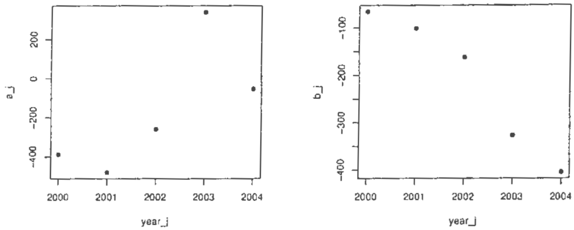

Fig. 25. Estimates of parameters a_j and b_j; model (4.2); Belgium.

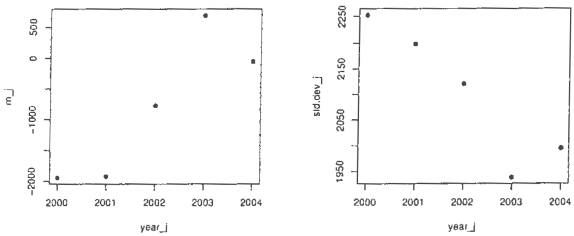

Fig. 26. Calculated values of m_j and σ_j; model (4.2); Belgium.

j	2000	2001	2002	2003	2004	mean	std
$\hat{a}_j$	-388.04	-480.5	-258.48	338.06	-52.45	-168.3	325.5
$\hat{b}_j$	-66.07	-100.8	-161.4	-326.1	-404.3	-211.7	146.8
m_j	-1940.21	-1921.99	-775.45	676.13	-52.45	$R^2 = 0.77$	
σ_j	2252.4	2198.5	2120.5	1938.9	1995.8	$R^2 = 0.96$	

Table 7. Estimates of parameters in the model (4.2); Belgium, data from the National Inventory Reports.

4.3.2 Netherlands

The linear model for σ_j

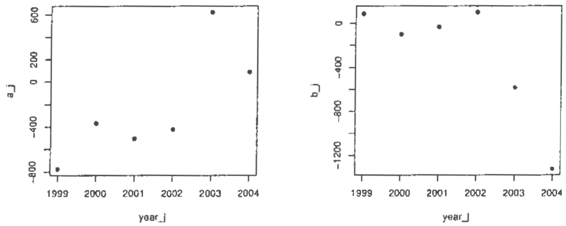

Fig. 27. Estimates of parameters a_j and b_j;model (4.1); Netherlands.

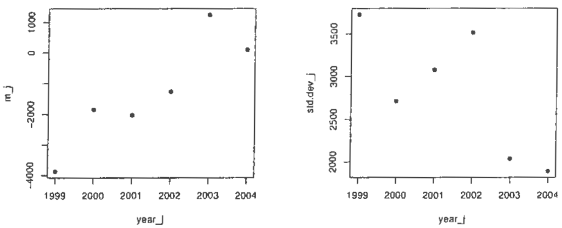

Fig. 28. Calculated values of m_j and σ_j; model (4.1); Netherlands.

j	1999	2000	2001	2002	2003	2004	mean	std
$\hat{a}_j$	-776.86	-370.23	-507.11	-425.31	619.04	88.69	-228.6	501.2
$\hat{b}_j$	85.88	-99.68	-33.39	101.6	-586.9	-1321.77	-309.0	556.3
m_j	-3884.32	-1851.16	-2028.45	-1275.94	1238.08	88.69	$R^2 = 0.65$	
σ_j	3725.78	2712.1	3076.95	3515.31	2036.71	1888.73	$R^2 = 0.59$	

Table 8. Estimates of parameters in the model (4.1); Netherlands; data from the National Inventory Reports.

The square root model for σ_j

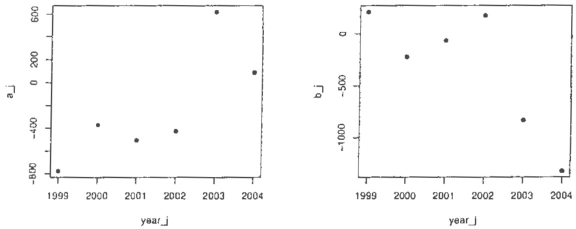

Fig. 29. Estimates of parameters a_j and b_j; model (4.2); Netherlands.

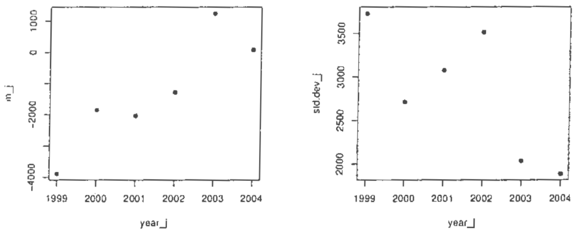

Fig. 30. Calculated values of m_j and σ_j; model (4.2); Netherlands.

j	1999	2000	2001	2002	2003	2004	mean	std
$\hat{a}_j$	-776.86	-370.23	-507.11	-425.31	619.04	88.69	-228.6	501.2
$\hat{b}_j$	210.4	-222.9	-66.78	175.9	-829.9	-1321.8	-342.5	610.1
m_j	-3884.32	-1851.16	-2028.45	-1275.94	1238.08	88.69	$R^2 = 0.65$	
σ_j	3725.78	2712.1	3076.95	3515.31	2036.71	1888.73	$R^2 = 0.59$	

Table 9. Estimates of parameters in the model (4.2); Netherlands; data from the National Inventory Reports.

4.3.3 Denmark

The linear model for σ_j

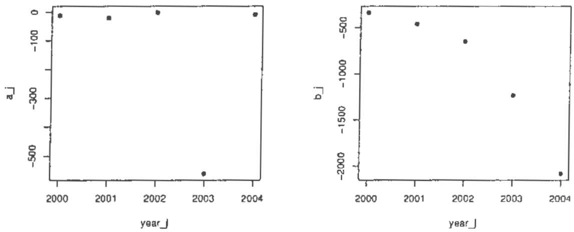

Fig. 31. Estimates of parameters a_j and b_j; model (4.1); Denmark.

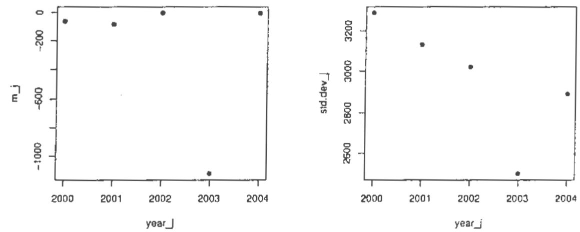

Fig. 32. Calculated values of m_j and σ_j; model (4.1); Denmark.

j	2000	2001	2002	2003	2004	mean	std
$\hat{a}_j$	-12.64	-21.32	-1.76	-560.9	-8.46	-121.0	246.0
$\hat{b}_j$	-336.2	-459.6	-649.8	-1235.1	-2079.8	-952.1	718.5
m_j	-63.2	-85.29	-5.27	-1121.9	-8.46	$R^2 = 0.10$	
σ_j	3288.5	3131.2	3020.02	2499.3	2889.7	$R^2 = 0.58$	

Table 10. Estimates of parameters in the model (4.1); Denmark, data from the National Inventory Reports.

The square root model for σ_j

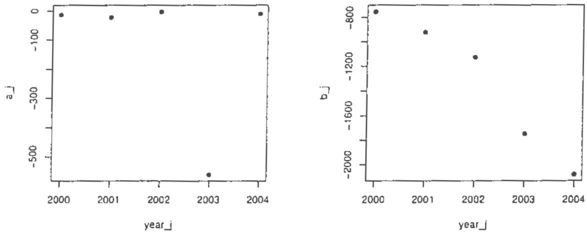

Fig. 33. Estimates of parameters a_j and b_j; model (4.2); Denmark.

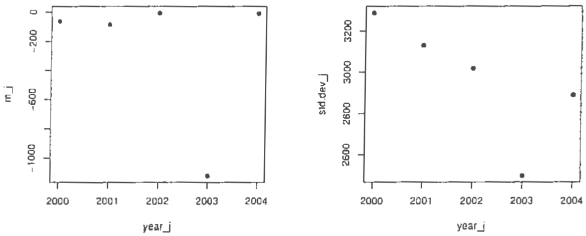

Fig. 34. Calculated values of m_j and σ_j; model (4.2); Denmark.

j	2000	2001	2002	2003	2004	mean	std
$\hat{a}_j$	-12.64	-21.32	-1.76	-560.9	-8.46	-121.0	246.0
$\hat{b}_j$	-751.8	-919.1	-1125.5	-1746.7	-2079.8	-1324.6	565.8
m_j	-63.2	-85.29	-5.27	-1121.9	-8.46	$R^2 = 0.10$	
σ_j	3288.4	3131.3	3020.02	2499.3	2889.6	$R^2 = 0.59$	

Table 11. Estimates of parameters in the model (4.2); Denmark; data from the National Inventory Reports.

4.3.4 Finland

The linear model for σ_j

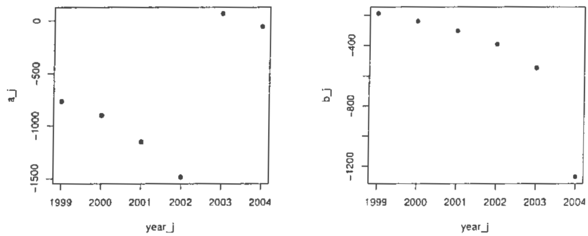

Fig. 35. Estimates of parameters a_j and b_j; model (4.1); Finland.

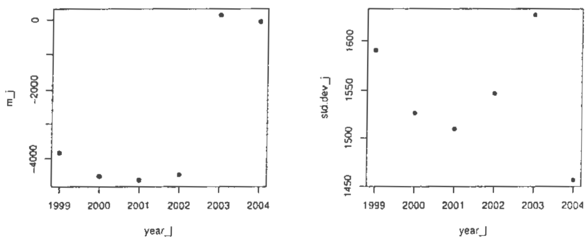

Fig. 36. Calculated values of m_j and σ_j; model (4.1); Finland.

j	1999	2000	2001	2002	2003	2004	mean	std
$\hat{a}_j$	-768.21	-902.93	-1154.69	-1487.88	67.41	-57.67	-717.3	612.0
$\hat{b}_j$	-189.17	-239.84	-303.95	-392.89	-549.09	-1268.7	-490.6	401.7
m_j	-3841.05	-4514.67	-4618.74	-4463.63	134.83	-57.67	$R^2 = 0.60$	
σ_j	1590.14	1525.97	1509.38	1546.52	1627.0	1456.49	$R^2 = 0.48$	

Table 12. Estimates of parameters in the model (4.1); Finland, data from the National Inventory Reports.

The square root model for σ_j

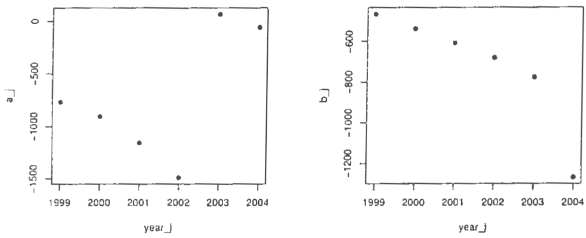

Fig. 37. Estimates of parameters a_j and b_j; model (4.2); Finland.

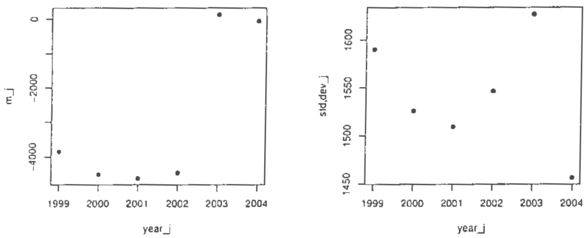

Fig. 38. Calculated values of m_j and σ_j; model (4.2); Finland.

j	1999	2000	2001	2002	2003	2004	mean	std
$\hat{a}_j$	-768.21	-902.93	-1154.69	-1487.88	67.41	-57.67	-717.3	612.1
$\hat{b}_j$	-463.4	-536.3	-607.9	-680.5	-776.5	-1268.7	-722,2	289.1
m_j	-3841.05	-4514.67	-4618.74	-4463.63	134.83	-57.67	$R^2 = 0.60$	
σ_j	1590.1	1525.9	1509.4	1546.5	1627	1456.5	$R^2 = 0.84$	

Table 13. Estimates of parameters in the model (4.2); Finland, data from the National Inventory Reports.

4.3.5 UK

The linear model for σ_j

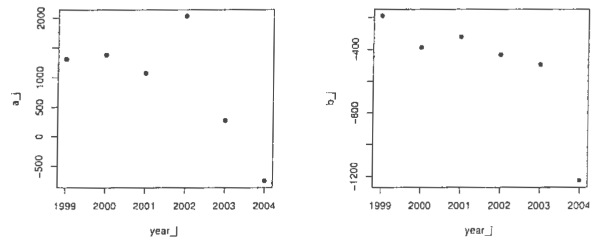

Fig. 39. Estimates of parameters a_j and b_j; model (4.1); UK.

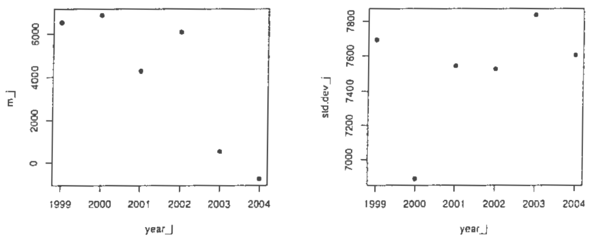

Fig. 40. Calculated values of m_j and σ_j; model (4.1); UK.

j	1999	2000	2001	2002	2003	2004	mean	std
$\hat{a}_j$	1304.6	1373.9	1068.08	2026.57	265.03	-749.99	881.4	980.7
$\hat{b}_j$	-189.17	-239.84	-303.95	-392.89	-549.09	-1268.7	-490.6	401.7
m_j	-3841.05	-4514.67	-4618.74	-4463.63	134.83	-57.67	$R^2 = 0.60$	
σ_j	1590.14	1525.97	1509.38	1546.52	1627.0	1456.49	$R^2 = 0.84$	

Table 14. Estimates of parameters in the model (4.1); UK; data from the National Inventory Reports.

The square root model for σ_j

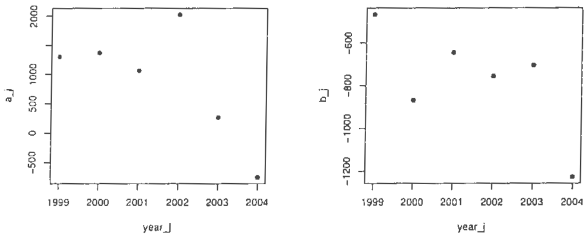

Fig. 41. Estimates of parameters a_j and b_j; model (4.2); UK.

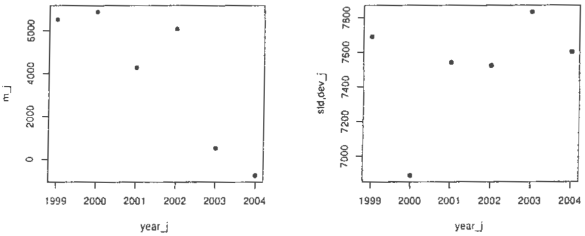

Fig. 42. Calculated values of m_j and σ_j; model (4.2); UK.

j	1999	2000	2001	2002	2003	2004	mean	std
$\hat{a}_j$	1304.6	1373.9	1068.08	2026.57	265.03	-749.99	881.4	980.7
$\hat{b}_j$	-465.4	-868.5	-644.6	-755.06	-703.9	-1227.4	-777.5	257.6
m_j	-3841.05	-4514.67	-4618.74	-4463.63	134.83	-57.67	$R^2 = 0.60$	
σ_j	7691.5	6889.5	7542.3	7523.7	7836	7604.16	$R^2 = 0.43$	

Table 15. Estimates of parameters in the model (4.2); UK, data from the National Inventory Reports.

Summing up the above results, the model (4.2) improves constance of the sequence $\hat{b}_j$ in almost all cases, except for the Netherland. However, the sequence $\hat{b}_j$ is still decreasing. This means that it is possible to find a better model, which will provide a sequence of $\hat{b}_j$ much closer to a constant sequence.

Chapter 5

First attempt: All revisions considered together

5.1 Maximum likelihood estimators

We assume now that the parameters in the expected value and standard deviation models do not depend on revisions, that is a and b do not depend on j. Thus, the model has the following form

$$m_j = a(Y - y_j) \qquad \sigma_j = \sigma_Y + b\,f(Y - y_j), \quad b \neq 0, \tag{5.1}$$

where f is a function, satisfying $f(Y - y_j) > -\frac{\sigma_Y}{b}$. Now, we consider all data, for all revisions, that is $\sum_j^J N_j$ data.

The likelihood function is now equal to

$$L(\mathrm{p}) = \Pi_{j=1}^{J} \frac{1}{\sqrt{2\pi}^{N_j} (\sigma_j)^{N_j}} \Pi_{n=1}^{N_j} \exp \frac{\left[v_j^n - m_j\right]^2}{2(\sigma_j)^2}$$

and its logarithm

$$\ln L(\mathrm{p}) = \sum_{j=1}^{J} \Big[- N_j \ln \sqrt{2\pi} - N_j \ln \sigma_j - \frac{1}{2(\sigma_j)^2} \sum_{n=1}^{N_j} \left[v_j^n - m_j\right]^2 \Big].$$

Taking into account model (5.1), we get the first order necessary optimality conditions of the form

$$\frac{\partial \ln L(\mathrm{p})}{da} = \sum_{j=1}^{J} \frac{f(Y - y_j)}{(\sigma_Y + b\,f(Y - y_j))^2} \sum_{n=1}^{N_j} \left[v_j^n - a(Y - y_j)\right] = 0, \tag{5.2}$$

$$\frac{\partial \ln L(\mathrm{p})}{db} = \sum_{j=1}^{J} \Big[\frac{-N_j f(Y - y_j)}{\sigma_Y + b\,f(Y - y_j)} + \frac{f(Y - y_j)}{(\sigma_Y + b\,f(Y - y_j))^3} \sum_{n=1}^{N_j} \left[v_j^n - a(Y - y_j)\right]^2 \Big] = 0, \tag{5.3}$$

$$\frac{\partial \ln L(\mathrm{p})}{d\sigma_Y} = \sum_{j=1}^{J} \Big[\frac{-N_j}{\sigma_Y + b\,f(Y - y_j)} + \frac{1}{(\sigma_Y + b\,f(Y - y_j))^3} \sum_{n=1}^{N_j} \left[v_j^n - a(Y - y_j)\right]^2 \Big] = 0. \tag{5.4}$$

Equations (5.3) and (5.4) are linearly dependent, so it is impossible to estimate both b and σ_Y. Only the sum $\sigma_Y + b\,f(Y-y_j)$ can be calculated from each of those equations. As before, we therefore assume that the estimate of σ_Y is known.

From equation (5.2) we can find the estimate of a of the form

$$\hat{a} = \frac{\sum_{j=1}^{J} \frac{Y-y_j}{(\sigma_Y + b\,f(Y-y_j))^2} \sum_{n=1}^{N_j} v_j^n}{\sum_{j=1}^{J} \frac{N_j (Y-y_j)^2}{(\sigma_Y + b\,f(Y-y_j))^2}}$$

To find an estimate of b it is necessary to solve one of nonlinear equations – (5.3) or (5.4), substituting the estimate $\hat{a}$ for a. Thus, to estimate the parameters a and b, one needs to solve numerically the system of equations (5.2) –(5.3) or (5.2) – (5.4).

To find the asymptotic parametric covariance matrix, we calculate the second derivatives.

$$\frac{\partial^2 \ln L(\mathbf{p})}{\partial a^2} = -\sum_{j=1}^{J} \frac{(Y-y_j) f(Y-y_j)}{[\sigma_Y + b f(Y-y_j)]^2} \tag{5.5}$$

$$\frac{\partial^2 \ln L(\mathbf{p})}{\partial a \partial b} = -2\sum_{j=1}^{J} \frac{f^2(Y-y_j)}{[\sigma_Y + b f(Y-y_j)]^3} \sum_{n=1}^{N_j} [v_j^n - a(Y-y_j)] \tag{5.6}$$

$$\begin{aligned} \frac{\partial^2 \ln L(\mathbf{p})}{\partial b^2} &= \sum_{j=1}^{J} \frac{N_j f(Y-y_j)}{[\sigma_Y + b f(Y-y_j)]^2} - \frac{3 f(Y-y_j)}{[\sigma_Y + b f(Y-y_j)]^4} \sum_{n=1}^{N_j} [v_j^n - a(Y-y_j)]^2 \\ &= \sum_{j=1}^{J} \frac{f(Y-y_j)}{[\sigma_Y + b f(Y-y_j)]^2} \left[N_j - \frac{3}{[\sigma_Y + b f(Y-y_j)]^2} \sum_{n=1}^{N_j} [v_j^n - a(Y-y_j)]^2 \right] \\ &= \sum_{j=1}^{J} \sum_{n=1}^{N_j} \frac{f(Y-y_j)}{[\sigma_Y + b f(Y-y_j)]^2} \left[1 - \frac{3[v_j^n - a(Y-y_j)]^2}{[\sigma_Y + b f(Y-y_j)]^2} \right] \end{aligned} \tag{5.7}$$

As before, there is no guarantee that the Hessian matrix is negative definite.

To find the covariance matrix, the maximum likelihood parameter estimates have to be inserted and the Hessian matrix inverted.

5.2 Data analysis

Let us apply the model

$$\begin{aligned} f(Y-y_j) &= Y - y_j, \qquad j = 1, \ldots, J, \\ m_j &= a(Y-y_j) \qquad \sigma_j = \sigma_Y + b(Y-y_j) \end{aligned} \tag{5.8}$$

to previously analyzed data, starting with Austria. The parameter estimates are depicted in Table 16. For comparison, the mean values of the sequences $\hat{a}_j$ and $\hat{b}_j$, together with their mean square deviations from $\hat{a}$ and $\hat{b}$, respectively, are given. The mean square deviations are calculated according to the following formulas.

$$s_a = \sqrt{\frac{1}{J}\sum_{j=1}^{J} (\hat{a}_j - \hat{a})^2} \quad s_b = \sqrt{\frac{1}{J}\sum_{j=1}^{J} (\hat{b}_j - \hat{b})^2}$$

Data	$\hat{a}$	mean	s_a	$\hat{b}$	mean	s_b
CDIAC data	-29.01	-29.7	94.4	-22.13	-61.4	95.5
Nat.Inv.Rep.	218.12	170.6	189.3	-387.7	-435.8	361.0

Table 16. Estimates of parameters a and b in the model (5.8); Austria.

To illustrate the results, we present them in Figures 43 and 44, together with sequences $\hat{a}_j$ and $\hat{b}_j$, calculated in section 4.1 and 4.2.

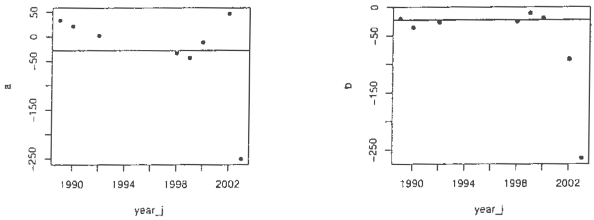

Fig. 43. Estimates of parameters a and b; model (5.8); Austria; CDIAC data.

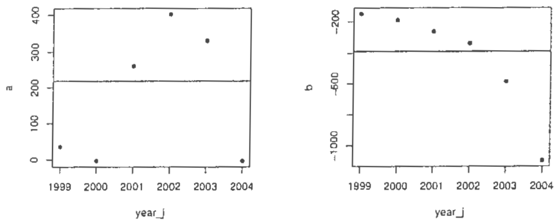

Fig. 44. Estimates of parameters a and b; model (5.8); Austria; data from the National Inventory Reports.

We also conduct the analysis for other data sets from section 4.2 (the results obtained are in Table 17 below).

Country	$\hat{a}$	mean	$\hat{b}$	mean
Belgium	-331.47	-168.3	-62.28	-161.6
Netherlands	-411.86	-228.6	-30.9	-309.0
Denmark	-36.98	-121.0	-513.83	-952.1
Finland	-734.04	-717.3	-562.6	-490.6
UK	1285.95	881.4	-328.17	-490.6

Table 17. Estimates of parameters a and b; model (5.8) for EU countries, considered in section 4.2.

Now we consider model of the form

$$f(Y - y_j) = \sqrt{Y - y_j}, \qquad j = 1, \ldots, J,$$
$$m_j = a(Y - y_j) \qquad \sigma_j = \sigma_Y + b\sqrt{Y - y_j}, \qquad (5.9)$$

and apply it to the CDIAC data for Austria and to the data from the Austrian National Inventory Reports. The results obtained are presented in Table 18, and in Figures 45 and 46.

Data	$\hat{a}$	mean	s_a	$\hat{b}$	mean	s_b
CDIAC	-28.8	-29.7	94.4	-79.92	-101.98	344.2
Nat.Inv.Rep.	217.1	170.6	189.0	-694.7	-435.8	282.4

Table 18. Estimates of parameters a and b in the model (5.9); Austria.

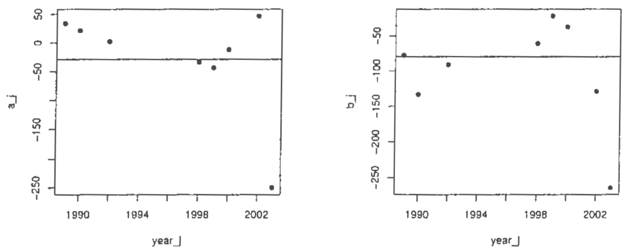

Fig. 45. Estimates of parameters a and b; model (5.9); Austria; CDIAC data.

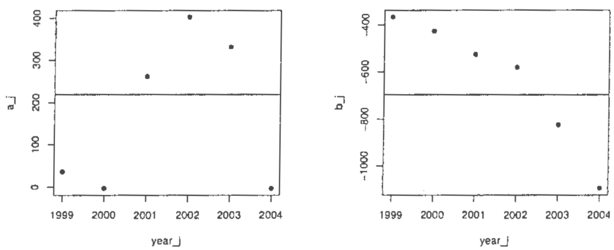

Fig. 46. Estimates of parameters a and b; model (5.9); Austria; data from the National Inventory Reports.

Finally, we use the model (5.9) to data from the National Inventory Reports for EU countries analyzed in section 4.2. Results of the analysis are presented in Table 19.

Country	$\hat{a}$	mean	$\hat{b}$	mean
Belgium	-301.6	-168.3	-122.0	-161.6
Netherlands	-396.6	-228.6	-261.1	-309.0
Denmark	-55.8	-121.0	-633.4	-952.1
Finland	-729.2	-717.3	-577.5	-490.6
UK	1156.5	881.4	-369.6	-490.6

Table 19. Estimates of parameters a and b; model (5.9) for EU countries, considered in section 4.2.

These are very preliminary results, just showing possibility of calculations. Much better interpretation of these models will be possible when the nonlinear model for the revisions treated independently is found.

Chapter 6

Conclusions

In this report an idea of a model describing the learning process in evaluation of the national emissions is presented. The maximum likelihood estimators of the model parameters and their covariance matrix have been derived. Preliminary calculations for data from several countries and a discussion on choice of the model have been done.

The following preliminary conclusion can be drawn at this stage.

- The square root model parameter estimates are closer to the mean values than those for the linear model. However, the sequence of estimates $\hat{b}_j$ have clear decreasing patterns. This suggests that a better model can still be found.
- Comparison of the results from the CDIAC and National Inventory Report data for Austria shows that they have quite different characters. In particular, while the variance for the latter data have a decreasing trend, the trend for the former data is rather increasing. As the decreasing character can easily be interpreted to be due to learning, it is rather difficult at this stage to interpret the increasing variance of the CDIAC data.
- The sequences of $\hat{a}_j$ do not show clear common trend for data from examined countries. Perhaps a model of the form $m_j = a_{1j}(Y - a_j) + a_{2j}(Y - y_j)^2$ could be tried, with very different estimates of parameters. The problem, which could be spotted, is connected with very small number of data in each revision (5-6 values). This may not be enough for good estimation of an additional parameter.

Bibliography

[1] Ch. D. Ghiliani (2010) Adjustment Computations: Spatial Data Analysis. Wiley.

[2] K. Hamal (2010) Reporting GHG emissions: Change in uncertainty and its relevance for detection of emission changes. Interim Report IR-10-003. IIASA, Laxenburg.

[3] G. Marland (2008), Uncertainties in accounting for CO_2 from fossil fuels. *Journal of Industrial Ecology*, 12(2):136-139.

[4] G. Marland, K. Hamal, M. Jonas (2009), How uncertain are estimates of CO_2 emissions? *Journal of Industrial Ecology*, 13(1):4-7.

[5] Z. Nahorski, W. Jęda (2007), Processing national CO_2 inventory emission data and their total uncertainty estimates, *Water, Air, and Soil Pollution: Focus*, 7(4-5):513-527.