52/2011

Raport Badawczy
Research Report

RB/39/2011

Modeling uncertainty of GHG inventories – CDIAC data for several EU countries

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 Zakładu zgłaszający pracę:
Prof. zw. dr hab. inż. Zbigniew Nahorski

Warszawa 2011

MODELING UNCERTAINTY OF GHG INVENTORIES – CDIAC DATA FOR SEVERAL EU COUNTRIES

JOLANTA JARNICKA, ZBIGNIEW NAHORSKI
SYSTEMS RESEARCH INSTITUTE
POLISH ACADEMY OF SCIENCES
NEWELSKA 6, 01-447 WARSAW, POLAND
EMAIL: JOLANTA.JARNICKA@IBSPAN.WAW.PL
EMAIL: ZBIGNIEW.NAHORSKI@IBSPAN.WAW.PL

ABSTRACT. This paper is devoted to the study of uncertainty in the greenhouse gases (GHG) emission inventories. Analyzing the CDIAC data from the Oak Ridge National Laboratory, USA, collected every few years between 1986 and 2004, and their revisions, made in 1989 – 2004, we model changes in uncertainty structure, occurring in consecutive years. This is achieved by a parametric model, applied earlier to investigate data from the National Inventory Reports (NIR). Results obtained for several EU countries, are presented in the form of figures and tables. They are also compared with those, obtained from the NIR data.

Keywords: greenhouse gases, uncertainty, model

1. Introduction

This paper is a continuation of studies on the problem of reducing the uncertainty in the inventories on greenhouse gases (GHG) emissions. We started to deal with this issue in [3], describing the idea and presenting preliminary results, and then continued the discussion in [2], developing a suitable model, and applying it to data on GHG emissions.

Reports on GHG emissions are being prepared regularly by the cosignatories of the United Nations Framework Convention on Climate Change (UNFCCC) and its Kyoto Protocol. Each of these countries is obliged to prepare annual reports – the so-called National Inventory Reports, as well as to prepare revisions of past data (if needed).

The National Inventory Reports (NIR), however, are not the only database on GHG emissions. Independent studies are conducted also by research centers, e.g. Carbon Dioxide Information Analysis Center (CDIAC), in Oak Ridge National Laboratory, USA, where the data on GHG are gathered and processed (e.g. with respect to their source or type of a gas). There are a few such well-known databases – in addition to NIR and CDIAC one should mention also IEA (International Energy Agency), EIA (U.S. Energy Information Administration), and many others. They not only describe various, sometimes overlapping, data, are often collected in different years, but are also expressed in different units. The biggest problem, however, is that these data are calculated in different ways, with various uncertainty, making it difficult to compare and characterize them. Discussion on this problem can be found e.g. in [6], [5], [4], and in many others.

In this article we are interested in modeling the uncertainty in such inventories. Therefore, we present a parametric model, introduced and applied earlier in [2] for the NIR data. This time we will focus on the CDIAC data, conducting analysis for those EU countries that were considered in [2] (i.e. for Austria, Belgium, Denmark, Finland, United Kingdom, Ireland, and Sweden), and then comparing the results obtained.

Section 2 provides a brief description of the data from three databases: NIR, CDIAC, and IEA, along with an example (based on CO_2 emissions from the year 2004 for the EU-15). In Section 3 we give description of the model introduced in [2], together with the method of estimation of the parameters. The CDIAC data for several EU countries, are analyzed in Section 4. Finally, Section 5 provides a summary of the results obtained.

2. Data on CO_2 emissions

Data on GHG emissions from different databases vary considerably – they refer to different sources of emissions, are often collected and revised in different years, may also be expressed in different units).

To illustrate these differences we present below examples, showing data on CO_2 emissions for countries of the EU-15, from the year 2004 (including revisions of past data made in 2004). At the beginning, in Figure 1 we present data from the year 2004, for each of these three data sets. This shows the huge differences in scale of values considered.

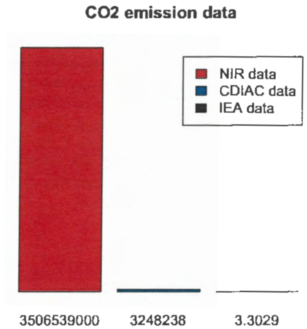


FIGURE 1. Data on CO_2 emissions in 2004, in [Mt], EU-15.

Now we consider data for countries of the EU-15, derived from the NIR (Figure 2 and 5), the CDIAC data (Figure 3 and 4), and IEA data (Figure 6), from the year 2004, and all the revisions of past data. Figure 2 presents data on total CO_2 emissions (excluding LULUCF) in [Gg], for the EU-15 countries. The data refer to emissions in the year 2004, and the annual emissions in 1986 – 2003, recalculated in 2004.

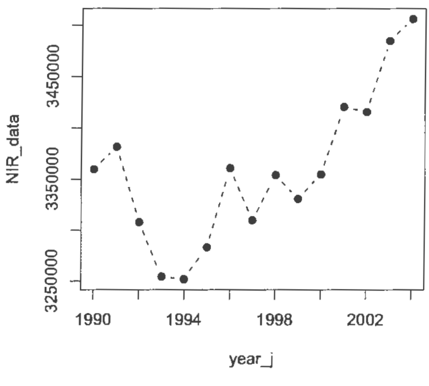


FIGURE 2. National Inventory Reports data on CO_2 emissions in 2004, in [Gg], EU-15.

The CDIAC data presented in Figure 3 express CO_2 emissions, in mass units of carbon. To convert the data to CO_2 mass units, we multiply each value by the ratio of the molecular mass of carbon dioxide to the atomic mass of carbon ($\frac{44}{12}$ or 3.667). The data converted, expressed in metric tons (Mt)[1] are presented in Figure 4. Both figures (Figure 3 and 4) show the data on CO_2 fossil-fuel emissions, for the countries of the EU-15, from the year 2004 and revisions of data from the years 1986, 1989, 1990, 1992, 1998, 1999, 2000, 2002, and 2003.

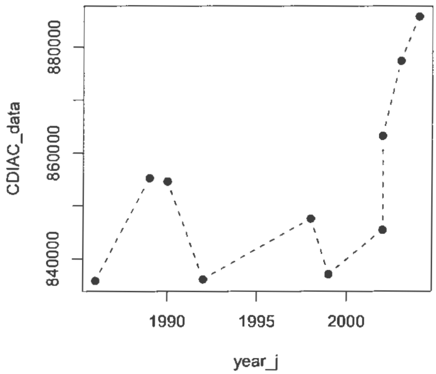


FIGURE 3. CDIAC data on CO_2 fossil-fuel emissions in 2004, in mass units of carbon, EU-15.

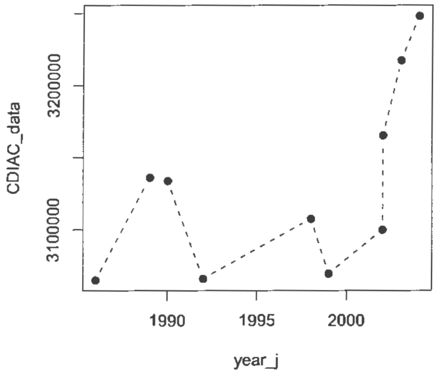


FIGURE 4. CDIAC data on CO_2 fossil-fuel emission in 2004, in [Mt], EU-15.

Comparing Figures 2 and 4, you may notice some similarity in trend. Visible differences result from different types of emissions, from the fact that the data were collected in different years,

[1] 1 Mt = 0.001 Gg

but also, slightly, from different scale of the data, and different units. The first two reasons are beyond our control, but for a better comparison, we show the NIR data in Figure 5 in [Mt]. These units will also be used later in this paper.

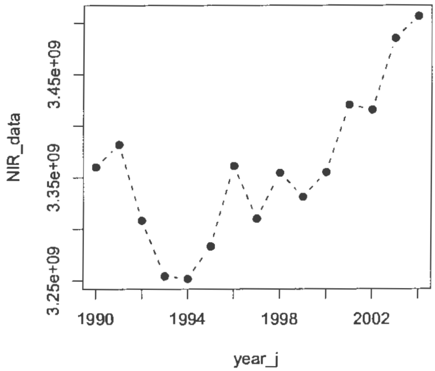


FIGURE 5. NIR data on CO_2 emission in 2004, in [Mt], EU-15.

Finally, Figure 6 presents the IEA data (energy related). These data concern CO_2 emissions (expressed in metric tons [Mt]) from fuel combustion, for the countries of the EU-15, in 2004. The revisions concerned were made in the years 1990, 1995, 2000, 2001, 2002, 2003, and 2004.

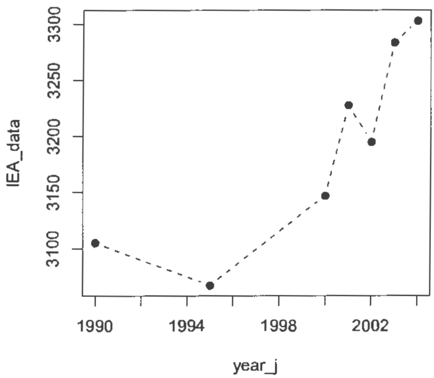


FIGURE 6. IEA data on CO_2 emissions from fuel combustion in 2004, in [Mt], EU-15.

For the convenience of the reader, the data from Figures 2, 5, and 6 in [Mt] are also presented together in one graph (Figure 7). Due to the large differences in the scale, we have prepared two Figures. The one on the left shows the NIR data (red), CDIAC data (blue) and IEA data (black). Because the difference between the CDIAC and IEA data is almost invisible, in the figure on the right are presented only the data from these two sets.

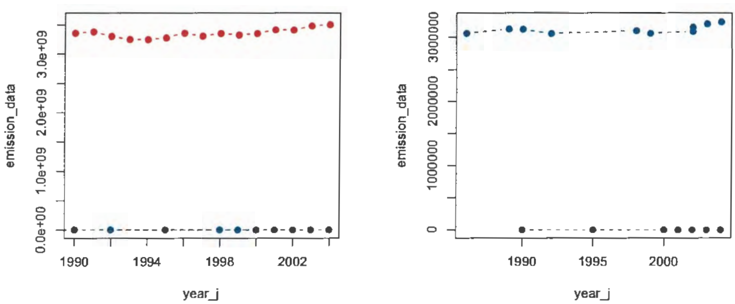


FIGURE 7. Data on CO_2 emissions in 2004, in [Mt], EU-15, (NIR data – red, CDIAC – blue, IEA – black).

Wider discussion of data conversion, and the ways to compare them, can be found, for example in [1].

3. Model

In this section we present the model and the way of interpreting the data (see [2] for details). Let $E^n_{y_j,i}$ – denote the inventory data for the country i, in the year n revised in the year y_j, $y_j < Y$, where Y – is the last year, when the revision is made.

We use the fact that, each revision data, for a given country i,

$$E^n_{y_1,i}, E^n_{y_2,i}, \ldots, E^n_{y_j,i}, \ldots, E^n_{Y,i},$$

forms a realization of a stochastic process. These stochastic processes for a fixed country form a bunch of stochastic processes.

For a given country i, we model any revision 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.

$$E^n_{Y,i} = D^n_{Y,i} + S^n_{Y,i}, \qquad S^n_{Y,i} \sim \mathcal{N}(0, \sigma_{Y,i}),$$

where E – stands for the emission inventory, D – for its deterministic fraction, S – for the stochastic fraction, and n – is the year, for which the revised data were recalculated. Similarly, if $y_j < Y$,

$$E^n_{y_j,i} = D^n_{Y,i} + S^n_{y_j,i}, \qquad \text{with} \quad S^n_{y_j,i} \sim \mathcal{N}\left(m^n_{y_j,i}, \sigma^n_{y_j,i}\right),$$

where the mean values $m^n_{y_j,i}$ and the standard deviations $\sigma^n_{y_j,i}$ are of the form

$$m^n_{y_j,i} = a_i(Y - y_j), \qquad \sigma^n_{y_j,i} = \sigma_{Y,i} + b_i\, f(Y - y_j), \quad b_i \neq 0,$$

and f is a given function, such that

$$f(Y - y_j) > -\frac{\sigma_{Y,i}}{b_i}.$$

The parameters a_i and b_i, for a country i, associated with the stochastic fraction $S^n_{y_j,i}$, can be estimated from the data together with $\sigma_{Y,i}$. Parameter a_i describes a shift in the accuracy of the inventory gathering, and b_i – a shift of the precision level. They both depend on the difference between the revision year y_j, and the most recent revision year Y, due to the learning. To find the deterministic fraction $D^n_{Y,i}$, the smoothing splines can be used, as presented in [7]. This approach, when applied to the most recently revised data $E^n_{Y,i}$ will give not only the estimate of the deterministic fraction, but also an estimate of the variance $\sigma^2_{Y,i}$.

3.1. **Algorithm for a fixed country** i. .

Fix i and consider all the inventory data $E^n_{y_j,i}$ in the year n for $n = 1, \dots, N_j$, revised in the year y_j, $j = 1, \dots, J$. For a fixed country i, the procedure can be describe as follows.

1. For the most recently revised inventory data E^n_Y **calculate the smoothing spline** Sp_Y **and estimate the variance** σ^2_Y of the stochastic fraction S^n_Y.
2. **Subtract the spline** Sp_Y, **built on the data from the year** Y, **from all earlier revisions** $E^n_{y_j}$, $y_j < Y$, calculating differences
$$v^n_j = E^n_{y_j} - \mathrm{Sp}_Y, \quad n = 1, \dots, N_j, \; j = 1, \dots, J.$$
For some years the difference v does not exist, due to lack of revised inventories in this year. These years are skipped from the sequence of N_j data.
We consider the following **model**:

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

where

(2)
$$m_j = a(Y - y_j), \quad \sigma_j = \sigma_Y - b\,(Y - y_j)^{c+1}, \quad b \neq 0.$$

Assume also that differences (1) are independent.

3. **Estimate parameters** a, b, **and** c (and hence m_j and σ_j, $j = 1, \dots, J$ in (2)) in the following three-step procedure.

(3.1). Estimate parameters α_j and β_j, $j = 1, \dots, J$ in the model

(3)
$$m_j = \alpha_j (Y - y_j),$$

(4)
$$\sigma_j = \widehat{\sigma}_Y + \beta_j\,(Y - y_j), \quad \beta_j \neq 0,$$

using Maximum Likelihood estimators

(5)
$$\widehat{\alpha}_j = \frac{1}{N_j (Y - y_j)} \sum_{n=1}^{N_j} v^n_j,$$

and

(6)
$$\widehat{\beta}_j = \Big(\sqrt{\frac{1}{N_j} \sum_{n=1}^{N_j} \left(v^n_j - \bar{v}_j\right)^2} - \widehat{\sigma}_Y\Big) / (Y - y_j),$$

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

(3.2). Use $\widehat{\alpha}_j$, $j = 1, \dots, J$, obtained in (3.1), to estimate parameter a in the first order autoregressive model

(7)
$$\alpha_{j-1} = \frac{1}{\widetilde{a}} \alpha_j + \varepsilon_j, \quad |\widetilde{a}| < 1, \; \widetilde{a} \neq 0,$$

where

$$\alpha_{J+1} := 0,$$

and ε_j are independent and $\varepsilon_j \sim N(0, \sigma)$. Estimator of the parameter a is then given by

$$\widehat{a} = \frac{1}{\widetilde{a}}$$

(3.3). Use the sequence $\widehat{\beta}_j$, $j = 1, \ldots, J$, obtained in (3.1), to estimate parameters b and c in the regression model

$$\beta_j := -b\,(Y - y_j)^c, \quad j = 1, \ldots, J, \quad \text{where} \quad b < 0. \tag{8}$$

Since $\beta_j > 0$, $j = 1, \ldots, J$, nonlinear model (8) can be converted into a linear one of the form

$$\ln \beta_j = \ln(-b) + c \ln(Y - y_j), \tag{9}$$

and the parameters $\widetilde{b} := \ln(-b)$ and c can now be estimated using the Least Squares method.

3.2. **CDIAC data for the EU-15.** To illustrate how it works in practice, we apply the model (1) – (2), and the procedure described in Subsection 3.1 to the CDIAC data (in [Mt]) from the year $Y = 2004$, and from the years 1989, 1990, 1992, 1998, 1999, 2000, 2002, 2003, including all the revisions. We start with building the smoothing spline Sp_{2004}, using the method described in [7]. We get $\widehat{\sigma}_{2004} = 28701$, and the estimation result (in blue) is presented in Figure 8, together with data from the year 2004 (in black).

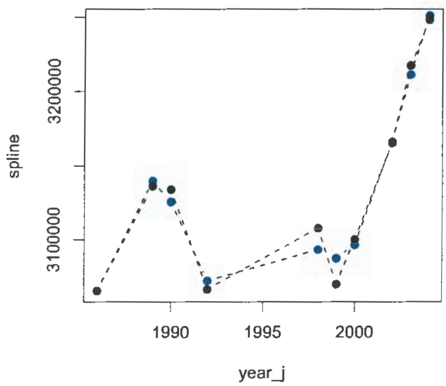


FIGURE 8. Smoothing spline Sp_Y for $Y = 2004$, CDIAC data for EU-15, $\widehat{\sigma}_Y = 28701$

The main problem that differs this analysis from those, carried out previously in [2], for the NIR data, is a smaller number of observations, but also the fact that in some years, data revisions were not performed. This means that some of the differences (1) do not exist and are ignored in the calculation (it can often be seen in figures).

We use the procedure, described in Subsection 3.1. First, we need to find sequences $\widehat{\alpha}_j$ and $\widehat{\beta}_j$, $j = 1, \ldots, J$ in (3) – (4). Using the formulas (5) and (6), we get $\widehat{\alpha}_j$ of the form

$$3279.46, 3615.51, 2006.73, 2230.63, 4180.43, 4108.21, 9222.48, 18661.79$$

and $\widehat{\beta}_j$

$$-642.43, -995.22, -1737.52, -2551.20, -3478.36, -5174.52, -9824.09, -21924.90.$$

The results are presented in Figure 9, below. It can be noticed that the values $\widehat{\alpha}_j$ are rather scattered, although one can also notice an increasing trend. On the other hand, the values $\widehat{\beta}_j$ are strictly decreasing (and also negative, what enables further analysis).

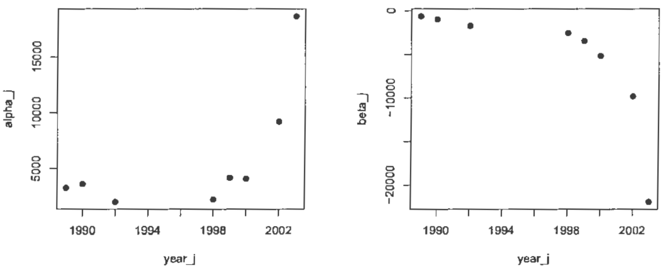


FIGURE 9. Estimates of parameters α_j and β_j, CDIAC data, EU-15.

Having obtained the results of the initial estimation, we can fit the parameters a, b, and c in the model (2). First, we estimate the parameter a in the first order autoregressive model (7). We get $\widehat{a} = \frac{1}{\tilde{a}} = 0.585$, where $\sigma^2 = 6487811$. Then, with an estimate of a, we can determine the mean values m_j, $j = 1, \ldots, J$ in the model (2): $8.78, 8.19, 7.02, 3.51, 2.93, 2.34, 1.17, 0.59$. The values m_j are also depicted in Figure 10.

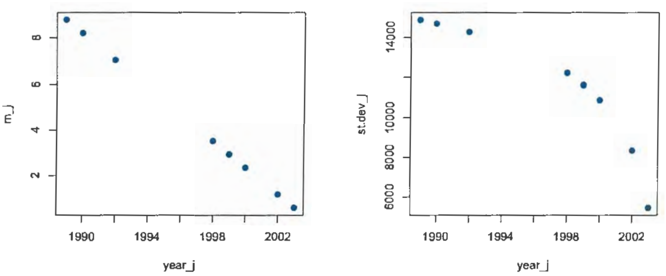


FIGURE 10. Estimated values of m_j and σ_j, CDIAC data, EU-15.

Let us now estimate parameters b and c, and hence standard deviations σ_j, $j = 1, \ldots, J$ in the model (2). Consider the regression function of the form (8). Since $\widehat{\beta}_j < 0$, $j = 1, \ldots, J$, we consider the regression model (9). We get the following results.

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 10.05438    0.16541   60.78 1.33e-09 ***
ly          -1.19199    0.08737  -13.64 9.63e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2231 on 6 degrees of freedom
Multiple R-squared: 0.9688,     Adjusted R-squared: 0.9636
F-statistic: 186.1 on 1 and 6 DF,  p-value: 9.63e-06
```

The parameter $\tilde{b} = \ln(-b) = 10.05$, so $\widehat{b} = 23257.43$. The estimate of c is equal -1.19. Using the model (2) and taking $\widehat{\sigma}_Y = 28701$ calculated before, when building the smoothing spline Sp_{2004}, we get estimates for standard deviations σ_j:

$$14873.02, 14688.64, 14267.73, 12213.30, 11625.94, 10878.52, 8341.6, \text{ and } 5443.73.$$

The values obtained are presented in Figure 10. In conclusion, we gather the results obtained for EU-15, in Table 1.

j	1989	1990	1992	1998	1999	2000	2002	2003
$\widehat{\alpha}_j$	3279.46	3615.51	2006.73	2230.63	4180.43	4108.21	9222.48	18661.79
$\widehat{\beta}_j$	-642.43	-995.22	-1737.52	-2551.20	-3478.36	-5174.52	-9824.09	-21924.90
m_j	8.78	8.19	7.02	3.51	2.93	2.34	1.17	0.59
σ_j	14873.02	14688.64	14267.73	12213.30	11625.94	10878.52	8341.6	5443.73

TABLE 1. Model parameters, CDIAC data for EU-15, $\widehat{a} = 0.585$, $\widehat{b} = 23257.43$, $\widehat{c} = -1.19$.

4. Analysis of the CDIAC data for a few EU countries

We apply the model, described in Section 3, to the CDIAC data for Austria, Belgium, Denmark, Finland, United Kingdom, Ireland, and Sweden. We consider the data from the year $Y = 2004$, and from the years: 1989, 1990, 1992, 1998, 1999, 2000, 2002, and 2003, including all the revisions made. The data refer to CO_2 emissions from fossil fuels, and are expressed in mass units of carbon. To convert them into the mass units of carbon dioxide, we multiply each value by $\frac{44}{12}$. The data converted are now expressed in metric tons (Mt) of CO_2.

4.1. **Austria.** We start the analysis with Austria. First, we build a smoothing spline Sp_Y, for $Y = 2004$. We get $\widehat{\sigma}_Y = 2342.01$. The result obtained is presented in Figure 11, where the smoothing spline Sp_{2004} is depicted in red.

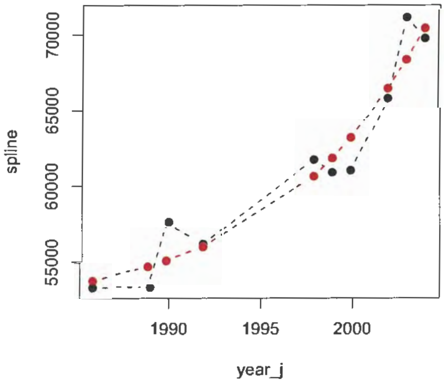


FIGURE 11. Smoothing spline Sp_Y for $Y = 2004$, Austria, CDIAC data, $\widehat{\sigma}_Y = 2342.01$

Then we estimate parameters α_j, and β_j, $j = 1, \dots, J$, according to formulas (5) and (6). The results are presented in Figure 12. Next step is the estimation of the parameters a, b, and c in the model (2), and hence the sequences m_j, and σ_j, $j = 1, \dots, J$. The results can be seen in Figure 13. All the results obtained are also gathered in Tables 2a nad 2b.

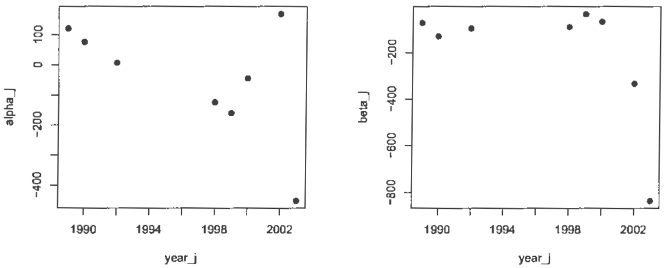


FIGURE 12. Estimates of parameters α_j and β_j, Austria, CDIAC data.

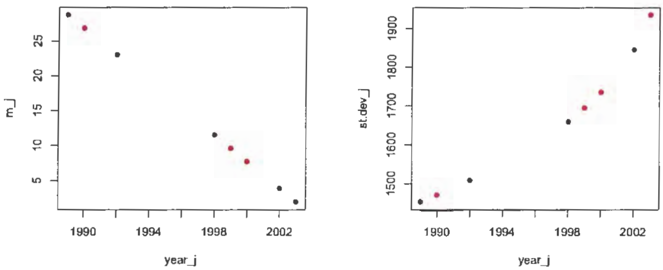


FIGURE 13. Estimated values of m_j and σ_j, Austria, CDIAC data.

j	1989	1990	1992	1998	1999	2000	2002	2003	mean	std
$\widehat{\alpha}_j$	121.7	76.6	7.4	-123.8	-160.4	-44.7	169.9	-451.45	-50.59	185.73
$\widehat{\beta}_j$	-73.61	-130.61	-96.75	-91.36	-35.81	-68.26	-334.08	-835.76	-208.28	252.29
m_j	28.77	26.85	23.01	11.51	9.59	7.67	3.84	1.92	$\hat{a} = 1.92$	
σ_j	1452.8	1470.3	1508.3	1659.44	1694.4	1734.8	1844.9	1935	$\hat{b} = 406.9$, $\hat{c} = -0.7$	

TABLE 2a. Estimates of parameters in the model (2) Austria, CDIAC data.

Parameter	Estimate	Model fit
$\hat{a}$	1.92	$\sigma^2 = 36636$
$\hat{b}$	406.9	St.error 0.5782, t-test: p-value=0.0000065, $R^2 = 0.49$
$\hat{c}$	-0.7	St.error = 0.3054, t-test: p-value= 0.00587

TABLE 2b. Estimates of parameters a, b, and c, Austria, CDIAC data.

Similar analysis is conducted for other EU countries mentioned – Belgium, Denmark, Finland, UK, Ireland, and Sweden. The results are presented in Figures 14 – 31 and Tables 3a, 3b – 8a, and 8b.

4.2. **Belgium.**

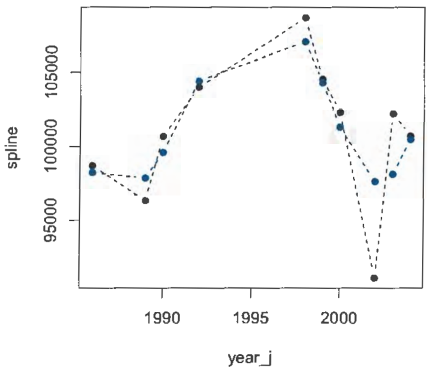


FIGURE 14. Smoothing spline Sp_Y for $Y = 2004$, Belgium, CDIAC data, $\widehat{\sigma}_Y = 5340.9$

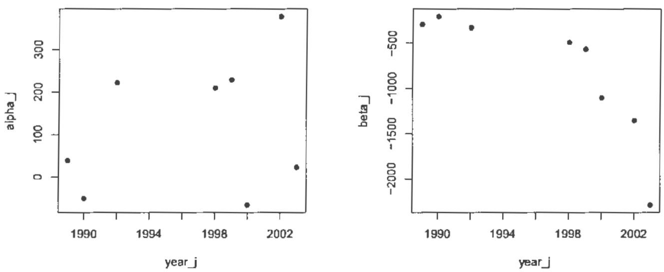


FIGURE 15. Estimates of parameters α_j and β_j, Belgium, CDIAC data.

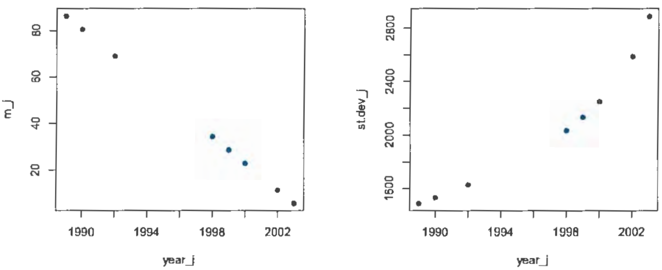


FIGURE 16. Estimated values of m_j and σ_j, Belgium, CDIAC data.

j	1989	1990	1992	1998	1999	2000	2002	2003	mean	std
$\widehat{\alpha}_j$	139.7	-49.86	222.6	210.3	228.9	-65.21	375.6	23.8	123.2	147.7
$\widehat{\beta}_j$	-296.9	-214.6	-330.9	-488.9	-563.9	-1102.6	-1351.8	-2288.3	-829.7	668.2
m_j	86.26	80.5	69	34.5	28.75	23	11.5	5.75	$\hat{a} = 5.75$	
σ_j	1489.5	1533.6	1630.4	2036	2135.2	2252.5	2590.1	2890.9	$\hat{b} = 2449.9$, $\hat{c} = -0.8$	

TABLE 3a. Estimates of parameters in the model (2), Belgium, CDIAC data.

Parameter	Estimate	Model fit
$\hat{a}$	5.752	$\sigma^2 = 40801$
$\hat{b}$	2449.9	St.error=0.149, t-test: p-value=0.0000000033, $R^2 = 0.95$
$\hat{c}$	-0.83	St.error = 0.079, t-test: p-value=0.0000043

TABLE 3b. Estimates of parameters a, b, and c, Belgium, CDIAC data.

4.3. Denmark.

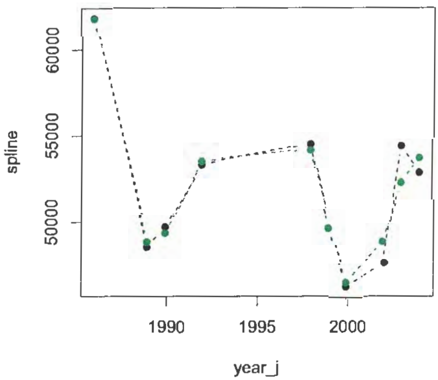


FIGURE 17. Smoothing spline Sp_Y for $Y = 2004$, Denmark, CDIAC data, $\widehat{\sigma}_Y = 1672.6$

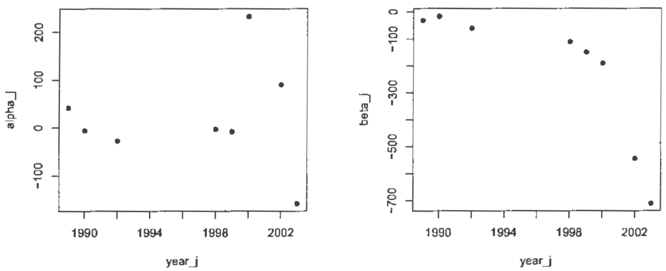


FIGURE 18. Estimates of parameters α_j and β_j, Denmark, CDIAC data.

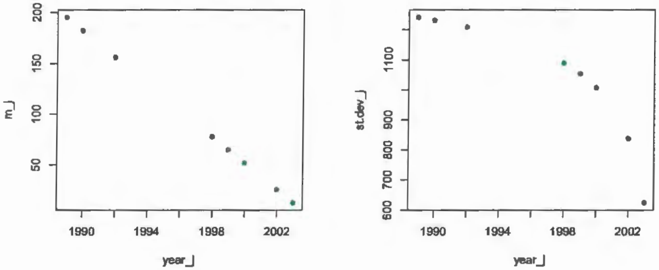


FIGURE 19. Estimated values of m_j and σ_j, Denmark, CDIAC data.

j	1989	1990	1992	1998	1999	2000	2002	2003	mean	std
$\widehat{\alpha}_j$	41.63	-5.76	-26.95	-3.04	-7.8	232.7	89.9	-157.5	20.4	103.8
$\widehat{\beta}_j$	-29.78	-14.6	-59.3	-109.9	-148.3	-189.3	-543.4	-709.4	-225.5	241.4
m_j	195.3	182.3	156.3	78.1	65.1	52.1	26	13.02	$\hat{a} = 13.02$	
σ_j	1241.1	1231.3	1208.4	1090	1054.2	1007.2	837.6	624.6	$\hat{b} = 1047.9$, $\hat{c} = -1.33$	

TABLE 4a. Estimates of parameters in the model (2), Denmark, CDIAC data.

Parameter	Estimate	Model fit
$\hat{a}$	13.02	$\sigma^2 = 12500$
$\hat{b}$	1047.9	St.error=0.312, t-test: p-value=0.000000053, $R^2 = 0.92$
$\hat{c}$	-1.33	St.error = 0.165, t-test: p-value=0.000195

TABLE 4b. Estimates of parameters a, b, and c, Denmark, CDIAC data.

4.4. **Finland.**

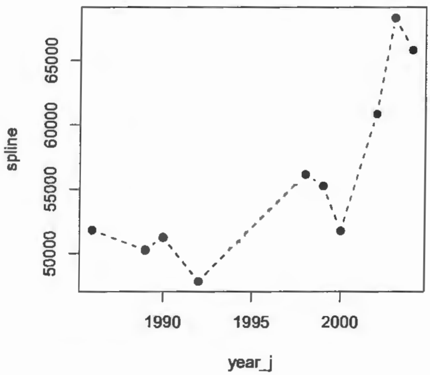


FIGURE 20. Smoothing spline Sp_Y for $Y = 2004$, Finland, CDIAC data, $\widehat{\sigma}_Y = 2271.8$

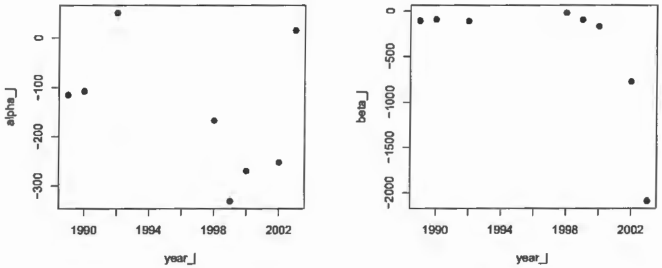


FIGURE 21. Estimates of parameters α_j and β_j, Finland, CDIAC data.

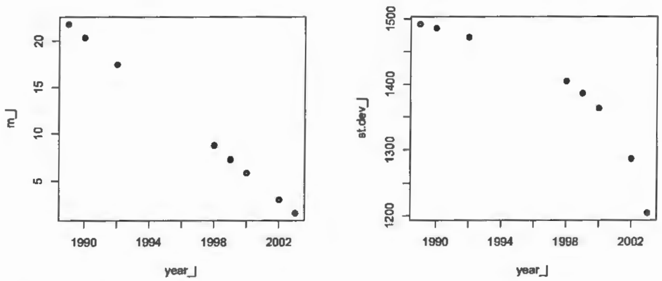


FIGURE 22. Estimated values of m_j and σ_j, Finland, CDIAC data.

j	1989	1990	1992	1998	1999	2000	2002	2003	mean	std
$\widehat{\alpha}_j$	-114.16	-106.86	50.95	-167.09	-330.35	-269.48	-252.08	15.19	-146.8	126.6
$\widehat{\beta}_j$	-107.94	-93.68	-111.77	-20.13	-96.39	-168.41	-771.89	-2086.9	-432.2	664.1
m_j	21.81	20.35	17.44	8.72	7.27	5.81	2.91	1.46	$\hat{a} = 1.46$	
σ_j	1492.33	1486.05	1471.85	1404.75	1386.19	1362.92	1286.69	1204.06	$\hat{b} = 1067.7$, $\hat{c} = -1.12$	

TABLE 5a. Estimates of parameters in the model (2), Finland, CDIAC data.

Parameter	Estimate	Model fit
$\hat{a}$	1.46	$\sigma^2 = 20765$
$\hat{b}$	1067.7	St.error=0.745, t-test: p-value=0.0000083, $R^2 = 0.52$
$\hat{c}$	-1.12	St.error = 0.393, t-test: p-value=0.00297

TABLE 5b. Estimates of parameters a, b, and c, Finland, CDIAC data.

4.5. United Kingdom.

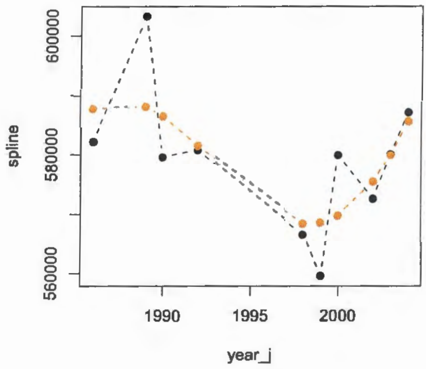


FIGURE 23. Smoothing spline Sp_Y for $Y = 2004$, UK, CDIAC data, $\widehat{\sigma}_Y = 11598.8$

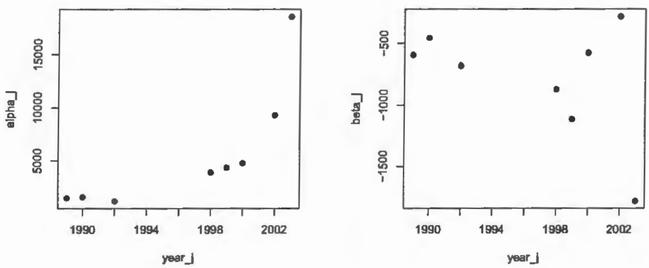


FIGURE 24. Estimates of parameters α_j and β_j, UK, CDIAC data.

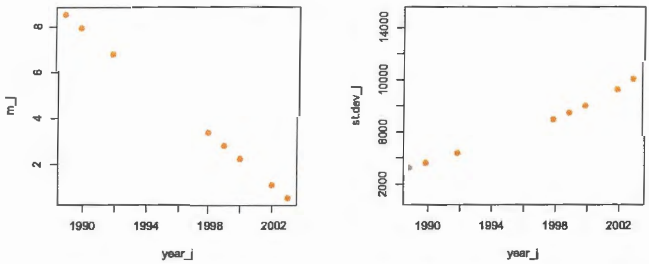


FIGURE 25. Estimated values of m_j and σ_j, UK, CDIAC data.

j	1989	1990	1992	1998	1999	2000	2002	2003	mean	std
$\widehat{\alpha}_j$	1468.2	1566.8	1165.3	3918.4	4361.7	4792.8	9311.4	18571	5644.4	5477.5
$\widehat{\beta}_j$	-593.1	-456.3	-677.6	-869.9	-1114.1	-576.4	-280.4	-1783.9	-793.9	442.6
m_j	8.52	7.95	6.81	3.41	2.84	2.27	1.14	0.57	$\hat{a} = 0.57$	
σ_j	3372.4	3734.2	4486	7071.6	7578.9	8123	9386.5	10191	$\hat{b} = 1408.1$, $\hat{c} = -0.348$	

TABLE 6a. Estimates of parameters in the model (2), UK, CDIAC data.

Parameter	Estimate	Model fit
$\hat{a}$	0.57	$\sigma^2 = 3904397$
$\hat{b}$	1408.1	St.error=0.216, t-test: p-value=0.000000081, $R^2 = 0.54$
$\hat{c}$	-0.348	St.error = 0.113, t-test: p-value=0.0022

TABLE 6b. Estimates of parameters a, b, and c, UK, CDIAC data.

4.6. **Ireland.**

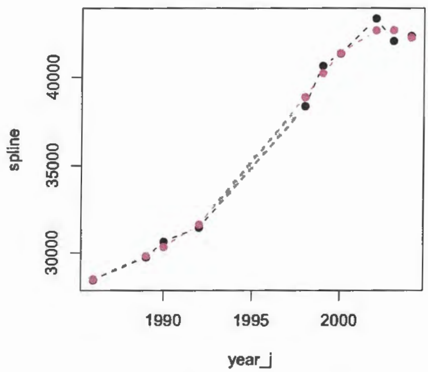


FIGURE 26. Smoothing spline Sp_Y for $Y = 2004$, Ireland, CDIAC data, $\widehat{\sigma}_Y = 773.1$

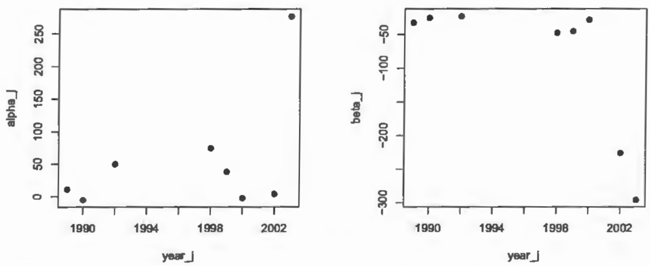


FIGURE 27. Estimates of parameters α_j and β_j, Ireland, CDIAC data.

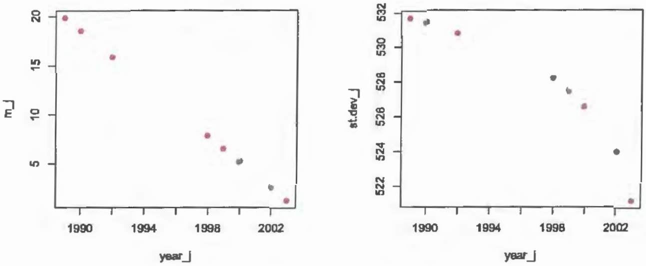


FIGURE 28. Estimated values of m_j and σ_j, Ireland, CDIAC data.

j	1989	1990	1992	1998	1999	2000	2002	2003	mean	std
$\widehat{\alpha}_j$	11.10	-5.38	50.05	74.52	38.23	-2.21	4.27	276.62	55.9	87.47
$\widehat{\beta}_j$	-32.92	-25.85	-23.39	-47.66	-45.36	-28.61	-226.06	-295.89	-90.72	100.2
m_j	19.83	18.50	15.86	7.93	6.61	5.29	2.64	1.32	$\hat{a} = 1.32$	
σ_j	531.7	531.5	530.9	528.2	527.5	526.7	523.9	521.2	$\hat{b} = 251.9$, $\hat{c} = -0.92$	

TABLE 7a. Estimates of parameters in the model (2), Ireland, CDIAC data.

Parameter	Estimate	Model fit
$\hat{a}$	1.32	$\sigma^2 = 11506$
$\hat{b}$	251.9	St.error=0.359, t-test: p-value=0.00000047, $R^2 = 0.76$
$\hat{c}$	-0.92	St.error = 0.1898, t-test: p-value=0.0029

TABLE 7b. Estimates of parameters a, b, and c, Ireland, CDIAC data.

4.7. Sweden.

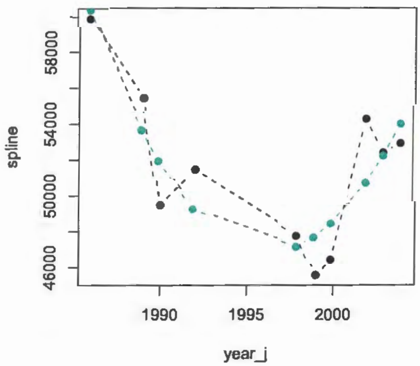
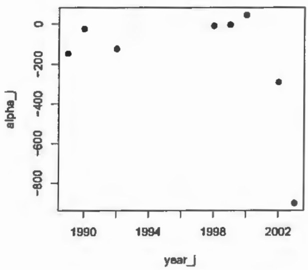


FIGURE 29. Smoothing spline Sp_Y for $Y = 2004$, Sweden, CDIAC data, $\widehat{\sigma}_Y = 2757.6$



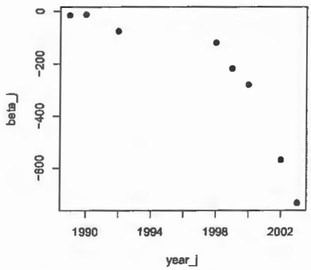
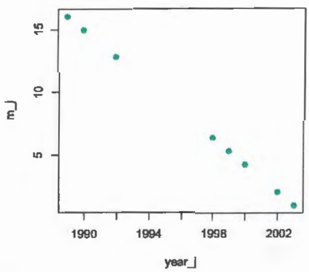


FIGURE 30. Estimates of parameters α_j and β_j, Sweden, CDIAC data.



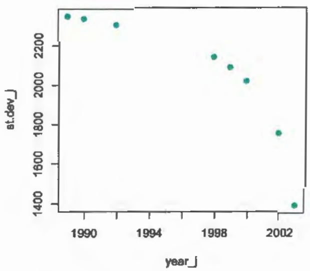


FIGURE 31. Estimated values of m_j and σ_j, Sweden, CDIAC data.

j	1989	1990	1992	1998	1999	2000	2002	2003	mean	std
$\widehat{\alpha}_j$	-147.1	-23.55	-123.3	-9.54	-3.88	44.59	-290.8	-901.9	-181.9	289.9
$\widehat{\beta}_j$	-15.6	-13.02	-76.4	-120.6	-218.9	-280.4	-568.9	-734.3	-253.6	249.1
m_j	16.04	14.97	12.83	6.42	5.38	4.28	2.14	1.07	$\hat{a} = 0.94$	
σ_j	531.7	531.5	530.9	528.2	527.5	526.7	523.9	521.2	$\hat{b} = 1359.4$, $\hat{c} = -1.45$	

TABLE 8a. Estimates of parameters in the model (2), Sweden, CDIAC data.

Parameter	Estimate	Model fit
$\hat{a}$	0.94	$\sigma^2 = 54940$
$\hat{b}$	1359.4	St.error=0.465, t-test: p-value=0.00000045, $R^2 = 0.83$
$\hat{c}$	-1.45	St.error = 0.246, t-test: p-value=0.00105

TABLE 8b. Estimates of parameters a, b, and c, Sweden, CDIAC data.

5. Conclusions

This report presented a model, describing the learning process in the CDIAC data on CO_2 emissions. This model has been introduced in [2], and applied to the NIR data for a few EU countries (Austria, Belgium, Denmark, Finland, United Kingdom, Ireland, and Sweden). In this paper it has been recalled in Section 3, together with an example of its application to the CDIAC data for countries from the EU-15 (detailed description of the analysis conducted, as well as Figures 8 – 10, and Table 1 are presented in Section 3.2).

Results of analysis conducted for the CDIAC data, from the year 2004, with revisions of past data carried out in the years 1989, 1990, 1992, 1998, 1999, 2000, 2002, 2003, and 2004, are presented in Section 4, in the form of figures (Figures 11 – 31) and tables (Tables 2a, 2b – 8a, and 8b). The data were considered for EU countries, studied earlier in [2] in the case of the NIR data. The results obtained are summarized and presented together in Figures 32 and 33.

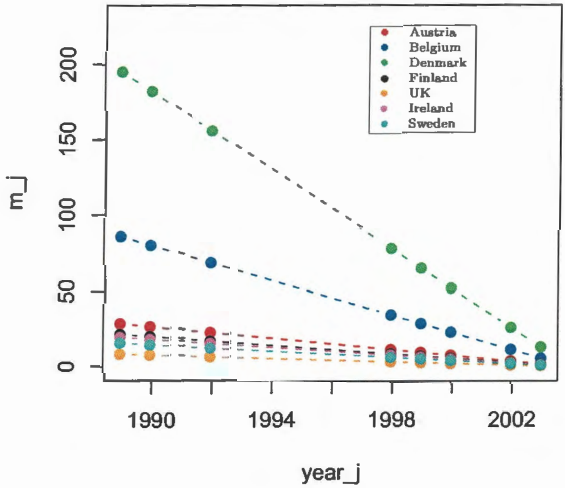


FIGURE 32. Mean values m_j n the model (2), CDIAC data for EU countries.

Let us take a closer look at Figure 32, comparing the results with those previously obtained in [2], for the NIR data. For a better comparison, Figure 34 recalls the mean values and standard deviations, calculated for the NIR data (in [Gg]).

It may be noted that the mean values differ slightly. Although the changes seem to be greater than it was in the case of the NIR data analysis, it must be remembered that we use a different units – [Mt] instead of [Gg] (where 1 Mt = 0.001 Gg). This means that analyzing the CDIAC data, we get smaller values of m_j. This fits our intuition, which suggests that they should not only converge to zero, but also be close to zero (how much, it depends on standard deviations). Observe that also this time the mean values m_j, calculated for Denmark are larger and protrude slightly from the others. On the other hand, m_j calculated for Sweden have values similar to the others obtained.

Figure 33 shows the standard deviations σ_j, calculated for all the countries considered.

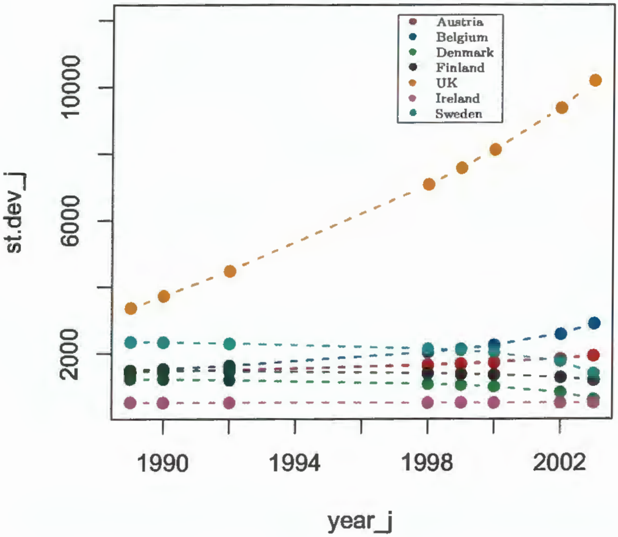


FIGURE 33. Mean values σ_j n the model (2), CDIAC data for EU countries.

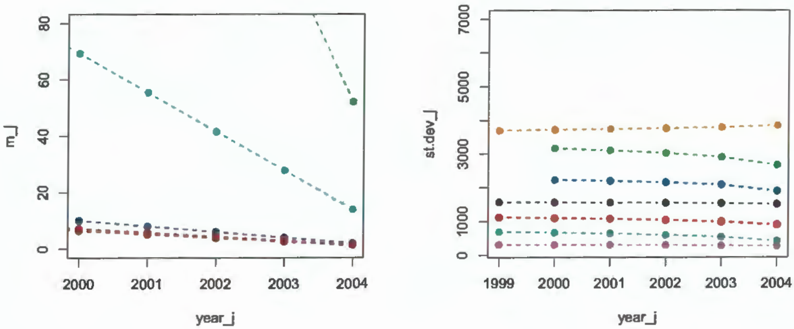


FIGURE 34. Estimated values of m_j and σ_j, NIR data in [Gg] for EU countries.

In the case of standard deviations σ_j, can be seen that, except for UK, the values obtained are quite similar. Comparing the results with those calculated for the NIR data (Figure 34, on the right), we can see that in most cases there is no decreasing trend (it was already noticeable in the figures in Section 4). The size of the values obtained may (except for UK) indicate slightly better results (in particular, taking into account the units). More reliable way to compare the results, will however be the comparison of relative values $\frac{\widehat{\sigma}_j}{\mathrm{Sp}_j}$. The results are shown in Figures 35 (for the CDIAC data) and 36 (for the NIR data, from [2]).

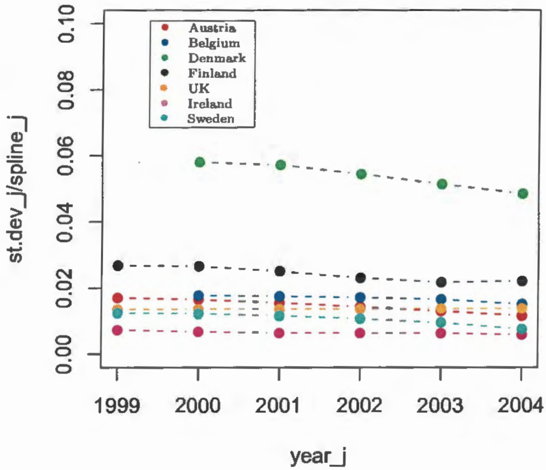


FIGURE 35. Relative values $\frac{\widehat{\sigma}_j}{\mathrm{Sp}_j}$, NIR data for EU countries.

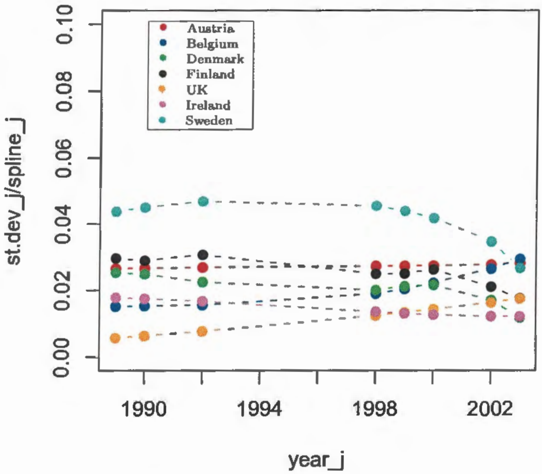


FIGURE 36. Relative values $\frac{\widehat{\sigma}_j}{\mathrm{Sp}_j}$, CDIAC data for EU countries.

Comparing the relative values $\frac{\widehat{\sigma}_j}{\mathrm{Sp}_j}$, $j = 1, \ldots, J$, for both types of data, we can see that in both cases we obtain numbers from a similar range. Important is that, dividing by Sp_j made it possible to get the same scale. You may notice some differences – results in Figure 35 give the impression of parallel lines, while Figure 36 shows some disruption in trend (e.g. a sequence $\frac{\widehat{\sigma}_j}{\mathrm{Sp}_j}$, $j = 1, \ldots, J$ in the case of Sweden is not monotonic).

In summary, it can be concluded that the model used, proved to work well in practice, for the different data types. However, it is necessary to test larger data sets (as they become available), and other databases (e.g., the previously mentioned IEA).

References

[1] J. Macknick (2011) Energy and CO_2 emission data uncertainties. *Carbon Management*, 13(2):189-205.

[2] Z. Nahorski, J. Jarnicka (2011), Modeling Uncertainty Structure of Greenhouse Gases Inventories – Revisited. Analysis of National Inventory Reports data for EU countries, RB, 2011, IBS PAN, Warszawa.

[3] Z. Nahorski, J. Jarnicka (2010), Modeling Uncertainty Structure of Greenhouse Gases Inventories, RB/11/2010, IBS PAN, Warszawa.

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

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

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

[7] 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.