Raport Badawczy

Research Report

RB/8/2005

Processing national CO_2 inventory emission data and their total uncertainty estimates

Z. Nahorski, W. Jęda

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) 8373578

fax: (+48) (22) 8372772

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

Warszawa 2005

INSTYTUT BADAŃ SYSTEMOWYCH PAN

Zbigniew Nahorski, Waldemar Jęda

Processing national CO_2 inventory emission data and their total uncertainty estimates

Warszawa 2005

Spis treści

Modelling CO_2 emissions and uncertainty variance estimation *

Zbigniew Nahorski and Waldemar Jęda
Systems Research Institute, Polish Academy of Sciences, Newelska 6, PL-01-447 Warsaw, Poland e-mail: Zbigniew.Nahorski@ibspan.waw.pl

5 grudnia 2005

Streszczenie. Uncertainty of reported greenhouse gases emissions obtained by aggregation of the partial emissions from all sources and estimated so far for several countries is very high. Independent calculation of the estimates could confirm or question the values obtained up to now. One of the aim of this paper is to propose statistical signal processing methods for doing it. They use the yearly reported observations while assuming temporal smoothness of the emission curve. The considered methods are: a spline function smoothing procedure, a time-varying parameter model, and the geometric Brownian motion model. They are verified on historical observations of the CO_2 emissions from combustion of the fossil fuels. The obtained estimates of variances agree in the range with those obtained from national inventories. As an additional result, some regularities in observed curves were noticed.

Keywords: modelling CO_2 emissions, nonparametric methods, parametric methods, geometric Brownian motion, estimation of variance.

1. Introduction

The Kyoto Protocol contains obligations to decrease the emission of the greenhouse gases of 5.2% below 1990 level by the first period up to 2008-2012. The greenhouse gas emission inventories of each country are monitored by the secretariat of the United Nations Framework Convention on Climate Change. However, the uncertainty ranges of the national accounts are big and in most times exceed, sometimes very considerably, the emission reductions agreed upon in the Annex I to the Protocol.

The parties who signed the Annex I have to monitor their emissions starting from the base year, which is mainly 1990. This way a dozen of emission observations for each country are already available. This redundancy in observations could be perhaps favorably used to improve estimates of individual emissions in the commitment period 2008-2012, using statistical inference. They may be also used to estimate the parameters of the statistical distribution of the observation errors, this

* Partial financial support from the Polish State Scientific Research Committee within the grant 3PO4G12024 is gratefully acknowledged.

way providing independent assessment of the range of errors estimated up to now by propagation of initial errors in calculations. Both these tasks are addressed in the paper. The methods used are verified on data on CO_2 emissions from combustion of the fossil fuels estimated for the years 1751 - 1998 (Marland et al., 1999).

We consider three methods for estimating emissions and their variances: the smoothing splines (Wahba, 1990; Gu, 2002), a parametric model with a time variable coefficient, and the Brownian motion model. Other methods could be tried to solve the problem. A simple alternative would be to use another smoothing method. Those based on the wavelets might be promising ones (Debnath, 2002; Walter, 1994). Popular methods in the automatic control literature use the parametric models with calculation of the state errors following an earlier phase of the parameter estimation. In some of them, like in the extended Kalman filter, the parameters and the states are estimated simultaneously. To use this kind of methods, the parametric model is needed. Apart of that, most of them require quite long data samples to converge.

Prediction of future emissions is of interest of many investigations, see e. g. (Kroeze et al., 2004; Manne & Richels, 2004; Riahi et al., 2004), as it is connected, for example, with estimation of the national greenhouse gas balances in the commitment period or, on the global scale, with estimation of the supply and demand to predict the price of the tradable emission permits. Although the results presented in the paper could be also used for the prediction purposes, our paper concentrates more on historical data and possibility of gathering useful informations from them.

Thus, we estimate the variance of the observation errors, this way obtaining independent values of the uncertainty level. Although uncertainties considered here are not exactly the same as those estimated as inaccuracies in inventories, the estimates turned out to be of similar range for both cases. This supports, in a sense, correctness of the methods applied for estimation of the inventory uncertainties up to now, at least for the fossil fuel emissions.

To estimate the variance, different methods of modelling the emissions were used. Applying them to the historical data some regularities were noticed. The emissions often follow piecewise exponential curves, particularly in the periods of steady growth. Much less regular data are observed in the decline periods. Highly irregular are periods of wars and that of change from growth to decline.

In Sec. 2 a basic notation is introduced. Sec. 3 presents the nonparametric method based on smoothing splines. In Sec. 4 application of a parametric method is discussed and some numerical results are presented. Sec. 5 shortly discusses possibility of using the Brownian motion model for describing evolution of national CO_2 emissions in time. Sec. 6 concludes.

2. Notation used

By $x(t)$, as a function of time, we denote the integral of the real emission calculated on the interval $(t-1, t]$, where t is expressed in years. Thus, the integral is calculated over the one-year-back period. In the sequel we call $x(t)$ the emission. The function $x(t)$, as the integral of a positive function, is continuous and positive. In the paper we assume that $x(t)$ is a smooth enough function. The emission balances provided by the Annex I Parties are prepared by inventory of emissions from all involved activities during a year. Due to uncertainties in assessing the exact quantities and coefficients, they are in errors. To properly handle this situation, integer values of t are assigned to the end of consecutive years, which means that the emissions can be only observed with errors in integer time instants t_i. We denote the observed (reported) values $y(t_i)$ or shortly y_i. The index i begins here at 0 and takes the consecutive integer values. The real emissions $x(t_i) = x_i$ are unknown and can be only estimated. Hats will mark the estimated values, so $\hat{x}_i$ is the estimated emission.

By δ we denote the fraction of the emission to be reduced within the Kyoto obligations until the commitment period. Thus at the commitment period the emission should be not greater than $(1-\delta)x_0$. Obviously, the percentage reduction required by the Kyoto protocol is 100δ, but we often refer directly to δ in per cents. The value of δ is not greater than few per cents.

As it is common to express obligations in percentages, it is useful to work not with the straight observations but with their logarithms. Let us denote $\hat{X}_i = \ln(\hat{x}_i/\hat{x}_0)$, thus $\hat{X}_i$ is the logarithm of the normalized emission. As in our case $\hat{x}_i/\hat{x}_0$ is close to 1, then it approximately holds

$$\hat{X}_i = \ln\frac{\hat{x}_i}{\hat{x}_0} \approx \frac{\hat{x}_i}{\hat{x}_0} - 1 = \frac{\hat{x}_i - \hat{x}_0}{\hat{x}_0} \qquad (1)$$

Thus, $\hat{X}_i$ may be interpreted as the relative change of $\hat{x}_i$ with respect to $\hat{x}_0$ and may be expressed in percentages.

3. A nonparametric method

3.1. Basic assumptions and simplifications

We assume that the real process x_i is observed with a multiplicative error $\varepsilon_i = u_i x_i$, where

$$E(u_i) = m_i, \qquad E[(u_i - m)^2] = \sigma_i^2, \qquad \mathrm{cov}(u_i, u_j) = \gamma_{ij}$$

Thus, the observation can be presented in the following way

$$y_i = x_i + u_i x_i = (1 + u_i)x_i, \qquad i = 0, 1, \ldots, N$$

where y_i is the observed emission, x_i the (unknown) real emission, and u_i its relative uncertainty.

The above dependencies are also true for $i = 0$. Dividing sides and taking the logarithms we get

$$Y_i = X_i + \ln \frac{1 + u_i}{1 + u_0}$$

where $Y_i = \ln y_i/y_0$ and $X_i = \ln x_i/x_0$. For small u_0 and u_i it approximately holds

$$\ln \frac{1 + u_i}{1 + u_0} \approx u_i - u_0$$

resulting in the expression

$$Y_i = X_i + u_i - u_0$$

The error $v_i = u_i - u_0$ has the zero mean, $E(v_i) = 0$, and the variance $\sigma_{v_i}^2 = \sigma_i^2 + \sigma_0^2 - 2\gamma_{i0} = \sigma_i^2 + \sigma_0^2 - 2\rho_{i0}\sigma_i\sigma_0$, where $\rho_{i0} = \gamma_{i0}/\sigma_i\sigma_0$ is the cross correlation of u_0 and u_i. The covariance is equal to

$$\mathrm{cov}(v_i, v_j) = E[(u_i - u_0)(u_j - u_0)] = \gamma_{ij} - \gamma_{i0} - \gamma_{j0} + \sigma_0^2$$

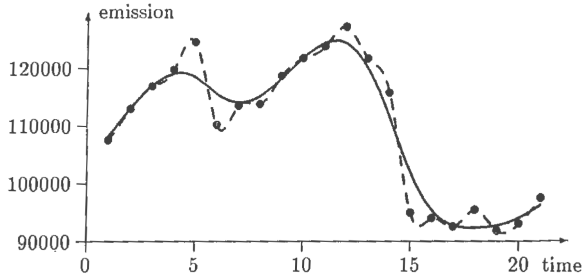
It equals zero, if all summands are equal. But generally the sequence is correlated, even if the original errors u_i are not. We assume, however, that the correlation is negligibly small. As noticed by Wahba ((Wahba, 1990), sec. 4.9), correlation of errors may considerably worsen the smoothing results, as far as reconstruction of the original function is considered.

3.2. Smoothing and uncertainty analysis

3.2.1. *Smoothing splines*

Let us consider some abstract data z_i generated by the following system

$$z_i = f(t_i) + e_i, \qquad i = 0, 1, 2, \ldots, N$$



Rysunek 1. The interpolating spline (dashed curve) and the smoothing spline (solid curve) to some emission data (dots).

The vector

$$\mathbf{e} = (e_0, \ldots, e_N) \propto \mathcal{N}(0, \sigma^2 \mathbf{I})$$

where $\mathbf{I}$ is the identity matrix, contains the set of observation errors. We want to recover the function $f(t)$, assumed to be smooth enough, knowing only the erroneous observations $z_i, i = 0, 1, \ldots, N$. For this we use splines.

In the interpolating splines an approximation $\hat{z}(t)$ to $f(t)$ is obtained assuming that $\hat{z}(t)$ is a polynomial of an order m (we use $m = 3$) on each segment $[t_i, t_{i+1}), i = 0, 1, 2, \ldots, N-1$, satisfying $\hat{z}(t) = z_i$ and having the continuous derivatives up to the order $m-1$ on the whole interval (t_1, t_N). In the presence of noise the interpolating spline generally quickly varies in time, overshooting and undershooting considerably the function $f(t)$.

Much better approximation can be achieved for noisy data using the smoothing splines. Their idea is to find the function $\hat{z}(t)$ that does not need to go directly through the observed points z_i, in order to get a function with a smaller $(m-1)$th derivative, see Fig. 1.

If we restrict our attention to the third order smoothing splines, then the task is to find a function $\hat{z}(t)$, which minimizes the sum

$$\frac{1}{N+1} \Sigma_{i=0}^{N} (z_i - \hat{z}(t))^2 + \lambda \int_{t_0}^{t_N} (\hat{z}^{(2)}(t))^2 dt \tag{2}$$

where

$$\hat{z}(t) = a_i + b_i(t - t_i) + c_i(t - t_i)^2 + d_i(t - t_i)^3$$

$$t \in [t_i, t_{i+1}), \qquad i = 0, 1, \ldots, N-1$$

The solution of the problem, for a given λ, is delivered e. g. in Wahba (Wahba, 1990) and can be written in a general form

$$\hat{z}(t) = a_i = \Sigma_j A_{ij}(N, \lambda) z_j, \qquad \frac{d\hat{z}(t)}{dt} = b_i = \Sigma_k B_{ij}(N, \lambda) z_k \qquad (3)$$

see also (Gu, 2002), where A_{ij} and B_{ik} are coefficients which do not depend on the data z_i and can be precomputed. Thus, both $\hat{z}(t)$ and $d\hat{z}(t)/dt$ are kernel estimators.

3.2.2. *Uncertainty analysis*

The solution depends on the value of λ. This value is estimated by the generalized cross validation method (Wahba, 1990) by minimizing over λ the criterion

$$V(N, \lambda) = \frac{\Sigma_{i=0}^{N}[z_i - \hat{z}_i(N, \lambda)]^2}{N + 1 - \Sigma_{i=0}^{N} A_{ii}(N, \lambda)} \qquad (4)$$

where $\hat{z}_i(N, \lambda)$ is the solution of the problem (2), in which the observation $\hat{z}_i$ is dropped. The optimal value will be denoted $\hat{\lambda}$. The optimal value of the criterion can be used as an estimate of σ^2, i.e.

$$\hat{\sigma}^2(N) = V(N, \hat{\lambda}) \qquad (5)$$

The expression in the denominator of (4) can be interpreted as the degrees of freedom of the noise, in analogy to the degrees of freedom in the regression analysis. However, in contrast to the regression analysis, only consistency of the estimate for the smoothing splines has been proved theoretically ((Gu, 2002), th. 3.4), while other good statistical properties have been checked only on numerical simulations.

The estimated variance of $\hat{z}(t_i)$ is now

$$\hat{\sigma}^2_{\hat{z}_i}(N) = \hat{\sigma}^2(N) A_{ii}(N, \hat{\lambda}) \qquad (6)$$

3.2.3. *Application to real data*

The above analysis was applied for smoothing the data $Y_i = \ln(y_i/y_0)$. Equation (6) has been used to calculate the estimates of the standard deviations $\hat{\sigma}_{\hat{z}_i}(N)$ for the emission from the fossil fuels provided by Marland et al. (Marland et al., 1999), in the periods 1950-1998 and 1970-1998. The value $\hat{\sigma}_{\hat{z}_i}(N)$ depends on the number of data used. This dependence is visible, although mostly not crucial, in the results presented in Table I for different time periods. For few cases, like e.g. Argentine, Canada, USA, reduction of the number of data caused big drop of the standard deviation value.

In ((Wahba, 1990), sec. 4.9) it is recommended to use at least 25-30 observations when applied the smoothing splines. The data used in calculating the values in the left side of Table I contained 29 points, just satisfying the recommendations. However, for many countries, the corresponding standard deviations differ for different length of data. At least in some cases this is correlated with extreme values of $\hat{\lambda}$, either very close to zero, like for Argentina, Canada and USA, or very high, like for Austria and Cuba. This phenomenon is also mentioned in ((Wahba, 1990), sec. 4.9). This may suggest that the data in the shorter sequence may be too short.

The estimated values agree quite well with the common idea on the magnitude of errors made in calculation of the fossil fuel emission, believed to be of few per cents. They also agree well with the estimates calculated by other methods for few countries and collected in ((Gugele at al., 2003), Tab. 6). A little bigger figures obtained in some of our calculations may be connected with some additional factors that might have influenced the calculated estimates, as year-to-year variations in the weather conditions or variations due to change in economic factors of the countries. However, application of statistical paired Student test gives no reason to reject the hypothesis on equality of means of data in any two columns in Table I (the biggest value of the t statistic is equal to 0.78). This supports the claim on similarity of results obtained by different discussed here methods.

4. Empirical parametric models

In the previous section we noticed that the consecutive values in the emission sequence might be correlated. To better model this property, in this section we consider a set of values x_i forming a time series consisting of N elements and introduce a difference model to describe the time evolution of the data. Then we motivate the choice of the model and finally present some results for fitting the model to the emission data for some countries.

As we assumed that x_i are positive we can define a new time series

$$g_i = \frac{x_{i+1}}{x_i} - 1 = \frac{x_{i+1} - x_i}{x_i}, \qquad i = 0, 1, \ldots, N-1$$

Each element g_i of a new time series can be interpreted as a relative difference of the two consecutive elements x_{i+1} and x_i.

From the latter relation we can now formulate the following difference equation

$$x_{i+1} - x_i = g_i x_i, \qquad x_0 = x(t_0) \tag{7}$$

Tablica I. Estimated standard deviations of observation errors for different countries and two time periods, for two methods, in [%].

Years	1950 - 1998		1970 - 1998		~ 2000
Country	smooth.	param.	smooth.	param.	reported
Argentina	2.3	0.7	0.4	0.1	
Australia	1.8	0.5	0.9	0.5	
Austria	2.7	0.9	1.1	1.0	1.0
Belgium	2.3	3.3	2.3	3.3	1.1
Brazil	1.9	1.1	1.3	1.7	
Canada	1.9	0.8	0.5	1.8	
China	4.7	7.1	1.4	1.7	
Cuba	6.6	2.2	1.9	1.4	
Egypt	3.4	1.4	2.6	1.1	
Finland	4.8	1.3	3.8	3.6	3.0
France	2.3	3.0	2.3	1.1	< 2.5
Greece	2.8	0.9	2.2	0.9	
Iceland	3.5	1.4	2.7	1.4	
Ireland	4.3	1.2	2.2	2.2	< 1.0
Israel	3.4	2.2	2.0	0.9	
Italy	1.6	2.3	1.3	0.7	
Japan	2.7	4.8	1.8	2.4	
Luxembourg	2.9	4.3	2.8	4.0	
Mexico	1.7	2.1	1.7	2.0	
Netherlands	2.8	0.9	3.7	1.4	1.5
New Zealand	1.8	0.8	2.9	2.1	
Norway	4.2	2.0	5.2	3.3	1.5
Poland	1.5	1.8	1.8	2.2	
Portugal	1.9	0.9	1.9	1.2	
Romania	1.9	2.4	2.1	2.9	
Spain	3.0	1.2	1.7	1.0	
Sweden	2.5	1.1	2.3	1.4	1.0
Switzerland	3.3	4.3	1.9	1.0	
Turkey	3.1	4.3	3.4	1.1	
U. K.	1.6	0.5	1.4	0.7	2.0
USA	1.8	0.5	0.4	2.1	1.5

Because $y_i = (1+u_i)x_i$, then (7) can be transformed to

$$y_{i+1} = (1+g_i)\frac{1+u_{i+1}}{1+u_i}y_i$$

Dividing both sides by y_0 and taking logarithms yields

$$Y_{i+1} = \ln(1+g_i) + \ln\frac{1+u_{i+1}}{1+u_i} + Y_i$$

or approximately

$$Y_{i+1} - Y_i \approx g_i + u_{i+1} - u_i$$

from where an estimator $\hat{g}_i$ can be designed as

$$\hat{g}_i = Y_{i+1} - Y_i \tag{8}$$

Under our assumption on u_i's we have

$$E(\hat{g}_i) = E(Y_{i+1} - Y_i + u_i - u_{i+1}) = X_{i+1} - X_i = \ln(1+g_i) \approx g_i$$

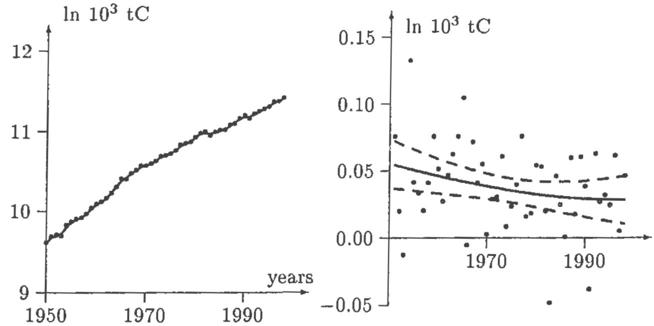
Thus the estimator is approximately unbiased. Its approximate variance is

$$\mathrm{var}(\hat{g}_i) = E(Y_{i+1} - X_{i+1} - Y_i + X_i)^2 =$$
$$= E(u_{i+1} - u_0 - u_i + u_0)^2 = E(u_{i+1} - u_i)^2 = \sigma_{i+1}^2 - 2\gamma_{i,i+1} + \sigma_i^2$$

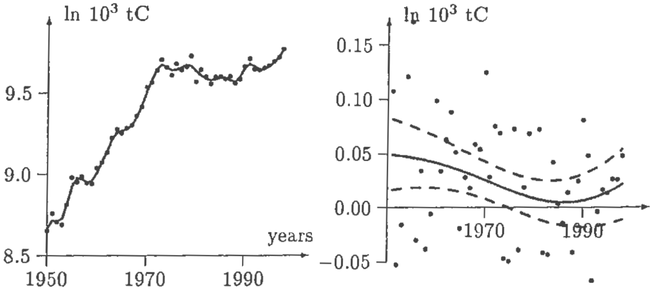
4.1. Estimation of the parameter g_i

The expression (8) was used to estimate the function g_i for few countries from the previously mentioned data of CO_2 emission from the fossil fuels (Marland et al., 1999). Some chosen results are presented in Figs. 2-5. The smoothing splines were used to smooth the points obtained from (8) with the formulae (3). For each country, in the left panel the observations (dots) and their smoothing spline approximations (solid lines) are depicted. The right panel shows the estimates of the function g_i. The dots represent the points calculated using the formula (8). The bold solid line is obtained by smoothing these points. The dashed lines show the 95% confidence intervals of the estimates.

Table I depicts also the estimates of the standard deviation of the errors $u_{i+1} - u_i$. From the comparison with the values obtained from smoothing it can be seen that both estimates of the standard deviations are of the same order, although not always very close to each other. Notice, however, that values from smoothing correspond to the standard deviations of the errors $u_i - u_0$, while those from the parametric model to $u_i - u_{i-1}$, what might partly cause the differences.



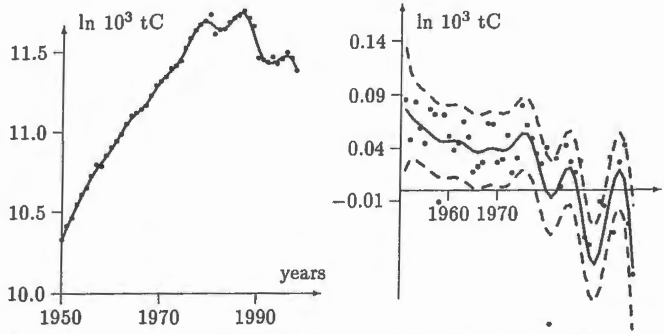
Rysunek 2. Results of smoothing and estimation of the function g for Australia in the years 1950-1998. Left panels: dots – logarithms of observations, solid lines – smoothed logarithms of observations. Right panels: dots – estimates of $\hat{g}_i$ from the formula (8), the bold solid lines – their smoothed continuous approximations, the normal thickness dashed lines – the 95% confidence intervals of these approximations.



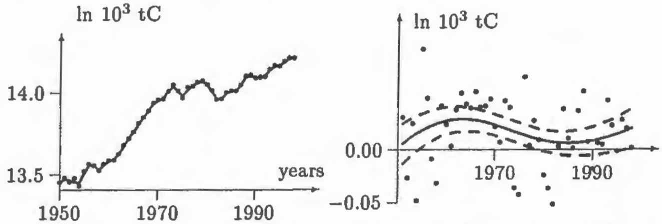
Rysunek 3. Results of smoothing and estimation of the function g for Austria in the years 1950-1998. Left panels: dots – logarithms of observations, solid lines – smoothed logarithms of observations. Right panels: dots – estimates of $\hat{g}_i$ from the formula (8), the bold solid lines – their smoothed continuous approximations, the normal thickness dashed lines – the 95% confidence intervals of these approximations.

4.2. Piecewise exponential model

Although the estimated functions $\hat{g}(t)$ in the previous section vary in time, in many periods their patterns resembles the constant value lines. To better investigate this question let us start with examining of few curves. Figs. 6 and 7 contain emission curves y_i and logarithmic



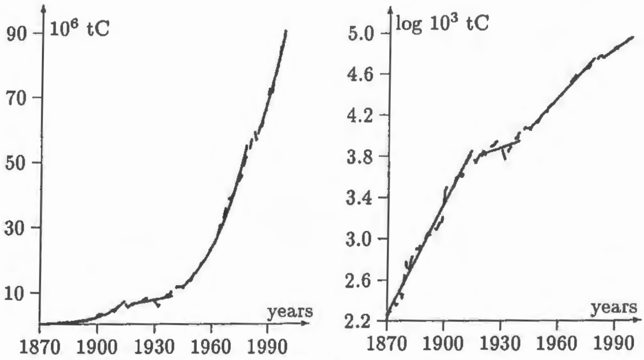
Rysunek 4. Results of smoothing and estimation of the function g for Poland in the years 1950-1998. Left panels: dots – logarithms of observations, solid lines – smoothed logarithms of observations. Right panels: dots – estimates of $\hat{g}_i$ from the formula (8), the bold solid lines – their smoothed continuous approximations, the normal thickness dashed lines – the 95% confidence intervals of these approximations.



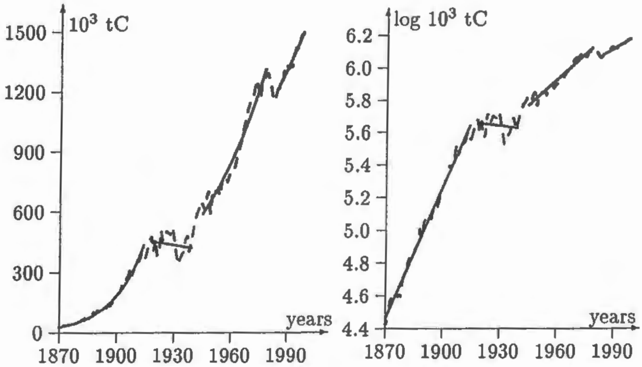
Rysunek 5. Results of smoothing and estimation of the function g for USA in the years 1950-1998. Left panels: dots – logarithms of observations, solid lines – smoothed logarithms of observations. Right panels: dots – estimates of $\hat{g}_i$ from the formula (8), the bold solid lines – their smoothed continuous approximations, the normal thickness dashed lines – the 95% confidence intervals of these approximations.

curves $Y_i = \ln(y_i/y_0), t_0 = 1990$, for the emission data (Marland et al., 1999) for Australia and USA. It can be seen that the data evolve approximately along piecewise exponential curve, and the logarithmic curves are approximately linear.

Thus, the exponential growth models describe quite well development of data only in some definite intervals. These intervals are the periods of constant development conditions. One can easily distinguish in the figures the period of the 19th century industrial revolution or

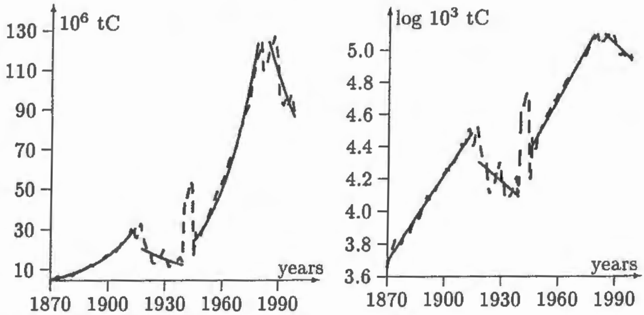


Rysunek 6. **Emissions for Australia with fitted piecewise exponential curve (left) and their logarithms with fitted straight lines (right), in millions of metric tons of C.**

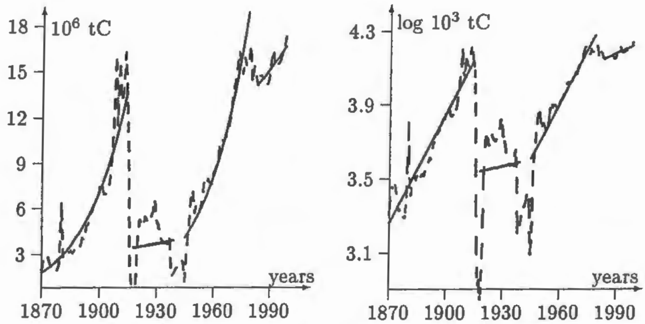


Rysunek 7. **Emissions for USA with fitted piecewise exponential curve (left) and their logarithms with fitted straight lines (right), in millions of metric tons of C.**

the period of the post-war prosperity of 1950s-1970s. However, even for USA, Fig. 7, and more visible for the European countries like Poland or Austria, Figs. 8 and 9, it can be easily noticed that there are periods where the assumption on the simple constant parameter g (and therefore the growth along the exponential curve) can not be true. This is particularly visible in the periods of the World Wars and Great Crisis of 1930s,and the energy shocks of 1970s-1980s. Also smaller ripples can be



Rysunek 8. Emissions for Poland with fitted piecewise exponential curve (left) and their logarithms with fitted straight lines (right), in millions of metric tons of C.

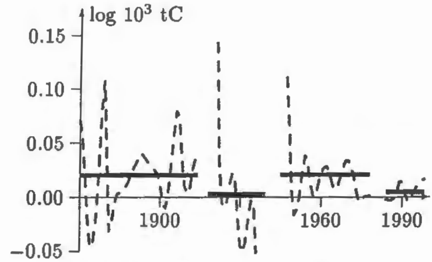


Rysunek 9. Emissions for Austria with fitted piecewise exponential curve (left) and their logarithms with fitted straight lines (right), in millions of metric tons of C.

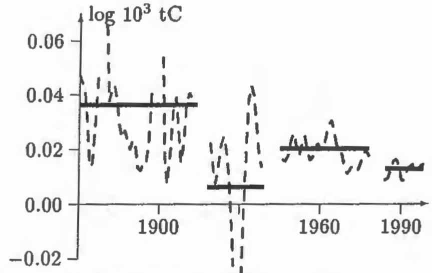
distinguished and explained, like for example in the case of the Polish transformation period.

The fit of this simple piecewise exponential model is quite good in the periods of growth or decay. In the period of steady growth it is almost perfect. In the decay periods the emission is often more volatile. War and transition periods, like those of 1970s in the West Europe or 1980s in Poland, are highly irregular and were skipped from fitting.

The results obtained are generally quite similar for both methods. The error variance estimates calculated by the regression method (parametric model) turn out to be usually greater than those calculated by the smoothing splines. This seems to be connected with too big



Rysunek 10. Estimates of g for Austria. Solid line - piecewise exponential model, dashed lines - smoothing.

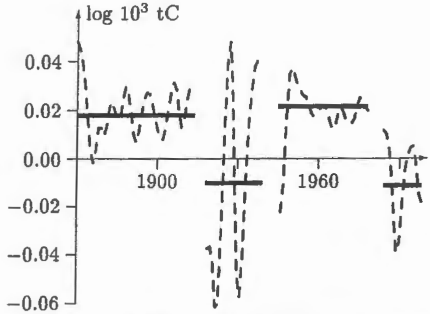


Rysunek 11. Estimates of g for Australia. Solid line - piecewise exponential model, dashed line - smoothing.

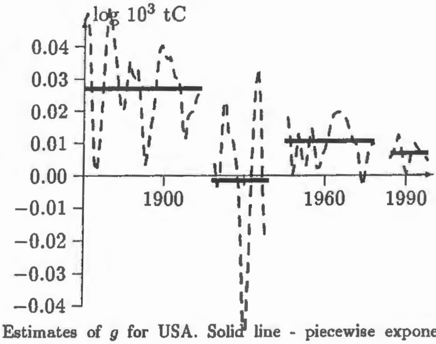
simplicity of the exponential model used. However, the good fit of the piecewise exponential model seems to be an important observation. It means that in the past the emissions have followed approximately the exponential functions in defined longer periods. The jump from one such segment to another is mostly connected with a big political or economic change.

5. Geometric Brownian motion

Geometric Brownian Motion is the most often used stochastic process in financial economics theory, and in our case may be considered as an useful alternative from a practical point of view. In several calculated cases it was found to be not a better model than others, even being a reasonable mapping of probabilities within the time.



Rysunek 12. Estimates of g for Poland. Solid line - piecewise exponential model, dashed line - smoothing.



Rysunek 13. Estimates of g for USA. Solid line - piecewise exponential model, dashed line - smoothing.

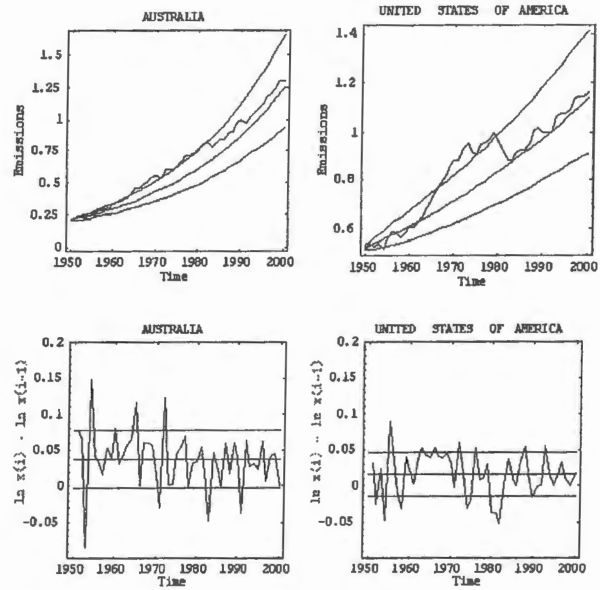
5.1. Geometric Brownian model for the emissions

For a signal $x(t)$ that follows a geometric Brownian motion, the stochastic equation for its variation in time t is

$$dx = gx + \sigma x dz \tag{9}$$

where $dz = \varepsilon dt^{1/2}$ is the Wiener increment, ε - standard normal distribution, g is the drift, and σ is the volatility of x.

In the above equation the first term on the right hand side is the expectation (trend) term and the second term is the variation term (deviation from the trend or uncertainty).



Rysunek 14. Illustrations of the considered stochastic process, showing sample paths, the 66% confidence intervals, and the forecasted expected values (exponential trend line) for two countries (upper panels) and estimates $\hat{g}_i = \ln x_i - \ln x_{i+1}$ with their means and 66% confidence intervals (lower panels).

The geometric Brownian motion is a log-normal diffusion process with the expected value of x at the time t (starting at $t_0 = 0$)

$$E[x(t)] = x_0 e^{gt} \tag{10}$$

and the standard deviation SD

$$SD[x(t)] = x_0 e^{gt} \sqrt{e^{\sigma^2 t} - 1} \tag{11}$$

This is illustrated in Fig. 14.

5.2. Arithmetic Brownian model for the logarithm of the emissions

Due to its simplicity, it is useful to work with the logarithmic diffusion equation. Letting $X = \ln x$, and using Itô's lemma we find that x follows the arithmetic (or ordinary) Brownian motion

$$dX = d\ln x = (g - \frac{1}{2}\sigma^2)dt + \sigma dz \tag{12}$$

so

$$dX = g'dt + \sigma dz$$

where $g' = g - \frac{1}{2}\sigma^2$. The variable X follows an arithmetic Brownian motion with the drift g' and volatility σ.

We should note here, that although the volatility term is the same in (12) as in of the geometric Brownian motion for x (9), the element $d(\ln x)$ is different from dx/x due to the different drift expression (so called Itô's effect).

The drift parameter g can be estimated as the average value of a set of differences of the logarithms $\ln x_i - \ln x_{i-1}$. Using the same historical series we can get an estimation of the volatility σ by taking the standard deviation of $\ln x_i - \ln x_{i-1}$, as for the parametric model of Sec. 4. They can be inserted in equations (10) and (11) to obtain the characterization of the process in time.

The calculations, not presented here, give bigger estimates of the standard deviations than those depicted in Table I, comparable to the piecewise exponential model of Sec. 4. These bigger values seem to be mainly caused by constant value of g in the model.

6. Conclusions

Nonparametric and parametric methods for modelling the greenhouse gas emission phenomena and for estimating the parameters are proposed in the paper. They differ in degree of smoothing and precision of fitting the observations. Comparison of the methods used reveals that the parametric method of Sec. 4 gives in many instances simpler, less volatile curves, although it is more sensitive to the smoothing interval. The smoothing method of Sec. 3 is more accurate and better emphasizes the ripples in data. The parametric piecewise exponential model gives the most rough but also most simple description, showing general trends in evolution of emission data.

One of the main goals of the paper was to estimate the standard deviation of the errors. Some signal processing methods are proposed and preliminary results are presented. They are based on the published observations of the emissions from the fossil fuels (Marland et al., 1999) and therefore do not cover the whole emissions reported within the Kyoto agreement. Moreover, the volatility of observations may be related not only to the observation errors but also to such factors as changing weather conditions and rapidly changing economic situation of the country. These phenomena might have contributed to increase of the estimated variance.

With these reservations, the calculations performed for the fossil fuels indicate that the empirical approach gives reasonable estimates, comparable to the estimates obtained so far by the methods recommended by IPPC (IPCC, 2000a). Or, to be more cautious, the partial results obtained here do not falsify the uncertainty estimation procedures applied up to now for the inventories. The present knowledge does not allow us to state definite conclusions as yet.

An interesting relation between the piecewise exponential growth of the CO_2 emission curve and the country economic development may well be also true for other gases. An open question is how removal of the greenhouse gases by sinks, also included in the full calculation of the greenhouse gas balance of countries, may behave. Evolution of this type of data in time will be possible to analyze when longer historical records will be available.

The proposed approach can be used to better estimate the real emissions, by filtering out errors, and possibly for prognosis. The latter application might be important as an alternative to the scenarios built on the basis of technological and economic assumptions. But such application it is still rather risky until more will be known on dependence of the emissions on the economic and weather conditions.

Literatura

Brandes L.J., Olivier J.C.J., van Oorschot M.H.P.: Validation, verification and uncertainty assessment for improving the Netherlands' emission inventory. Proc. Workshop *Uncertainty in Greenhouse Gas Inventories: Verification, Compliance & Trading.* SRI PAS & IIASA, Warsaw 2004, 19–33. http://www.ibspan.waw.pl/GHGUncert2004/papers/Brandes.pdf.

Debnath L.: *Wavelet Transforms and Their Applications.* Birkhäuser, 2002.

FCCC. Report of the Conference of the Parties on its third session, held at Kyoto from 1 to 11 December 1997. Addendum. Document FCCC/CP/1997/7/Add.1. United Nations Framework Convention on Climate Change (FCCC), 1998. http://www.unfccc.de/ http://unfccc.int/index.html.

FCCC. Implementation of the Buenos Aires plan of action: Adoption of the decisions giving effect to the Bonn agreements. Draft decisions forwarded for elaboration, completion and adoption. National systems, adjustments and guidelines under

Articles 5, 7 and 8 of the Kyoto Protocol. Document FCCC/CP/2001/L.18. United Nations Framework Convention on Climate Change (FCCC), 2001. http://www.unfccc.de/

IPCC guidelines for national greenhouse gas inventories. Vols. 1-3, IPCC, London 1996.

IPCC. Good practice guidance and uncertainty management in national greenhouse gas inventories. J. Penman, D. Kruger, I. Galbally, T. Hiraishi, B. Nyenzi, S. Emmanuel, L. Buendia, R. Hoppaus, T. Martinsen, J. Meijer, K. Miwa and K. Tanabe (eds.), Intergovernmental Panel on Climate Change (IPCC) National Gas Inventories Program, Technical Support Unit, Institute for Global Environmental Strategies, Hayama, Kanagawa, Japan, 2000a.

IPCC. Land use, land-use change, and forestry. Special report of the Intergovernmental Panel on Climate Change (IPCC). R.T. Watson, I.R. Noble, B. Bolin, N.H. Ravindranath, D.J. Verardo and D.J. Dokken (eds.), Cambridge University Press, Cambridge, United Kingdom, 2000b.

Gu Ch.: *Smoothing Spline ANOVA Models.* Springer, 2002.

Gugele, B., Huttunen K., Ritter M.: Annual European Community greenhouse gas inventory 1990–2001 and inventory report 2003. Technical Report No. 95, European Environment Agency, Copenhagen, Denmark, 2003. http://reports.eea.eu.int/technical_report_2003_95/en.

Gupta J., Oltshoorn X., Rotenberg E.: The role of scientific uncertainty in compliance with the Kyoto Protocol to the Climate Change Convection. *Environmental Science & Policy*, **6**, 2003, 475–486.

Kroeze C., Vlasblom J., Gupta J., Boudri C., Blok K.: The power sector in China and India: greenhouse gas emission potential and scenarios for 1990-2020. *Energy Policy*, **32**, 2004, 55–76.

Manne A., Richels R.: US rejection of the Kyoto Protocol: the impact on compliance cost and CO_2 emissions. *Energy Policy*, **32**, 2004, 447–454.

Marland G., Andres R. J., Boden T. A., Johnston C. A., Brenkert A. L.: Global, regional, and national CO_2 emission estimates from fossil fuel burning, cement production, and gas flaring: 1751-1996 (revised March 1999). Carbon Dioxide Information Analysis Center, 1999. http://cdiac.esd.ornl.gov/trends/emis/em_cont.htm

Monni S., Syri S., Pipatti R., Savolainen I.: Comparison of uncertainty in different emission trading schemes. Proc. Workshop *Uncertainty in Greenhouse Gas Inventories: Verification, Compliance & Trading.* SRI PAS & IIASA, Warsaw 2004, 106–115. http://www.ibspan.waw.pl/GHGUncert2004/papers/Monni.pdf.

Nahorski Z., Jęda W., Jonas M.: Coping with uncertainty in verification of the Kyoto obligations. In: Studziński J., Drelichowski L., Hryniewicz O.: *Zastosowania informatyki i analizy systemowej w zarządzaniu*, IBS PAN, Warszawa 2003, 305-317.

Nilsson S., Jonas M., Obersteiner M., Victor D.: Verification: The gorilla in the struggle to slow global warming. *The Forestry Chronicle*, **77**, 2001, 475–478.

Nilsson, S., M. Jonas, Obersteiner M.: COP 6: A healing shock. *Climatic Change*, **52**, 2002, 25–28.

Riahi K., Rubin E. S., Taylor M. R., Schrattenholzer L., Houndshell D.: Technological learning for carbon capture and sequestration technologies. *Energy Policy*, **26**, 2004, 539–564.

Rypdal K., Winiwarter W.: Uncertainty in greenhouse gas emission inventories – evaluation, comparability and implications. *Environmental Science & Policy*, **4**, 2001, 104–116.

Wahba G.: *Spline Models for Observational Data.* SIAM, 1990.
Walter G.G.: *Wavelets and Other Orthogonal Systems With Applications.* CRC Press, 1994.