Comparison of three signal analysis methods for modelling of GHG emissions and uncertainties

Zbigniew Nahorski, Waldemar Jęda

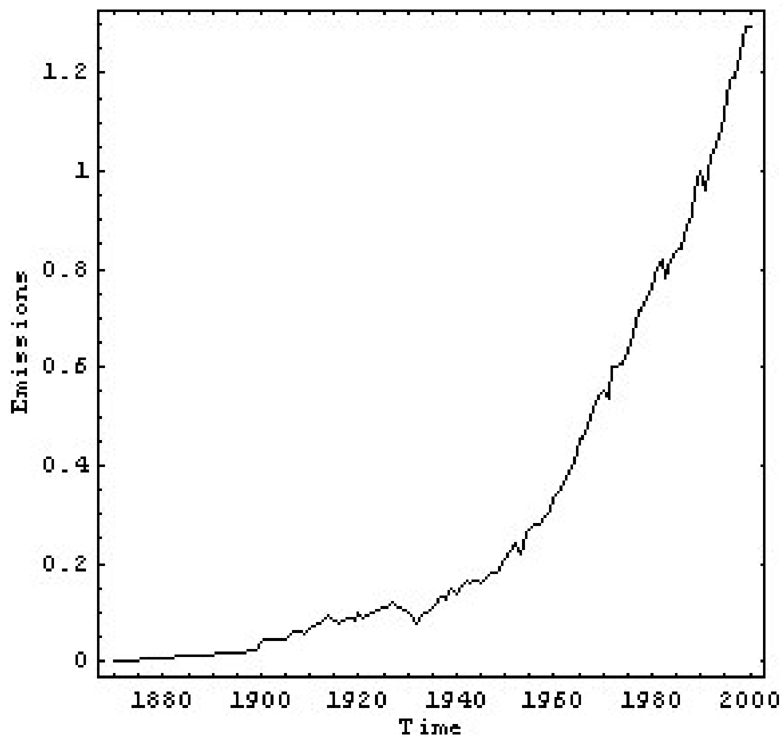
Systems Research Institute
Polish Academy of Sciences

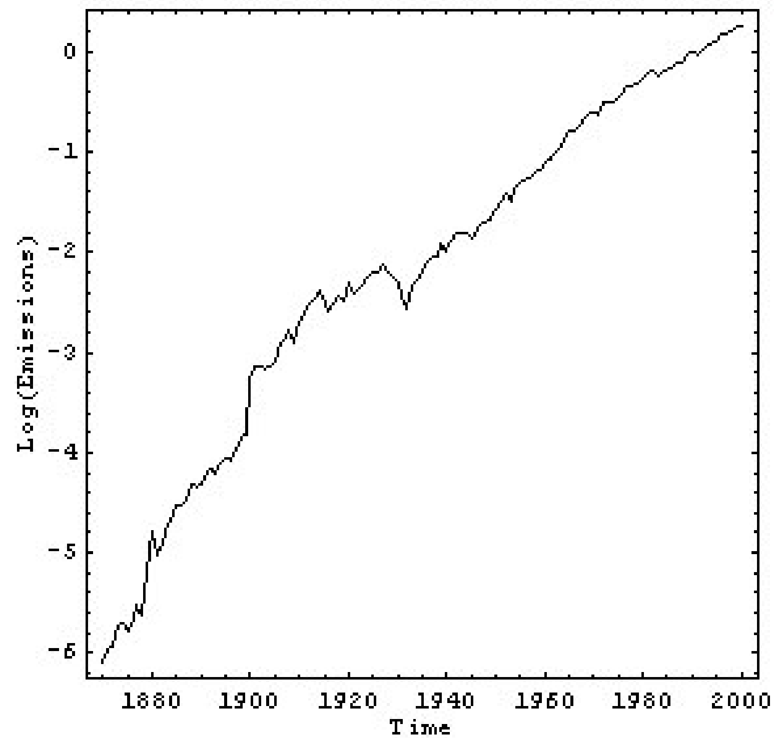
GHGUncert 2004   24-25 September 2004

Aims

- To investigate procedures for independent calculation of uncertainty estimates.
- To consider methods of signal processing:
  - Smoothing procedure based on spline functions
  - Parametric model with a time-varying parameter
  - Geometric Brownian motion model

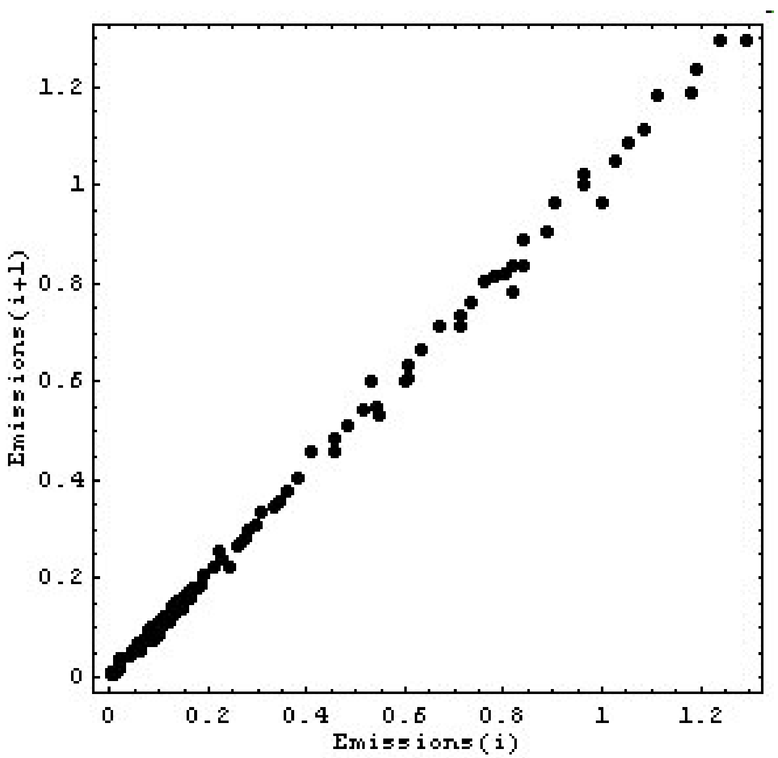
CO_2 emissions data (1)

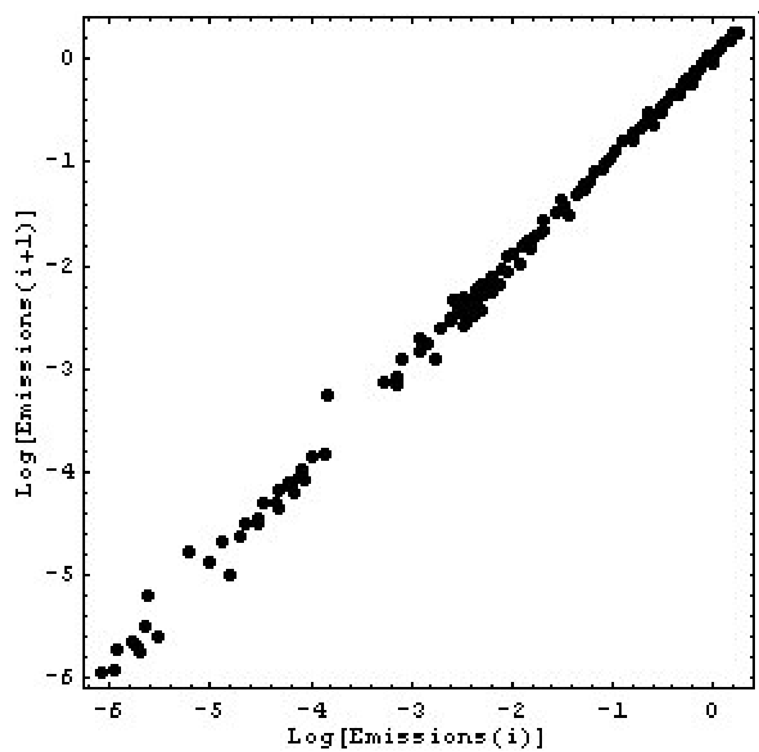




GHG emissions from fossil fuels combustion: Australia

CO_2 emissions data (2)





GHG emissions from fossil fuels combustion: Australia

Notation (1)

Emissions:

$x(t)$ - real („true”)

$y(t)$ - observed (reported)

$\hat{x}(t)$ - estimated

$$\hat{X}(t) = \ln\frac{\hat{x}(t)}{\hat{x}(t_0)} \approx \frac{\hat{x}(t)}{\hat{x}(t_0)} - 1 = \frac{\hat{x}(t) - \hat{x}(t_0)}{\hat{x}(t_0)}$$

Notation (2)

Uncertainty:

Real process $x_i \equiv x(t_i)$ is observed with errors...

$$y_i = x_i + \varepsilon_i \ , \ i = 0,1,\ldots,N$$

... which are of multiplicative character:

$$\varepsilon_i = u_i \, x_i \qquad\qquad y_i = x_i + u_i \, x_i = (1 + u_i)\, x_i$$

$$E(u_i) = m_i$$

$$E[(u_i - m)^2] = \sigma_i^2$$

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

Empirical nonparametric method (1)

Problem:

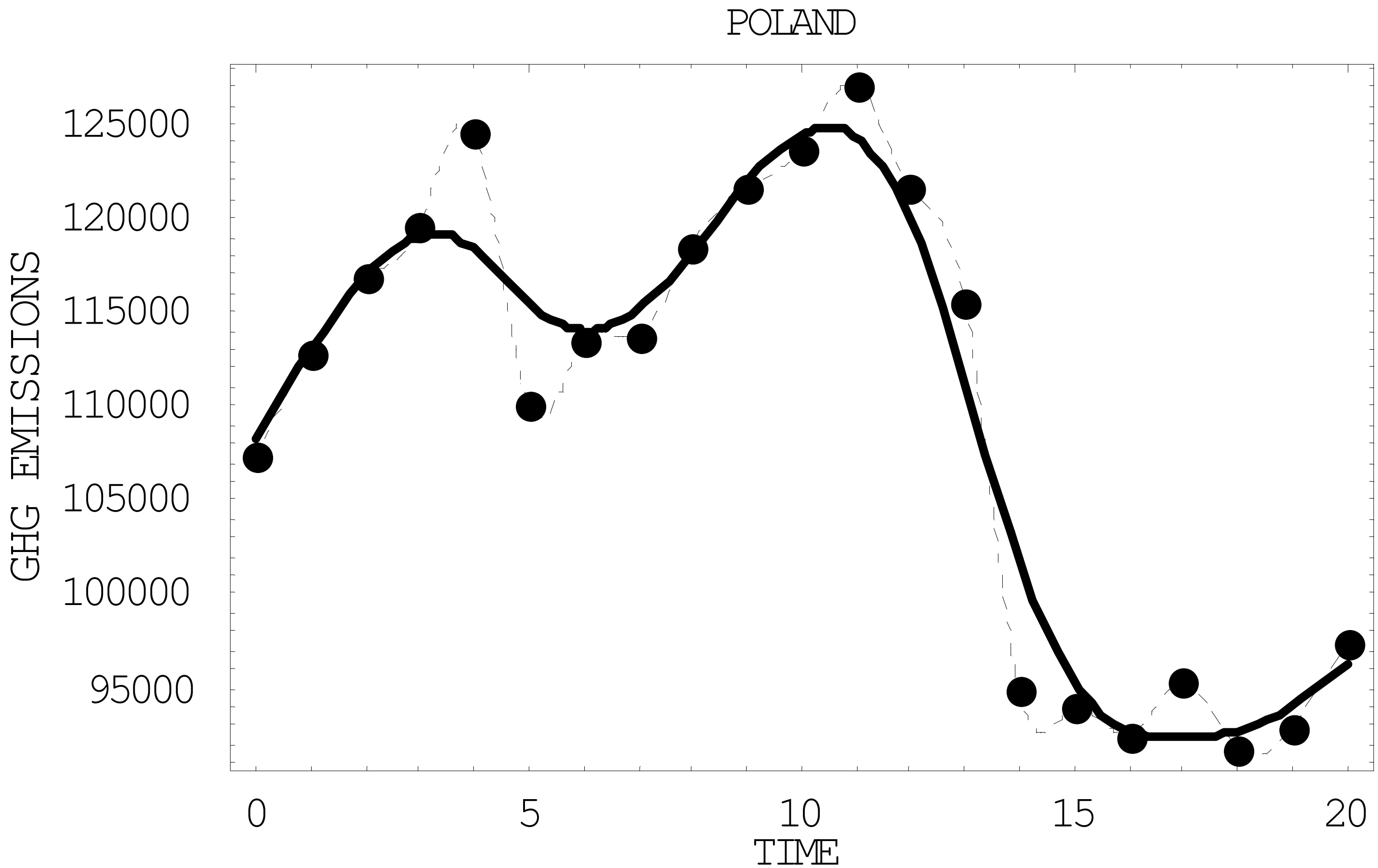
How to recover the function $f(t)$, assumed to be smooth enough, knowing only the erroneous observations z_i, $i = 0, 1, ..., N$

Smoothing Splines:

smooth function $\hat{z}(t)$, which minimises the sum:

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

Empirical nonparametric method (2)



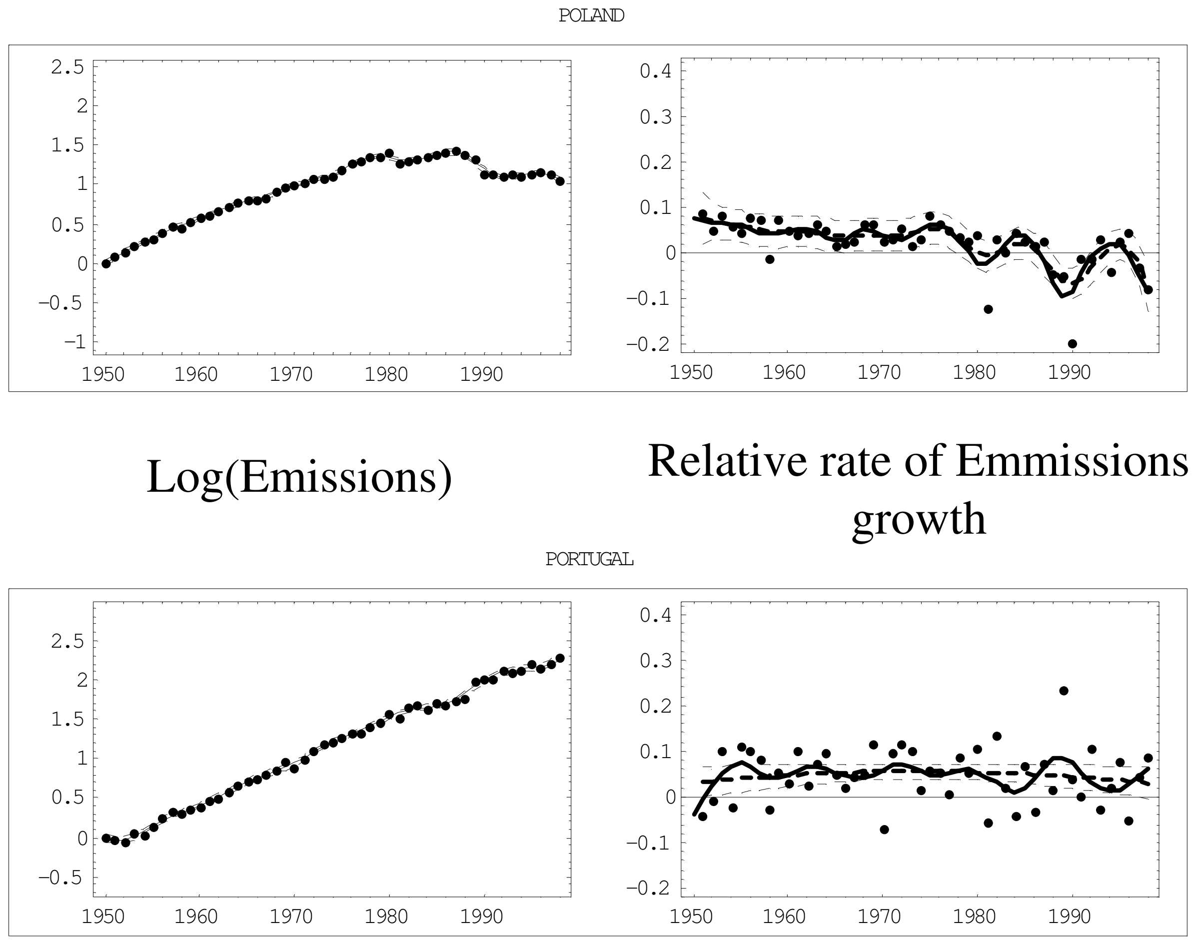
The interpolating spline (dashed curve) and the smoothing spline (solid curve)

Empirical nonparametric method (3)

- Number of observations (at least 25-30)
- Choice of λ (generalized cross validation method)
- Estimator of σ^2 is consistent
- Other good statistical properties checked numerically

Empirical nonparametric method (4)

Country	$\hat{\lambda}$	Std. Dev. [%]
ARGENTINA	0.06	2.4
AUSTRALIA	0.06	1.8
AUSTRIA	0.15	2.7
BELGIUM	0.07	2.4
BRAZIL	0.31	2.0
CANADA	0.03	2.0
CHINA	0.03	4.8
CUBA	0.16	6.7
EGYPT	1.16	3.5
FINLAND	0.03	4.9
FRANCE	0.14	2.4
GREECE	0.14	2.9
ICELAND	1.64	3.7
IRELAND	0.11	4.4
ISRAEL	0.03	3.5
ITALY	0.10	1.7
JAPAN	0.01	2.8
LUXEMBOURG	0.05	3.0
MEXICO	0.77	1.8
NETHERLANDS	0.08	2.9
NEW ZEALAND	5.11	2.0
NORWAY	3.44	4.6
POLAND	0.71	1.5
PORTUGAL	3.35	2.1
ROMANIA	0.20	1.9
SPAIN	0.03	3.1
SWEDEN	3.69	2.8
SWITZERLAND	0.11	3.4
TURKEY	0.11	3.2
UNITED KINGDOM	0.15	1.6
USA	0.02	1.7



Empirical parametric models (1)

Variable parameter model

$$g_i := \frac{x_{i+1} - x_i}{x_i}$$

$$x_{i+1} - x_i = g_i \, x_i$$

$$x_0 = x(t_0)$$

$$x_i = x_0 \prod_{j=0}^{i-1} (1 + g_j)$$

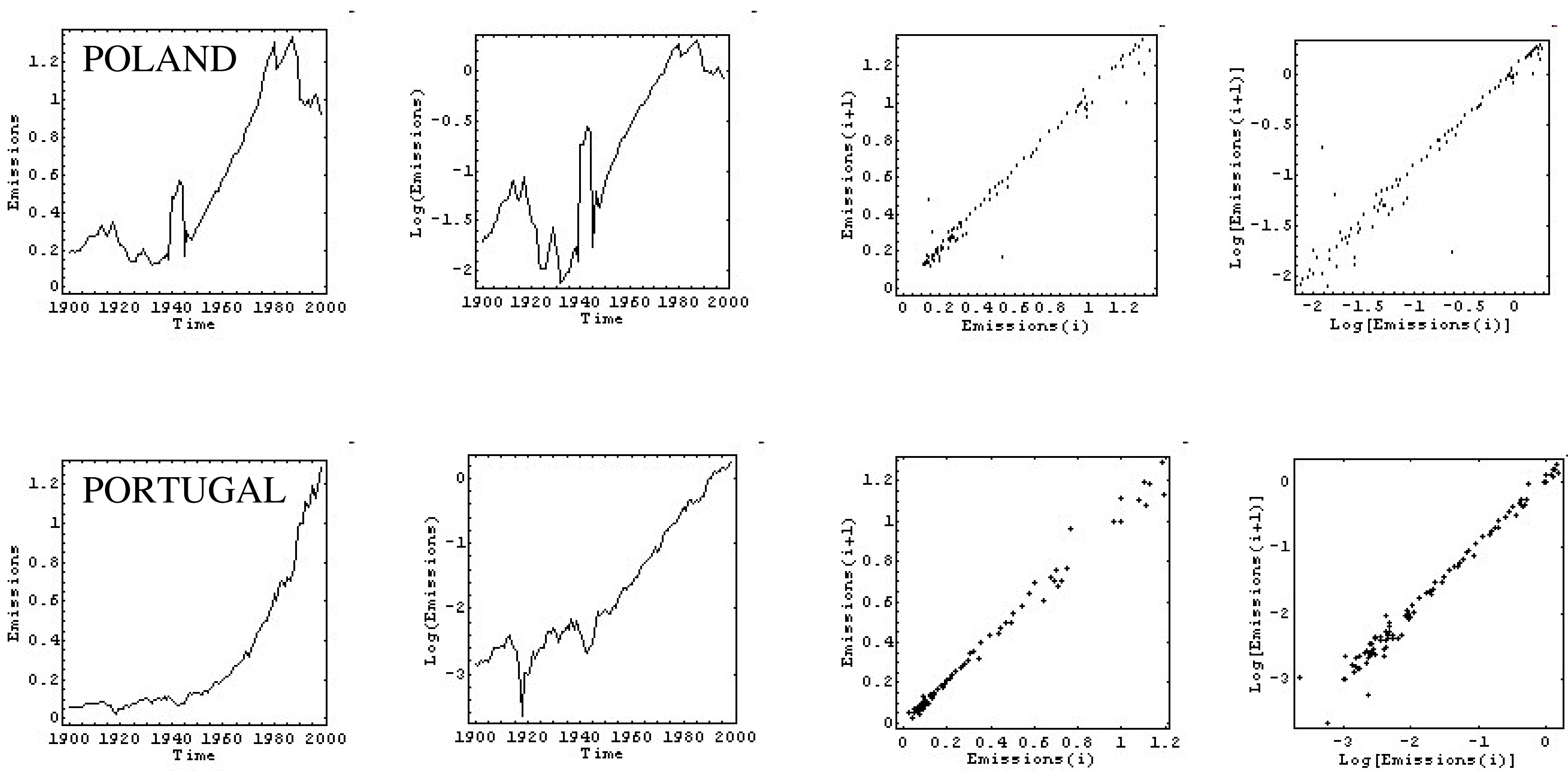
$$X_i := \ln \frac{x_i}{x_0} = \sum_{j=0}^{i-1} \ln(1 + g_j) \cong \sum_{j=0}^{i-1} g_j$$

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

$$\begin{cases} E(\hat{g}_i) = \ln(1 + g_i) \approx g_i \\ \operatorname{var}(\hat{g}_i) = \sigma_{i+1}^2 - 2\gamma_{i,i+1} + \sigma_i^2 \end{cases}$$

Empirical parametric models (2)

Piecewise exponential model



Empirical parametric models (3)

POLAND: 1870-1998

YEARS 1870–1914

	Estimate	SE	TStat	PValue
1	−29.9226	0.535618	−55.8655	0.
t	0.0179755	0.000283089	63.4976	0.

E.Var. 0.000608259

YEARS 1918–1938

	Estimate	SE	TStat	PValue
1	25.2475	5.74289	4.3963	0.000310319
t	−0.0109162	0.00297866	−3.66481	0.00164665

E.Var. 0.00683176

YEARS 1947–1978

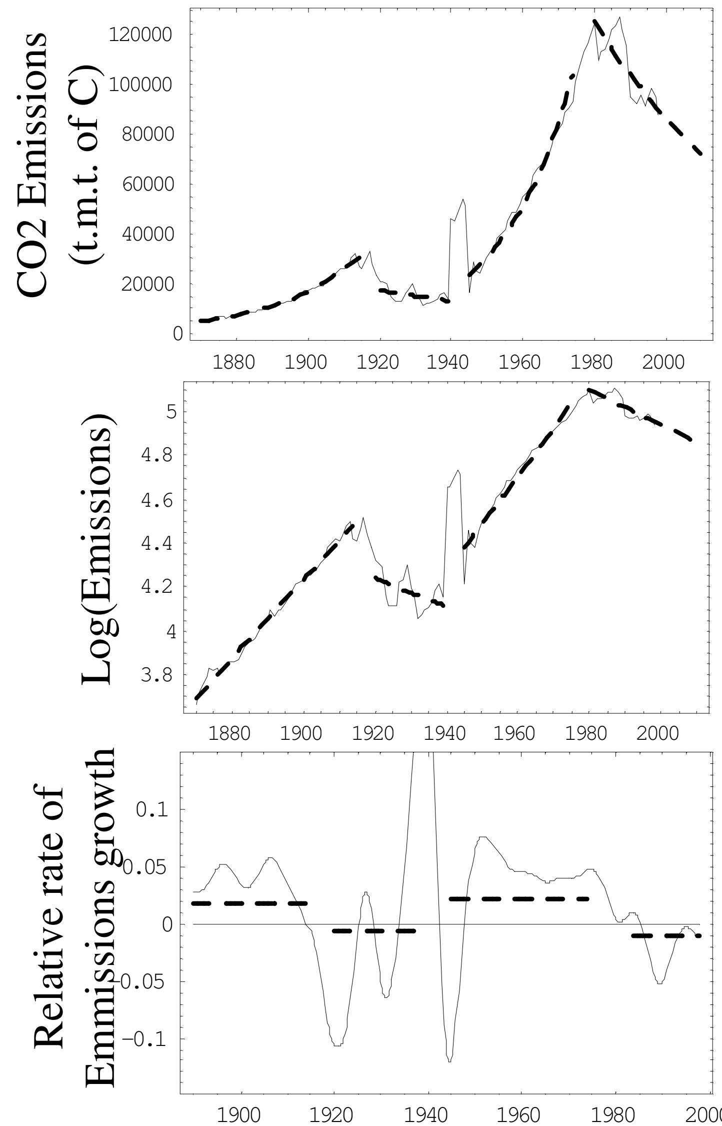
	Estimate	SE	TStat	PValue
1	−35.7734	0.935288	−38.2486	0.
t	0.0206553	0.000476575	43.3411	0.

E.Var. 0.000619593

YEARS 1979–1998

	Estimate	SE	TStat	PValue
1	20.1063	2.60431	7.72036	4.05251×10^{-7}
t	−0.00758245	0.00130968	−5.78953	0.0000173922

E.Var. 0.00114065



Smoothing vs. parametric model

Uncertainty measure

Years	1950-1998		1970-1998		~ 2000
Country	smooth.	param.	smooth.	param	reported
Belgium	2.3	3.3	2.3	3.3	1.1
Finland	4.8	1.3	3.8	3.6	3.0
France	2.3	3.0	2.3	1.1	<2.5
Ireland	4.3	1.2	2.2	2.2	<1.0
Netherlands	2.8	0.9	3.7	1.4	2.5
Sweden	2.5	1.1	2.3	1.4	1.0
U.K	1.6	0.5	1.4	0.7	1.1

Comparison of standard deviation estimates, in percents

Geometric Brownian motion (1)

- Stochastic equation for GBM process:

$$dx = g\,x\,dt + \sigma\,x\,dz$$

$$dz = \varepsilon\,dt^{1/2}$$

$x(t)$ – signal

dz – Wiener increment

ε – standard normal distribution

g – drift (rate of growth)

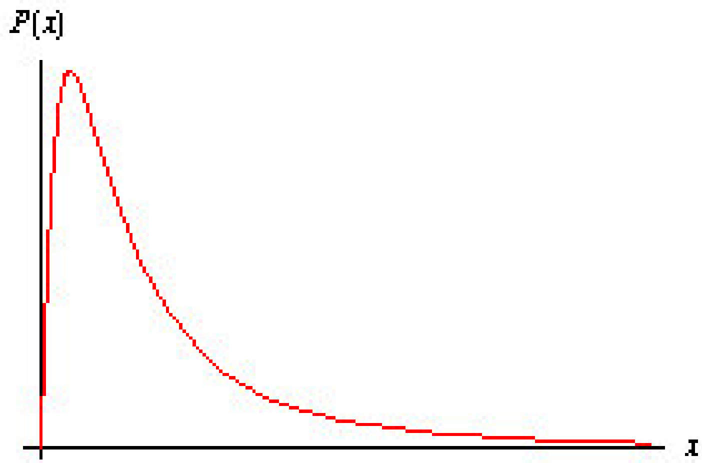
σ – volatility of x

Geometric Brownian motion (2)

- GBM - lognormal diffusion:

$$E[x(t)] = x(t_0) e^{g t}$$

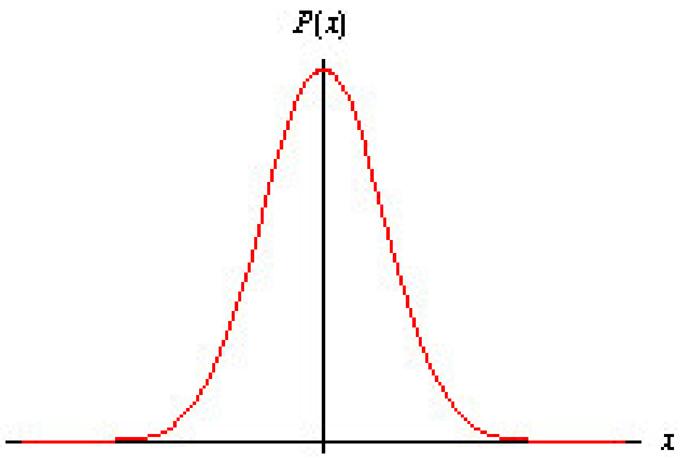
$$SD[x(t)] = x(t_0) e^{g t}[e^{\sigma^2 t} - 1]^{1/2}$$



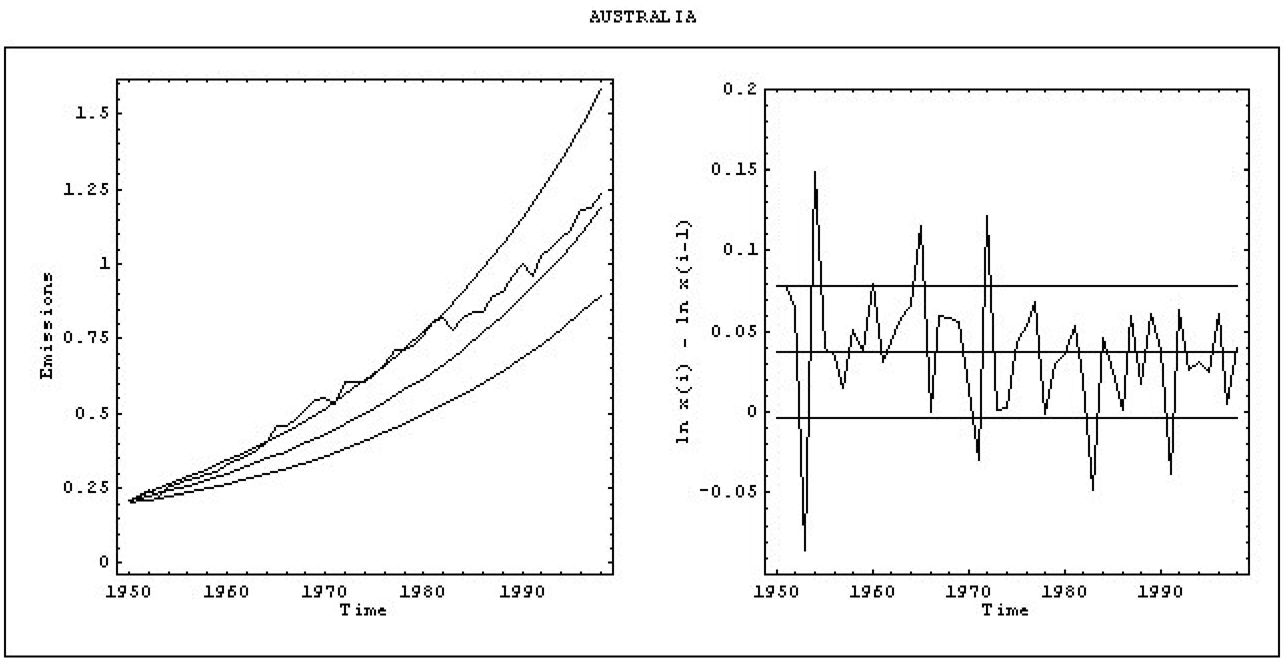
- BM - normal diffusion:

$$E[\ln x(t)] = \ln x(t_0) + (g - \tfrac{1}{2}\sigma^2) t$$

$$SD[\ln x(t)] = \sigma t^{1/2}$$



Geometric Brownian motion (3)



CO_2 emissions as illustration of the considered GBM process

Conclusions

- Empirical approach gives reasonable estimates of uncertainties, comparable to the aggregated ones.
- The methods proposed estimates the stochastic part of the error.
- Emissions have the piecewise exponential character, related to the economic development.
- The nonparametric method gives more smooth curves in many instances, but it is more sensitive to the smoothing interval.
- The parametric piecewise exponential model gives more rough but also simple description, showing general trends in evolution of emission data.