Experiences from the Application of a Parameter Estimation and Identifiability Analysis Methodology to the Operational Street Pollution Model (OSPM)

Thor-Bjørn Ottosen[1,2] Matthias Ketzel[3] Ole Hertel[3] Henrik Skov[2,3] Jørgen Brandt[3] Konstantinos Kakosimos[1]

[1]Department of Chemical Engineering, Texas A&M University at Qatar

[2]Department of Chemical Engineering, Biotechnology and Environmental Technology, University of Southern Denmark

[3]Department of Environmental Science, Aarhus University

16th International Conference on Harmonisation within Atmosperic Dispersion Modelling for Regulatory Purposes

Overview

1 Motivation

Why uncertainty and sensitivity analysis?

- Role of models has changed – need for uncertainty and sensitivity analysis

Why uncertainty and sensitivity analysis?

- Role of models has changed – need for uncertainty and sensitivity analysis
- Guide research directions

1 Motivation

Why uncertainty and sensitivity analysis?

- Role of models has changed – need for uncertainty and sensitivity analysis
- Guide research directions
- Improve scientific communication – transparency

1 Motivation

Why uncertainty and sensitivity analysis?

- Role of models has changed – need for uncertainty and sensitivity analysis
- Guide research directions
- Improve scientific communication – transparency
- Has not been performed for air quality models before

1 Motivation

Only little tradition for uncertainty and sensitivity analysis within atmospheric dispersion modelling.

Possible explanations:

- Is computational intensive

1 Motivation

Only little tradition for uncertainty and sensitivity analysis within atmospheric dispersion modelling.
Possible explanations:

- Is computational intensive
- Originated in other branches of science (ecological modelling, econometrics, engineering etc.)

Only little tradition for uncertainty and sensitivity analysis within atmospheric dispersion modelling.

Possible explanations:

- Is computational intensive
- Originated in other branches of science (ecological modelling, econometrics, engineering etc.)
- Lack of demonstration of feasibility and advantage of approach

2 Objectives

Methodology of choice: The iterative parameter estimation and identifiability analysis of Brun et al. (2001). It has been widely applied in other branches of science.

Aim: To analyse the applicability of applying the parameter estimation and identifiability analysis methodology to a model within atmospheric science, in this case the Operational Street Pollution Model (OSPM^{TM}).

The method consists of ten steps (Brun et al., 2002):

1. Define model

3 Methods

The method consists of ten steps (Brun et al., 2002):

1. Define model
2. Specify experimental layout

The method consists of ten steps (Brun et al., 2002):

1. Define model
2. Specify experimental layout
3. Prior analysis

3 Methods

The method consists of ten steps (Brun et al., 2002):

1. Define model
2. Specify experimental layout
3. Prior analysis
4. Calculate model output

3 Methods

The method consists of ten steps (Brun et al., 2002):

1. Define model
2. Specify experimental layout
3. Prior analysis
4. Calculate model output
5. Compute sensitivities

3 Methods

The method consists of ten steps (Brun et al., 2002):

1. Define model
2. Specify experimental layout
3. Prior analysis
4. Calculate model output
5. Compute sensitivities
6. Produce parameter importance ranking

The method consists of ten steps (Brun et al., 2002):

1. Define model
2. Specify experimental layout
3. Prior analysis
4. Calculate model output
5. Compute sensitivities
6. Produce parameter importance ranking
7. Assess identifiability of parameter subsets

3 Methods

The method consists of ten steps (Brun et al., 2002):

1. Define model
2. Specify experimental layout
3. Prior analysis
4. Calculate model output
5. Compute sensitivities
6. Produce parameter importance ranking
7. Assess identifiability of parameter subsets
8. Choose parameter subsets for manual tuning or parameter estimation

The method consists of ten steps (Brun et al., 2002):

1. Define model
2. Specify experimental layout
3. Prior analysis
4. Calculate model output
5. Compute sensitivities
6. Produce parameter importance ranking
7. Assess identifiability of parameter subsets
8. Choose parameter subsets for manual tuning or parameter estimation
9. Tune selected parameters manually or perform parameter estimation

Iterate until convergence

The method consists of ten steps (Brun et al., 2002):

1. Define model
2. Specify experimental layout
3. Prior analysis
4. Calculate model output
5. Compute sensitivities
6. Produce parameter importance ranking
7. Assess identifiability of parameter subsets
8. Choose parameter subsets for manual tuning or parameter estimation
9. Tune selected parameters manually or perform parameter estimation
10. Explore potential bias problems

Iterate until convergence (steps 4–9)

3 Methods

Identifiability analysis: Search for parameters identifiable from data

For a parameter to be identifiable:

- The model have to be sensitive to small changes in the parameter $\left(\left|\frac{\partial f(u,\theta)}{\partial \theta}\right| \gg 0\right)$
- Only small compensation effects in the model output. Analysed via linear dependency of the sensitivity matrix.

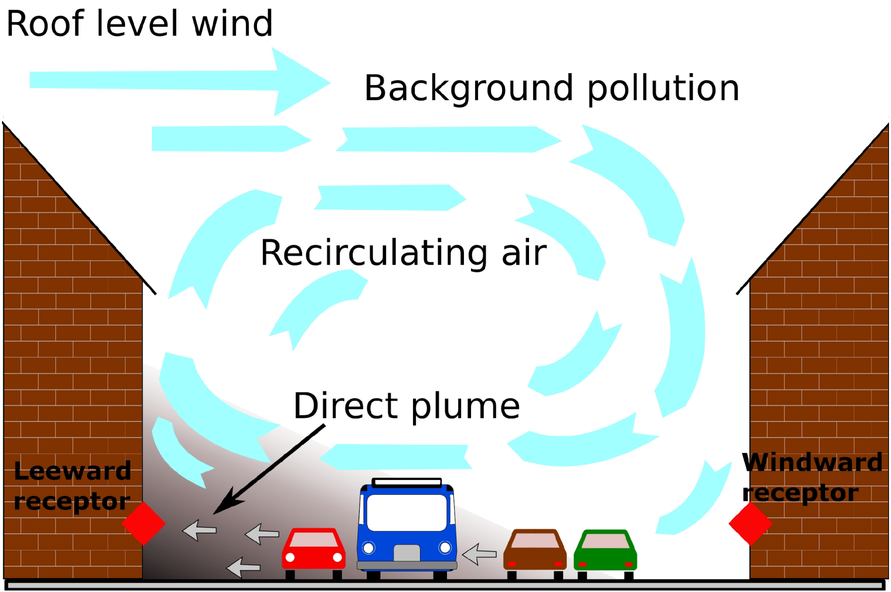
Parameter estimation: Estimation of parameters from data

- Used least-squares minimization criterium
- Weighted to balance species

Model input:

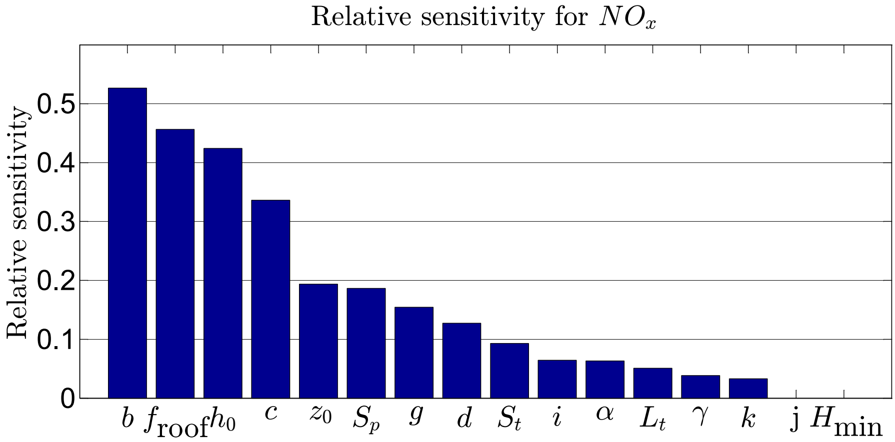
- Species: NO_x and NO_2
- Years: 1994–2010
- Five streets in major cities in Denmark
- Meteorology and urban background concentration from rooftop stations
- Measurements from Ellermann et al. (2013)
- Two data splitting approaches: DUPLEX and Seasonal

4 The Operational Street Pollution Model (OSPM^{TM})

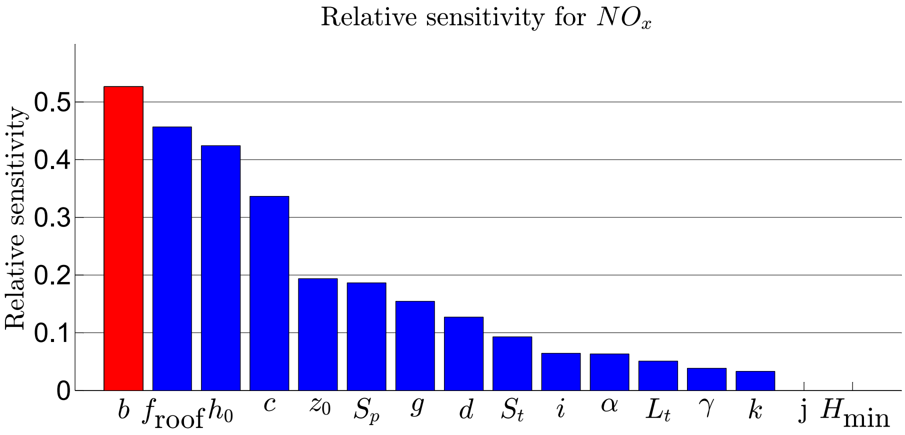


5 Results

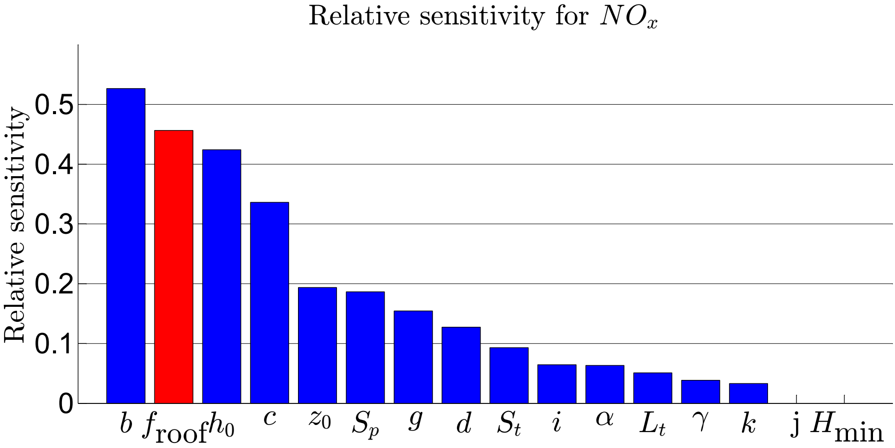
Local sensitivity analysis



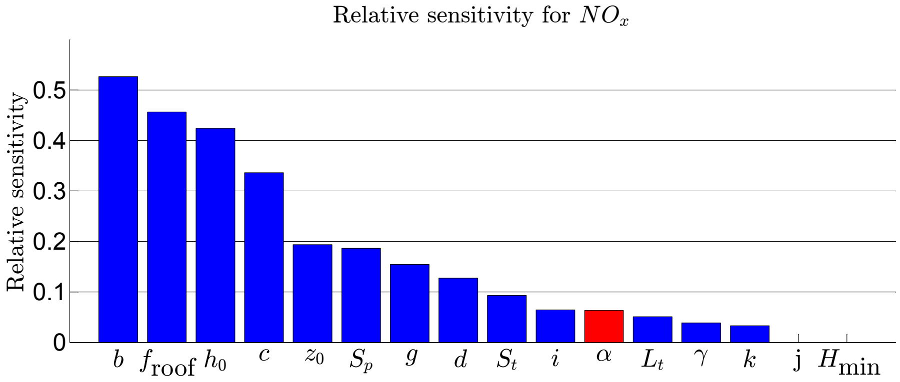
Distinct trend in sensitive and non-sensitive parameters – dependent on parameter values



Scale-factor for traffic-produced turbulence



$$f_{\mathrm{roof}} = \frac{u_{\mathrm{mast}}}{u_{\mathrm{roof}}}$$



Slope of emission plume

Size	Combinations	γ_K range	$\gamma_K < 10$ (%)	Parameters subset for $\gamma_{\min}$

5 Results

Collinearity analysis

Size	Combi-nations	γ_K range	$\gamma_K < 10$ (%)	Parameters subset for $\gamma_{\min}$
2	91	0.71–9.29	100.0	c, L_t
3	364	0.78–42.28	97.8	c, L_t, i
4	1001	0.84–42.30	93.3	c, L_t, i, k
5	2002	1.00–42.30	86.8	d, g, i, k, γ
6	3003	1.07–42.33	78.4	α, L_t, g, i, k, γ
7	3432	1.13–42.43	68.4	α, L_t, d, g, i, k, γ
8	3003	1.53–42.48	56.6	α, L_t, d, g, i, S_t, k, γ
9	2002	2.08–42.50	43.7	α, L_t, d, g, i, S_p, S_t, k, γ
10	1001	2.66–42.52	30.2	α, c, L_t, d, g, i, S_p, S_t, k, γ
11	364	4.92–42.52	17.3	α, c, L_t, d, z_0, g, i, S_p, S_t, k, γ
12	91	8.41–42.53	6.6	α, c, L_t, d, f_{roof}, z_0, g, i, S_p, S_t, k, γ
13	14	20.47–42.54	0.0	α, c, L_t, d, f_{roof}, h_0, z_0, g, i, S_t, b, k, γ
14	1	42.54–42.54	0.0	α, c, L_t, d, f_{roof}, h_0, z_0, g, i, S_p, S_t, b, k, γ

Size	Combi-nations	γ_K range	$\gamma_K < 10$ (%)	Parameters subset for γ_{min}
2	91	0.71–9.29	100.0	c, L_t
3	364	0.78–42.28	97.8	c, L_t, i
4	1001	0.84–42.30	93.3	c, L_t, i, k
5	2002	1.00–42.30	86.8	d, g, i, k, γ
6	3003	1.07–42.33	78.4	α, L_t, g, i, k, γ
7	3432	1.13–42.43	68.4	α, L_t, d, g, i, k, γ
8	3003	1.53–42.48	56.6	α, L_t, d, g, i, S_t, k, γ
9	2002	2.08–42.50	43.7	α, L_t, d, g, i, S_p, S_t, k, γ
10	1001	2.66–42.52	30.2	α, c, L_t, d, g, i, S_p, S_t, k, γ
11	364	4.92–42.52	17.3	α, c, L_t, d, z_0, g, i, S_p, S_t, k, γ
12	91	8.41–42.53	6.6	α, c, L_t, d, f_{roof}, z_0, g, i, S_p, S_t, k, γ
13	14	20.47–42.54	0.0	α, c, L_t, d, f_{roof}, h_0, z_0, g, i, S_t, b, k, γ
14	1	42.54–42.54	0.0	α, c, L_t, d, f_{roof}, h_0, z_0, g, i, S_p, S_t, b, k, γ

Size	Combi-nations	γ_K range	$\gamma_K < 10$ (%)	Parameters subset for $\gamma_{\min}$
2	91	0.71–9.29	100.0	c, L_t
3	364	0.78–42.28	97.8	c, L_t, i
4	1001	0.84–42.30	93.3	c, L_t, i, k
5	2002	1.00–42.30	86.8	d, g, i, k, γ
6	3003	1.07–42.33	78.4	α, L_t, g, i, k, γ
7	3432	1.13–42.43	68.4	α, L_t, d, g, i, k, γ
8	3003	1.53–42.48	56.6	α, L_t, d, g, i, S_t, k, γ
9	2002	2.08–42.50	43.7	α, L_t, d, g, i, S_p, S_t, k, γ
10	1001	2.66–42.52	30.2	α, c, L_t, d, g, i, S_p, S_t, k, γ
11	364	4.92–42.52	17.3	α, c, L_t, d, z_0, g, i, S_p, S_t, k, γ
12	91	8.41–42.53	6.6	α, c, L_t, d, f_{roof}, z_0, g, i, S_p, S_t, k, γ
13	14	20.47–42.54	0.0	α, c, L_t, d, f_{roof}, h_0, z_0, g, i, S_t, b, k, γ
14	1	42.54–42.54	0.0	α, c, L_t, d, f_{roof}, h_0, z_0, g, i, S_p, S_t, b, k, γ

Parameter estimation

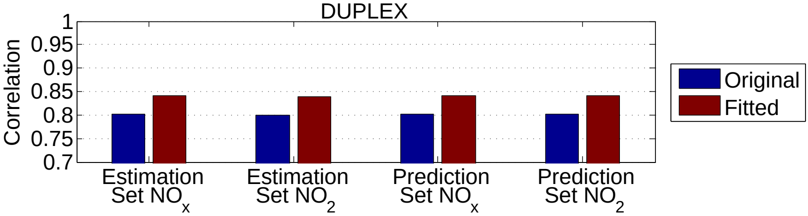
θ	Original value	Limits	Estimate Seasonal	Estimate DUPLEX	% Difference	95 % CL % of mean θ

5 Results

Parameter estimation

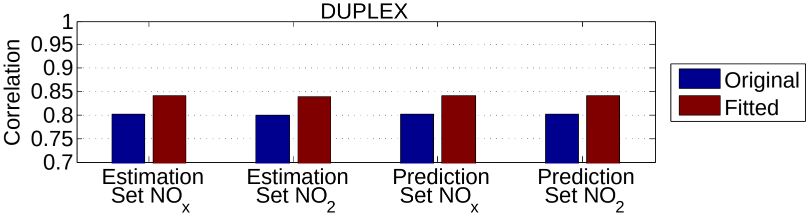
θ	Original value		Limits	Estimate Seasonal	Estimate DUPLEX	% Difference	95 % CL % of mean θ
b	0.3		$[10^{-5} : 0.999]$	0.212	0.288	30.5	± 0.6
f_{roof}	0.4		$[10^{-5} : 0.999]$	0.427	0.422	1.2	± 1.2
h_0	2.0	m	$[0.6 : 10]$	2.177	1.422	42.0	± 0.9
c	2.0		$[0.25 : 10.00]$	6.607	5.685	15.0	± 1.5
S_P	2.0	m^2	$[10^{-5} : 10]$	1.198	0.360	107.5	± 2.5
g	2.0	$\frac{\mathrm{m}}{\mathrm{s}}$	$[10^{-5} : 10]$	0.116	0.085	31.5	± 18.5
d	0.5		$[10^{-5} : 2\pi]$	1.873	1.105	51.6	± 0.3
i	1.0	$\frac{\mathrm{m}}{\mathrm{s}}$	$[10^{-5} : 10]$	0.455	0.713	44.3	± 1.4
α	0.1		$[0.05 : 2.00]$	0.277	0.292	5.5	± 1.0
L_t	0.5		$[10^{-5} : 0.999]$	0.008	$1.335 \cdot 10^{-5}$	199.3	$\pm 1.3 \cdot 10^5$
γ	0.2		$[10^{-5} : 0.999]$	0.789	0.017	191.5	± 52.1
k	0.4		$[0.04 : 0.999]$	0.999	0.999	0.0	± 8.7

In general: More sensitivity $\rightarrow$ less uncertainty

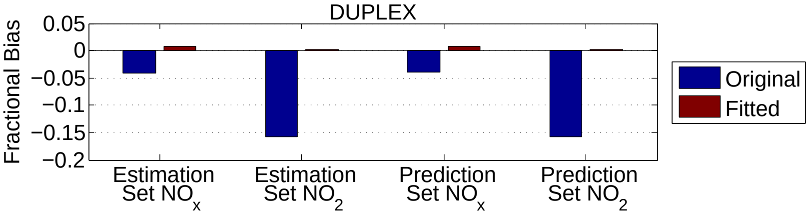


Approximately homogeneous correlation before and after

Validation of estimated parameters – Statistics

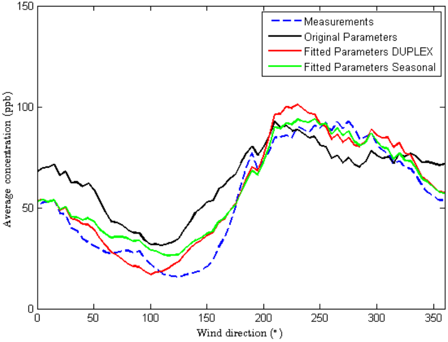


Approximately homogeneous correlation before and after

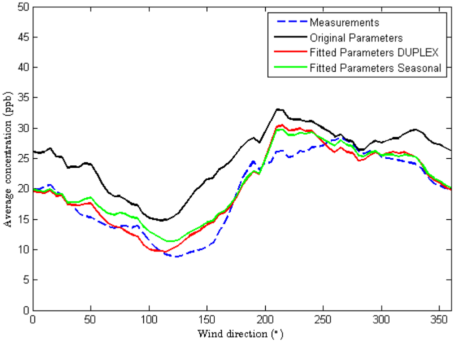


More homogeneous and reduced fractional bias after fit

Wind direction plot: Albanigade, Odense

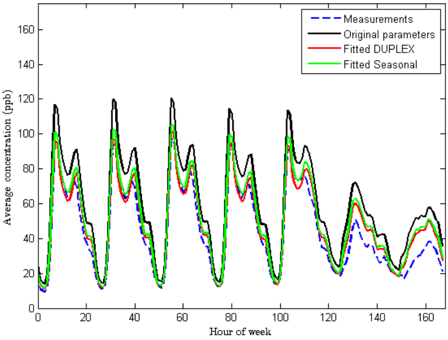


(a) NO_x

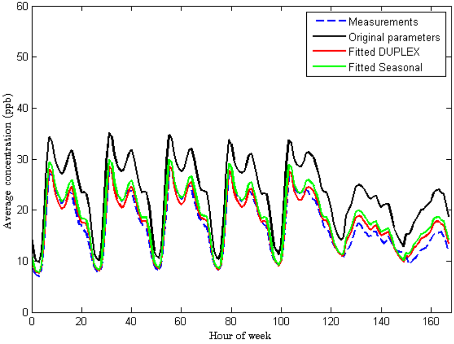


(b) NO_2

Diurnal average plot: Albanigade, Odense



(c) NO_x



(d) NO_2

The methodology:

- Successfully estimates a set of parameters with better performance statistics compared to the original model parameters

The methodology:

- Successfully estimates a set of parameters with better performance statistics compared to the original model parameters
- Balances the fractional bias among the individual streets and species

6 Conclusions

The methodology:

- Successfully estimates a set of parameters with better performance statistics compared to the original model parameters
- Balances the fractional bias among the individual streets and species
- Converts other types of uncertainty into parameter uncertainty

6 Conclusions

The methodology:

- Successfully estimates a set of parameters with better performance statistics compared to the original model parameters
- Balances the fractional bias among the individual streets and species
- Converts other types of uncertainty into parameter uncertainty
- Have significantly increased the transparency about the model performance and model uncertainty

6 Conclusions

The methodology:

- Successfully estimates a set of parameters with better performance statistics compared to the original model parameters
- Balances the fractional bias among the individual streets and species
- Converts other types of uncertainty into parameter uncertainty
- Have significantly increased the transparency about the model performance and model uncertainty

Other experiences:

- The 95% confidence intervals underestimates uncertainties

Further reading:

Thor-Bjørn Ottosen, Matthias Ketzel, Henrik Skov, Ole Hertel, Jørgen Brandt, and Konstantinos Kakosimos (2014). "A Parameter Estimation and Identifiability Analysis Methodology Applied to a Street Canyon Air Pollution Model". In: *Environmental Modelling & Software* (in preparation)

Acknowledgements

- Thomas Ellermann for access to measurements
- Jeremy D. Silver for constructive feedback
- Research computing facilities for access to high performance computing resources

Bibliography

Brun, Roland, Peter Reichert, and Hans R. Künsch (2001). "Practical identifiability analysis of large environmental simulation models". In: *Water Resources Research* 37.4, pp. 1015–1030.

Brun, Roland, Martin Kühni, Hansruedi Siegrist, Willi Gujer, and Peter Reichert (2002). "Practical identifiability of AMS2d parameters—systematic selection and tuning of parameter subsets". In: *Water Research* 36, pp. 4113–4127.

Ellermann, Thomas, Jakob Klenø Nøjgaard, Claus Nordstrøm, Jørgen Brandt, Jesper Christensen, Matthias Ketzel, Stefan Jansen, Andreas Massling, and Steen Solvang Jensen (2013). *The Danish Air Quality Monitoring Programme. Annual Summary for 2012*. Tech. rep. Scientific Report from DCE. no. 67. Arhus University, DCE – Danish Centre for Environment and Energy.

Ottosen, Thor-Bjørn, Matthias Ketzel, Henrik Skov, Ole Hertel, Jørgen Brandt, and Konstantinos Kakosimos (2014). "A Parameter Estimation and Identifiability Analysis Methodology Applied to a Street Canyon Air Pollution Model". In: *Environmental Modelling & Software* (in preparation).