RUNNING A LARGE SCALE AIR-POLLUTION MODEL ON THE EPCC SUPERCOMPUTERS

Tzvetan Ostromsky

Central Laboratory for Parallel Processing, Bulgarian Academy of Sciences, Acad. G. Bonchev str., bl. 25-A, 1113 Sofia, Bulgaria
e-mail: *tto@epcc.ed.ac.uk*

A joint work with: **Zahari Zlatev**
Nat. Environmental Research Institute, Frederiksborgvej 399, DK–4000 Roskilde, Denmark
e-mail: *zz@dmu.dk*

EPCC Contact: **Mario Antonioletti**
The Univ. of Edinburgh, Mayfield Rd, EH9 3JZ
e-mail: *mario@epcc.ed.ac.uk*

Hosted by: **Asif Usmani**
Dept. of Civil and Environmental Engineering, The Univ. of Edinburgh, Mayfield Rd, EH9 3JZ
e-mail: *asif.usmani@ed.ac.uk*

Mathematical background of the air pollution modeling

A system of PDE for calculating the concentrations of a number of chemical species (pollutants*) in the atmosphere (3-D domain!).

The main physical and chemical processes should be adequately represented in the system:

1. horizontal wind *(advection)*;
2. *horizontal diffusion*;
3. *chemical reactions and emissions*;
4. *deposition (dry and wet)*;
5. *vertical transport* (convection and diffusion).

A huge computational task, because of the:

- size of the domain;
- dynamics of the processes – small time step;
- complexity of the equations – decomposition

*Some of these are natural components of the air.

The Danish Eulerian Model

$$\frac{\partial c_s}{\partial t} = -\frac{\partial(uc_s)}{\partial x} - \frac{\partial(vc_s)}{\partial y} - \frac{\partial(wc_s)}{\partial z}$$

$$+ \frac{\partial}{\partial x}\left(K_x \frac{\partial c_s}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial c_s}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_z \frac{\partial c_s}{\partial z}\right)$$

$$+ E_s + Q_s(c_1, c_2, \dots c_q) - (k_{1s} + k_{2s})c_s$$

- q – the number of chemical species;
- c_s – the concentrations $(s = 1, \dots q)$;
- $u,\ v,\ w$ – the wind components along the coordinate axes;
- $K_x,\ K_y,\ K_z$ – diffusion coefficients;
- E_s – the emissions;
- $k_{1s},\ k_{2s}$ – dry / wet deposition coefficients;
- $Q_s(c_1, c_2, \dots c_q)$ – non-linear functions describing the chemical reactions between the species (the condensed CBM IV scheme, proposed by Gery et al. (1979) is used).

Splitting into submodels

$$\frac{\partial c_s^{(1)}}{\partial t} = -\frac{\partial (u c_s^{(1)})}{\partial x} - \frac{\partial (v c_s^{(1)})}{\partial y}$$ advection

$$\frac{\partial c_s^{(2)}}{\partial t} = \frac{\partial}{\partial x}\left(K_x \frac{\partial c_s^{(2)}}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial c_s^{(2)}}{\partial y}\right)$$ horizontal diffusion

$$\frac{\partial c_s^{(3)}}{\partial t} = E_s + Q_s(c_1^{(3)}, c_2^{(3)}, \ldots c_q^{(3)})$$ chemistry & emissions

$$\frac{\partial c_s^{(4)}}{\partial t} = -(k_{1s} + k_{2s}) c_s^{(4)}$$ deposition

$$\frac{\partial c_s^{(5)}}{\partial t} = -\frac{\partial (w c_s^{(5)})}{\partial z} + \frac{\partial}{\partial z}\left(K_z \frac{\partial c_s^{(5)}}{\partial z}\right)$$ vertical transport

Related work:

Strang - 1968

Marchuk - 1982

McRay, Goodin and Seinfeld - 1982

Lancer and Verwer - 1999

Dimov, Farago and Zlatev - 1999

Space discretization and numerical treatment of the five submodels

— Five large ODE systems*:

$$\frac{dg^{(i)}}{dt} = f^{(i)}(t, g^{(i)}) , \quad i = 1, 2, 3, 4, 5$$

$$g^{(i)} \in \mathcal{R}^{N_x \times N_y \times N_z \times q}$$
$$f^{(i)} \in \mathcal{R}^{N_x \times N_y \times N_z \times q}$$

- N_x, N_y, N_z – the number of grid-points along the coordinate axes (different versions);
- $q = 35$ – the number of chemical species considered in the model;
- the components of the $g^{(i)}$ – vector-functions, approximating the concentrations of the different species throughout the spatial domain.

* In general, these systems are simpler. For example, the chemical submodel system splits into $N_x \times N_y \times N_z$ independent ODE systems of size q.

Size of the different versions presented

- Coarse grid 2-D version (step 50 km.):
  $N_x = N_y = 96\ ,\ \ N_z = 1$
  — 4 ODE systems of order 322 560 each (no vertical exchange);

- Refined grid 2-D version (step 10 km.):
  $N_x = N_y = 480\ ,\ \ N_z = 1$
  — 4 ODE systems of more than 8 million equations;

- Coarse grid 3-D version (step 50 km.):
  $N_x = N_y = 96\ ,\ \ N_z = 10$
  — 5 ODE systems of order 3 225 600.

More than 20 000 primary time-steps are to be carried out to cover a time period of 1 month.

Smaller time-step must be used in the chemical submodel.

Numerical methods used in the chemical submodel

QSSA† (Hesstevedt et al. - 1978)

Backward Euler algorithm

Trapezoidal Rule

Runge-Kutta methods (Zlatev - 1981)

Rosenbrock methods (Verwer et al. - 1998)

Native parallel tasks: The calculations in a single grid-point.

†Quazi Steady-State Algorithm

Numerical methods used in the advection-diffusion part

Pseudo-spectral discretization (Zlatev - 1984)

Finite elements (Pepper et al. - 1979)

Finite differences (up-wind)

"Positive" methods (Bott - 1989, Holm - 1994)

Semi-Lagrangian methods (Neta - 1995)

Native parallel tasks: The calculations for a given pollutant.

Input data sets

Emission data: More or less constant over a long time period. Loaded at the begining, updated with internal constants if necessary.

$\leftarrow$ Antropogenic emissions: SO_2, NO_x,
Antr. hydro-carbons, NH_3;

$\leftarrow$ Natural hydro-carbons emissions

Geographical data: Constant. Loaded at the begining.

$\leftarrow$ Latutudes and longitudes

$\leftarrow$ Land and sea arreas

Meteorological data: Quickly changing, updated every 6 hours. New data must be read regularly after certain number of time steps.

$\Leftarrow$ Wind velocities (horizontal wind)

$\Leftarrow$ Precipitations

$\Leftarrow$ Mixing hights

$\Leftarrow$ Temperatures

$\Leftarrow$ Humidities

$\Leftarrow$ Cloud cover

$\Leftarrow$ Surface temperatures

$\Leftarrow$ Vertical wind

Numerical results

- Versions of the model, used in the experiments:
  - OpenMP and MPI codes for $(96 \times 96 \times 1)$ grid;
  - MPI code for $(480 \times 480 \times 1)$ grid;
  - OpenMP code for $(96 \times 96 \times 10)$ grid.

- All experiments are for a time period of 1 month.

- Machines at EPCC, used in the experiments:
  - SUN HPC cluster: HPC 3500 UltraSPARC, 400 MHz (shared memory)
  - CRAY T3E-900 (distributed memory)

- Compiler options:
  - On the SUN cluster: *"-fast -O4 -xarch=v8plusa -xparallel"*.
  - On the T3E: *"-O 3 "*.

- Some former results obtained on an IBM SP2 (distributed memory) and a SGI Origin-2000 (shared memory) computers at INI-C‡ are also given for comparison.

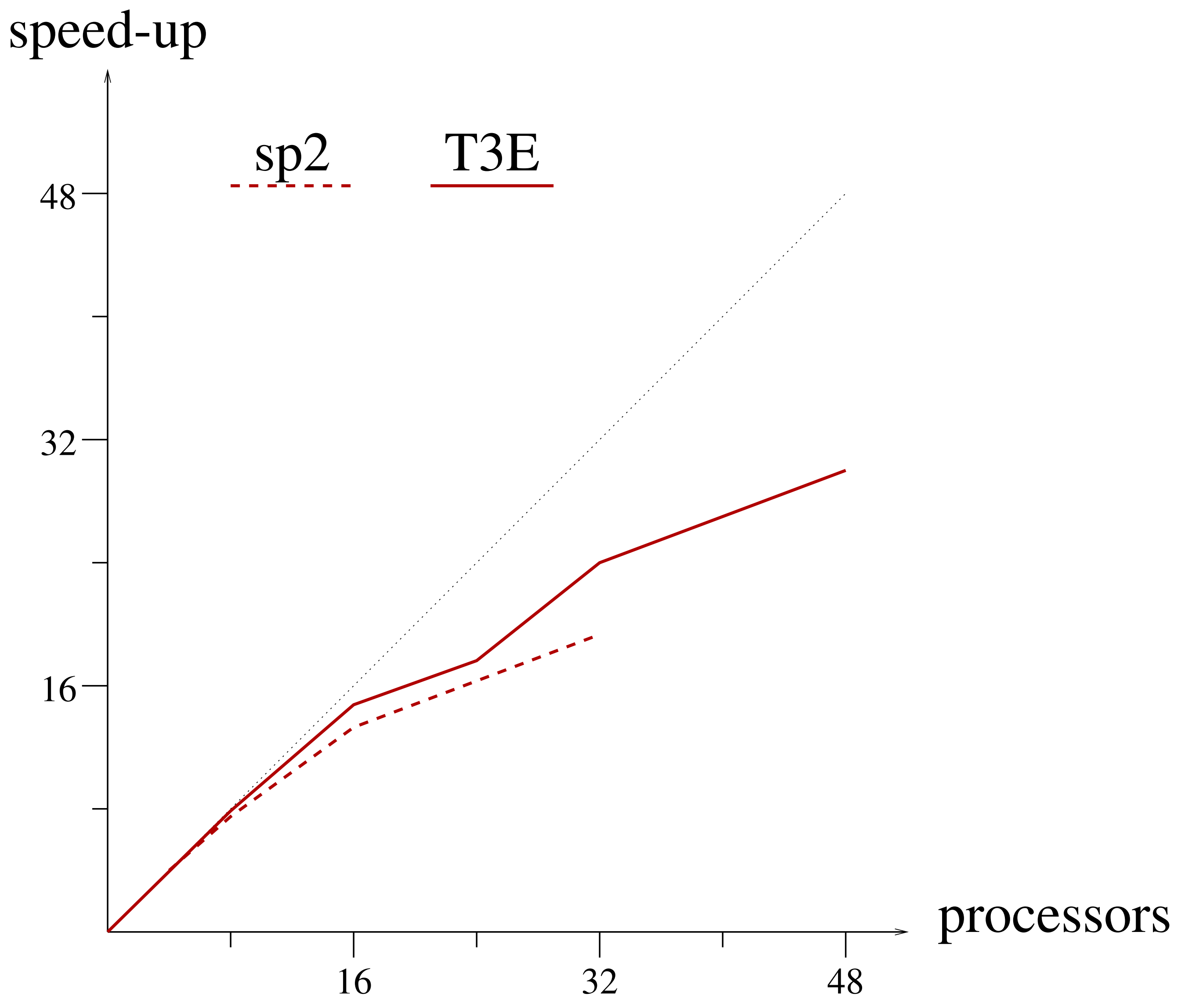
‡Danish Computing Centre for Research and Education

Times and performance of the 2-D coarse-grid codes on the SUN cluster

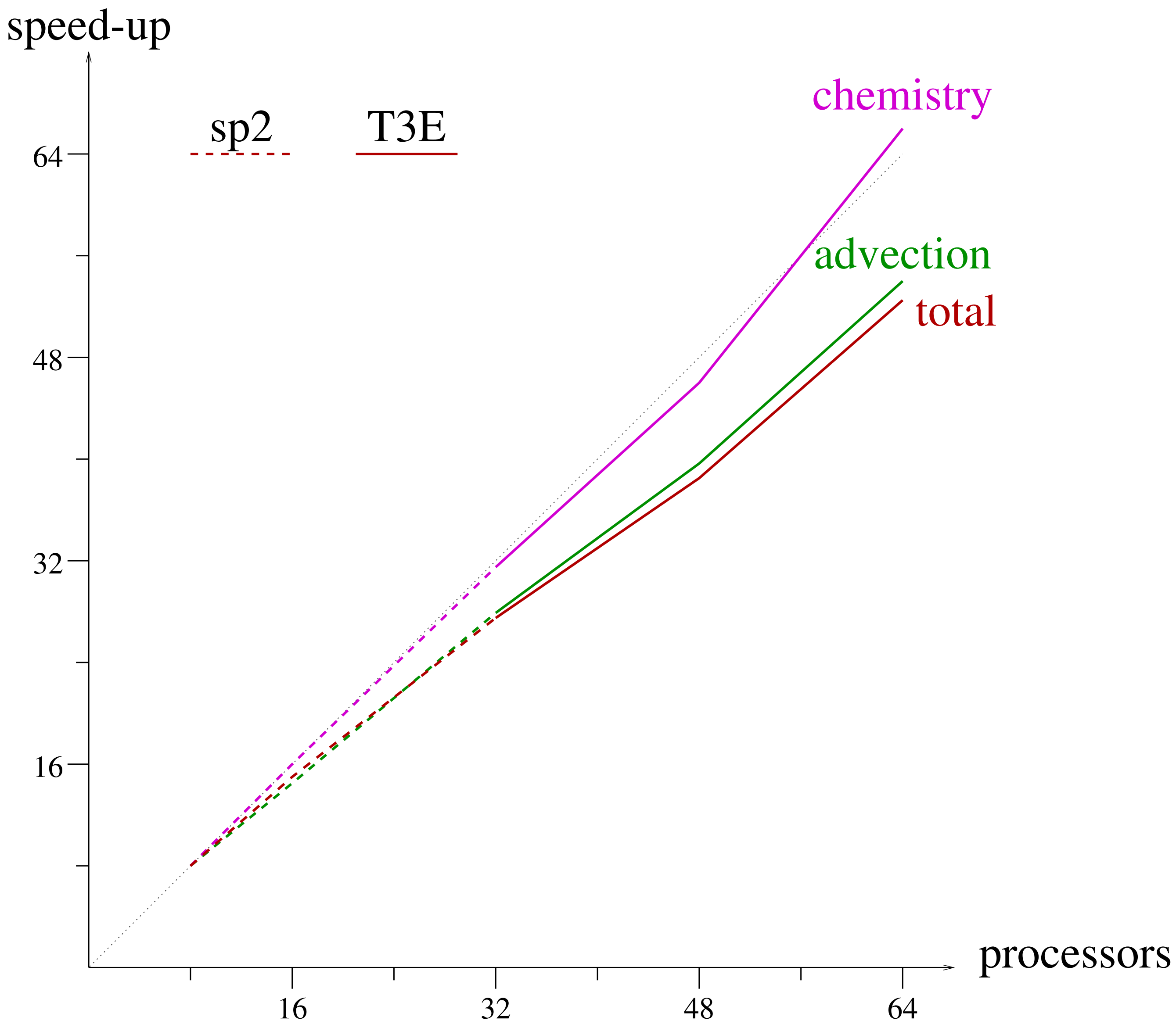
no. of proc.	OpenMP code		MPI code	
	Time [s]	Speed -up	Time [s]	Speed -up
1	4356		5067	
4	1199	3.6	1293	3.9
8	636	6.9	629	8.0
16	391	11.2	363	14.0

Time and speed-up of the OpenMP and the MPI versions of DEM for $(96 \times 96 \times 1)$ grid on the SUN HPC cluster at EPCC.

Scalability of the MPI code for the $(96 \times 96 \times 1)$ grid



Scalability of the MPI code for the $(480 \times 480 \times 1)$ grid



Times and performance of the 3-D code

Stage	SUN E3500		SGI Origin-2000	
	Time [s]	Part [%] of Total	Time [s]	Part [%] of Total
W+S (Input)	105	0.3 %	59	0.1 %
Adv.+Diff.	8602	22.9 %	10458	21.1 %
Chem.+Dep.	25381	67.6 %	35232	71.2 %
Vertical tr.	3307	8.8 %	3465	7.0 %
Output	202	0.5 %	277	0.6 %
Total	37565	100 %	49483	100 %

Time profile for sequential runs of the 3D version of DEM on two shared-memory supercomputers

Stage	Time [s]	% of Total	Speed-up
W+S (Input)	200	3.5 %	0.52
Adv.+Diff.	1244	22.0 %	6.91
Chem.+Dep.	3674	65.0 %	6.91
Vertical tr.	416	7.4 %	7.95
Output	219	3.9 %	0.92
Total	5656	100 %	6.64

Time profile of the 3D version of DEM, running on 8 proc. SUN E3500 at EPCC

Scalability of the 3-D OpenMP code

nu. of PE's	Speed-up			
	Total	Adv.+Diff.	Chem.+Dep.	Vert. tr.
2	1.96	1.91	1.98	2.09
4	3.71	3.62	3.81	4.23
8	6.64	6.91	6.91	7.95
16	10.27	10.71	11.53	18.90

Some applications

Calculating the excess ozone (AOT40) over Europe under different scenarios:

- basic_normal - the man-made (anthropogenic emissions are as in 1995 provided by EMEP (Oslo), while the biogenic (natural) emissions are as in Simpson et al., 1995
- basic_high - the man-made emissions are as in the previous scenario, while the natural emissions are as in Anastasi et al., 1991 (roughly speaking, about five times higher than in the previous scenario)
- s2010_normal - the expected man-made emissions in 2010 (which are considerably lower than the man made emissions for 1995) and the normal natural emissions (as in the first scenario)
- s2010_high - the man-made emissions as in the previous scenario, while the natural emissions are as in the second scenario.

SCALED AOT40 VALUES IN EUROPE

The period of time is May – July 1995.

AOT40 values in percentages: 100(AOT40/3000)

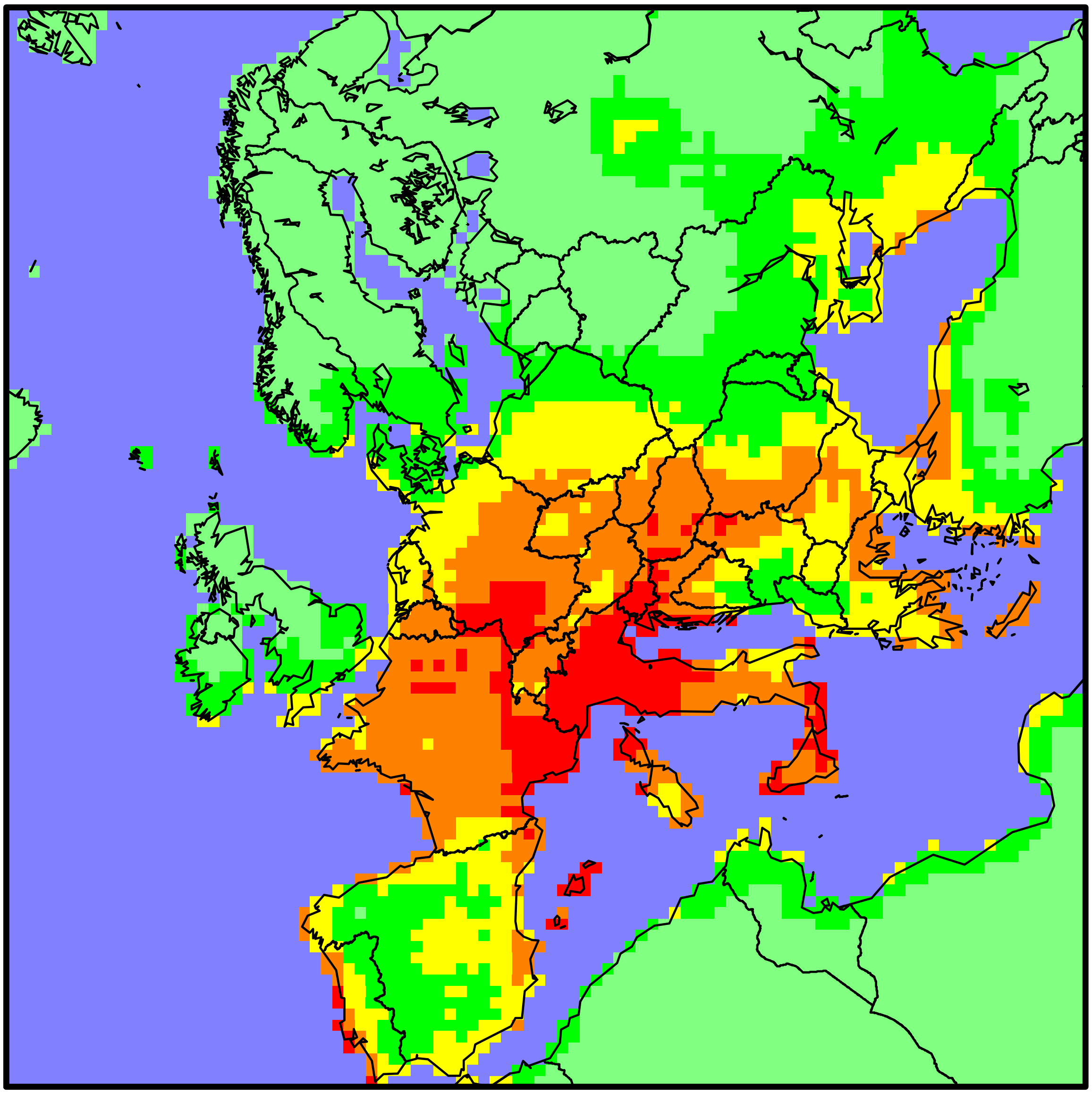
Run on a (96x96) grid / (50 km x 50 km) cells

Basic Scenario with Normal Biogenic emissions

Maximum value in the domain: 730

Minimal value in the domain: 2

Above 400
301 – 400
201 – 300
101 – 200
0 – 100
Water areas



SCALED AOT40 VALUES IN EUROPE

The period of time is May – July 1995.

AOT40 values in percentages: 100(AOT40/3000)

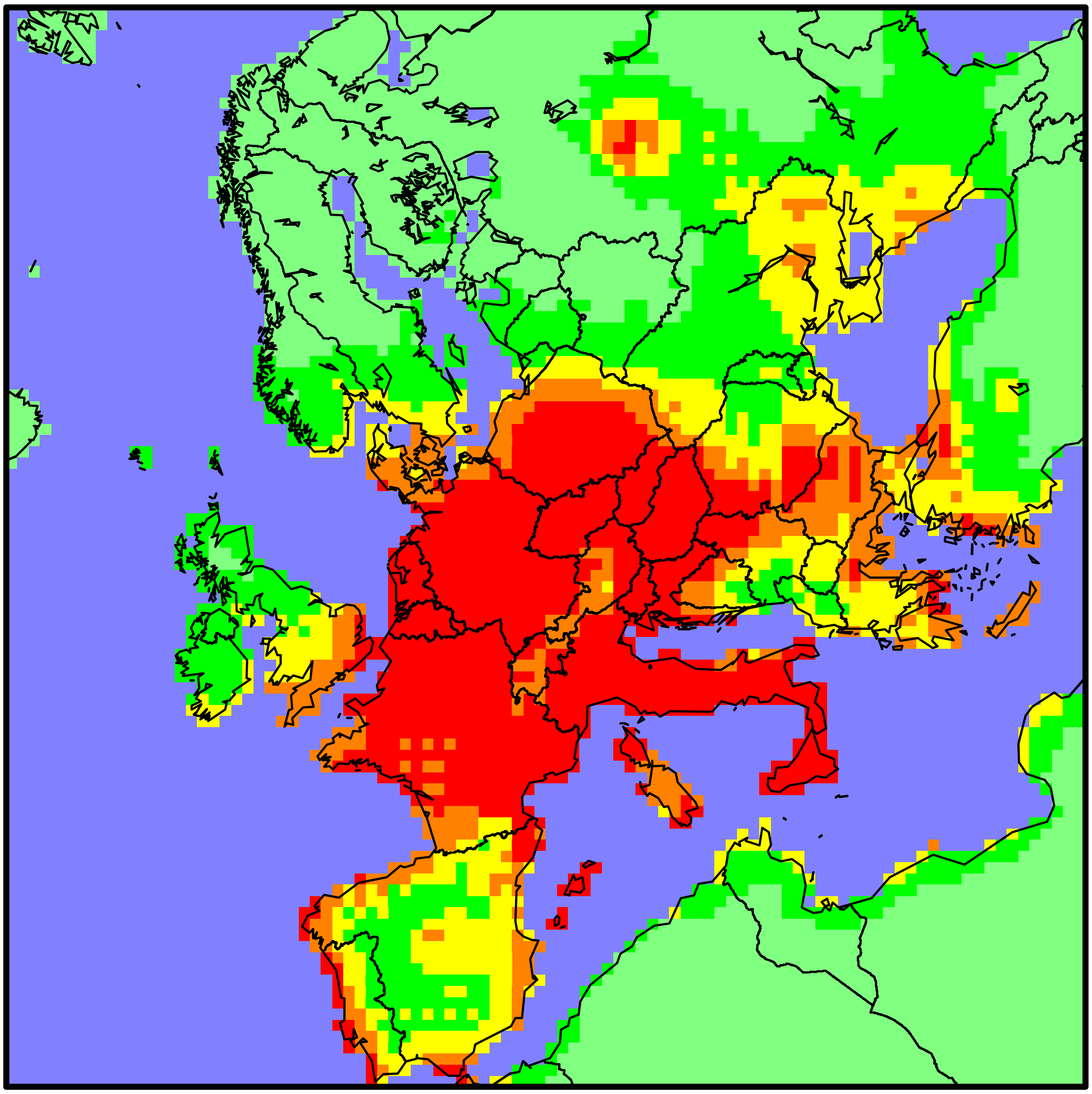
Run on a (96x96) grid / (50 km x 50 km) cells

Basic Scenario with High Biogenic emissions

Maximum value in the domain: 1105

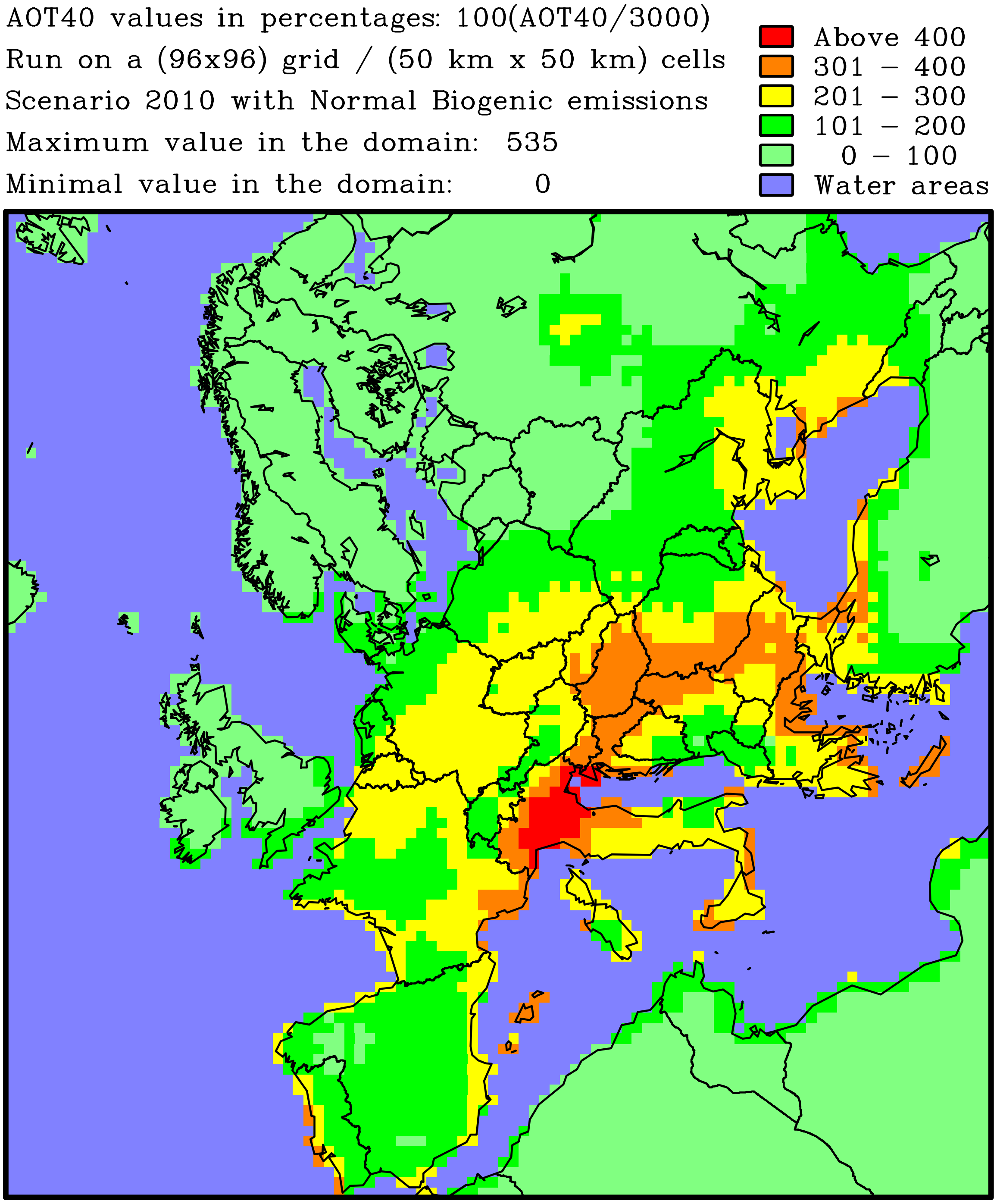
Minimal value in the domain: 3

Above 400
301 – 400
201 – 300
101 – 200
0 – 100
Water areas



SCALED AOT40 VALUES IN EUROPE

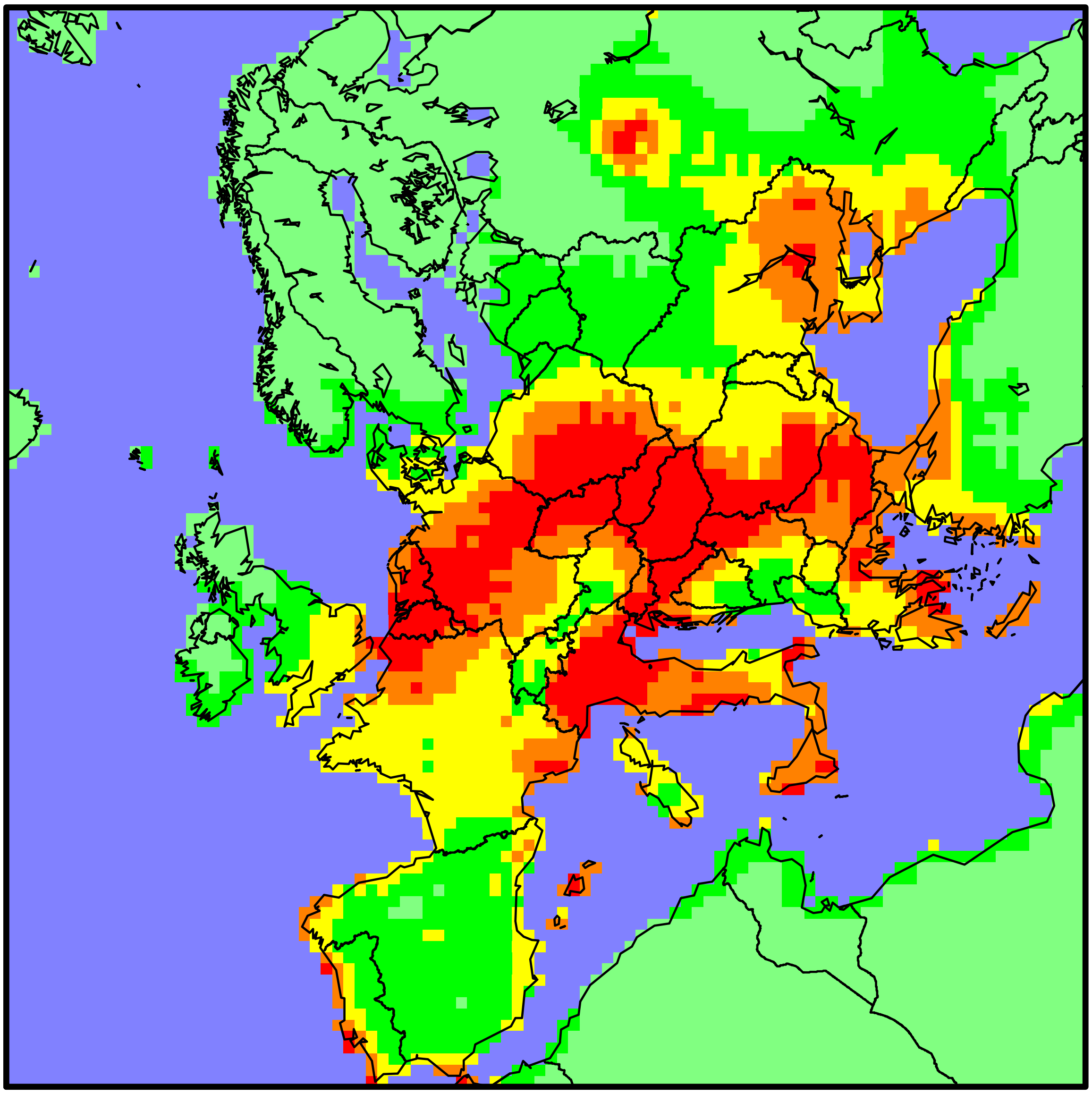
The period of time is May – July 1995.
AOT40 values in percentages: 100(AOT40/3000)
Run on a (96x96) grid / (50 km x 50 km) cells
Scenario 2010 with Normal Biogenic emissions
Maximum value in the domain: 535
Minimal value in the domain: 0



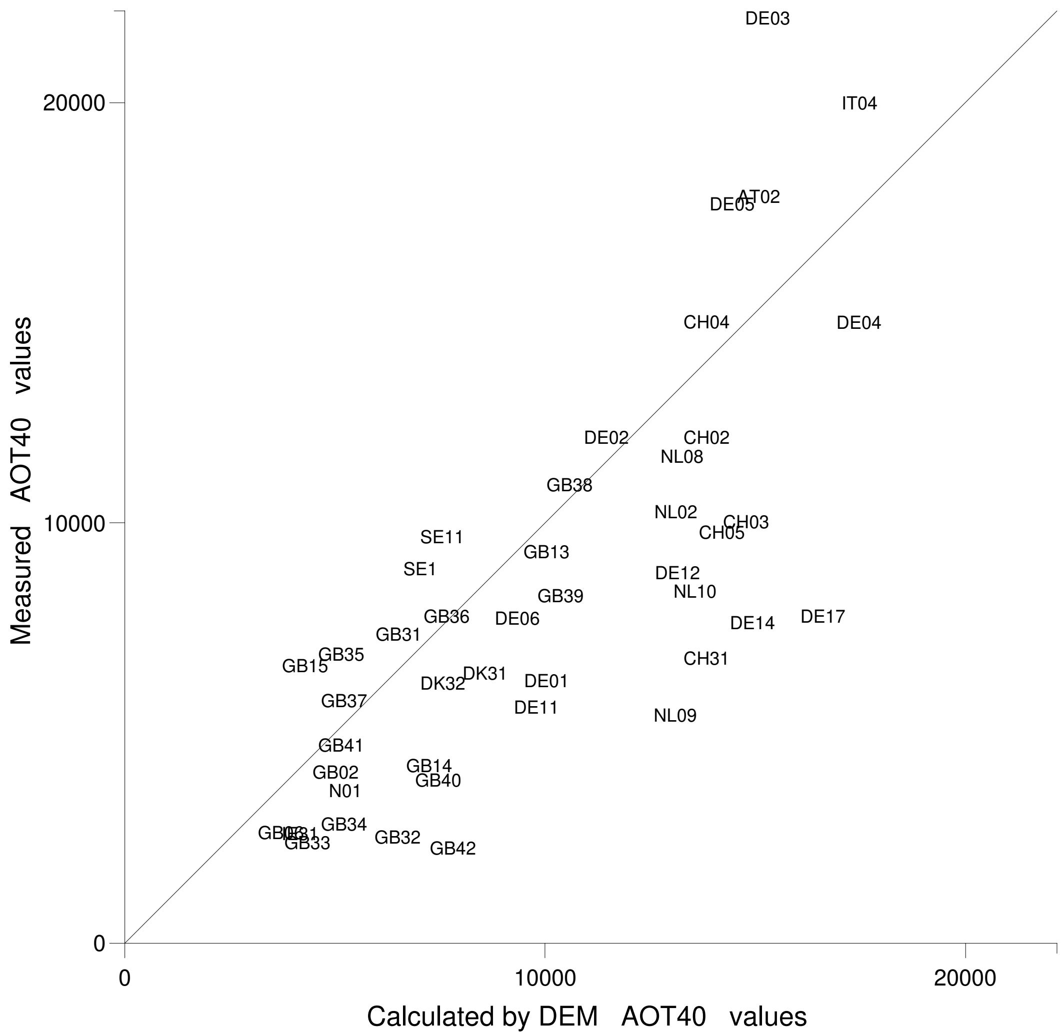
SCALED AOT40 VALUES IN EUROPE

The period of time is May – July 1995.
AOT40 values in percentages: 100(AOT40/3000)
Run on a (96x96) grid / (50 km x 50 km) cells
Scenario 2010 with High Biogenic emissions
Maximum value in the domain: 760
Minimal value in the domain: 0

Above 400	
301 – 400	
201 – 300	
101 – 200	
0 – 100	
Water areas	



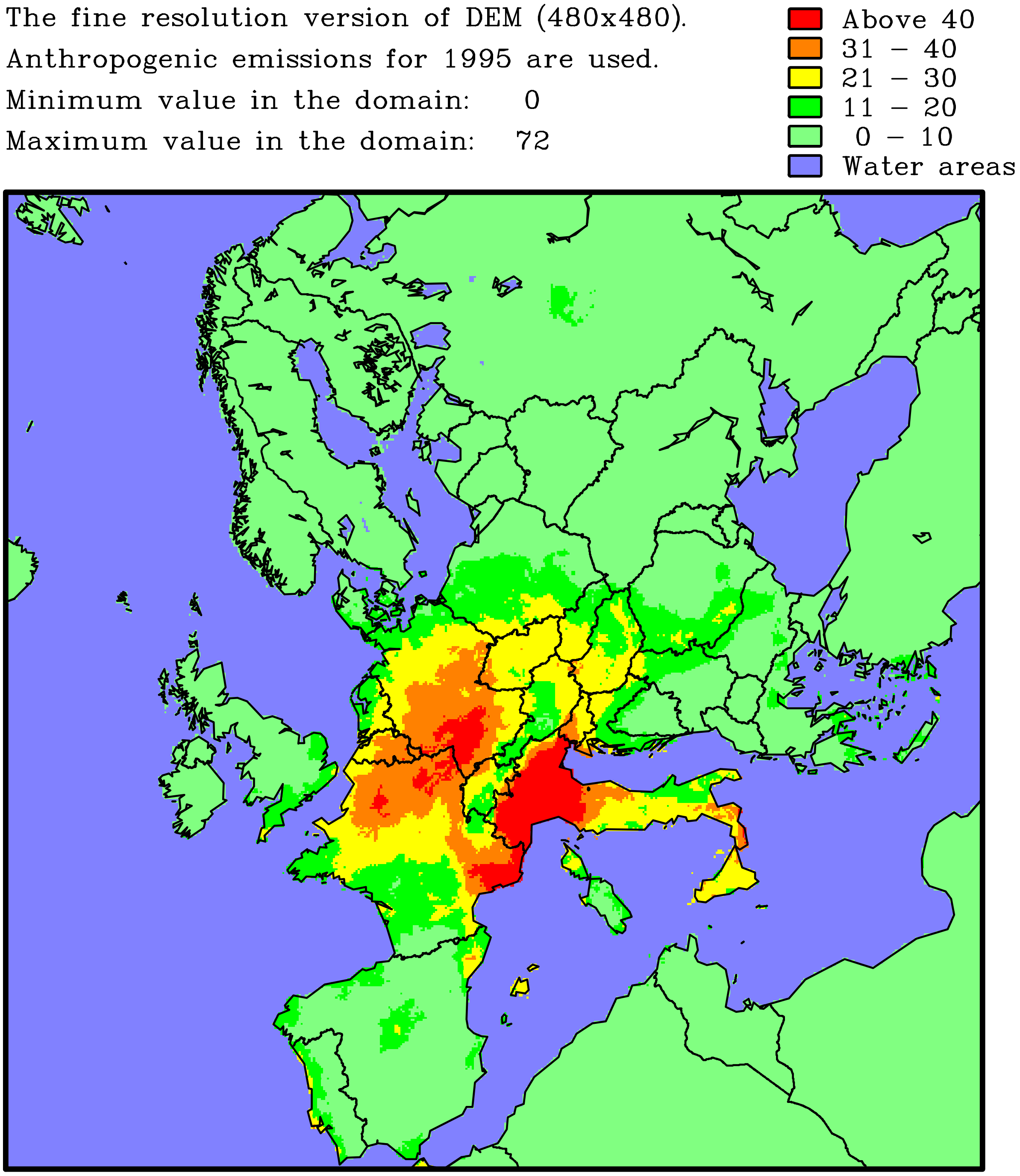
Measured vs calculated AOT40 values



COMPUTED MEAN, OBSERVED MEAN, CORRELATION FACTOR 10024.23 8496.89 0.75

EXPOSURE TO HIGH OZONE CONCENTRATIONS

Numbers of days in which 8−hour rolling averages
of ozone concentrations exceeded 60 ppb.
The fine resolution version of DEM (480x480).
Anthropogenic emissions for 1995 are used.
Minimum value in the domain: 0
Maximum value in the domain: 72



Plans for future work

- Optimization of the 3-D OpenMP code for $(96 \times 96 \times 10)$ grid .
- Development of a MPI version of the 3-D model.
- Obtaining results with new meteorological data sets and various emission scenarios.
- Development of a 3-D refined grid version of the model for $(480 \times 480 \times 10)$ grid for massively parallel computers (MPI).

Acknowledgments

This research was supported by EC via Access to Research Infrastructure action of the Improving Human Potential Programme (contract No. HPRI-1999-CT-00026). The results presented were obtained during a TRACS research visit at EPCC, the University of Edinburgh, where the author spent 3 months in the autumn 2000.

Special thanks to my EPCC contact Mario Antonioletti and to the whole TRACS team, who were always very kind and helped me a lot in this work.

REFERENCES:

1. K. Burrage, Parallel and sequential methods for ordinary differential equations, Oxford University Press, Oxford-New York, 1995.

2. M. W. Gery, G. Z. Whitten, J. P. Killus and M. C. Dodge, A photochemical kinetics mechanism for urban and regional modeling, J. Geophys. Res., **94**, 1989, 12925-12956.

3. G. I. Marchuk, Mathematical modeling for the problem of the environment, Studies in Mathematics and Applications, No. 16, North-Holland, Amsterdam, 1985.

4. G. J. McRae, W. R. Goodin and J. H. Seinfeld, Numerical solution of the atmospheric diffusion equations for chemically reacting flows, J. Comp. Phys., **45,** 1984, 1-42.

5. V. Alexandrov, A. Sameh, Y. Siddique and Z. Zlatev, Numerical integration of chemical ODE problems arising in air pollution models, Environmental Modelling and Assessment, to appear.

6. S. Skelboe and Z. Zlatev, Exploiting the natural partitioning in the numerical solution of ODE systems arising in atmospheric chemistry, In: "Numerical Analysis and Its Applications" (L. Vulkov, J. Wasniewski and P. Yalamov, eds.), pp. 458-465, Springer, Berlin, 1997.

7. Z. Zlatev, Computer treatment of large air pollution models, Kluwer Academic Publishers, Dordrecht-Boston-London, 1995.

8. Z. Zlatev, Partitioning ODE systems with an application to air pollution models, Internal report. National Environmental Research Institute, Department of Atmospheric Environment, Frederiksborgvej 399, P. O. Box 399, DK-4000 Roskilde, Denmark, 1997.