Computational Challenges in Air Pollution Modelling

Z. Zlatev

National Environmental Research Institute

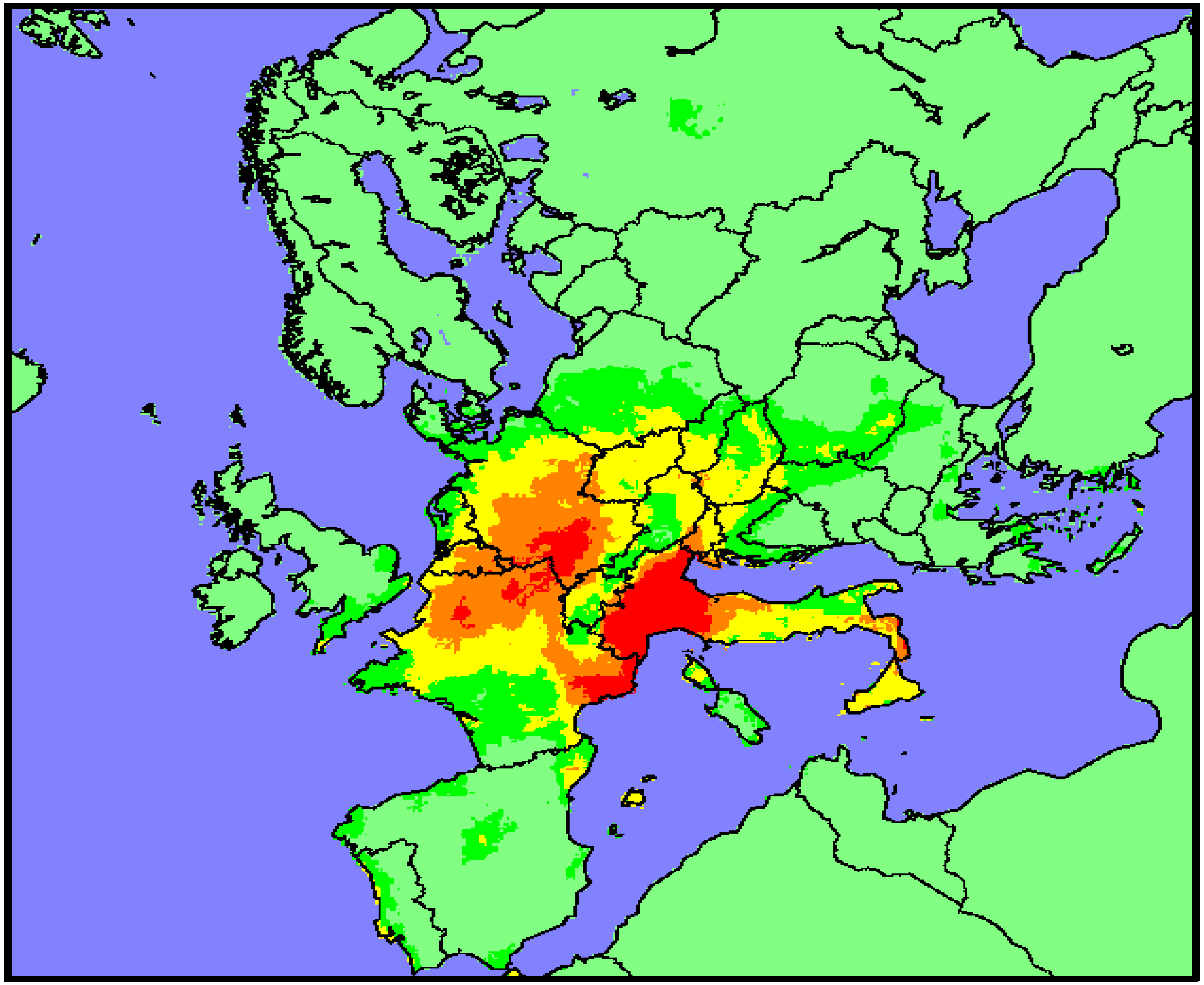
1. *Why air pollution modelling?*
2. *Major physical and chemical processes*
3. *Need for splitting*
4. *Computational difficulties*
5. *Need for faster and accurate algorithms*
6. *Different matrix computations*
7. *Inverse and optimization problems*
8. *Unresolved problems*

1. Why air pollution models?

- **Distribution** of the air pollution levels
- **Trends** in the development of air pollution levels
- Establishment of **relationships** between air pollution levels and key parameters (emissions, meteorological conditions, boundary conditions, etc.).
- **Predicting** appearance of high levels

EXPOSURE TO HIGH OZONE CONCENTRATIONS

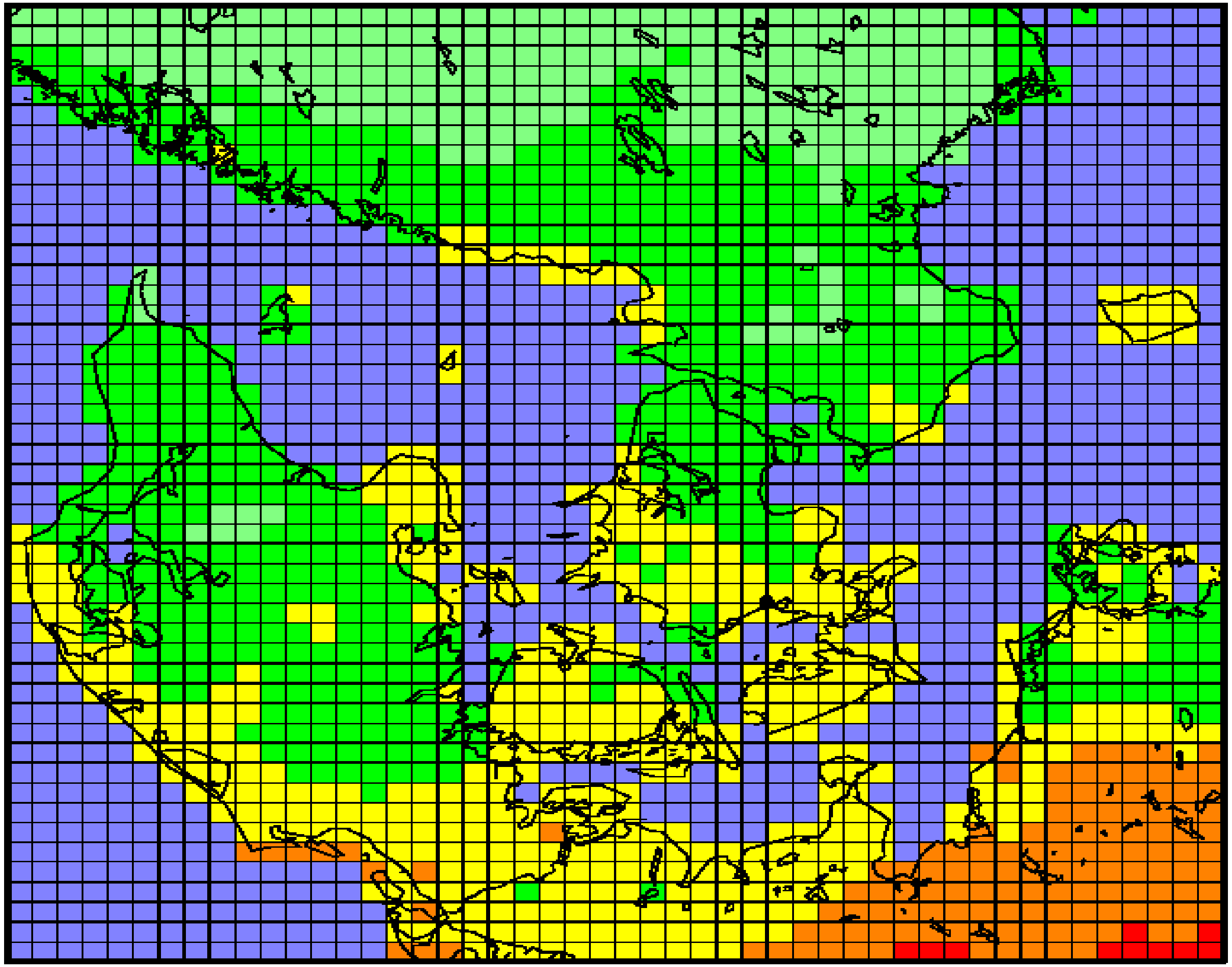
Numbers of days in which 8–hour rolling averages
of ozone concentrations exceeded 60 ppb.
The fine resolution version of DEM (480x480).
Anthropogenic emissions for 1995 are used.
Maximum value in the domain: 72



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.
Maximum value in the domain: 23

🟥 Above 20
🟧 16 – 20
🟨 11 – 15
🟩 6 – 10
🟩 0 – 5
🟦 Water areas



EXPOSURE TO HIGH OZONE CONCENTRATIONS

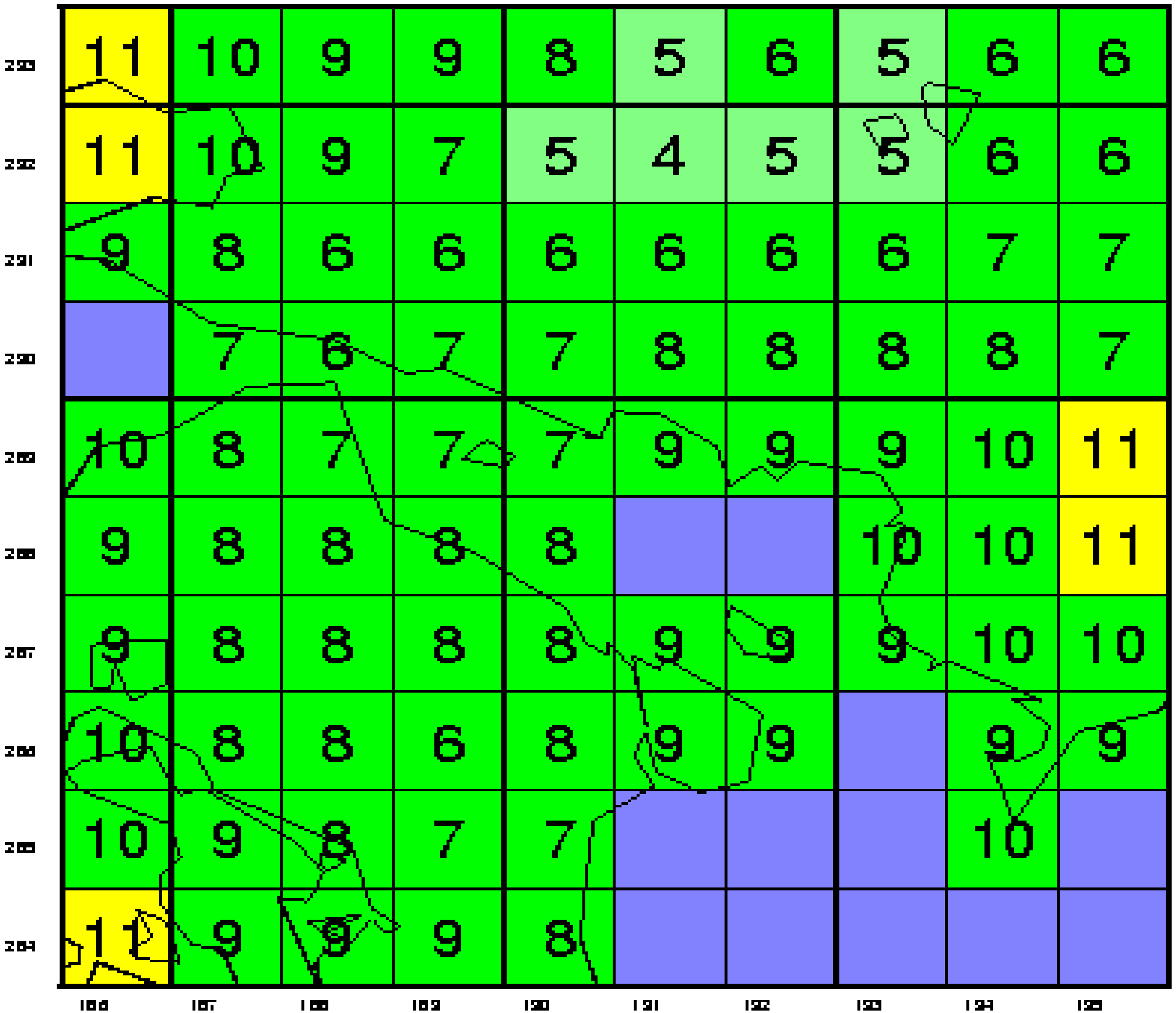
Numbers of days in which 8-hour rolling averages of ozone concentrations exceeded 60 ppb.

The fine resolution version of DEM (480x480).

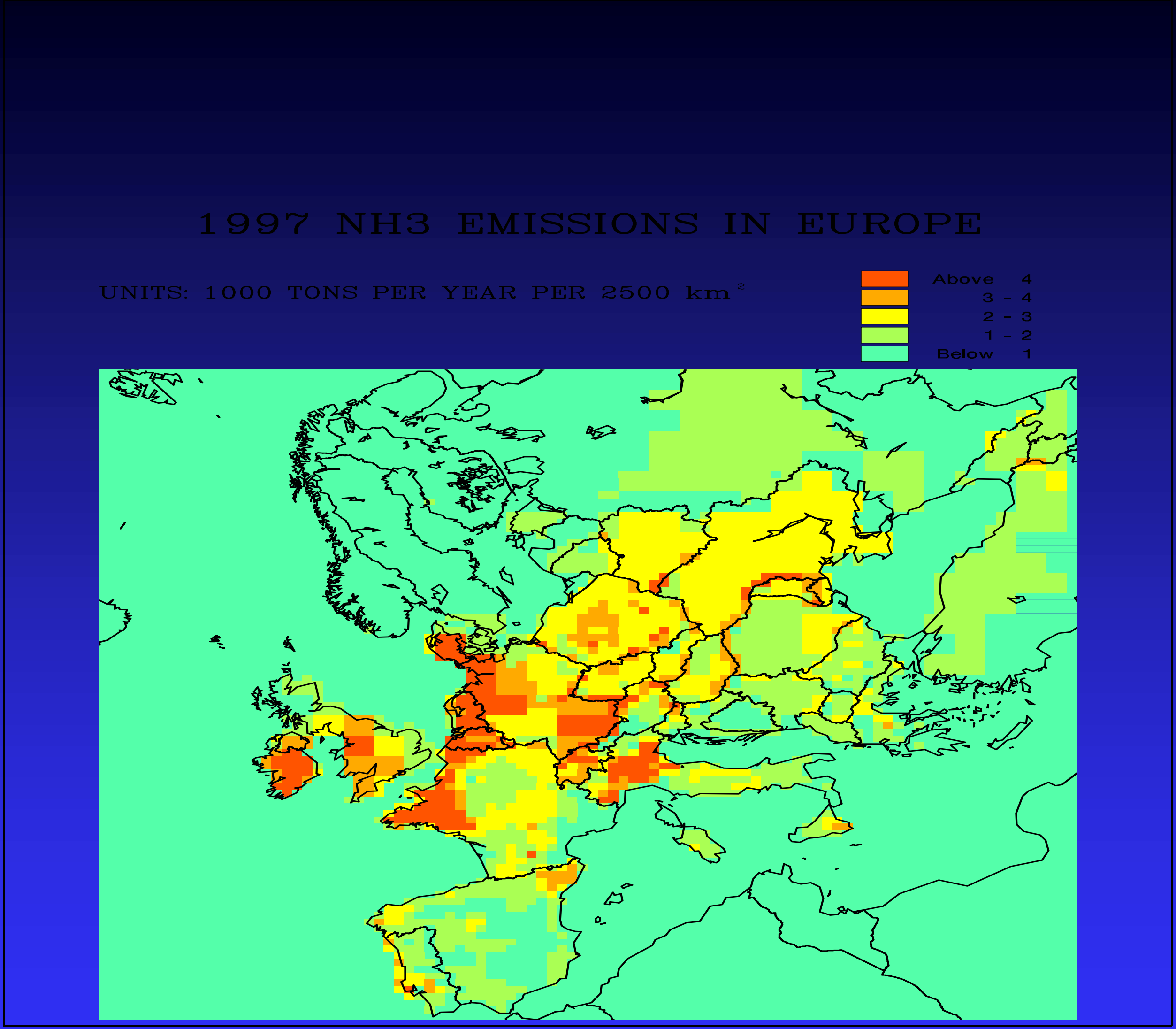
Anthropogenic emissions for 1995 are used.

Maximum value in the domain: 11

- 🟥 Above 20
- 🟧 16 – 20
- 🟨 11 – 15
- 🟩 6 – 10
- 🟩 0 – 5
- 🟦 Water areas



	186	187	188	189	190	191	192	193	194	195
293	11	10	9	9	8	5	6	5	6	6
292	11	10	9	7	5	4	5	5	6	6
291	9	8	6	6	6	6	6	6	7	7
290		7	6	7	7	8	8	8	8	7
289	10	8	7	7	7	9	9	9	10	11
288	9	8	8	8	8			10	10	11
287	9	8	8	8	8	9	9	9	10	10
286	10	8	8	6	8	9	9		9	9
285	10	9	8	7	7				10	
284	11	9	9	9	8					

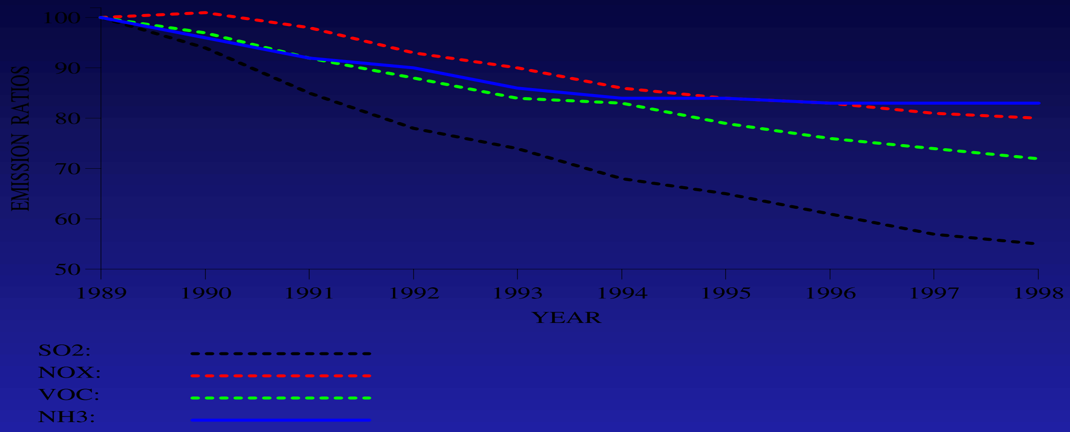

1997 NH3 EMISSIONS IN EUROPE

UNITS: 1000 TONS PER YEAR PER 2500 km^2

THE EUROPEAN EMISSIONS

IN THE PERIOD FROM 1989 TO 1998

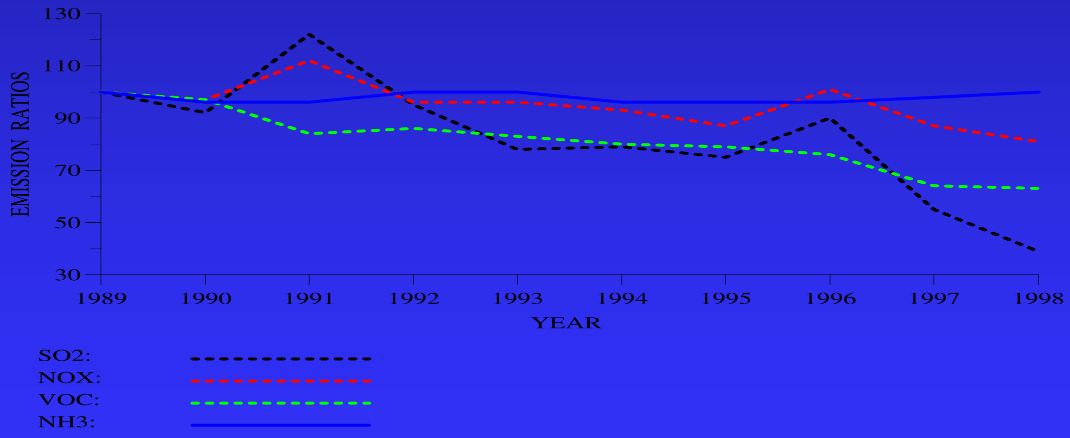
CHANGES (RELATIVE TO 1989) IN PERCENT

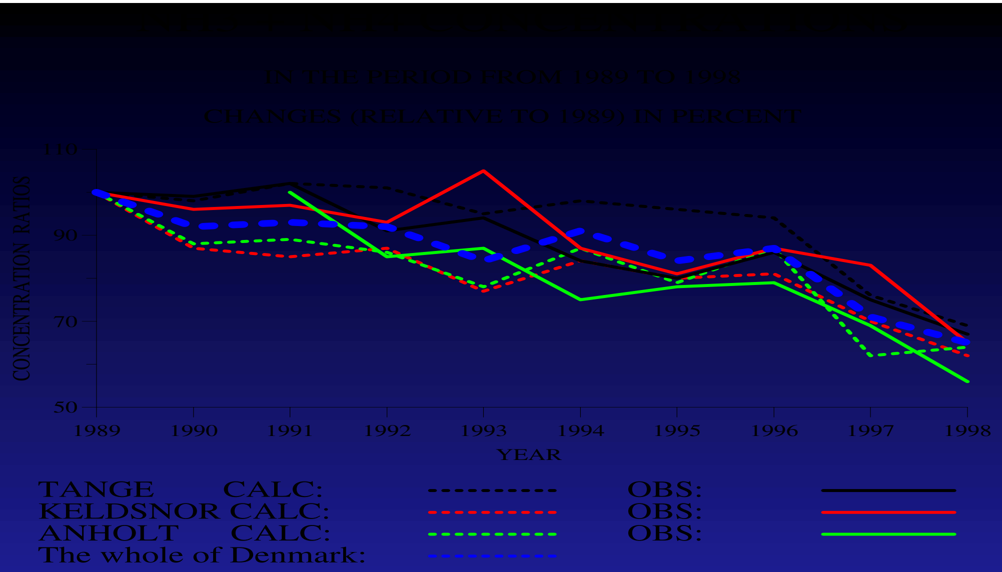


THE DANISH EMISSIONS

IN THE PERIOD FROM 1989 TO 1998

CHANGES (RELATIVE TO 1989) IN PERCENT





Country	1989	1998	Reduction
Germany	661	502	24%
The Netherlands	232	171	24%
Denmark	104	104	0%

Table 3 - NH3+NH4 concentrations in 1989 and 1998 in three European countries

2. Major physical processes

- **Horizontal transport** (advection)
- Horizontal diffusion
- Deposition (dry and wet)
- **Chemical reactions** + emissions
- **Vertical transport and diffusion**

Describe these processes mathematically

3. Air Pollution Models

$$\frac{\partial c_s}{\partial t} = -\frac{\partial (u c_s)}{\partial x} - \frac{\partial (v c_s)}{\partial y}$$ hor. transport

$$+\frac{\partial}{\partial x}\left(K_x \frac{\partial c_s}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial c_s}{\partial y}\right)$$ hor. diffusion

$$-(k_{1s} + k_{2s}) c_s$$ deposition

$$+E_s + Q_s(c_1, c_2, \ldots, c_q)$$ chem. + emis.

$$-\frac{\partial (w c_s)}{\partial z} + \frac{\partial}{\partial z}\left(K_z \frac{\partial c_s}{\partial z}\right)$$ vert. transport

$$s = 1, 2, \ldots, q$$

4. Need for splitting

- Bagrinowskii and Godunov 1957
- Strang 1968
- Marchuk 1968, 1982
- McRay, Goodin and Seinfeld 1982
- Lancer and Verwer 1999
- Dimov, Farago and Zlatev 1999

- Zlatev 1995

4. Criteria for choosing the splitting procedure

- Accuracy
- Efficiency
- **Preservation of the properties** of the involved operators

5. Resulting ODE systems

$$\frac{dg^{[i]}}{dt} = f^{[i]}\left(t, g^{[i]}\right) \qquad 1 \le i \le m,$$

$$g^{[i]} \in \Re^N, \qquad m > 1,$$

$$f^{[i]} \in \Re^N,$$

$$N = (NX \times NY \times NZ) \times NC.$$

6. Size of the ODE systems

- (480x480x10) grid and 35 species results in ODE systems with more than **80 mill.** equations (**8 mill.** in the 2-D case).
- More than **20000 time-steps** are to be carried out for a run with meteorological data covering one month.
- Sometimes the model has to be run over a time period of up to **10 years**.
- Different **scenarios** have to be tested.

7. Chemical sub-model

- Parallel tasks

  The calculations at a given grid-point

- Numerical methods

  QSSA (Hesstvedt et al., 1978)

  Backward Euler (Alexandrov et al., 1997)

  Trapezoidal Rule (Alexandrov et al., 1997)

  Runge-Kutta methods (Zlatev, 1981)

  Rosenbrock methods (Verwer et al., 1998)

Criteria for choosing the numerical method?

8. Advection sub-model

- Parallel tasks

  The calculations for a given compound

- Numerical methods

  Pseudo-spectral discretization (Zlatev, 1984)

  Finite elements (Pepper et al., 1979)

  Finite differences (up-wind)

  “Positive” methods (Bott, 1989; Holm, 1994)

  Semi-Lagrangian algorithms (Neta, 1995)

  Wavelets (not tried yet)

9. Discretization of the derivatives

$$f(x) \in \Re, \quad x \in [0, 2\pi], \quad f(x + 2\pi) = f(x)$$

$$X_N = \left\{ x_n \,/\, x_n = \frac{2n\pi}{2N+1}, \quad n = 0(1)2N \right\}$$

$$F_N = \{ f(x_0), f(x_1), \ldots, f(x_{2N}) \}$$

$$G_N = \left\{ \frac{df(x_0)}{dx}, \frac{df(x_1)}{dx}, \ldots, \frac{df(x_{2N})}{dx} \right\}$$

10. Pseudo-spectral discretization

$$T_N = A + \sum_{k=1}^{N} \left[a_k \cos(kx) + b_k \sin(kx) \right]$$

$$A = \frac{1}{2N+1} \sum_{m=0}^{2N} f(x_m)$$

$$a_k = \frac{1}{2N+1} \sum_{m=0}^{2N} f(x_m) \cos(k x_m)$$

$$b_k = \frac{1}{2N+1} \sum_{m=0}^{2N} f(x_m) \sin(k x_m)$$

11. Convergence of the Fourier series

If f(x) is continuous and periodic and if f´(x) is piece-wise continuous, then the Fourier series of f(x) converges uniformly and absolutely to f(x).

Davis (1963)

12. Accuracy of the Fourier series

It can be proved (Davis, 1963) that

$$|a_k| \le \frac{M}{k^{\mu+1}} \quad \text{and} \quad |b_k| \le \frac{M}{k^{\mu+1}}, \quad M \text{ is a constant,}$$

if

$$f^{(v)}(0) = f^{(v)}(2\pi), \; v = 0, 1, \ldots, \mu$$

13. Drawbacks of the pseudo-spectral method

Drawback	Removing	Reference
Equidistant grids	?	
Periodicity for convergence	Yes	Lyness, 1974
Periodicity for accuracy	Yes	Roache, 1971, 1978

14. Finite elements

The application of finite elements in the advection module leads to an ODE system:

$$P \frac{dg}{dt} = H g$$

Choice of method

$$P^{-1}, \quad (P - \Delta t \beta H)^{-1}$$

P is a constant matrix,
H depends on the wind

15. Matrix Computations

- Fast Fourier Transforms
- Banded matrices
- Tri-diagonal matrices
- General sparse matrices
- Dense matrices

Typical feature: The matrices are not large, but these are to be handled many times in every sub-module during every time-step

16. Major requirements

- Efficient performance on a single processor
- Reordering of the operations

What about parallel tasks?

"Parallel computation actually reflects the concurrent character of many applications"

D. J. Evans (1990)

17. Chunks on one processor

SIZE	Fujitsu	SGI	IBM SMP
1	76964	14847	10313
48	2611	12114	5225
9216	494	18549	19432

First line: the straight-forward call of the box routine

Last line: the vectorized option

Second line: using 192 chunks

Owczarz and Zlatev (2000)

18. “Non-optimized” code

Module	Comp. time	Percent
Chemistry	16147	83.09
Advection	3013	15.51
Initialization	1	0.01
Input operations	50	0.26
Output operation	220	1.13
Total	**19432**	**100.00**

IBM SMP computer, one processor

19. Parallel runs on IBM SMP

Processors	Advection	Chemistry	Total
1	933	4185	5225
2	478	1878	2427
4	244	1099	1405
8	144	521	799
16	62	272	424

IBM *“Night Hawk”* (2 nodes); NSIZE=48

20. Scalability

Process	(288x288)	(96x96)	Ratio
Advection	1523	63	24.6
Chemistry	2883	288	10.0
Total	6209	432	14.4

IBM *"Night Hawk"* (2 nodes); **NSIZE=48**

22. Why is a good performance needed?

Grid	Comp. Time
(96x96)	424 (45.8)
(288x288)	6209 (3.1)

Non-optimized code: 19432

IBM *“Night Hawk”* (2 nodes); NSIZE=48

23. PLANS FOR FUTURE WORK

- **Improving** the spatial **resolution** of the model used to obtain information.
- **Object-oriented code**
- **Predicting** occurrences where the critical levels will be **exceeded.**
- **Evaluating** the **losses** due to long exposures to high pollution levels.
- **Finding optimal** solutions.

24. Unresolved problems

- 3-D models on fine grids
- Local refinement of the grids
- Data assimilation
- Inverse problems
- Optimization problems

Important for decision makers