Grid Computing for Air Quality and Environmental: Studies in Bulgaria

Kostadin Ganev¹, Dimiter Syrakov², Maria Prodanova², Emanouil Atanasov³, Todor Gurov³, Aneta Karaivanova³, Nikolai Miloshev¹, Hristo Chervenkov¹

¹Geophysical Institute, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl.3, Sofia 1113, Bulgaria
²National Institute of Meteorology and Hydrology, Bulgarian Academy of Sciences, 66 Tzarigradsko Chausee, Sofia 1784, Bulgaria
³Institute for Parallel Processing, Bulgarian Academy of Sciences Akad. G. Bonchev str., bl. 25A, Sofia 1113, Bulgaria

kganev@geophys.bas.bg
Dimiter.Syrakov@meteo.bg
Maria.Prodanova@meteo.bg
emanouil@parallel.bas.bg
gurov@parallel.bas.bg
anet@parallel.bas.bg
miloshev@geophys.bas.bg
Hristo.Tchervenkov@meteo.bg

Abstract

Comprehensive atmospheric composition studies require multi-scale numerical experiments to be carried out, which to clarify to some extend different scale processes interaction, but also to further specify requirements for input data (emissions, boundary conditions, large scale forcing). Model interfaces from synoptic through meso- to local scale have to be tailored. Shortly speaking, extensive sensitivity studies have to be carried out, tailoring the model set-up and parameters – a possible forerunner of single model ensemble forecasts.

Performing extensive simulations of this kind with up to date highly sophisticated numerical models obviously requires computer resources of the order of magnitude of those provided by the so-called supercomputers. Using supercomputers, however, is rather expensive and far beyond what most of the research groups can afford. Luckily an alternative technology – the grid computing, is recently very intensively developing, which makes it already quite relevant to formulating and solving problems absolutely unthinkable several years ago.

Some examples of environmental problems which are recently developed / tested / treated as grid applications are given in the present paper.

Keywords: air pollution modelling, multi-scale modelling, country-to-country pollution exchange, emergency response, grid computing

1. Introduction

Air Quality (AQ) is a key element for the well-being and quality of life of European citizens. Bulgaria also faces AQ problems. It should be noted that, while in Western Europe the photo-oxidant and PM air pollution is at present the major environmental problem, in Bulgaria the classic acidifying pollutants (SO₂, NOₓ), the heavy metals (Hg, Cd, Pb) and the persistent organic pollutants are still a serious problem and so the study of their environmental impact is absolutely necessary. The reduction of the emissions of these compounds is a major task in the environmental policy of the country.

AQ modeling methods include statistical, deterministic and hybrid systems. The approach presented in the present paper is deterministic - a combination of meteorological input, weather diagnosis/forecasting, additional meteorological pre-processing and chemical composition simulations. This is a fruitful approach
for indicating exceedances of limit and target values, forecasting potential exceedances and assessing possible emergency measures to abate exceedances. The deterministic approach will also help the better understanding the role of different transport scales and phenomena in the formation of the air pollution model, thus contributing to the model validation.

The main scientific challenge of local to regional atmospheric composition pattern modeling probably is the accounting for the strong dependence of concentrations on fluctuations of local and regional meteorological conditions, the complex interaction of transport scales (different life times of the pollutants make it even more complex), uncertainties and responses to emission forcing and boundary conditions, both introducing information noise.

Multi-scale numerical experiments have to be carried out, which to clarify to some extend different scale processes interaction, but also to further specify requirements for input data (emissions, boundary conditions, large scale forcing). Model interfaces from synoptic trough meso- to local scale have to be tailored (two-way nesting effects have to be checked up?). Shortly speaking, extensive sensitivity studies have to be carried out, tailoring the model set-up and parameters – a possible forerunner of single model ensemble forecasts.

Performing extensive simulations of this kind with up to date highly sophisticated numerical models obviously requires computer resources of the order of magnitude of those provided by the so-called super-computers. Using supercomputers, however, is rather expensive and far beyond what most of the research groups can afford. Luckily an alternative technology – the grid computing, is recently very intensively developing, which makes formulating and solving problems absolutely unthinkable several years ago already quite relevant.

2. Grid Computing

2.1 What is a Computational Grid?

A Computational Grid is a computing environment which enables the unification of widely geographically distributed computing resources into one big (super)computer (Atanassov et al., 2006, Foster and Kesselmann, 1998). The individual computing resources commonly consist mostly of computer clusters and several individual computers which are interconnected by a high speed very wide area network. The Grid is a computer system which is a viable solution for supporting e-Science. The major goal of the Grid is to enable the clustering and unification of distributed computing and data processing resources, as to collect as much computing power usable to applications necessitating high computer strength as possible. Some of scientific application examples necessitating the Grid are applications from the fields of particle physics, climate analysis, biomedical research, meteorology etc.

2.2 Why use parallel computing and cluster?

New possibilities which are being offered by the enormously enhanced computing power of distributed systems open up new application areas for computation. It is possible to simulate natural (and thought-out) processes and phenomena in more detail, using much larger spatial and temporal resolution. New scientific methods, and even new scientific fields, specifically multidisciplinary research, are being ‘born’ and evolved, based on vast sets of data and massive parallel computing.

The following is a short list of some applications of parallel computing:
- Modeling and simulation – commonly based on the principles of stepwise approximation: higher amount of computing allows higher precision (i.e. modeling of climate, simulations in seismology, fluid flow and turbulence simulation, ocean stream movements etc.)
- Processing of large data quantities – computer vision, multimedia data processing, real-time video, large data base searches, data mining etc.

2.3 Algorithms and the Grid

To effectively use the Grid infrastructure, not all algorithms are equally viable. Firstly, the Grid is a parallel execution infrastructure, but it is also more complex for programming (and therefore algorithm selection) than Clusters.

However, though in Cluster execution one of the realistic prerequisites for each algorithm is to be parallel (except in rare cases when algorithm parallel streams execution does not necessitate any data/results interdependency), the Grid Infrastructure is quite more complex. On a Grid a series of independent and fully serial programmes may be very effectively used, sometimes even more effectively than parallel programs. This is due to the fact that inter-cluster communication on the Grid is very inefficient, and if using the Grid not only to find the best Cluster on the Infrastructure to execute the programme, but to realistically use a huge number of resources widely geographically distributed, i.e. resources which are not grouped in a specific Cluster, it is often almost impossible to use a tightly coupled algorithm spread throughout several Clusters.

The usage of a Grid Infrastructure therefore, on one hand, gives higher freedom of algorithm selection, but, on the other hand, necessitates a much stricter and more careful proper application execution planning.

2.4 SEE-GRID-SCI project

The active development of the Grid technology is strongly supported by the European Commission through projects in the 6th and 7th FP. Some of these projects are intended to build and maintain a sustainable grid infrastructure in South Eastern Europe (SEE-GRID). The current project, SEE-GRID eInfrastructure for regional eScience (SEE-GRID-SCI), started on May 2008, stimulates widespread eInfrastructure uptake by new user groups extending over the region, fostering collaboration and providing advanced capabilities to more researchers, with an emphasis on strategic groups in seismology, meteorology and environmental protection. The initiative thus aims to have a catalytic and structuring effect on target user communities that currently do not directly benefit from the available eInfrastructures. In parallel, it enlarges the regional eInfrastructure to cater for demands of the communities by increasing the computing and storage resources and involving new partner countries in the region. Finally, SEE-GRID-SCI helps to mature and stabilise the National Grid Initiatives in the region, allowing them to join the new era of longer-term sustainable Grid infrastructure in Europe. In longer term, SEE-GRID-SCI aspires to contribute to the stabilisation and development of South-East Europe, by easing the digital divide and stimulating eInfrastructure development and adoption by new user communities, thus enabling collaborative high-quality research across target scientific fields.

The current SEE-GRID infrastructure consists of 35 grid clusters with more than 2200 CPUs and 57 TB storage, distributed in 14 countries. A number of applications from diverse end-user communities are already running on the regional eInfrastructure, as a direct development of two phases of SEE-GRID project and of the various national-level initiatives. Most of these applications are focused on a user community within a single country, with some regional and European collaborations. Within the SEE-GRID-SCI
3. Some examples of Multi-scale Atmospheric Composition Modeling

3.1 Application description and main features

Sustainable development at local, regional and global scales represents perhaps the most pressing challenge that humanity faces. In the case of air pollution, the rapid growth of emissions in certain regions coupled with the recognition that pollutants can be transported over very large distances between continents means that this is a problem of global proportions. High levels of atmospheric pollutants can impact on global change (greenhouse gases, aerosols) as well as regional air quality. Our ability to develop successful strategies for mitigation of pollutants depends not only on sound scientific knowledge about local and regional processes but also on the quantification of import and export fluxes into/out of a region. Recent recognition of the scale of long-range transport and transformation of atmospheric pollutants and the potential for increasing emissions in developing regions has combined to convert this into a major global problem. In order to achieve a sustainable atmospheric composition and to design and implement strategies for mitigation, it is necessary both to understand and to quantify the export and import fluxes of hazardous pollutants, greenhouse gases, aerosols, and their precursors, particularly with respect to Europe. Addressing this problem requires the development of an integrated view of the atmosphere at different spatial and temporal scales.

An attempt will be made in this application to develop and promote a modeling system able to interface the scales of the problem from emissions on the urban scale to their transport and transformation on the local and regional scales. The application aims at developing of an integrated, multi-scale Balkan region oriented modeling system, which would be able to:

- Study the atmospheric pollution transport and transformation processes (accounting also for heterogeneous chemistry and the importance of aerosols for air quality and climate) from urban to local to regional (Balkan) scales;
- Track and characterize the main pathways and processes that lead to atmospheric composition formation in different scales;
- Account for the biosphere-atmosphere exchange as a source and receptor of atmospheric chemical species;
- Provide high quality scientifically robust assessments of the air quality and its origin, thus facilitating formulation of pollution mitigation strategies at national and Balkan level.

All the environmental applications are based on the US EPA Model-3 system. The system consists of three components:

- MM5 - the 5th generation PSU/NCAR Meso-meteorological Model MM5 (Dudhia, 1993, Grell et al., 1994) used as meteorological pre-processor;
- CMAQ - the Community Multiscale Air Quality System CMAQ (Byun and Ching, 1999);

3.2 Example of regional scale simulations: Country-to-country pollution exchange in the Balkans

Detailed description of integration domains, input data, model setup and nesting procedures can be seen in a number of papers (see for example Ganev et al., 2008, Prodanova et al., 2008.a). It will be only men-
tioned here that the NCEP Global Analysis Data with 1º×1º resolution was used as meteorological forcing and the MM5 and CMAQ nesting capabilities were used for downscaling to a horizontal resolution of 10 km over the Balkans.

As far as the aim of this simulation is a country-to-country (CtC) study, i.e. to estimate the influence of pollution emissions of each of the three countries (Bulgaria, Romania and Greece) on the pollution levels of the others, four emission scenarios were prepared: basic scenario with all emission sources (scenario All), scenario with Bulgarian emissions set to zero (noBG), scenario with Romanian emissions set to zero (noRO) and scenario with Greek emissions set to zero (noGR).

High ozone concentrations can cause damages on plants, animals and human health. In fact, when the effects from high ozone levels were studied, one should look not at the ozone concentrations but on some related quantities – ozone indices, like AOT40C - Accumulated over threshold of 40 ppb in the day-time hours during the period from May 1 to July 31 concentrations), or NOD60 - Number of days in which the running 8-hour average over ozone concentration exceeds at least once the critical value of 60 ppb. If the limit of 60 ppb is exceeded in at least one 8-hour period during a given day, then the day must be classified as “bad”. People with asthmatic diseases have difficulties in “bad” days.

The calculated NOD60 fields are given in Fig. 1. From the scenario noBG, one can see that switching off Bulgarian sources leads to considerable decrease of this index not only over the territory of the country itself but over the European part of Turkey and the northern Greece. The fact that the NOD60 over Romania is not so much influenced by the elimination of Bulgarian emissions is due to the prevailing NW transport of air masses in the domain. From its side, Romania contribute essentially to the ozone pollution not only in Bulgaria and Moldova but even in Turkey and part of northern Greece, as can be seen from scenario noRo. The last graph in Fig. 1 shows the results of scenario noGR. The exclusion of emissions over the Greek territory decreases the ozone pollution mainly in the country itself. Only European Turkey is influenced to some extent by Greek NOx and VOC pollution.

![Figure 1: Number of days with averaged O3 concentration over 40 ppb for May-July 2000](image)

The EMEP methodology for determining the impact of each country in the Sulfur and Nitrogen depositions over all countries in the region (Bulgaria, Romania and Greece, in our case) is applied. According to this methodology the Oxidized sulfur, Oxidized and reduced nitrogen loads – wet, dry and total (total = wet + dry) deposition and the corresponding country-to-country pollution budget matrixes were also calculated (example for oxidized sulphu total deposition shown on Table 1).
The analysis of different scenarios of switching off the emissions from Bulgaria, Greece or Romania shows, that the impact is most prominent on the sulfur or nitrogen loads in the respective country. The impact on the neighboring countries however can also be significant. For example, the exclusion of Bulgarian sources leads to substantial decrease of oxidized sulfur loads in northeast Greece and European Turkey. The exclusion of Romanian sources also leads to substantial decrease of oxidized sulfur loads in northeast Bulgaria.

### 3.3 Example of local scale simulations: Some cases of extreme air pollution in the city of Stara Zagora

Stara Zagora is one of the biggest towns in Bulgaria (300 000 inhabitants) located in the middle of the country. In the summer of 2004 two very high level SO2 pollution events happened there, leading to serious discontent among the population. Analogous events happened in 2005, too. An attempt for numerical study of one of these episodes – from 8 to 11 of July 2005 was described in (Prodanova et al, 2006, 2008b).

The “Maritza-Iztok” TPPs are the main sulphur polluters not only in Bulgaria but in all SE Europe. The total SOx emission from the TPP is about 700 000 tones per year, i.e. about 2000 tones daily.

The main impression from the analysis of time and space variation of the simulated wind fields during the period is that calm and non-oriented winds prevail. There is a very fast change of wind directions in the different points from the region and at different levels.

As a whole, the calculated SO2 concentration fields have a reasonable behavior from physical point of view (an example is shown in Fig. 2). In night hours, in relatively stable PBL the pollution released from elevated sources (gases with high temperature and release velocity) keeps aloft and the whole 1-km domain is not polluted at all. In the day hours the fast development of turbulent mixing drains the pollution to the surface at distances not far from the stacks, so forming well expressed plumes with high concentrations.

The calculated concentrations however, in spite of the numerous runs with different parameters, do not agree well with the measurements. A suitable direction of the wind from TPPs to the town of Stara Zagora is simulated each afternoon. These flows form pollution spots in different places around the town, but not over it. Obviously the meandering of the plume is not simulated precisely enough in this episode with small and none-oriented winds. Study of the surface concentration plots, which can not be demonstrated here due to the lack of space, shows that deviation of the wind direction by several degrees or change of
the wind speed by several m/s could form spots over the town in the right periods. Small changes in the PBL height and turbulent mixing could lead to the same results. Here, all difficulties faced when trying to model local scale phenomena in complex conditions emerge. Most probably the main shortcomings come from the MM5 simulations. The reality can be and often is much more complicated than any model, which is the case, here. Several reasons for the ill-simulation can be identified:

- The episode is characterized by weak and unstable winds. The small wind speed makes the use of PBL parameterizations very uncertain – the geostrophic wind is, possibly, the most important parameter in every PBL scheme. The lack of synoptic forcing leads to the fact that the PBL properties are determined mainly by the other local parameters. These parameters change their values from grid point to grid point, thus forming different winds in neighboring points. In such a way false divergence and convergence fields appear. The adaptation ability of the model eliminates this discrepancies but at the price of modification of wind and turbulent properties of the PBL that can lead to deviations from their real values.

- The unstable winds can also lead to mis-reproducing of the real variables. It is clearly seen from the meteorological and concentration graphs that the wind direction changes rapidly from almost all points of the compass during a day. As far as the PBL parameterization schemes present stationary state of the PBL, the achievement of this steady state needs some time after the governing parameters has been changed. When the wind changes its characteristics rapidly there is not enough time for full adaptation and this is an additional source of discrepancies between the simulated and real fields.

- Finally, it is quite possible that the used vertical resolution of MM5 is not enough to reproduce accurately the complex character of the local ABL, with its layers, complex stability profiles and their evolution.

Figure 2: Evolution of the surface SO2 fields for July 9, 2005

The applied MM5/CMAQ model system is quite complex and needs validation for each area of simulation. Therefore, when important air pollution management decisions are needed, not only modeling but measurements of the vertical structure of the atmosphere are crucial at least for short periods.


4.1 Application description and main features

Scientific and technical information is critical for helping emergency managers to make sound decisions with regards to response to critical threats. The main features of a dispersion modeling systems that are critical for emergency management needs probably are adequate, proper and detailed treatment of the underlying surface heterogeneities – topography, land-sea interactions, urban effects. The system should
have high enough speed for the “fast decision phase” and at the same time produce accurate enough estimates for the preparedness and recovery phases.

The current application aims at developing of unified Balkan region oriented modeling system for operational response to accidental releases of harmful gases in the atmosphere (as a result of terrorist attack or industrial accident), which would be able to:
- Perform highly accurate and reliable risk analysis and assessment for selected “hot spots”;
- At a warning signal from the measuring network, by using the adjoin functions technique, to detect (if not known) the harmful release location and evaluate the nature and the amount of the released harmful gases;
- Provide the national authorities and the international community with short-term regional scale forecast of the propagation of harmful gases;
- Perform, in an off-line mode, a more detailed and comprehensive analysis of the possible longer-term impacts on the environment and human health in the Balkan region and make the results available to the authorities and the public.

4.2 Example of risk analysis simulations: Instantaneous chlorine release

The application is in the stage of developing and testing the grid-note, so some very preliminary results will be demonstrated.

The simulations are made for the case of accidental release of chlorine in the “VEREA-HIM” factory in Jambol. The accidental scenario is of an instantaneous release of 25 t of chlorine near the earth surface in 3 o’clock am. The simulations were made by series of nesting procedures, downscaling to a horizontal resolution of 1 km. The concentration evolution for 17.07.2007 at earth surface is shown in Fig. 3. As it should be expected the concentration has a very steep “Gaussian” shape in the beginning and gradually the pollutants spreads in horizontal and vertical direction.

5. Conclusions

It can be concluded that the air pollution studies in Bulgaria are currently in a stage mature enough, so that development of a national AQ Forecast and Information System (AQFIS) is already a logical and realistic objective for the Bulgarian AQ modeling community. The general idea of the AQFIS is briefly described in Syrakov et al (2009). The other prerequisite for creating such an operational system is the availability of powerful platform for data processing and transfer. The grid computing appears to be a very promising and up-to-date technology to be applied for the purpose.
Acknowledgement

The present work is supported by EC through 6FP NoE ACCENT (GOCE-CT-2002-500337), IP QUANTIFY (GOGE-003893), SEE-GRID-SCI project, contract № FP7 –RI-211338, COST Action 728, as well as by the Bulgarian National Science Fund (grants № Д002-161/16.12.2008, Д002-1115/16.12.2008 and Д002-146/2008).

Deep gratitude is due to US EPA, US NCEP and EMEP for providing free-of-charge data and software. Special thanks to the Netherlands Organization for Applied Scientific research (TNO) for providing us with the high-resolution European anthropogenic emission inventory.

References


