GEM-PP: a GIS EMissions Pre-Processor to ingest European emission inventory (EMEP/CORINAIR) into photochemical transport models

Dario Conte[(i)], Gian Paolo Marra[(i)], Cristina Mangia [(i)], Umberto Rizza [(i)], Ilenia Schipa[(i)], Annalisa Tanzarella[(i)]

[(i)] Institute of Atmospheric Sciences and Climate, CNR, Str.Prov.Lecce-Monteroni km 1.2, 73100, Lecce, Italy {d.conte, gp.marra, c.mangia, u.rizza, i.schipa, a.tanzarella}@isac.cnr.it

Abstract

Air quality simulations and forecasting using photochemical transport models require detailed information of the emissions in the area under investigation. Due to the wide range of space and time scales involved in the physical and chemical processes of air pollution phenomena, a need to perform air quality simulations over larger regions, at finer grid resolutions, using more emissions sources, and for more purposes (e.g., ozone, particulate, toxics) has developed. Geographic information systems (GIS) have become important decision-making tools for location information of real-world objects and phenomena in the context of air quality, as they allow for the management of geographic information in a flexible and efficient manner. In the past few years, several open source (OS) GIS solutions have emerged, permitting the development of complex GIS projects without requiring commercial closed-source software.

In this paper, we present GEM-PP, an emissions pre-processor for air quality models developed using OS-GIS technologies. This tool elaborates the emission inventories according to the coordination of information on air emissions (CORINAIR) methodology, and generates gridded, hourly emissions in a format consistent with air quality models. To test the functionality of the developed GEM-PP, the pre-processor has been applied to reproduce emission patterns and the distribution of air pollutants over the Apulia region in southeast Italy. The data has been pre-processed to be input into two commonly used dispersion models, WRF-Chem and CALGRID.

1. Introduction

Over the last few decades, the management of air quality has become a significant challenge for decision makers, in particular with respect to aerosol and photochemical pollution, wherein the relationships between emissions and atmospheric concentrations can be quite complex due to the non-linear interactions among emissions, meteorology, chemistry, and land use patterns. By combining meteorology and chemistry, air quality models are fundamental to understanding the different mechanisms that lead to atmospheric pollution, and consequently, to support present and future emission control strategies.

A fundamental input to grid air quality models are the pollutant emissions for each cell of the simulation domain in the unit time. The estimation of the emissions over a certain area can be made in two major ways, by taking either a “bottom-up” or a “top-down” approach (Lindley et al, 1996).

In Europe, national emission data are based on the CORINAIR air emission inventory program, coordinated by the European Environment Agency for annual sector-wise national emission estimates (EEA, 2007).

The aim of this investigation is to describe the emissions pre-processor GEM-PP, "GIS emissions pre-processor", which, starting from a national emission inventory based on CORINAIR methodology, provides the chemical compounds at each point of a numerical grid for two commonly used air quality models, the WRF-Chem (Grell et al, 2005) and CALGRID (Yamartino et al. 1992), using a "top-down" methodology. It is based on the GIS approach and has been fully developed using open source components. The pre-processor acts as a workspace for the numerical grid generator, and provides a suitable solution for disaggregation of administrative boundary shaped inventories into small cells using an appropriate computation algorithm.

The methodology for the estimation of the emissions is based on the works by Monforti and Pederzoli (2002), and Carnevale (Carnevale et al. 2006). As shown in recent papers (Naoum et al., 2005; Dalvia et al., 2006), the presence of GIS tools to manage emissions assumes a strategic role, that is, administrative boundary-shaped inventories can be represented in a GIS vectorial format, and subsequently superimposed over a vectorial-type geographic grid (map-overlapping). In this way, we can gather pollutant emission data from the sum of the contributions associated to the administrative boundaries. The use of GIS tools is the ideal choice when it is necessary to manipulate geographic units.

2. The GEM-PP project

The objective of the project is to develop a suitable GIS tool to provide gridded emission input data using a top-down approach for air quality models. The tool has been developed using open source software for all GIS components, such as geo-processing libraries and digital mapping tools, and a database management system (RDBMS with GIS extensions). The GIS components are valuable for data manipulation of emissions over geographic entities, while the data management system is used as a backbone for the proposed processing emissions system (Janssen et al., 2003). It is worth noting that the GIS approach does not improve the accuracy of the emissions inventory, but does assist in obtaining gridded data and any other manipulations over geographic information. The utilization of a RDBMS system that can provide GIS functionalities improves the efficiency in such operations because it provides a structured workspace for the storage and retrieval of information. This permits the creation of a variety of input emissions for many different photochemical transport models, each with different requirements and file formatting. In the past few years, several commercial and open-source software solutions have become available to the GIS industry. In particular, the open-source GIS projects provide several powerful tools that can satisfy many major requirements of both scientific and the industrial applications. These open-source tools permit the development of affordable, low-cost software projects that can be utilized by the scientific community. Figure 3 shows how GEM-PP acts as interface between input data, coming from European emission inventory, and various photochemical models.

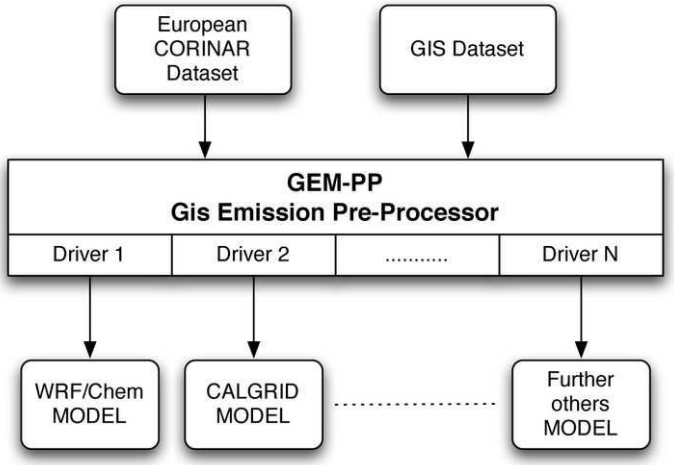

Figure 1: GEM-PP overview.

The GEM-PP was originally designed to provide several functionalities that attempt to satisfy the already described objectives, and in particular, permits:

- Emissions data processing of spatial disaggregation, time modulation, chemical speciation, and lumping,
- the preparation of a standard input for photochemical models for air quality simulations,
- an automatic re-projection of input/output files with respect to the user specified geographic reference systems and projections,
- post-processing to produce input files for a given photochemical model, and to produce both digital maps and shape (a very common ESRI format) the model output.

Among the available open-source projects, we selected components that matched the imposed functional requirements and that were able to satisfy technical constraints, such as flexibility, robustness, and interoperability with other software packages. In particular, we used the following key tools: PostgreSQL (http://www.postgresql.org) with PostGIS extensions (PostGIS documentation); Geospatial Data Abstraction Library/Simple Features Library (GDAL/OGR, http://www.gdal.org/); (UMN) MapServer (http://mapserver.gis.umn.edu/).

Finally, the GEM-PP has been developed for UNIX-like systems to satisfy the requirements of the most commonly used photochemical models. It was implemented in Ruby (http://www.rubylang. org), an open-source object-oriented scripting programming language, because this language allows the development of robust applications in a time efficient manner, thanks to its object oriented approach.

3. System architecture

GEM-PP has been designed using the model view controller (MVC) pattern (Buchman et al., 1996) and has been implemented using the Ruby scripting language. The MVC design pattern has been selected to better organize the application, while Ruby permits efficient object-oriented code development. Due to the MVC, we designed the user interface, the models, and the data structures

separately from the controller, which has been developed to act as a middleware between them. This is done by specifying the logic flow between user commands and processing.

Figure 2 depicts an overview of GEM-PP architecture. The two interfaces, the command line gempp-CLI-view and the utility logger-view, exploiting their controllers (gempp-CLI-controller and logger-controller) provide full access to all tool functionalities implemented by the Disaggregation model, gridding emissions model, photochemical drivers model, and post-processing-model. The robot model and the monitor model allow the automatic processing and monitoring of job execution.

The disaggregation model provides spatial disaggregation, time modulation, chemical speciation and lumping, as defined by user input.

The gridding emission model allocates the disaggregated emissions into a numerical grid by defining the geographic boundaries and dimensions of the domain.

At user request, the output of the photochemical drivers model can be formatted following the requirements of the air quality models supported by GEM-PP. The post-processing model provides routines for post-processing graphics operations, such as the exportation of disaggregated and gridded emissions as ESRI shapefiles, or geographic image maps obtained through digital mapping processes.

GEM-PP can operate in both interactive and batch mode by means of the special class robot model. The choice is left to the user by a command line option. The presence of the this option is because in some case the entire task to obtain input for the photochemical models may be computationally expensive, making the execution of the GEM-PP in batch mode more convenient for users.

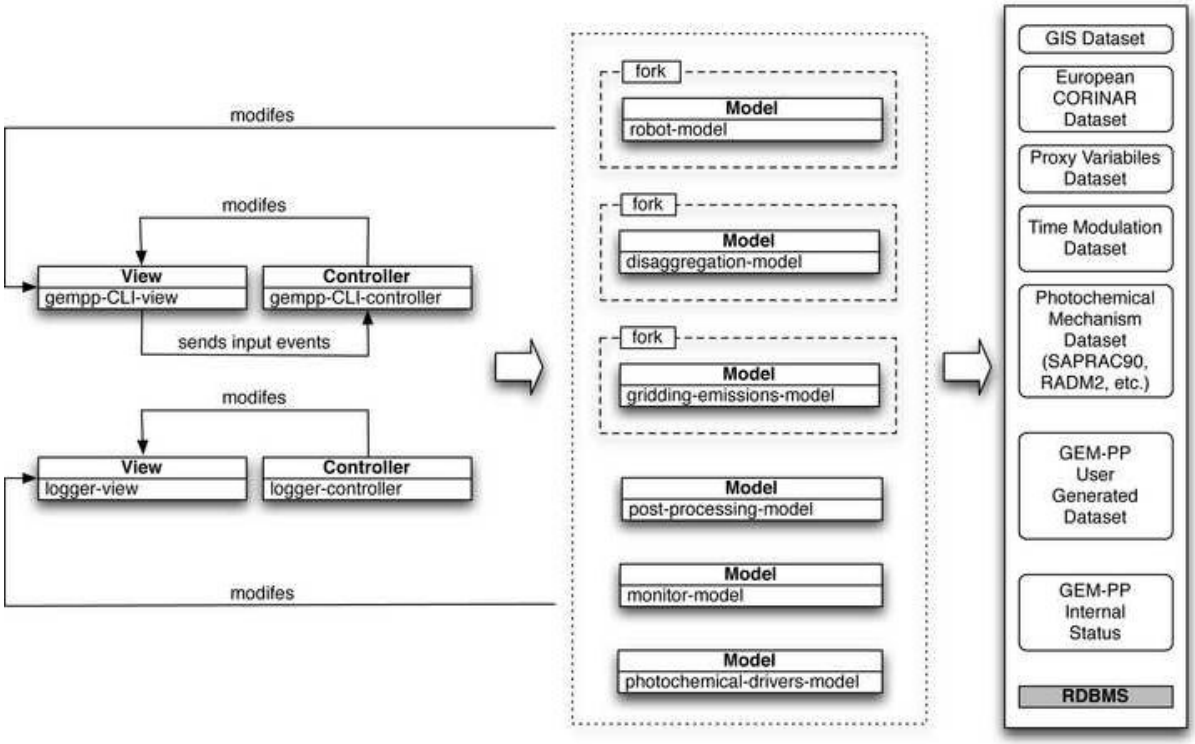

Figure 2: Overview of GEM-PP architecture.

4. Case study - Apulia Region

GEM-PP has been applied to pre-process emission data from the national emission inventory developed by the Agency for Environmental Protection and Technical Services (APAT) of Italy in 2000, which reports the annual emission data for CO, NMVOC, CH4, NOX, SOX, N2O, CO2, NH3.

Emissions in the Apulia region derive primarily from urban, shipping, and industrial activities. The largest industrial emissions are located on opposite coastlines in the southern part of the region, specifically in the Taranto and Brindisi areas. A domain 486 X 396 km^2 in size has been subdivided according to a grid system having a 3 km mesh size.

The first step in the application of GEM-PP on a specific domain is the preparation of the disaggregation datasets, which contain all surrogate variables for space disaggregation, time modulation and lumping classes.

Surrogate variables were determined from different sources, i.e., the Italian National Institute of Statistics (ISTAT), the Italian National Automobile Association (ACI), the Italian highways Bureau, national electricity companies, and the military meteorological service. All data were integrated with GIS and land use data.

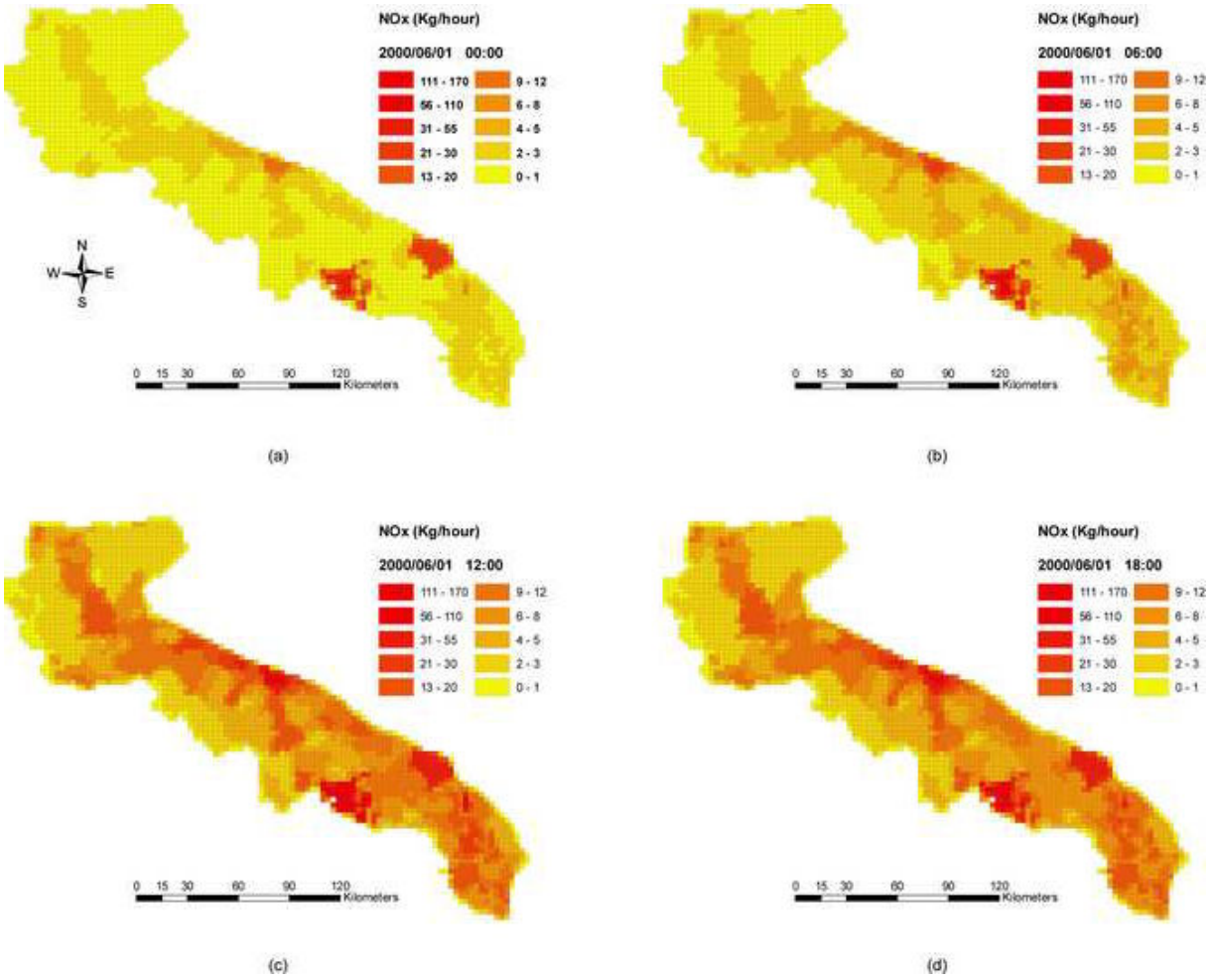

Figure 3: Distribution of NOx emissions at times: 00:00, 06:00, 12:00, 18:00, in the year 2000 over the Apulia Region.

As an example of the spatial Disaggregation and temporal modulation, figures 3a-d depict the total NOX emissions on a 3km x 3 km grid at 00 06 12 18 hour time steps for a typical working day. Here, it is evident that the modulated emission due to the transport sectors are aggregated to the constant emissions derived from the largest industrial activities.

5. Conclusions

In the past few years, environmental scientists have developed advanced numerical air quality models to better understand the interactions between meteorology, anthropogenic and biogenic emissions, and pollutant chemistry and dynamics. Emissions data are one of the most important inputs for air quality models. The actual trend is to model larger regions, at a finer grid resolution, with a more detailed description of the emissions sources, and for more purposes (e.g., ozone, particulates), using a computationally efficient, and flexible emissions data processing system. In this paper, we have presented the GEM-PP emission pre-processor that can be utilized to convert the resolution of the emission inventory data to the resolution needed by an air quality model.

Currently, GEM-PP elaborates the emission inventories developed according to the CORINAIR methodology, following a top-down approach. The novelty of the GEM-PP approach consists of the combined use of GIS tools with a robust database management system (RDBMS with GIS extensions) without the use of any commercial, closed-source software. This combination of technologies has several advantages. GIS tools are the best solution when it is necessary to manipulate and elaborate geographic information. GIS tools have become important decision making tools involving location information for real-world objects and phenomena in the air quality domain, and allowing the management of geographic information in a flexible and efficient manner.

Additionally, the use of the RDBMS as a backbone for the proposed emissions processing system makes GEM-PP a robust and efficient tool. Finally, the use of open-source software permits an open and inexpensive platform for the processing application-specific constraints for any complex domain of interest and, in particular, for air quality systems.

From the user point of view, GEM-PP is particularly useful when several simulations are required to analyze model results as a function of initial parameters (tuning of the parameters). Moreover, GEM-PP can also be used as a graphical post-processor for outputs of photochemical models through the digital mapping processes. Finally, due to its UNIX-like basis, GEM-PP can facilitate the inter-operability of a great number of existing modeling systems for understanding air quality phenomena.

As a preliminary test, GEM-PP has been applied to the Apulia Region in southern Italy. The data has been pre-processed for input into the WRF-Chem and CALGRID dispersion models. More stringent evaluation will be performed by the application of the air quality models. Further developments will focus on improving the disaggregation by including the land use in the gridding emission phase, setting up other drivers for other photochemical models, and integrating other emission inventories presently available for air quality studies.

References

Lindley, S.J., Longhurst, J.W.S., Watson, A.F.R. and Conlan, D.E. 1996, 'Procedures for the estimation of regional scale atmospheric emissions — an example from the NW region of England', *Atmospheric Environment*, vol. 30, no. 17, pp. 3079–3091.

European Environment Agency (EEA) 2007, *EMEP/CORINAIR Emission Inventory Guidebook*, Viewed 25 July 2007, <http://reports.eea.europa.eu/EMEPCORINAIR5/en/page002.html>.

Grell, G. A., S. E. Peckham, S. A. McKeen, J. Wilczak, and B. Eder 2005, 'Fully coupled online chemistry within the WRF model', *Atmospheric Environment*, vol. 39, no. 37, pp. 6957–6975.

Yamartino, R.J., J. Scire, G.R. Carmichael, and Y.S. Chang 1992, 'The CALGRID mesoscale photochemical grid model', *Atmospheric Environment*, vol. 26A, pp. 1493-1512.

Monforti F., Pederzoli A. 2005, 'THOSCANE: a tool to detail CORINAIR emission inventories', *Environmental Modelling & Software*, vol. 20, no. 5, pp. 505–508.

Carnevale C, Gabusi V., Volta M. 2006, 'POEM-PM: An emission model for secondary pollution control scenarios', *Environmental Modelling & Software*, vol. 21, no. 3 , pp. 320-329.

Naouma S., Tsanisb I.K., Fullartona M. 2005. 'A GIS pre-processor for pollutant transport Modelling', *Environmental Modelling & Software*, vol. 20, no. 1, pp. 55-68.

Mohit Dalvia, Gufran Beigb, Uday Patilc, Akshara Kaginalkara, C. Sharmad, A.P. Mitrad 2006, 'A GIS based methodology for gridding of large-scale emission inventories: Application to carbon-monoxide emissions over Indian region', *Atmospheric Environment*, vol.40, no. 16, pp. 2995–3007.

Janssen, M. and Z. Wang 2003. *Open Emissions Model (OPEM): Re-Thinking the Emissions Modeling Paradigm*, Report prepared by Lake Michigan Air Directors Consortium and University of California at Riverside. March 5, Viewed 25 July 2008, <http//www.ladco.org/tech/emis/misc/OPEM.pdf>.

Drake Joshua D. 2008, *PostgreSQL,v. 8.2.*, object-relational database management system (ORDBMS), Viewed 25 July 2008, <http://www.postgresql.org>.

Refractions Research 2008, *PostgGIS*, v. 1.3, software program that adds support for geographic objects to the PostgreSQL, Viewed 25 July 2008 <http://www.postgis.org>.

OSGeo (Open Source Geospatial Foundation) 2008, *GDAL - Geospatial Data Abstraction Library, v. 1.4*, library for reading and writing raster and vectorial (OGR) geospatial data formats, v. 1.4, Viewed 25 July 2008, http://www.gdal.org.

Lime Stephen 2008, *UMN MapServer, v. 5.0*, open source development environment for building spatially-enabled internet applications, Viewed 25 July 2008 <http://mapserver.gis.umn.edu>.

Matsumoto Yukihiro 2008, *Ruby programming language, v. 1.8*, general purpose object-oriented programming language, Viewed 25 July 2008 <http//www.ruby-lang.org>.

Buchman F., Meunier R., Rohnert H., Sommerlad P., Stall M. 1996, *Pattern-Oriented Software Architecture: A System of Patterns*, John Wiley & Sons Publishing.

The Italian Agency for Environmental Protection and Technical Services (APAT) 2008, Viewed 25 July 2008, <http://www.apat.gov.it/site/en-GB/default.html>.