Brigham Young University
BYU ScholarsArchive

International Congress on Environmental Modelling and Software

6th International Congress on Environmental Modelling and Software - Leipzig, Germany - July 2012

Jul 1st, 12:00 AM

GIS-based procedure for evaluation of performances of the Italian atmospheric modelling system simulated data versus observed measurements

Antonio Piersanti

Gaia Righini

Giuseppe Cremona

Luisella Ciancarella

Irene Cionni

See next page for additional authors

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference

Piersanti, Antonio; Righini, Gaia; Cremona, Giuseppe; Ciancarella, Luisella; Cionni, Irene; D'Isidoro, Massimo; Mircea, Mihaela; and Vitali, Lina, "GIS-based procedure for evaluation of performances of the Italian atmospheric modelling system simulated data versus observed measurements" (2012). *International Congress on Environmental Modelling and Software*. 172.
https://scholarsarchive.byu.edu/iemssconference/2012/Stream-B/172

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Presenter/Author Information
Antonio Piersanti, Gaia Righini, Giuseppe Cremona, Luisella Ciancarella, Irene Cionni, Massimo D'Isidoro, Mihaela Mircea, and Lina Vitali

This event is available at BYU ScholarsArchive: https://scholarsarchive.byu.edu/iemssconference/2012/Stream-B/172

International Environmental Modelling and Software Society (iEMSs)
2012 International Congress on Environmental Modelling and Software
Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany
R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.)
http://www.iemss.org/society/index.php/iemss-2012-proceedings

GIS-based procedure for evaluation of performances of the Italian atmospheric modelling system simulated data versus observed measurements

Antonio Piersanti, Gaia Righini, Giuseppe Cremona, Luisella Ciancarella, Irene Cionni, Massimo D'Isidoro, Mihaela Mircea, Lina Vitali
ENEA, Via Martiri di Monte Sole 4, Bologna, Italy
antonio.piersanti@enea.it

Abstract: Air quality assessment and management is usually performed using atmospheric integrated modelling systems and comparisons with data observed by monitoring stations. In Italy, the Project MINNI (National Integrated Model in support to the International Negotiation on Air Pollution) has been supporting Italian Ministry of Environment for the last ten years within the international negotiation process on Air Pollution and assessing Air Quality policies at national/local level. In order to evaluate the performances of the MINNI atmospheric modelling system, a GIS approach was applied to compare simulated and observed concentrations of different atmospheric pollutants. The geoprocessing workflow integrates GIS tools with external procedures. The study described here includes the collection of detailed information on monitoring networks, their geolocation, the development of a relational database for storage and management of all information collected, the implementation of dedicated procedures for extraction and statistical analysis of observed and modelled data time series and 2D maps. Moreover, management and interpretation of output maps from the air quality model were implemented in GIS environment for specific purpose analyses. This geoprocessing workflow, allows to investigate different aspects of air pollution, evaluating model performances in reproducing real atmospheric processes at local or regional scales and assessing the impact of air quality measures on human health and environment on a national basis.

Keywords: relational database, atmospheric modelling system, air quality, monitoring network, GIS.

1 INTRODUCTION

Integrated Assessment Modelling Systems are used to simulate the transformations and transport of atmospheric pollutants in support of air quality evaluation and monitoring.

ENEA, with the air quality laboratory of the Technical Unit on Models, Methods and Technologies for the Environmental Assessment, is a primary actor for air quality evaluation using MINNI (www.minni.org), which is the Italian Integrated Assessment Modelling System for supporting the International Negotiation Process on Air Pollution and assessing Air Quality Policies at national/regional level, sponsored by the Italian Ministry of the Environment [Zanini et al. 2005, Mircea et al. 2010a]. The system MINNI is composed of an Atmospheric Modelling System (AMS) and the Greenhouse Gas and Air Pollution Interactions and Synergies Model over Italy (GAINS-Italy) as shown in D'Elia et al., [2009]. The main components of AMS are the meteorological model [Cotton et al. 2003] for simulating the meteorological conditions, the Emission Manager for apportioning emission inventories data to grid cells, and the air quality model (Flexible Air quality Regional Model) FARM for simulating atmospheric chemistry and transport [Silibello et al. 2008]. Yearly simulations were performed using national emission inventories for Italy and the European Monitoring and Evaluation Programme (www.emep.int) emission inventories for the other countries included in the computational domain. AMS simulates 3-dimensional meteorology and air quality fields for the entire Italian region, on a national domain (20 km resolution on horizontal grid) and on 5 nested regional domains (4 km resolution on horizontal grid). Yearly simulations are available for several years: 1999, 2003, 2005, 2007. The model uses an integrated and multi-pollutant approach for calculating hourly concentrations of SO_2, NO_2, O_3, PM10, PM2.5, NH_3, etc., and depositions of sulphates and nitrates.

Simulated results are compared to measurements from monitoring networks distributed over the country, aiming at the validation of AMS results, since ground based in-situ measurements represent the reference for analysis of air composition and pollution levels.

The BRACE national database (www.brace.sinanet.apat.it), managed at national level by the Institute for Environmental Protection and Research and described in Caricchia et al. [2003], includes information on the air quality monitoring stations distributed over the Italian territory, along with measured data.

In order to use the measurements stored in BRACE and to perform the analysis of simulated versus measured data in GIS environment, a geoprocessing workflow that integrates GIS tools with external procedures is defined here. The work performed includes the collection of detailed information on monitoring stations, their geolocation, the development of a relational database for storage and management of the information, the implementation of dedicated procedures for extracting and processing measured and modelled data time series, and finally the visualization and analysis of results.

2 TOOLS AND METHODS

2.1 ASQUA geodatabase

A relational geodatabase named ASQUA (*Anagrafica Stazioni QUalità dell'Aria,* information registry on air quality stations) was built, including monitoring stations metadata to manage information such as geographical location, a classification based on the type of prevailing emissions (background, traffic, industrial), population density (urban, suburban, rural) and analysis of pollutants. ASQUA does not include measurements as they are stored off-line in BRACE.

As a first step, a wide collection of records related to national monitoring stations was composed focusing on coordinates, stations typology, type of data measured, recording period and local specifications. This information was collected from regional environmental agencies (ARPA) that officially manages the air quality data in each regional administration, and can therefore guarantee the best consistency and update of the information gathered. Only in few cases, when information were not available from ARPA, it has been necessary to retrieve them directly from the BRACE database. An accurate revision of the evidences collected regarding location, definition, and characterization of the network was performed for selecting a reliable set of stations in terms of spatial accuracy and quality assurance.

The data collected were managed in GIS using ArcGIS™ 9.2 software and projected in UTM WGS84 system Zone 32. More than 1000stations were registered in ASQUA, while about thirty stations were rejected due to lack of reliable information. As clearly shown in Figure 1, the majority of the stations stored in ASQUA are located in the North of Italy where high urbanization and industrialization causes frequent exceedances of EU limit values on air pollution. However, Southern Italy represents also an important area for monitoring the significant influence that natural sources, such as Saharan dust, fires, volcanoes and biogenic emissions have on the regional air quality.

The second step pertained to the harmonization and organization of the information in a set of tables linked by relationships and joined through a primary key numerical code called “station_code”.

Figure 1. Spatial distribution of the 1081 monitoring stations recorded in ASQUA (yellow points).

Two main sets of tables (Table 1) were created to better summarize the stations information: the registry information including location, typology and metadata; the measurements information including the type of pollutants measured and the years of activity.

Table 1. Tables in ASQUA database.

	Table name	Description
Registry info	ASQUA_OK	Main features of reliable stations
	ASQUA_KO	Features of not reliable stations
	Type of station	Classification related to the type of prevailing emission (background, traffic, industrial) and the population density (urban, suburban, rural)
	Coordinates	Coordinates in meters WGS84 UTM32 and decimal degree WGS84, elevation.
Measurements info	FILE_DATI	Join table to the file of measurements classified per pollutant and per year
	Pollutants_year	Information on sensors and pollutants measured
	Pollutants_period	Information on period of activity per sensors and pollutants measured

The relations between the measurement files , stored in the BRACE database, and the ASQUA tables needed to be built with special attention. In fact it was often difficult to establish an unique relationship since many measurement files were found to be named following their local original classifications. The table “FILE_DATI” was therefore created using a set of multiple and structured queries to relate the “station_code” to the measurement files. Dedicated fields indicating pollutant types, years of recording and other specifications were introduced to cope with any need of data enquiry. The availability of GIS then allows to run queries on metadata fields such as type, zone, location, availability of pollutants in order to drive the analysis of air quality measurements.

As the measurements file of interest has been selected, a dedicated procedure, described in the next paragraph, can be run to extract observed data, perform statistical analysis with the final goal of comparing the data with AMS simulated outputs.

2.2 Procedure for data extraction and processing

In order to evaluate the AMS performance, a numerical code has been implemented to compare simulations against observed quantities at air quality monitoring sites. The code was developed under UNIX environment, written both in shell scripting and FORTRAN90 language. It is installed under the ENEA CRESCO (Centro Computazionale di RicErca sui Sistemi Complessi, Computational Research Center for Complex Systems, http://www.cresco.enea.it/CRESCOengl.htm) system, which hosts also the MINNI datasets (about 80TB) that hosts all the records. CRESCO represents the major node of the ENEA GRID computing facility which links among themselves the HPC (High Performance Computing) platforms located in the different ENEA Research Centers throughout Italy, providing a huge storage capability. It allows the exploitation of parallel computing in many of the MINNI modelling system applications.

The code operation is divided in three main sequential steps. Firstly, access to the data stored in BRACE is performed and measurements at stations, selected using ASQUA database, are extracted; secondly, the simulated data are extracted at the same locations and finally the two datasets are compared to assess model performance.

More specifically, the first step operates the access to the measurements and the rewriting of the data in a predefined common format. The monitoring stations are sorted by pollutant, year and station type (i.e.: urban background, urban traffic, rural background, etc.). During this phase, a rough check is done in order to eliminate stations which result having a number of records below a given threshold. In addition, locations with unrealistic or negative concentration values are excluded as unreliable. Then the remaining observations are stored to be used for the next steps. The second step based on the geographical coordinates of the monitoring stations recorded in ASQUA extracts the corresponding simulated time series from AMS simulations using a bilinear interpolation which is a standard procedure to compare grid point values gathered from models with not overlapping point measurements by monitoring stations.

Finally, the last step of the procedure uses the extractions from the previous processes to perform model evaluation. In this phase stations with data time coverage not in compliance with the Council EU Directive 2008/50/CE [2008] are filtered out of the ensemble. The percentage of valid data on daily, monthly, seasonal or yearly basis is also accurately checked according with the type of pollutant and the period under investigation. Finally the code calculates several statistic indicators, specific to the evaluation of atmospheric models [Chang and Hanna 2004, Boylan and Russel 2006], to compare observed and simulated time series. A table is produced as an output for every pollutant reporting general information about every station with the set of statistical indicators satisfying the state-of-art model performance assessment and EU guidance (http://fairmode.ew.eea.europa.eu/guidance-use-models-wg1).

2.3 Procedure for data mapping

A further branch of the tool chain is dedicated to generating 2D maps and single point time-series plots of meteorological variables, pollutant emissions and concentration parameters on the native model raster grid. The 3D output fields are stored in NetCDF format and NCL scripting (http://www.ncl.ucar.edu/) has been implemented in order to fully exploit NCAR tools for data analysis and visualization, freely accessible and widely used in environmental modelling (e.g.

Jonson et al [2010], Zender [2008]). Individual scripts have been developed for different purposes, like the single point time-series plots and 2D mapping of results for a selected period: in the latter case, once the statistical parameter to be calculated is chosen (e.g. monthly average), a raster grid is produced along with quick-look maps, useful to rapidly assess the results. The raster grid is imported into ArcGIS™ (or other GIS/mapping software) and mapped as color-filled blocks or contoured areas.

3 RESULTS

An example of the application of the complete procedure is here described by illustrating a comparison of simulated data versus measurements for the average ozone concentration during July 1999, previously investigated by Briganti et al., [2010].

ASQUA database is used for the preliminary analysis on location, distribution and characteristics of monitoring stations. The overlay of several thematic layers in GIS environment gives a valuable contribution to identify the most useful stations according to the investigated pollutant. In this work these layers are: CORINE land cover (2006 updates), Digital Terrain Model derived from SRTM (Shuttle Radar Topography Mission) data with 90 meters spatial resolution, main hydrography, highways and main roads, distribution of urban areas, administrative boundaries. Using ASQUA, the expert user can identify, from the monitoring network, the most suitable stations for the specific analysis to be performed, according to type of pollutant, period of recording, characteristics of the stations and availability of measurements. Figure 2 shows an example of the visualization of selected background rural stations that measured the ozone in the year 1999, overlaid on the CORINE land cover theme with first level of classification.

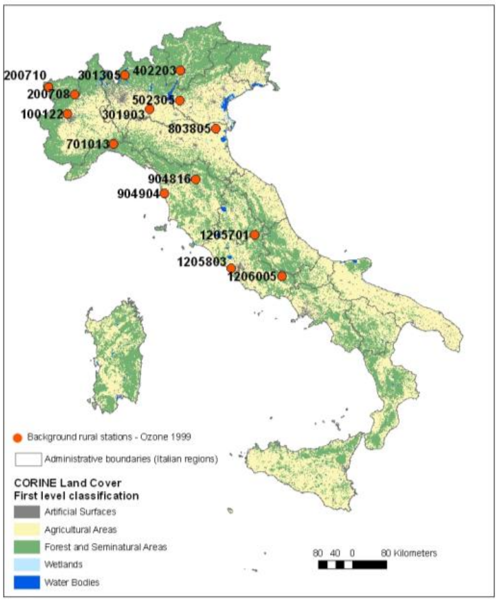

Figure 2. Selection of background rural station that monitored ozone in 1999, overlaid on CORINE land cover. The numerical codes identify each station.

The data extraction procedure (paragraph 2.2) creates the datasets of interest and produces statistical analysis of observed and simulated data. As high values of ozone are recorded during summer, due to high solar radiation promoting chemical reactions producing ozone, July is selected as period of interest. Table 2 shows output values of data extraction for this period. Reported parameters are the observed average (OBS_AVG) and standard deviation (OBS_SD), the simulated average (SIM_AVG) and standard deviation (SIM_SD) plus three statistical indicators of model performance, selected from the set reported in EU guidance: correlation index (CORR), index of agreement (IOA), fractional bias (FB). As shown in grey records, stations are excluded if data coverage is under EU guidance threshold.

Table 2. Statistical comparison between observed and modelled data for stations selected as in Figure 2. Grey records are related to stations where data coverage is not sufficient.

STATION_CODE	OBS_AVG	OBS_SD	SIM_AVG	SIM_SD	CORR	IOA	FB
100122	-999	-999					
200708	88	36	78	23	0.51	0.63	-0.12
200710	90	20	73	14	0.32	0.27	-0.21
301305	113	49	95	30	0.65	0.69	-0.17
301903	73	50	99	29	0.77	0.70	0.31
402203	-999	-999					
502305	-999	-999					
701013	72	44	97	26	0.35	0.42	0.29
803805	54	42	91	26	0.59	0.39	0.52
904816	96	33	93	18	0.61	0.70	-0.03
904904	106	22	101	29	0.49	0.68	-0.04
1205701	-999	-999					
1205803	92	47	87	30	0.48	0.66	-0.05
1206005	79	47	85	16	0.39	0.45	0.08

The data mapping procedure (paragraph 2.3) allows to draw, for example, a time-series plot (Figure 3) extracted at a selected station (e.g. 301305 - Varenna), showing hourly measures against model values during the whole period, highlighting daily concentrations cycles. These outputs are analysed by modelers in order to evaluate the model performances and drive further developments and improvements.

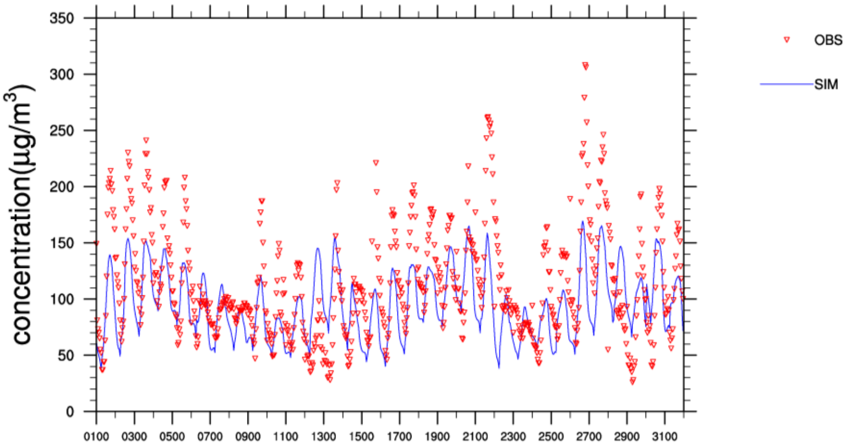

Figure 3. Ozone hourly measured values (red dots) and model results (blue line) at 301305-Varenna station, during July 1999.

The study on the distribution of pollutants is held in GIS environment where both observed and simulated data are imported and managed for spatial analysis and processing. Figure 4 shows modelled ozone concentrations at ground level on the national domain (resolution 20 km x 20 km horizontal grid), averaged over July 1999, as color-filled blocks. An additional layer is added, showing measured values of ozone concentrations at rural background monitoring stations. Starting with hourly values, the same statistical parameter is calculated, a vector point theme is produced and then represented with the same color palette. Then, ancillary layers can help the geographical interpretation of the map. Datasets are classified according to their distribution and relative frequencies. A histogram special stretching is applied in order to enhance the variability of data and facilitate the visualization in complex situations, such as the different variability over the land and the sea. Moreover, the most suitable coloured scale is chosen for visual interpretation. Overlaying model output raster cells and monitoring stations vector points allows immediate evaluation of the agreement of measurements with simulated data. In this application the most significant discrepancies were found for the station_code 200710 (alpine region), caused by the model limitations in simulating complex orography situations, and for the station code 803805, probably due to low detail in the representation of the local emission sources. In similar cases, simulations using the 4km resolution grid allow to improve model predictions, as shown in Mircea et al., [2010b] where a preliminary approach of the operational evaluation of MINNI concerning ozone concentrations is presented.

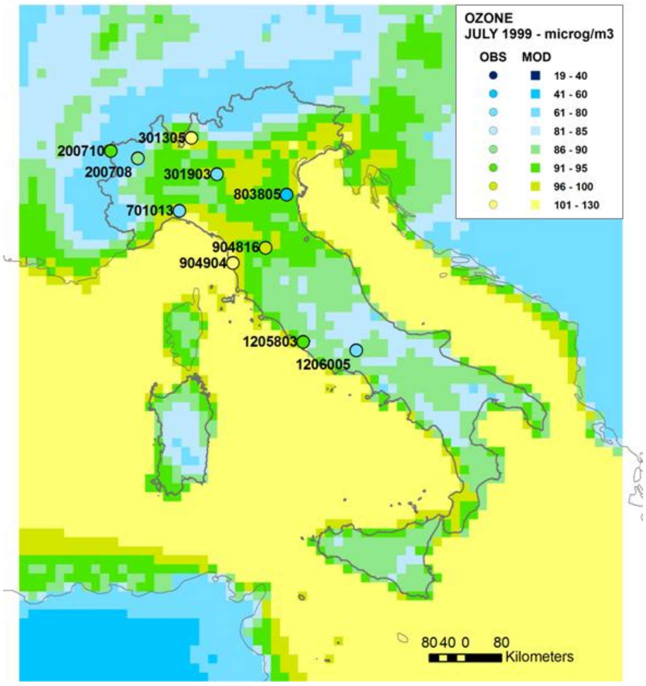

Figure 4. Comparison between observed (circles) and simulated data (AMS gridded output) for ozone hourly concentrations averaged over July 1999.

4 CONCLUSIONS

In the framework of MINNI's activities on atmospheric pollution modelling on Italian territory, a geoprocessing workflow that integrates GIS tools with external procedures was developed for flexible browsing of model results, with rapid production of data maps and plots, and a complete analysis of model performances in comparison with air quality measurements.
The first module of the workflow is the ASQUA relational database, a useful tool to retrieve and investigate the information regarding the Italian monitoring network, that are stored at the national level. The information is managed, processed and updated in GIS and is easily available to drive the analysis on air quality measurements.
Processing modules include data extraction and elaboration, in order to compare simulations against observed quantities at air quality monitoring sites, and to generate 2D maps and single point time-series plots of pollutant emission and concentration parameters on the native model raster grid.
The tool chain described here is currently operational for routine exploration and validation of the atmospheric modelling system, for a wide set of actions: mapping of real emission sources, modelled pollutant distribution on whole Italian territory and identification of hot spots, overlay and thematic mapping of measurement networks, quantitative statistical analysis of model performance following EU guidelines, selection of data subsets for further analysis and research. A clear workflow and procedure for data extraction and processing was defined by virtue of the availability of ENEA GRID computing facility, and further implementations are forthcoming in order to fully operate these procedure in GIS environment and provide web-public access to MINNI results.

ACKNOWLEDGMENTS

This work is part of the MINNI (Integrated National Model in support to the International Negotiation on Air Pollution) project, funded by the Italian Ministry for Environment and Territory and Sea and carried out by ENEA.
The authors would like to thank Felicita Russo for the revision of the paper.

REFERENCES

Boylan, J.W. and Russel, A.G. PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. *Atmospheric Environment*, 40, 4946-4959, 2006

Briganti, G., A. Cappelletti, M. Mircea, A. Pederzoli, L. Vitali, G. Pace, P. Marri, C. Silibello, S. Finardi, G. Calori, G. Zanini. Testing the capability of the MINNI atmospheric modelling system to simulate air pollution in Italy, *In: HARMO13 proceedings*, Paris,1-4 June 2010, pp.161-165. ISBN:2-8681-5062-4.

Caricchia, A.M, L. Merluzzi, A. Scaramella and L. Vaccaro. La banca dati nazionale di qualità dell'aria. *Proc. of 7th National Conference of Environmental Agencies*. APAT-ARPAL. ISBN – 88-448-0131-0, 2003.

Chang, J.C. and Hanna, S.R. Air quality model performance evaluation. Meteorol Atmos Phys 87, 167–196, 2004. DOI 10.1007/s00703-003-0070-7.

Cotton, W.R., R.A. Pielke, R.L. Walko, G.E. Liston, C.J. Tremback, H. Jiang, R.L. Mc Anelly, J. Y. Harrington, M.E. Nicholls, G.G. Carrio, and J.P. McFadden,. RAMS 2001: Current status and future directions. *Meteorology and Atmospheric Physics*,82, 5-29, 2003Council Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008: On ambient air quality and cleaner air for Europe (The Framework Directive). From the *Official Journal of the European Communities, 11.06.2008, En. Series, L152/51*, 2008.

D'Elia, I., M. Bencardino, L. Ciancarella, M. Contaldi and G. Vialetto. Technical and Non-Technical Measures for air pollution emission reduction: The integrated assessment of the regional Air Quality Management Plans through the Italian national model. *Atmospheric Environment*, 43, 6182-6189, 2009.

Jonson, J. E., A. Stohl, A. M. Fiore, P. Hess, S. Szopa, O. Wild, G. Zeng, F.J. Dentener, A. Lupu, M.G. Schultz, B.N. Duncan, K. Sudo, P. Wind, M. Schulz, E. Marmer, C. Cuvelier, T. Keating, A. Zuber, A. Valdebenito, V. Dorokhov, H. De Backer, J. Davies, G.H. Chen, B. Johnson, D.W. Tarasick, R. Stubi, M.J. Newchurch, P. von der Gathen, W. Steinbrecht, H.A. Claude. Multi-model analysis of vertical ozone profiles *Atmos. Chem. Phys.*, 10, 5759–5783, 2010.

Mircea, M., G. Zanini, G. Briganti, A. Cappelletti, A. Pederzoli, L. Vitali, G. Pace, P. Marri, C. Silibello, S. Finardi, G. Calori. Modelling Air Quality over Italy with MINNI Atmospheric Modelling System: from regional to local scale. In: *Air Pollution Modelling and its Application*, D.G. Steyn, S. Trini Castelli Editors. ISBN 978-94-007-1358-1. DOI 10.1007/978-94-007-1359-9_82, 2010a.

Mircea, M., A. Cappelletti, G. Briganti, L. Vitali, G. Pace, P. Marri, C. Silibello, S. Finardi, G. Calori, G. Zanini. Impact of horizontal grid resolution on air quality modeling: a case study over Italy In: *HARMO13 proceedings*, Paris,1-4 June 2010, pp.161-165. ISBN:2-8681-5062-4.2010b.

Silibello, C., G. Calori, G. Brusasca, A. Giudici, E. Angelino, G. Fossati, E. Peroni and E. Buganza. Modelling of PM_{10}concentrations over Milano urban area using two aerosol modules. *Environmental Modelling & Software*, 23, 333-343, 2008.

Zanini, G., T. Pignatelli, F. Monforti, G. Vialetto, L. Vitali, G. Brusasca, G. Calori, S. Finardi, P. Radice and C. Silibello. The MINNI Project: an integrated assessment modelling system for policy making. *Proceedings of MODSIM 2005 Int. Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand, 2005-2011. ISBN: 0-9758400-2-9, 2005.

Zender, C.S. Analysis of self-describing gridded geoscience data with netCDF Operators (NCO). *Environmental Modelling & Software*, Vol. 23, Issues 10–11, 1338–1342, 2008.