Preparation of emission data for modeling with CMAQ from Spanish emission inventories and emission projections

R. Borge, J. Lumbreras & M.E. Rodríguez.
Departamento de Ingeniería Química Industrial y del Medio Ambiente, Universidad Politécnica de Madrid (UPM). C/ José Gutiérrez Abascal, 2. 28006- Madrid. Spain.
rborge@etsii.upm.es

ABSTRACT

The Technical University of Madrid (UPM) is currently studying the different alternatives to reduce air pollution in Spain. The policy design and decision-making for improving Air Quality are based on emission projections up to 2020 under several scenarios. The following step is to evaluate the impact of these scenarios using an Air Quality Model. The US EPA CMAQ modeling system has been selected and therefore the focus is on providing SMOKE all the Emission Inventory information to generate model-ready emission data.

Emission inventory preparation is a critical stage in Air Quality Modeling and there is a great lack of quality input information in the Spanish case.

In spite of having a very flexible design, CMAQ and SMOKE modeling systems are developed taking into account the inventory formats and particular conditions of the United States. The National Spanish Inventory is made according to the European CORINAIR methodology, so a specific methodology to adapt it to the modeling system is needed. This paper shows briefly the main issues of this process.

INTRODUCTION

Our research group within the Department of Chemical and Environmental Engineering of the is currently studying the different alternatives to reduce air pollution in Spain. In order to evaluate the alternatives to improve air quality, the emissions for 2010 and 2020 have been projected. The base data is the emission inventory for Spain up to 2000. The emission projections that are being developed provide a very useful tool to support the planning of measures intended to meet the legal requirements and international commitments and protocols. Once the projections are finished it would be very valuable to obtain the related air quality values through an Air Quality Model. Namely, the selected AQM is the US EPA CMAQ (Linux version 2.1 platform). Modeling future scenarios will supply the possibility of evaluating the incidence of each activity in air quality as well. Therefore, this could improve the assistance for policy assessment and decision-making significantly.

At this moment the National Inventory is being prepared for modeling according to the standards required for the CMAQ modeling system.

BACKGROUND

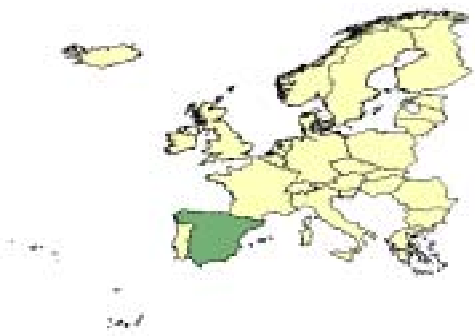
The CORINAIR methodology has been used and the associated nomenclature called SNAP (Selected Nomenclature for Air Pollution) has been selected to complete this study, in a way consistent with the CORINE-AIRE Spanish National Inventory[1] (hereinafter Inventory). All European inventories are developed in the same framework.

The inventory is calculated in a yearly basis. As for the spatial resolution, the geographical reference is provided by the European Statistical Office through the European administrative classification (NUTS, Nomenclature of Territorial Units for Statistics). This is a hierarchical classification scheme with 4 levels, as can be seen in Table 1. National Inventories are disaggregated at level 3.

Table 1. NUTS classification. Application to Spain

Level	Denomination	Spanish units	
		Denomination	**Number**
0	Country	-	1
1	European Community Regions	Large Region	7
2	Base Administrative Units	Autonomous Region	19
3	Base Administrative Subunits	Province	52

Figure 1. Administrative division



NUTS 0 (Countries)

NUTS 2 (Autonomous Regions)

NUTS 3 (Provinces)

This study considers all the pollutants under the Geneva Convention on long-range cross-border pollution (SO_2, NO_X, NH_3, NMVOCs, CO, PM, Pb, Cd and Hg) as well as the Kyoto protocol compounds (CO_2, CH_4, N_2O, PFCs, HFCs and SF_6).

SNAP nomenclature

The CORINAIR inventory covers practically all the activities considered in the most up to date version (SNAP 97) of the nomenclature known as Selected Nomenclature for Air Pollution (SNAP). It has been developed as part of the CORINAIR project and harmonized with the nomenclature included in the IPCC/OECD (the Intergovernmental Panel on Climate Change and the Organization for Economic Co-operation and Development) and with that of the EMEP (European Modeling and Evaluation Program) of the United Nations Economic Commission for Europe (UNECE). SNAP-97 has a structure based on the following three level hierarchy: Group, Sub-group, and Activity.

- The highest level, **Group**, comprises 11 divisions reflecting the largest categories of anthropogenic and natural activities.
- The middle level, **Sub-group**, divides the previous level into 76 classes reflecting the structure of the emission-producing activities in accordance with their technological and socio-economic specifications.
- The most disaggregated level, **Activity**, includes the core activities and those of some sub-groups (16) which do not appear disaggregated in activities. In total, SNAP-97 incorporates 414 core activities which, together with the 16 sub-groups without any further breakdown into activities, add up to 430 emitter activities/sub-groups.

It should be borne in mind that this classification does not follow a purely economic or technological criterion, but is rather the result of a merger of both. Each level is identified by a 2-digit code, so every activity has a unique 6-digit code.

METHODOLOGY

Preliminary studies developed by our research group showed that the EPA Community Multiscale Air Quality Model (CMAQ) is one of the most flexible and reliable Air Quality Modeling system available nowadays. Taking into account the Spanish features this model seems to fit perfectly to address the need for modeling dispersion and photochemical processes. Moreover, there are some references of successful applications to some Spanish regions[2].

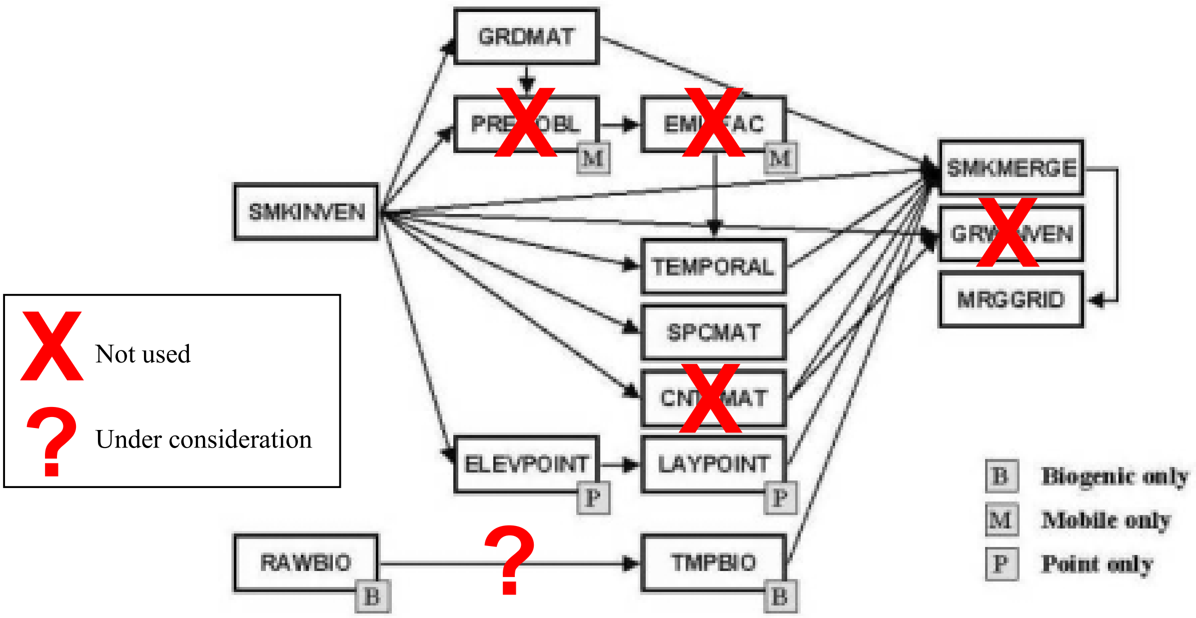
The choice of this model (stand-alone version) leads to the usage of the Sparse Matrix Operator Kernel Emissions (SMOKE) system for preparing model-ready emission data. This Modeling System created by MCNC-North Carolina Supercomputing Center permits both rapid and flexible processing of emissions data, but needs a great amount of information from the inventory. In order to provide all required data, the user manual has been followed step by step. The software version used is SMOKE V.1.4.2 for PC (running from scripts).

In spite of having a very flexible design, both modeling systems (CMAQ and SMOKE) are developed taking into account the inventory formats and particular conditions of the United States. This means an obstacle for its application to other countries. Therefore, the first step in the Spanish case is to develop a specific methodology. The main issues of this process are shown as follows.

The SMOKE approach

The SMOKE modeling system is built from a group of processors, capable of transforming inventory data set into CMAQ emission input data. This transformation is produced through several stages (data structuring, temporal and spatial processing, chemical speciation processing, etc.). Not all processors have been used for this application, since some of the data given as a result from the model have been provided as input data.

Figure 2. Information flow through SMOKE core programs



In all the files, the Country/State/County code has been replaced by an analogous Country/Autonomous Community/Province code. In the same way, all references to SCC codes have been turned into 6-digit SNAP codes.

Types of sources

Only two basic classes of sources for emissions have been considered: point and area sources. All files containing the information about point and area sources have been generated in IDA format (ASCII files).

Area sources

Area sources generally comprise various emission-producing units (activities in the primary sector such as agriculture and cattle breeding, mining, industrial plants, commercial and residential centers and sites, natural spaces). These sources have no individual significance or are inherently linked to a particular geographical area, and thus must be dealt with aggregately.

Besides these traditional sources, all road mobile sources have been taken into account.

The processors for mobile sources allow SMOKE to calculate and manage emissions from activity data (vehicle-miles traveled) using MOBILEb5. However, the emissions and projections for Spain have been calculated through COPERT III (COmputer Programme to calculate Emissions from Road Transport), developed by the European Environment Agency (EEA) according to the CORINAIR methodology. This is a complex system that calculates mobile source emissions on national, yearly basis, for each SNAP activity within the group 7 (road transport). Once all the emissions from mobile sources have been calculated, they have been introduced into SMOKE in the same way as any other area source.

As far as biogenic sources are concerned (SNAP group 11), in this first approach they have not been taken into account yet.

Point sources

Point sources are those which, because of their significance for the Inventory, must be addressed individually. The information about the exact location of these point sources (geographical coordinates) is provided in the Inventory. The criteria for identifying large point sources have been those proposed in the EMEP/CORINAIR methodology. Those criteria are specified below in Table 2.

Table 2. Criteria for the identification of large point sources

Sector	Criterion
Power Plants	All conventional power plants are included if their capacity is larger than or equal to 300 MW, regardless of the fuel they use.
Oil Refineries	All oil refinery plants are included.
Sulphuric Acid Plants	All plants, even those installations for the reduction of SO_2 in the non-ferrous metal industry and in any other type of industry.
Nitric Acid Plants	All plants.
Integrated Iron/Steel Plants	All those plants producing more than 3 million tons per year.
Paper Pulp Plants	Plants with a production capacity of more than 100,000 tons per year of pulp are included (9% of dry matter, prepared for use and dispatch).
Vehicle Painting Plants	Centers with a capacity of more than 100,000 vehicles per annum on their painting lines, or an equivalent volume.
Airports	All those airports with more than 100,000 LTO cycles (Landing and Take-Off turnarounds) per annum. Traffic at heights of more than 3,000 feet are not included in the LTO cycle.
Any plant where the emissions of:	SO_2, NO_X and NMVOC are larger than 1,000 tons per annum; or those of CO_2 which are larger than 300,000 tons per annum.
Any site where the emissions	Are vented using stacks greater than or equal to 100 meters in height (of relevance for EMEP).

As has been mentioned before, the Department of Chemical and Environmental Engineering of the Technical University of Madrid is currently working in the SEP (Spain's Emission Projections) project. One of the main tasks of this study has been the development of the program called EmiPro 1.0. This software is a Microsoft Access database application prepared to store and manage the projections. This tool is able to export the point source database files (both past and future years) to SMOKE-ready IDA ASCII format file. Therefore, it is possible to add or remove any point source in any year in a quite straightforward way.

The application of a specific methodology and software to create emission projections is the reason for not using CNTLMAT or GRWINVEN processors. Instead, the emissions for any future year are calculated and exported and then SMOKE is run from scratch with those new data as input.

SMOKE considers also biogenic sources. In this first approach this type of source has not been addressed. However, once the vegetation covertures are available and are reclassified into BEIS categories, we intend to generate the BGUSE file applying the methods used to create surrogate data or any other gridded information.

Temporal processing

For temporal processing, specific temporal profiles and cross-references files had been generated to convert the inventory information on annual basis to values from day and hour-specific emissions data sets. Cross-references for area and mobile sources are in the same file.

Though Spain is in GMT +1, the OUTZONE variable is set to 0 to avoid negative values. This issue is taken into account in the temporal references.

Chemical speciation

In principle, the photochemical mechanism CB-IV is profitable for this application, so the chemical speciation files included in SMOKE have been used. The speciation profile for this mechanism and cross-reference files have been adapted to the pollutants of interest. For this first stage of the inventory preparation the default split factors and no source-specific profiles have been used.

Spatial allocation

The spatial processing operation or gridding is a key issue in the transformation of inventory emissions to gridded emissions useful for the Air Quality Model. The main considerations for this process are the following:

The grid

As has been mentioned before, European inventories are developed in the framework of the CORINAIR methodology. This implies that the spatial resolution of the National Inventory is based on administrative units NUTS-3. The use of these units as reference is perhaps the most convenient both from the standpoint of the collection of basic data and from the point of view of the exploitation of the results for the decision-making process on environmental policy.

However, in order to provide useful information for the requirements of analysis of air pollution processes, it is indispensable to breakdown the inventories using a regular projection grid covering the territory under study. To this end, all the European inventories are re-calculated on the so-called EMEP grid (EMEP/OECD). Figure 3 shows an example of geographical projection of Spanish emissions onto the EMEP grid.

The EMEP grid system is based on a polar-stereographic projection with real area at latitude 60º N. The y-axis is oriented parallel to 32º W defined as a negative longitude if west of Greenwich. The EMEP 50x50 km^2 domain includes 132 x 111 cells (with x-axis varying from 1 to 132 and y-axis varying from 1 to 111), as it is shown in Figure 4. The Spanish domain is covered by 311 cells (The Canary Islands are not included in this study).

Figure 3. Projection of SO_2 emissions for the year 2000.

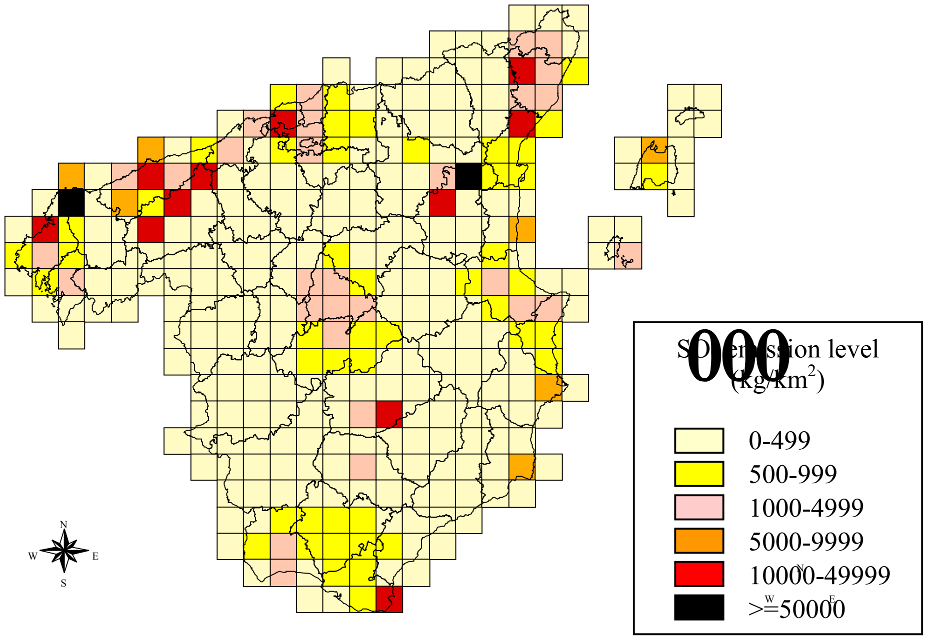
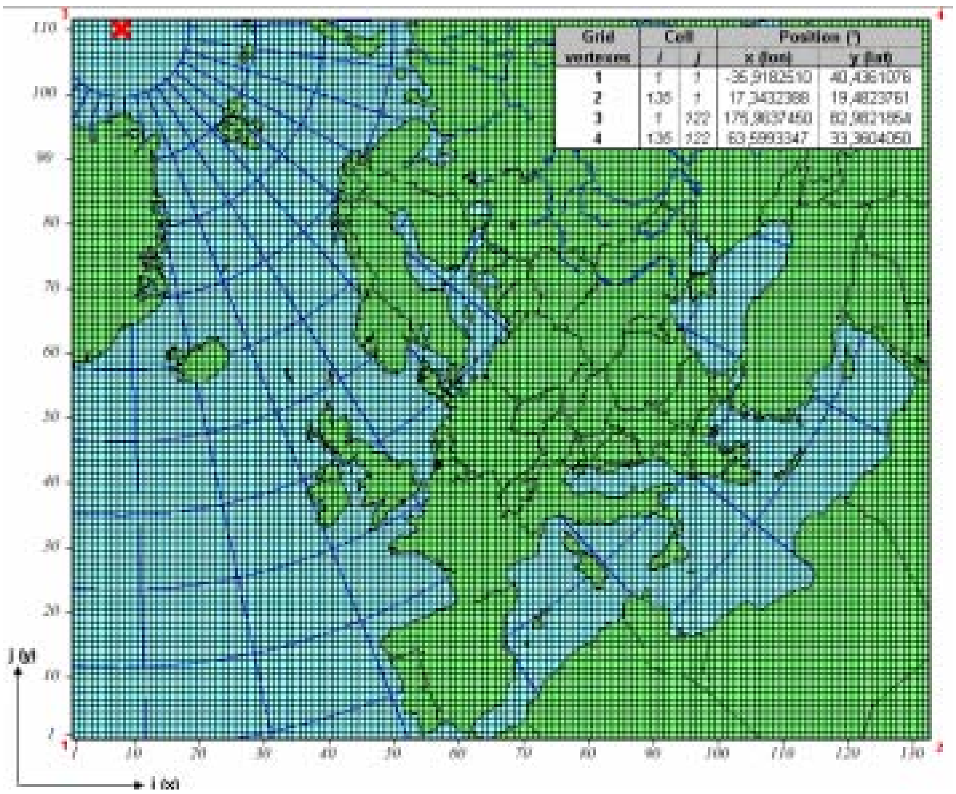


Figure 4. EMEP 50x50 grid system.



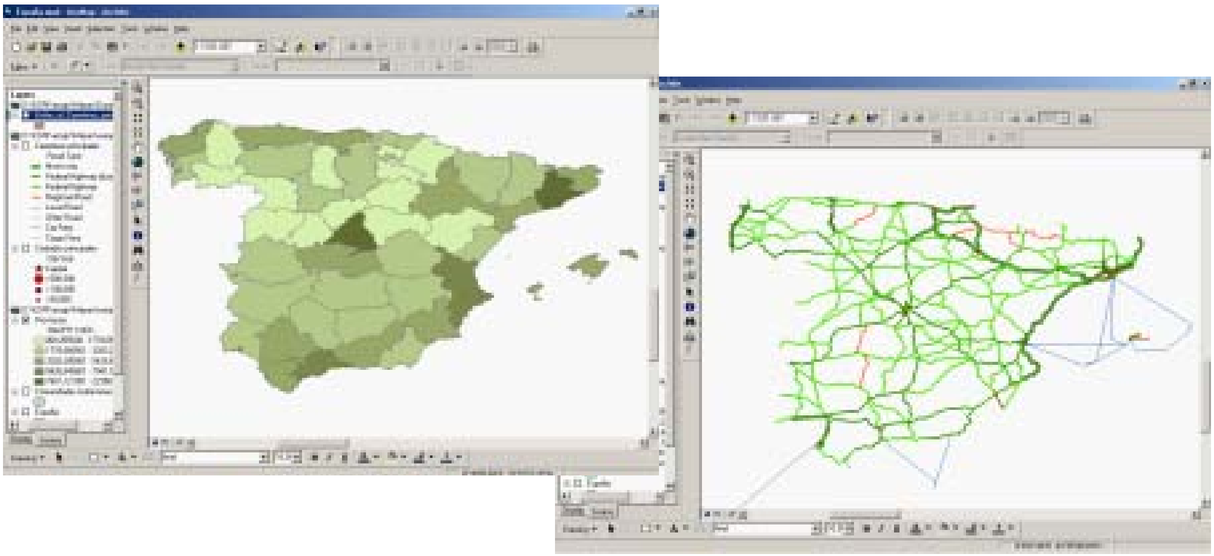
To guarantee the compatibility and comparability with CORINAIR procedures, a grid definition based on EMEP grid has been chosen. On the other hand, a finer grid resolution is needed to get better results in the modeling process. Therefore, each EMEP grid cell has been split into 100 cells of 5x5 km^2, thus obtaining the regular reference grid used to conduct the spatial allocation process.

Spatial Surrogates

The definition of spatial surrogates allows allocating Province-wide emissions. The CORINAIR database includes information of traditional surrogate data (population, etc.) for NUTS-3 level. This and more specific information (e. g. city population) have been use to generate spatial surrogates. An effort is being done to identify the most relevant set of surrogates for each source type or group of sources related, but the procedures to perform the overlay and surrogate calculation are already defined. The spatial allocation process is supported by ESRI ArcGIS 8.1.2, and it is based in scripts written in Visual Basic. Those programs perform the overlay and calculation of surrogate ratios by grid cell over the entire domain. Within each Province the sum of these ratios is equal to one. Finally that information is saved according to the AMGPRO file format.

The raw information is the emission data to be allocated (Inventory database) and the ArcInfo / ArcView shapefiles, covertures or layers containing the administrative division and the location of the surrogates used in the allocation process, as it is shown in Figure 5.

Figure 5. SNAP 7.1 (road transport – passenger cars) Province-wide emission and layer of Mayor Roads; one of the surrogates employed to spatial allocation of traffic emissions.



The definition of suitable spatial surrogates allows to get a more realistic spatial allocation and better results in Air Quality Modeling, so this is a feature still under development.

CONCLUSIONS

Emission inventory preparation for modeling is a critical step on Air Quality Modeling. US EPA CMAQ is a powerful tool to evaluate potential measures for air quality improvement, but as important as to use a good model is to have the proper input information ("Junk In Equals Junk Out"). This paper shows the initial efforts to achieve a high quality emission data set suitable for modeling from the Spanish Emission Inventory. The application of CMAQ and therefore SMOKE outside the United States implies an additional obstacle, since there is a lack of data and documentation.

The process of adapting the inventory is still starting, and there is a need to keep working in this line to improve the methodology exposed in this paper and to reduce the uncertainty about the whole modeling procedure, mainly in the issues related with the compatibility with MM5 and CMAQ.

REFERENCES

1- AED. "CORINAIR Spain 1997, 1998, 1999 and 2000 Inventories of Emissions of Pollutants into the Atmosphere" Ministry of the Environment. Secretariat-General for the Environment Directorate-General for Environmental Quality and Assessment. June 2002

2- S. José, R; Pérez, J.L.; Pleguezuelos, C.; Camacho, F. and González, R.M. (2002). MM5-CMAQ Air Quality Modeling process analysis: Madrid case. Air Pollution X. Wessex Institute of Technology.

BIBLIOGRAPHY

Adelman, Z. and M. Houyoux, 2001. "Processing the National Emissions Inventory 96 (NEI96) version 3.11 with SMOKE". Presented at the Emission Inventory Conference: One Atmosphere, One Inventory, Many Challenges, 1-3 May, Denver, CO, U.S. Environmental Protection Agency.

Boulton, J. W.; Siriunas, K. A.; Lepage, M.; Schill, S. 2002 "Developing Spatial Surrogates for Modeling Applications". Rowan Williams Davies and Irwin, Inc. Proceedings of the Eleventh annual symposium on emission inventories, Atlanta, GA. Sponsored by the US EPA.

EMEP/CORINAIR. "Atmospheric Emission Inventory Guidebook". Third Edition. 2001. UNECE-Convention on long-range transboundary air pollution & European Environmental Agency.

Houyoux, M., R. and J. M. Vukovich, 1999. "Updates to the Sparse Matrix Operator Kernel Emissions (SMOKE) Modeling System and Integration with Models-3". Presented at The Emission Inventory: Regional Strategies for the Future, 26-28 October, Raleigh, NC, Air & Waste Management Association.

Kinnee E.; Beidler A.; Possiel N.; Stella, G. 2001. "Revised Methodology for the Spatial Allocation of VTM and Mobile Source Emission Data". US. EPA. Presented at the Emission Inventory Conference: One Atmosphere, One Inventory, Many Challenges, 1-3 May, Denver, CO, U.S. Environmental Protection Agency.

Lumbreras, J.; Urquiza, G. and Rodríguez M.E. 2002. "Emission inventory and emission factor projections for modeling air pollution in the CAM (Spain)". Proceedings of the Eleventh annual symposium on emission inventories, Atlanta, GA. Sponsored by the US EPA.

Posch M.; de Smet P.A.M.; Hettelingh J.P.; Downing R. J. 1999. "Calculation and Mapping of Critical Thresholds in Europe: Status Report 1999". RIVM Report No. 259101009. Coordination Center for Effects. National Institute of Public Health and the Environment. Bilthoven, Netherlands.

Stella, G. 2001. "Inventory Preparation for Emissions Modeling". US. EPA. One Atmosphere, One Inventory, Many Challenges; Conference Emission Inventory Training Session April 30, 2001.

KEY WORDS

Emission Inventories
Preparation for modeling
CORINAIR methodology
SMOKE

ACKNOWLEDGMENTS

The authors would like to thank the staff of the *Dirección General de Evaluación y Calidad Ambiental (MIMAM)* for their collaboration.