EMEP/MSC-W Technical Report 1/2008
Date: July 2008

EMEP Co-operative Programme for Monitoring and Evaluation of the Long-Range Transmission of Air Pollutants in Europe

Towards the development of a common EMEP global modeling framework

Edited by Leonor Tarrasón and Alexey Gusev

Joint MSC-E, CCC and MSC-W Technical report

EMEP/MSC-E: Alexey Gusev, Oleg Travnikov and Valery Sokovych

EMEP/CCC Kjetil Tørseth

EMEP/MSC-W: Michael Gauss, Jan Eiof Jonson, Ágnes Nyíri David Simpson, Peter Wind and Leonor Tarrasón

Norwegian Meteorological Institute
ISSN 1504-6079 (print)
ISSN 1504-6206 (online)

Contents

Executive Summary **5**

1. Towards a common EMEP global model **9**
1.1 Introduction 9
1.2 Requirements for the common EMEP global modeling framework 10
1.3 Structure of this report 12
1.4 References 12

2. Co-operation for a hemispheric measurement database **13**
2.1 Planned work and expected results 14
2.1.1 WP1 Generic set of data standards 14
2.1.2 WP2 Core dataset of integrated observations 16
2.1.3 WP3. Recommendations for future work 18

3. Emission data in global scale **19**
3.1 The HTAP emission inventory on hemispheric scale 19
3.2 Emissions of SOx, NOx, NH3, VOC, CO and particulate matter 20
3.3 Heavy metals and mercury (Hg) 22
3.3.1 Mercury 23
3.3.2 Lead 24
3.3.3 Cadmium and other heavy metals 25
3.4 Available emissions of Persistent Organic Pollutants 26
3.4.1 PCBs 26
3.4.2 HCHs 27
3.4.3 HCB 29
3.4.4 PCDD/Fs 29
3.4.5 PAHs 30
3.5 References 30

4. Geophysical input data: Land cover and LAI **33**
4.1 Description of available land cover data 33
4.1.1 UMD land cover dataset 33
4.1.2 USGS Global Land Cover Characteristics database (GLCC) 34
4.1.3 Global Land Cover 2000 database (GLC 2000) 35
4.1.4 MODIS Land Cover products 36
4.2 Comparison of land cover datasets 37
4.3 Available Leaf Area Index information 41
4.4 Concluding remarks on geophysical data 42
4.5 References 43

5. Comparison of meteorological input data **45**
5.1 Description of the GEM and the IFS meteorological drivers 45
5.1.1 Short description of the IFS at ECMWF 45
5.1.2 Short description of the GEM model at CMC 47
5.2 Selected sites and fields used for meteorological driver evaluation 50
5.3 Statistics used 52
5.4 IFS and GEM output data comparison for 2001 53
5.4.1 Precipitation 53

5.4.2 Air Temperature 56
5.4.3 Wind Speed 60
5.5 Spatial variability of validated vertical profiles 61
5.6 Concluding remarks 64
5.7 References 65

6. Comparison of atmospheric dispersion by the EMEP global models 67
6.1 Global models at MSC-W and MSC-E 67
6.1.1 The Unified global model at MSC-W 67
6.1.2 The GLEMOS global model at MSC-E 68
6.2 The TP1x experiment under TFHTAP 70
6.3 Comparison of tracer dispersion by the EMEP/MSC global models . 72
6.4 Comparisons in the free troposphere, including vertical profiles 73
6.5 Source receptor relationships 77
6.6 Conclusions from the comparison of tracer dispersion 79
6.7 References 80

Executive Summary

This report was prepared for the 32nd session of the Steering Body to EMEP. It presents the initial progress in the development of the common EMEP global modeling framework by the two modeling centers in EMEP, MSC-East and MSC-West, as recommended by the Steering Body on its 31st session.

In recent years, both MSC-East and MSC-West have developed their modeling capacities to include air pollution transport in global scale. The challenge now is to streamline these efforts in order to better serve the needs of the Convention, in particular with regard to establishing the state and origin of pollution in Eastern European, Caucasus and Central Asian (EECCA) countries.

This is a first step in a joint co-operative process that will result by 2012 in an operational and validated global modeling framework able to provide timely long-range transboundary pollution information for all Parties and all pollutants under the LRTAP Convention. The main goal is to develop an improved framework under EMEP to describe transboundary air pollution in EECCA countries, including information on pollution transport from contiguous countries like Pakistan, India and China.

The report presents progress in the compilation and evaluation of input data. The input data considered here are: emissions, geophysical information on land use and vegetation characteristics, and meteorological fields. In addition, the report includes, for the first time, a comparison of the two global models respectively developed by MSC-W and MSC-E. Since the focus of the two models is on different pollutants, the comparison is made for a passive tracer, and involves the analysis of the transport characteristics derived by both models.

The work for the development of the common EMEP global modeling framework benefits to a great extent from the co-operation, information exchange and expertise under the Task Force on Hemispheric Transport of Air Pollution (TFHTAP), where the EMEP modeling centers participate as two of many. There are in particular three initiatives under TFHTAP that are of special relevance for this work. First is the compilation of validated monitoring data in a hemispheric scale to help establish the quality of modeling results. Second is the compilation of updated, consistent and complete emission inventories for the Northern Hemisphere. Third, and not least, is the on-going series of model inter-comparison tests to determine the extent of inter-continental transport that can also serve as benchmarking for the individual participating models.

The first version of the HTAP observational database will be launched in September 2008. By autumn 2008, the first version of the HTAP harmonized emission inventories for main pollutants is also expected to be available. This will enable further evaluation of the emission input data for main pollutants and particulate matter. For heavy metals, mercury and selected persistent pollutants, however, the HTAP emission database will still take some time to be completed. Therefore, work at MSC-E is progressing to compile existing data on these compounds and will also rely on the updates from the Artic Monitoring and Assessment Programme (AMAP).

A fist compilation of land-use information from different satellite sources has been carried out. However, none of the land-use datasets investigated here contains information on the species-composition of forests. Such information on species can be very valuable for biogenic VOC emission estimates, the same as information on the Leaf Area Index. Further compilation and evaluation of relevant geophysical information is presently under progress.

The core comparison of this report is on the use of meteorological input data. The goal was not to choose a single meteorological driver, but to secure that the framework is flexible enough to allow the use of different drivers. The comparison carried out in this report shows that the different candidates to be included as meteorological drivers are well suited for the purpose.

The final chapter of this report compares the results of tracer experiments carried out separately under the TFHTAP inter-comparison by MSC-East and MSC-West. The results show indeed the comparability of transport calculations carried out by the two different global models. The comparability of these results is reassuring and opens for further harmonization of the two global models as planned in the second stage of this EMEP global framework development.

The choices for input data presented in this report are not final. The compilation, review and improvement of input data is a continuous and inherent process to model development. Still, this common effort by MSC-E and MSC_W is a good indication that the first steps are taken to harmonize the use and criteria for input data in the new global modeling framework for EMEP.

Acknowledgments

We are thankful to John van Aardenne (DG JRC IES) for his contribution to this report , informing on his work on emissions carried out under HTAP (section 3.1).Our colleagues at MSC-E, Ilia Ilyin and Irina Strizhkina, and at MSC-W, Per Helmer Skaali, are greatly acknowledged for their technical support with the elaboration of this report. Special thanks are also due to Bjørg Rognerud and Ivar Isaksen for facilitating access to the IFS meteorology at ECMWF.

This work has been funded by the EMEP Trust Fund. For MSC-W, in addition, work with the development of the Unified EMEP global model version is also partly funded by the Norwegian Ministry of the Environment. The work at MSC-W has also received support from the Research Council of Norway through computer resources at the Norwegian Meta Center for Computational Science (NOTUR).

1. Towards a common EMEP global model

Leonor Tarrasón and Oleg Travnikov

1.1 Introduction

The development of a modeling system that can trace the long-range transport of air pollution in a global scale is a new priority in the work of EMEP.

The requirement for the development of a common modeling system within EMEP has emerged from the increased focus of the Convention on Long-range Transboundary Air Pollution on Eastern European, Caucasus and Central Asian (EECCA) countries. The task of EMEP within the LRTAP Convention is to provide information on the origin of air pollution in the area and contribute to enhance their competence to account and solve their air pollution problems.

Already in 2004, the Convention established the Task Force on Hemispheric Transport of Air Pollution (TFHTAP). The main focus of the TFHTAP is to determine the extent and impacts of intercontinental transport of air pollution and foster international co-operation to reduce air pollution.

This year, for the first time, information on the status of transboundary air pollution and source-receptor information has been provided by EMEP for 5 new EECCA countries: Kazakhstan, Turkmenistan, Uzbekistan, Kyrgyzstan and Tajikistan.

This achievement has been challenging from the technical point of view as it has required the extension of the traditional EMEP modeling to an area where there is little background information on emission sources. Still, the present extension of the EMEP modeling area has recognized drawbacks. In the first place, the current extended EMEP domain only partly covers territory of one of the EMEP countries – the Russian Federation. It is also a recognized drawback that results on air pollution in these central Asian countries are highly dependent on sources outside the calculation domain. Countries in Central Asia are contiguous with other Asian countries, like China, India, Pakistan and Iran, that significantly affect pollution levels over the EECCA territories but are not included directly in the calculations. Consequently, the current EMEP modeling capacity for EECCA countries is only an interim solution up to 2011.

The necessary further extension of the EMEP modeling capacity to better account for pollution levels in EECCA countries will cover the global scale. This is because many of the pollutants considered under the Convention, like ozone, mercury, some persistent organic compounds, and particulate matter are subject to intercontinental transport. For long-lived components, like mercury and some persistent organic compounds, dispersion of pollution in the atmosphere takes place on a global scale. For more short-lived components, like particulate matter and ozone, episodic intercontinental transport is of relevance to determine their atmospheric levels in the free troposphere. For all these pollutants, convection transport in the tropics becomes relevant

process to take into account for an accurate description of the pollution levels in the free troposphere. Consequently, a hemispheric modeling approach would lead to uncertainties in the description of pollution levels near the equator because convection processes become part of the boundary conditions. This drawback in hemispheric modeling is particularly important in Southeast Asia where growing economic development and significant increases of emission levels in recent years will have an impact in pollution levels in neighboring EECCA countries. Thus, global approaches are recommended to determine air pollution levels in the EECCA countries.

The Steering Body of EMEP recommended last year the development of a common EMEP global modeling system and urged the two modeling EMEP centers (MSC-E and MSC_W) to co-operate in this task. The work of the EMEP modeling centers should link and benefit from the work within the Task Force on Hemispheric Transport of Air Pollution and should result in the end in a better framework to describe transboundary air pollution in EECCA countries.

This report presents the initial progress in the development of the common EMEP global modeling system. This is a first step in a common development process that is expected to culminate in 2012 with the operationalization of a new global system to provide timely long-range transboundary pollution information for all Parties and all the pollutants under the LRTAP Convention.

1.2 Requirements for the common EMEP global modeling framework

There exist at present a considerable number of global modeling systems able to trace the transport of air pollution, many of them already contributing to the work under the TFHTAP. Also the EMEP modeling centers have recently developed their capabilities to include global scale modeling. Last year, MSC-W presented for the first time results from the global version of the Unified and a comparison with the global OsloCTM2 model developed at the University of Oslo. (Jonson et al., 2007). This year, MSC-E has extended their hemispheric model version (Gusev et al, 2005) to produce first results with the atmospheric version of their new Global EMEP Multi-Media Modelling System (GLEMOS). First results from this MSC-E global model are presented in Chapter 6 in this report. The challenge is now to combine efforts in order to achieve a common development for a system useful for EMEP activities. The main requirements for a common global modeling framework are:

1) ***Unified multi-pollutant.*** The common EMEP global modeling framework should include all the pollutants considered within the CLRTAP such as sulfur and nitrogen compounds, ozone, PM, heavy metals and POPs. A multi-pollutant approach has the advantage that interactions between these pollutants would be better represented, as for instance, it would allow the description of the impact of particulate matter acting as carrier in the concentrations of heavy metals and POPs.
2) ***Multi-media approach.*** Since the common EMEP global modeling framework should describe the long-range transport of mercury and POPs, it should include their cycling and accumulation in various environmental media: atmosphere, soil, ground water, vegetation, seawater.
3) ***Flexible model domain.*** The common EMEP global model should be flexible to the choice of model domain (global, hemispheric, regional, local) so that the same model can be used to feed information in the boundaries to limited-area applications. This will secure the consistency of the results in different applications.
4) ***Computationally efficient.*** The common EMEP global model should be computationally effective to allow source-receptor calculations. Allocation of the origin of pollution is a

priority within EMEP and demands a large number of model runs carried out in a short period of time. Therefore, EMEP models need to be highly parallelized and computationally efficient.

5) ***Modular model structure.*** The common EMEP global modeling should be flexible to include or exclude certain processes for particular model applications as appropriate. For example, the use of multi-media approach can be unnecessary for modeling atmospheric transport of photo-oxidants, whereas detailed ozone chemistry or aerosol dynamics can be unnecessary for long-term simulations of POP accumulation in the environment.
6) ***Flexible choice of model resolution.*** Additional flexibility in the choice of spatial and vertical resolution is also a requirement in order to allow for targeted applications. For instance, calculations on the accumulation of mercury and POPs in the environment can be performed with relatively low spatial resolution because of lack of detailed historical input data.
7) ***Allow trend analysis.*** The common EMEP global modeling framework should allow calculations for many different years, because the calculation of trends in air pollution is a priority for EMEP. This imposes significant constraints to the framework especially in what concerns the availability of input data for long-term trend simulations.

Figure 1.1 provides an overview of the envisaged general modular structure of common EMEP global modeling framework.

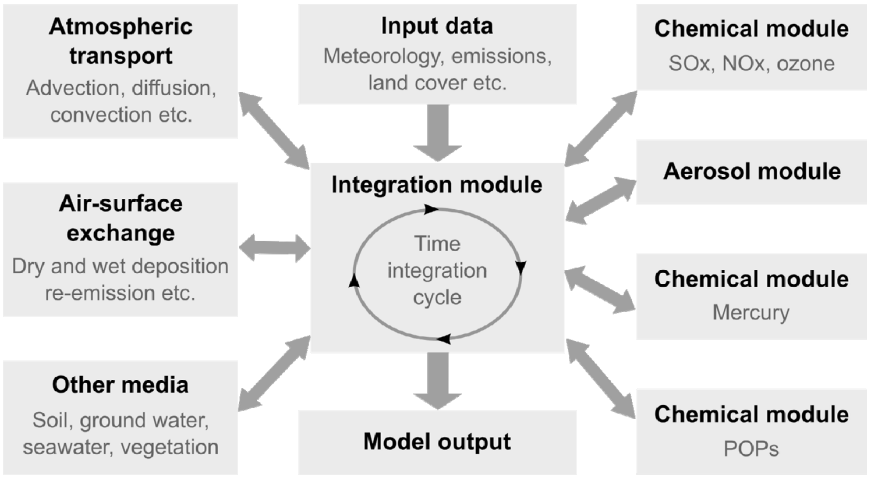


Figure 1.1. *Envisaged general structure of the global modeling framework*

The development of the common EMEP global system involves an extension of the existing global MSC-W and MSC-E modeling systems to include new compounds as well as the especial physical and chemical processes associated to these (full photochemistry, aerosol dynamics, transport over different media, etc.). A stepwise approach in four stages, involving the present capabilities at both centers, has been suggested for the development of the common global framework. These are:

- **Stage 1:** Joint compilation and evaluation input data (emissions, land-use, leaf area index, soil properties). Evaluation the meteorological drivers and driving meteorological input.
- **Stage 2**: Harmonization of description of common physical processes in the models. Identification of modules that need to be re-formulated to allow their introduction in the common framework.
- **Stage 3:** Unification of modeling approaches. Implementation of the common framework and initial implementation of module exchange. Identification of processes that need improved treatment.

- **Stage 4:** Testing and evaluation of the new common framework. Identification of processes that require further study.

1.3 Structure of this report

This report presents the progress of work with the first stage in the development of the common EMEP global modeling framework, namely the stage on the compilation and evaluation of input data.

For many years, the modeling EMEP centers over Europe have been compiling, validating and improving the accuracy of input information to their models: emissions, land cover, land use information, meteorological data. A similar task has not been carried out to the same extent in all areas of the world, so the compilation of good quality input data in the global scale is a challenging task. Data quality control has been always been central in EMEP, with special emphasis in validated monitoring data and observations to establish the quality of the EMEP results. Therefore, the next chapter in this report, after this introduction, deals with the compilation of monitoring data in a hemispheric/global scale. This is an initiative under the Task Force on Hemispheric Transport of Air Pollution (TFHTAP) intended to benefit the whole CLRTAP community and also other users.

The next chapter, Chapter 3, deals with the compilation of global emission information. Again, the new initiative under the TFHTAP to produce a harmonized set of global emissions is most welcome and a valuable contribution for the work of the air pollution community. We look forward to include and analyze these data in the common global modeling system as soon as they are available.

Chapter 4 evaluates different land-cover datasets available from remote sensing and sets the premises for the requirements on physical data information also from biosphere- atmosphere exchange models.

Chapter 5 presents the core comparison of this report, introducing the basis for a harmonization on the use of meteorological drivers for the new EMEP global modeling framework. The goal is not to choose a single meteorological driver, but to secure that the framework is flexible enough to allow the use of different drivers.

The final chapter, Chapter 6, compares the atmospheric transport as derived from the global MSC-E model (GLEMOS) and the global MSC-W model (Unified EMEP global version), based on their contributions to the model inter-comparison exercise under TFHTAP. Since the focus of the two models is on different pollutants, the comparison is made for a passive tracer, and involves the analysis of the transport characteristics derived by both models.

1.4 References

Gusev A., Mantseva E., Shatalov V., Travnikov O. (2005) Hemispheric model (MSCE-Hem) of persistent toxic substances dispersion in the environment. MSC-E Technical Report 11/2005, Meteorological Synthesizing Centre East of EMEP, Moscow. Available from www.msceast.org.

Jonson, J.E., Tarrasón, L., Wind, P., Gauss, M., Valiyaveetil, S. S., Tsyro, S, Klein, H., Isaksen, I.S.A. and A. Benedictow (2007) *First evaluation of the global EMEP model and comparison with the global OsloCTM2 model*. EMEP-MSC-W Technical report 1/2007, Norwegian Meteorological Institute, Oslo, Norway. Available from www.emep.int

2. Co-operation for a hemispheric measurement database

Kjetil Tørseth

Since the initiation of the Task Force on Hemispheric Transport of Air Pollution (TFHTAP) co-chaired by the European Commission and the US EPA, several workshops and meetings have concluded on the need to improve the access and inter-comparability of observational data to allow the assessment of long-range transport of pollutants at intercontinental scale. While several efforts are ongoing, there are still major gaps to be filled before a rational system that integrates observational data is available.

Last year, the European Commission issued a service contract for 'Integration of observational databases for evaluation of hemispheric transport of air pollution (**N° 070307/2007/481644/MAR/C5**) that was intended to enhance progress in this area. The Norwegian Institute for Air Research (NILU) and the Commiseriat á l'Energie Atomique (CEA) are contracted now to carry out the work and preliminary results are expected already by September 2008.

This chapter describes the work plan and expected results from this project, which is expected to provide added value for the modeling community in terms of:

(1) exploring of and access to all relevant data in one single portal
(2) availability of measurement data which were previously difficult to access.
(3) data format conversion and metadata harmonization

In the first place, the project will provide access to ground-based observational data and further facilitate the integration of these with other types of observations as well as with model results. This will be achieved through developing a set of generic standards on how to combine data from various original data repositories. Secondly, the project will compile and provide access to a core set of data in response to the needs expressed by the TFHTAP modeling experiments. Finally, the experiences made will be used as a basis for recommending further actions

The project is divided into three major areas of work, and these are to provide:

- A set of generic data standards
- A core data set of integrated observations
- Recommendations for further work

The efforts to be made will build on the operational activities handling data storage and dissemination for the technical work undertaken under the CLRTAP and the EMEP-program. Further, it will prepare for a direct link to similar programs in other regions and as such contribute to the integration on the larger scale (i.e. outside the EMEP area). This will be enabled through developing the EMEP core dataflow infrastructure with a direct interface to other

original regional data repositories. A major benefit of such an approach is that it secures the necessary links with the networks contributing to TFHTAP. Further, it will secure a long-term commitment and form the basis for a succeeding operational system. Figure 2.1 presents an overview of elements to be included, while a detailed description is given in section 2.2.

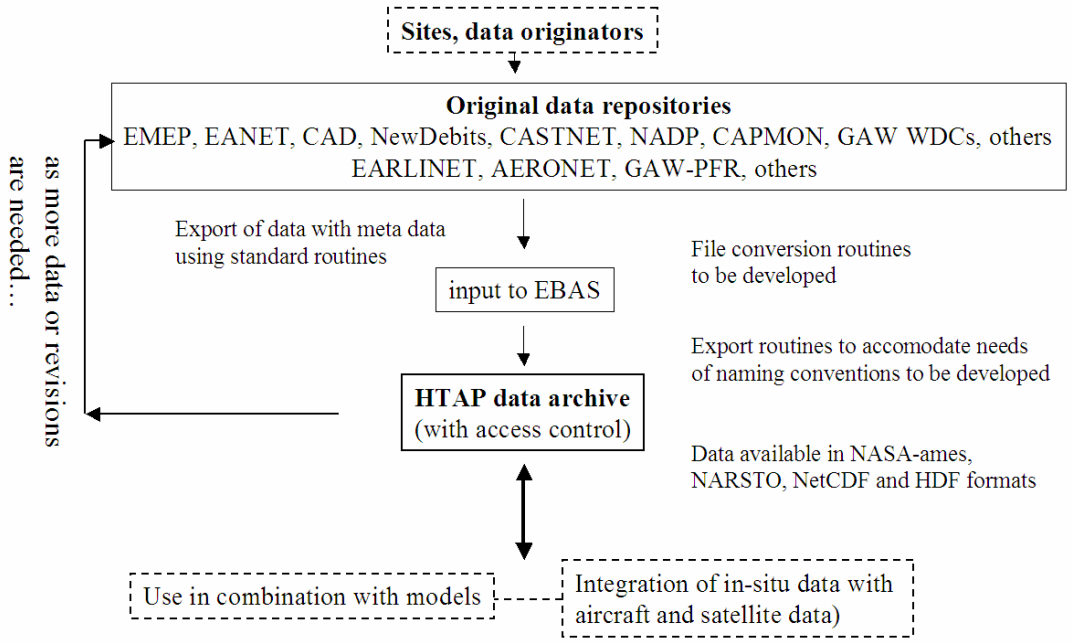


Figure 2.1: *Schematic diagram showing the planned structure for the achievement of enabling a core data set in support of HTAP. Actual efforts to be spent are limited to the central part while the boxes with dashed lines indicate important links to be developed through the interaction with contributing data providers, with other types of observations and with the user community.*

It should be noted that the proposed activities under this work are mainly targeting observations available from various original data repositories rather than trying to compile data from independent research activities (which would not be achievable within the resources available). The focus is on surface observations (including in-situ observations and remote sensing from the ground), and does not include compilation of observational data from aircrafts, satellites or similar. However, an interface to data centers hosting such observational data will be established.

2.1 Planned work and expected results

2.1.1 WP1 Generic set of data standards

This work aims to define a set of standard data formats that are flexible enough to fulfill (a) the needs of the data providers and (b) the requirements of the data users. Data users may, in this context, be those using data for model or satellite validation and those using the observations directly for assessments. In order to identify a common set of data standards, it is essential to first establish an overview of the currently existing standards and recent advances in information technology that may be used (or are being used) to improve data archiving and dissemination. Building on existing standards and practices, a set of standards will be proposed, including routines for QA/QC of the observational data. This will be done through consultation with potential data users and will, after consultation with the Commission, be finalized into a common set of standards for data exchange. WP1 is therefore divided into two sub-tasks, (1) a survey of existing standards and technologies and (2) definition of a common set of data standards. The

latter includes definition of interfaces and development of technical specifications for a system to handle data with such standards.

WP1.1. Review of data and metadata standards used for observational data

Data from observational networks are typically stored in data repositories owned by the network (or a participating partner) and standards for data sharing and archiving have commonly been defined within the network. Trying to create an integrated data set based on such sources, it needs to be accepted that:

- Each network typically have their own data format definition
- Metadata content and definitions are not standardized
- Internal standards for data format and metadata definitions are not always followed
- There are no common routines for data sharing
- QA/QC routines and documentation of these are different between each network
- Data quality is sometimes only to a limited degree known

The current task is therefore focusing on getting an overview of all commonly used data format, metadata standards and quality specifications. This will include standards used by the main observing networks under WMO (including GAW), NASA, EMEP, NARSTO, EANET, GMD (former CMDL under NOAA), ESA (especially Envisat Cal/Val activity), EIONET (i.e. AIRBASE), etc. Standards used for model data exchange and validation such as those developed under the EU projects TRADEOFF, QUANTIFY and RETRO will be assessed together with the upcoming CF naming convention currently being developed for model output on the NetCDF format.

Information on the above-mentioned formats and others, which are mentioned in the frameworks of GEOSS, GMES and INSPIRE as well as OGS (Open Geospatial Standards) will be assembled. A publicly available report will be prepared which summarizes the format and meta-data standards in order to define overlaps and gaps. Compatibility will be discussed in order to establish the properties of the tools to be developed to convert from one format to another.

WP1.2. Definition of data format and metadata standards

Based on the survey of existing standards and on requirements of such standards from the modeling community, a draft set of standards for data sharing will be proposed The set of standards will include rules and guidelines for (a) data file formats (e.g. ASCII, NetCDF, HDF, etc.), (b) metadata definitions (including comprehensive lists of what metadata information needs to be provided for each standard), (c) QA/QC routines (providing a set of procedures for use of flags to better harmonize how data quality is described) and (d) data exchange procedures (including both technical description of how data transfer could occur, and a draft data policy describing how issues such as IPR and data dissemination is handled).

The different networks that are planned to contribute to the TFHTAP observational database differ substantially in terms of meta-data documentation. As the task of proposing a standard for sharing of observational data in principle can be very extensive, a pragmatic approach will be taken as a first step, on which further recommendations can be built. The meta-data ambitions will need to be defined in close consultation with the modelling community.

A first meeting with the contributing networks was arranged in January 2008 (with financial support by the ACCENT network of Excellence). The outline and concept of the service contract was presented and further discussed with the regional programmes. It was well received, and

contact persons to serve as a node of interaction was nominated. ACCENT will further organise a follow-up activity to allow a close interaction throughout the completion of the service contract (ACCENT will fund direct costs for network representatives to attend one or two additional workshops). Dates and venue are still to be defined.

2.1.2 WP2 Core dataset of integrated observations

While WP1 focuses on defining data standards and developing specifications for a data centre able to handle an integrated data set, WP2 is focusing on building an actual database and to populate it with observations. Through collaboration with potential users of the measurements (mainly the modeling community of TFHTAP) the needs of observations for model evaluation will be defined. In addition to initial consultations with this group, users will later on be involved in testing the actual data provided by the proposed service. This will be done with the aim of making further improvements of the data set and the database. Based on user consultations and on technical specification provided through WP1.2, the TFHTAP database for observational data will be developed to host the identified core data sets. The data centre will thus be able to import data from major observational archives defined in WP1.1 and will be able to export the data in a set of standards defined in WP1.2. Finally, the database will be populated with the selected data and made available to the TFHTAP observational data user community. WP2 is thus divided into three sub-tasks for (1) definition of requirements and interaction with users, (2) development of the TFHTAP database and (3) population of the database with selected observations.

The first version of the data base will be launched in September 2008.

WP2.1. User needs and selection of observations, availability of preliminary datasets

Defining user needs and maintaining close collaboration, especially with the modeling community, is essential for the proposed project and WP2.1 will provide the main instruments of the project for handling such interaction.

Selection of sites
Due to the fact that meta-data information (i.e. Data quality flags, site representativeness etc) is poorly described (or at least not easily available in a quantifiable way) in many databases, the selection of sites to be included in the core dataset will also need to be based on direct dialog with the original repositories. The selection of representative sites providing the requested data quality will be done in close interaction with both, the modeller community and the repository operators.

Selection of parameters
At least the following types of data are to be included; inorganic gases and particulates in air, aerosol mass concentration, elemental carbon and organic carbon in aerosols, precipitation chemistry (inorganic compounds), surface ozone, NO_x, VOCs, CO, Hg. Special focus should be on datasets which are otherwise not easily accessible (e.g. the Asian monitoring data) The final list of parameters will be decided in dialogue with the participants to the HTAP modeling activities.

As the efforts required with extending the dataset is fairly low (after routines for importing data from selected networks into the TFHTAP database to be developed through WP2.2 are operational), one can later easily extend the data set. Data covering the years 2000-2006 will be made available in the core dataset. Data will be made available at the time resolution of the individual observations, i.e. no data aggregation will be conducted. The final database will however have functionality to provide data at various level of aggregation (monthly seasonal or annual means as well as frequency distributions).

Availability of preliminary datasets
In order to facilitate benchmarking efforts, preliminary datasets will be available by September 2008. Parallel to the process of selecting sites and parameters, individual datasets will be identified which are qualified and valuable for preliminary deployment. The selection of this datasets will be subject to usefulness for benchmarking (knowledge provided by modeler community) and data availability.
The data will be provided through NILUs NADIR database (the files will be available as provided by the data source, without any conversion or metadata harmonization)

Evaluation of database
CEA will provide a publicly available report, which describes the selection of data and the scientific motivation for this selection. The initial core data set will be used to test its usefulness for the purposes of the TFHTAP model inter-comparison. Since the LSCE is external to NILU, such testing will provide valuable feedback to the developers at NILU, who will construct and populate the database. Initial tests comprise data retrieval, data quality checks, accessibility and completeness with respect to the evaluation of the THTAP model simulations. A report will be elaborated by consulting the TFHTAP modeling community. The report will describe the design, the possible use and the structure of the core datasets.

WP2.2 Development of an HTAP observational data centre

As described above, the main concept of the TFHTAP data centre is to base it on the currently operational EMEP database system, EBAS, and to develop import and export routines to enable collection of data from external networks and export of data according to standards to be defined in WP1.2.

The various original data repositories (of e.g. regional networks) have to a varying extent developed export routines to share data with external users. It is not seen as realistic that these repositories, on a voluntary basis, develop routines to export data into a common format. Instead the aim is to build on their standard routines of export and accept whatever format these networks may have for their various parameters. Assuming that these are routines that will be sustained in the years to come, conversion routines will be developed to convert the data into a format which can be fed into EBAS (EBAS is the database platform which serve a number of databases on atmospheric composition, incl. EMEP, OSPAR, HELCOM, AMAP, EUSAAR and GEOMON, and a natural choice for also hosting the HTAP database). As a first step we will bilaterally, through consultations with the contributing networks, investigate how data best can be provided. As the different networks have different standards with respect to naming of parameters, units etc., major challenges remain before data can be easily combined – making this interpretation between different standards is the main task of creating import routines for the TFHTAP database.

Similar to the routines for importing data in various formats, routines to export data according to the standards defined in WP1.2 will be developed. This will be based on the experience from the CREATE, and RETRO-project where similar functionality have been created to export data into the NARSTO and Envisat Cal/Val HDF format directly from EBAS.

Through the EC-projects EUSAAR, EUCAARI and GEOMON, an extended functionality to access and combine data in EBAS is being developed and will be available also for HTAP towards the end of this service contract. This will further facilitate the access to data for the user community, by allowing querying data on a “one-stop shop”. As such, the major resources for these tasks are already in place and can thus be omitted in the resource planning. Still, the

evaluation and guidance by various users is a key to success, and a close interaction is thus foreseen. The system for identifying data will be based on a web-interface and as data are all compiled in the same database, it will offer a similar functionality to the Giovanni tool developed by NASA. In short, a user can with ease get information about which data are available for a given region or time period, and further easily generate time series plots, get associated air mass trajectories as well as download data in batch mode for post processing locally. The system will be able to differentiate between users and provide access to the various data sets in EBAS based on individual user privileges. As the data to be provided by the participating organizations of TFHTAP will be subject to various sorts of restrictions, the access control will be enabled.

In addition to searching for data stored locally in the TFHTAP database, it will be possible to locate remotely archived data through the same web portal. This will be accomplished by transfer of meta-datasets from major hosts of observations such as space agencies and networks performing atmospheric monitoring from aircraft (IAGOS and CARIBIC). The effort will build on the developments achieved through the ongoing GEOMON project where similar activities will be taking place.

WP2.3. Population of the HTAP observational data centre

The main part of the work to be accomplished include (a) regular contact with contributing data centers and their network, (b) collection and archiving of all incoming data in their original format and (c) running the database import tools developed under WP2.2. The latter will have to be done in parallel with the development of the import tools under WP2.2 and the software will likely need to be adapted not only to the standard formats, but will also need to take into account missing metadata information and erroneous use of standards. Such tailor fitting of the tools can only be achieved once actual inserting in the database is attempted.

The success of this project will ultimately rely on a close link with the contributing observational networks as well as with the user community (i.e. the HTAP modeling team). Only data actually being made available by contributing data repositories are to be included. Also the modeling teams should provide its support in-kind to the proposed activities in order to maximize the benefits of setting up the TFHTAP database.

2.1.3 WP3. Recommendations for future work

The review will be conducted as part of two main lines of activity, one being the bilateral discussions with contributing networks, the second from the feed-back received from the user community. Throughout the project, a log will be kept to facilitate the summary of recommendation to be presented towards the end of the project, which in turn will guide the future operational work related to the handling of observational data in EMEP.

3. Emission data in global scale

Leonor Tarrasón, Oleg Travnikov,
Jan Eiof Jonson and Agnes Nyiri

Emission data in global scale is presently available from many different sources, covering different pollutants and different years. The Global Emission Inventory Activity (GEIA) under IGAC and the Emission Database for Global Atmospheric Research (EDGAR) initiatives have both worked for many years to compile existing emission data information on global scale. Still, the reliability of the emission inventories varies widely from region to region and from pollutant to pollutant, in particular when the emissions are split into sectors.

This chapter briefly reports on the progress with the HTAP initiative to compile consistent and complete emission inventories in the Northern Hemisphere. It summarises the experience from MSC-W when evaluating available global emissions inventories of main pollutants and particulate matter (PM). Further, it presents available global emission inventories for lead, cadmium and mercury. It also provides an overview of existing information on global emissions of selected organic pollutants (POPs), as compiled by MSC-E.

This is only a first step of an iterative process to improve global emission estimates that will rely heavily on the co-operation with individual countries and research groups. Improving emission estimates is a long-term process that requires communication of methods for an adequate quantification of relevant emission sources and continuous evaluation of the emission estimates through model comparisons with validated and available observations.

3.1 The HTAP emission inventory on hemispheric scale

The Directorate General Joint Research Centre, Institute for Environment and Sustainability (DG-JRC-IES) is supporting the work of CLRTAP Task Force on Hemispheric Transport of Air Pollutants (HTAP) by providing improved emissions data relevant for the assessment of intercontinental transport of air pollution. The work involves combining up to date national emission inventories that are of known quality (e.g. North American, European and Asian inventories) with calculations from the Emission Database for Global Atmospheric Research (EDGAR) to provide the modeling efforts of HTAP with consistent gridded emission inventories of air pollutants.

This work is carried out under a special project aiming at improving in a transparent way the emission database EDGAR to be complete, up-to-date and relevant for assessment of intercontinental transport of air pollution. The work under this project started in October 2007. It involves three main tasks:

(1) Update and complement a Northern Hemispheric emission inventory
(2) assessment of uncertainty in EDGAR-HTAP inventory through comparison with observational data
(3) Outreach to countries with limited capacity/info to build emission inventories.

The emission inventory would in first instance contain the components SO_2, NO_x, NMVOC, NH_3, CH_4, BC, OC, PM and CO. In a later phase the inclusion of Hg and POPs will be considered. The inventory will be provided with a spatial resolution of 1ºx1º degree. It will provide information per different sectors. The project follows the recommendations from the HTAP emissions workshop in Beijing and the HTAP interim assessment report (UNECE, 2007). That is, it aims to provide up to date emission inventories (year 2000 to 2005) and to improve the quality of emission methodology and inventories for sources that are poorly known such as: biomass burning (including agricultural waste, biomass for heating and cooking and forest fires), small and medium scale industry/energy, transport and domestic use of coal.

The project is currently in phase 1 where the activities have focused on (i) analysis of availability of emission inventory data for the period 2000-2005 from National Emission Inventories and other inventories, and preparation of a data system to present and analyse the EDGAR-HTAP emission inventory and allocation of national emissions by sector to grid. Currently data from National Communications and EMEP emission inventory have been included in the dataset and are under review.

3.2 Emissions of SOx, NOx, NH3, VOC, CO and particulate matter

The emissions presently used in the global version of the EMEP Unified model are aggregated from several sources. This is far from an optimal set of emissions. Because of the different origins of the data for the different pollutants, the consistency of the emission data is not guaranteed. Another difficulty, in addition to the lack of consistency is that the emission data set is not complete. There are recognized missing sources in the inventory: volcanic emissions of sulphur are included only within the EMEP domain and the same happens with shipping emissions of sulphur dioxide and with natural DMS emissions from the oceans.

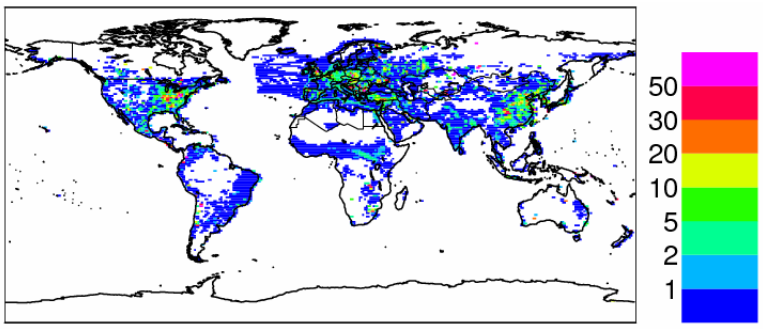
Table 3.1. *Sources of emissions aggregated for the use in the OsloCTM2 model. Muller92 is Müller et al. (1992), Muller02 is Müller et al. (2002), Edgar is the EDGAR v3.2 http://mnp.nl/edgar/ for 1997 (Olivier and Berddowski, 2001), GEIA emissions available from http://www.geiacenter.org/ based on Yienger and Levy (1995), AM is Andreae and Merlet (2001), ATSR is data from ATSR satellite (1997), V&A is Endresen et al. (2003) and Bouwman is Bouwman et al. (1997).*

	SOx	**CO**	**NOx**	**NMVOC**	**NH3**
Soil	GEIA		GEIA		
Ocean		Muller92			
Biogenic		Muller02			
Biomass burning	Edgar/AM/ATSR	Edgar/AM/ATSR	Edgar/AM/ATSR	Edgar/AM/ATSR	
Bio fuels	Edgar	Edgar	Edgar	Edgar	
Fossil fuels	Edgar	Edgar	Edgar	Edgar	
Industry	Edgar	Edgar	Edgar	Edgar	
Agriculture					Bouwman
Ships		V&A	V&A	V&A	

The global emission estimates are considered to be valid for year 2000 although data from different years close to 2000 have been used in the compilation. Over the EMEP domain, EMEP emissions are used. For the most other areas in the model domain emissions are adopted from the

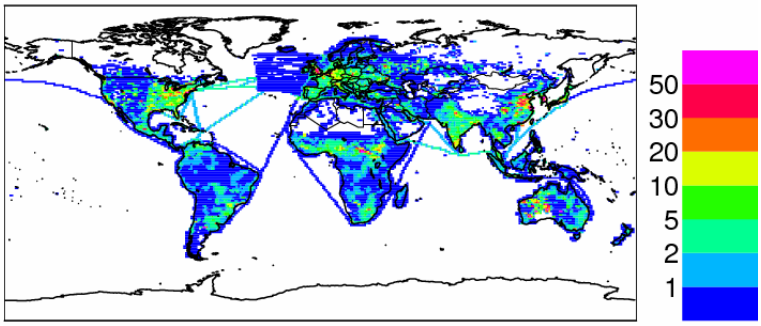
OsloCTM2 model, and originate from different research projects (RETRO, QUANTIFY, AEROCOM…). The main references for these emissions are summarized in Table 3.1. Over Asia, the OsloCTM2 emissions have been replaced by a bottom up inventory developed for East Asia for the year 2000 (Streets et al.2003) available through ACESS (Ace-Asia and Trace-P Modeling and Emission SupportSystem) at http://www.cgrer.uiowa.edu/ACESS/acess index.htm.

For biogenic VOC emissions, the OsloCTM2 model uses prescribed values. However, in the global Unified EMEP model, temperature dependent biogenic VOC emissions from forests are used instead, following the methodology of the EMEP regional model.



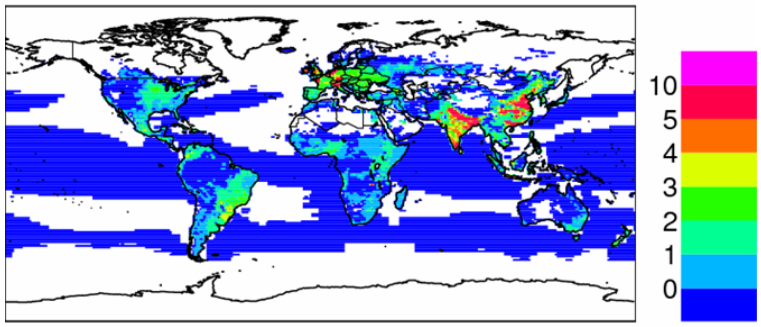
Emission of SO2 in g m-2

a)



Emission of NO2 in g m-2

b)



Emission of NH3 in g m-2

c)

Figure 3.1*. Spatial distribution of a) sulphur dioxide emissions in g(SO_2) m^{-2} yr^{-1} b) nitrogen oxides emissions in g(NO_2) m^{-2} yr^{-1} and c) ammonia emissions in g(NH_3) m^{-2} yr^{-1} ,as used in the global Unified model*

The data are provided in a spatial resolution of 1°x1°. Examples of the spatial distribution of the emissions used in the Unified EMEP global model are given in Figure 3.1, where some of the

problems of inconsistencies and lack of completeness of the data are clearly visible, as for instance with shipping routes.

The EMEP model makes use of yearly gridded emissions for different emission sectors. For each country and sector, the model uses time factors that determine the daily, weekly and monthly variations. For the global model applications, we use OsloCTM2 monthly emissions coupled with daily variations outside the EMEP area. Inside the EMEP area, the EMEP temporal emission variations are used.

A first evaluation of these emissions against satellite measurements and ground level observation was presented in Jonson et al. (2007).The evaluation identified a series of open questions on the input emission and measurement data, especially over Asia. It is envisaged to enhance this type of evaluation in the near future, co-operating closely with the HTPA initiatives on measurement and emission data compilation.

3.3 Heavy metals and mercury (Hg)

Heavy metals differ significantly in their properties and ability to long-range atmospheric transport. In particular, mercury is well known as a global scale pollutant with distinct ability to intercontinental transport. On the other hand, such particle bound heavy metals as lead and cadmium are mostly transported regionally. Therefore, mercury has the first priority for modeling on a global scale and requires global emissions data. Nevertheless, in many cases modeling other heavy metals also requires emissions data covering territories wider than the region of the primary interest. For example, lead and cadmium pollution levels in the EMEP countries and, in particular, in the EECCA countries can be affected by emission sources located in neighboring non-EMEP countries of Northern Africa, Meddle East, Eastern and Southern Asia.

Although, there are detailed heavy metal anthropogenic emissions inventories developed for a number of regions (e.g. Europe, North America), global emission datasets are quite rare. Table 3.2 contains characteristics of available global emission datasets and inventories for some heavy metals. Global spatially resolved data on mercury emission are available for a number of years (1990, 1995, and 2000) and new dataset for 2005 is under development at the moment. The only gridded emission dataset for lead relates to 1990. The lead emissions inventory has been updated for mid-nineties but these data have no spatial distribution. Only aggregated emission estimates without spatial distribution of emission sources are available for cadmium and relate to mid-nineties.

Table 3.2. *Available data on heavy metals anthropogenic emissions on a global scale*

Chemical	Period of time	Global emissions, t/y	Spatial resolution	Dataset location	Reference
Mercury	1990	2 144	1º×1º	CGEIC [(a)]	CGEIC website
	1995	2 317	0.5º×0.5º	AMAP [(b)]	Pacyna et al., 2003
	2000	2 190	0.5º×0.5º	AMAP [(b)]	Pacyna et al., 2006
	2005	in progress in AMAP			
Lead	1989	167 889 – 206 435	1º×1º	CGEIC [(a)]	Pacyna et al., 1995
	1995	119 259	n/a	n/a	Pacyna and Pacyna, 2001
Cadmium	1995	2 983	n/a	n/a	Pacyna and Pacyna, 2001

[(a)] http://www.ortech.ca/cgeic/index.html
[(b)] http://amap.no/Resources/HgEmissions/

The existing global scale datasets of heavy metal anthropogenic emissions are briefly described below.

3.3.1 Mercury

First global dataset of mercury anthropogenic emissions relates to 1990 and is available at website of the Canadian Global Emissions Interpretation Centre (CGEIC). It covers the globe with spatial resolution 1º×1º and distinguishes emissions from point sources (power plants, smelters and other large sources with stack height and a sufficiently buoyant plume that the emissions are transported into the atmosphere above 100 m) and area sources (e.g. residential and commercial boilers, mobile sources etc.)

Later two newer global mercury emissions datasets were released for 1995 and 2000. They are available at the website of the Arctic Monitoring and Assessment Programme (AMAP) and contain considerably more detailed information on mercury emissions. The datasets have spatial resolution 0.5º×0.5º and include the following parameters:

- Gridded anthropogenic emissions separately from point and area sources;
- Three emission height levels (below 50 m, 50-150 m, above 150 m);
- Speciation of mercury emissions into elemental gaseous mercury (GEM), reactive gaseous mercury (RGM) and particulate mercury (TPM).

Figure 3.2 shows spatial distribution of mercury anthropogenic emissions in 2000. Areas with elevated mercury emissions correspond to highly industrialized regions: Europe, Eastern United States, Southeastern China, Japan, India and South Africa.

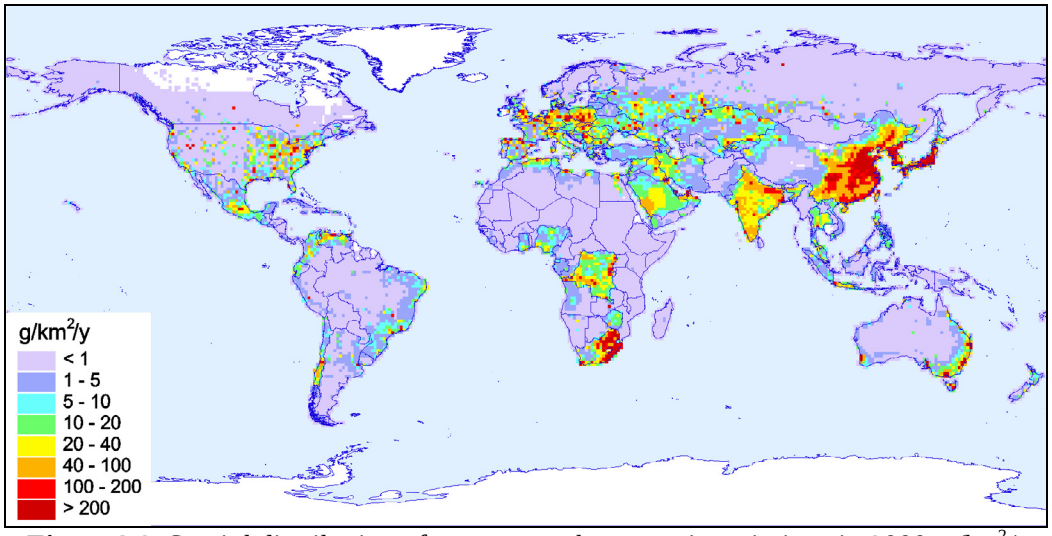


Figure 3.2. *Spatial distribution of mercury anthropogenic emissions in 2000, g/km^2/y*

Comparison of total anthropogenic mercury emissions in different continents and for different reference years is presented in Figure 3.3. As seen there, total mercury emission in Asia in 1990 was comparable with total emission in Europe and 2-3 times higher than that in North America. During the period 1990-2000 decreased in Europe and North America but significantly increased in Asia and Africa. Some increase also took place in Australia and South America. Therefore, in 2000 Asian emission sources contributed more than 50% to global mercury emissions whereas contribution of European and North American sources did not exceeded 20% in sum.

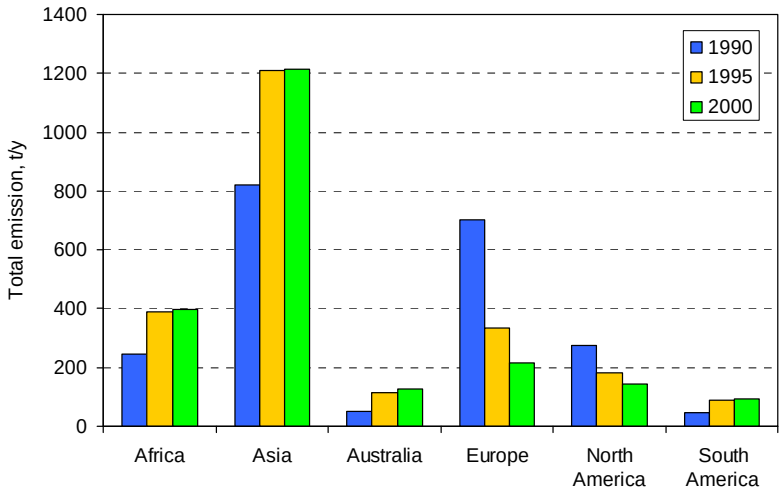


Figure 3.3. *Change of global anthropogenic emissions of total mercury to the atmosphere from 1990-2000, t/y*

Currently an update of the global mercury emissions inventory is being developed for 2005 within the framework of a joint AMAP and UNEP project. Release of the new dataset is expected by the end of 2008.

3.3.2 Lead

The only global spatially resolved emissions inventory for lead relates to the year 1989 and is available at the CGEIC website. This dataset includes gridded information on lead emissions from point and area sources with spatial resolution 1°×1°. Besides, it provides low and high estimates of the emission rates based on emission factors that were assumed to be the highest and the lowest values within the emission factor ranges for a given source category. Figure 3.4 shows global distribution of anthropogenic emissions of lead in 1989. According to this inventory the most significant emission sources were located in Europe, Southeastern Asia, Mexico and South Africa. Global emission estimates varied from 168 to 206 kt/y.

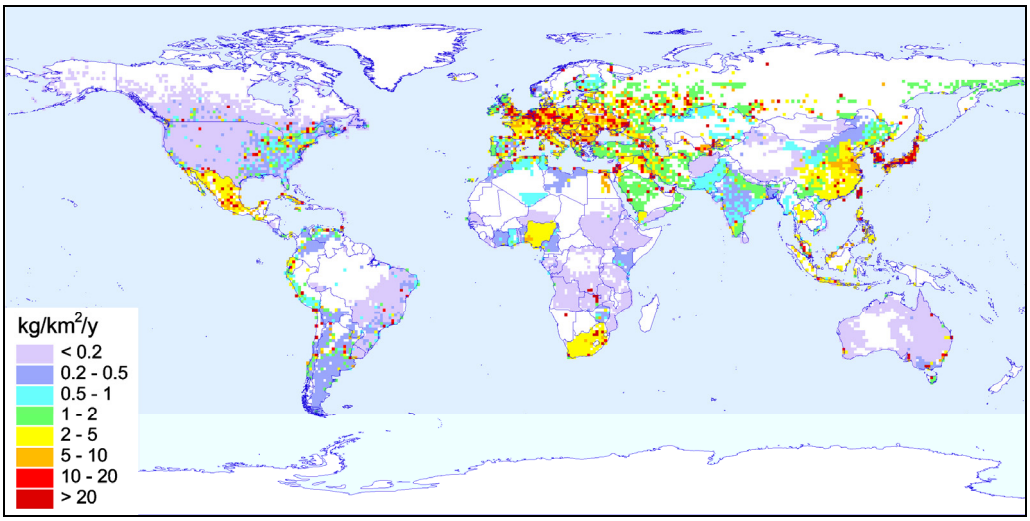


Figure 3.4. *Spatial distribution of lead anthropogenic emissions in 1990, kg/km^2/y*

The global emissions inventory for lead has been updated by *Pacyna and Pacyna* (2001) for mid-nineties. Estimates of lead emissions from anthropogenic sources have been revised based on new information from national experts in combination with statistical data on the consumption of raw material, industrial goods production. However, the updated inventory did

not include spatial distribution of emissions. According to the inventory combustion of gasoline continued to be the major source of atmospheric emissions of lead, contributing about 74% of global emission in 1995.

Comparison of the two mentioned above global inventories for lead is presented in Figure 3.5 for different continents. The largest emissions are estimated in Asia and Europe. Estimates of lead emissions for 1995 are significantly lower those for 1989, particularly for Europe. However, it should be noted that direct comparison of these two datasets should be made with certain caution since the differences can be caused not only by emissions reduction but also by distinctions in methodologies of data processing.

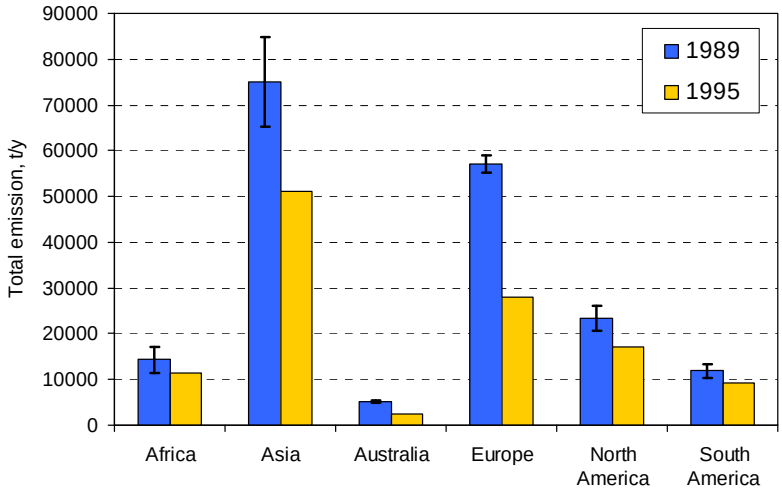


Figure 3.5. *Comparison of global anthropogenic emission inventories of lead to the atmosphere for 1989 and 1995, t/y*

3.3.3 Cadmium and other heavy metals

There is no spatially resolved global emission datasets available so far for cadmium and other heavy metals and metalloids (arsenic, chromium, copper, nickel, selenium and zinc). The most recent inventory for these metals includes total emission estimates for different continents and various source categories and relates to 1995 (*Pacyna and Pacyna,* 2001). The inventory contains estimates of heavy metal emissions from heat and power production, non-ferrous metal production, iron and steel industry, cement production and waste disposal. Emission of heavy metals from major anthropogenic source categories is presented in Table 3.2. The estimates indicated that stationary fossil fuel combustion was the major source of chromium (69%), nickel (90%), and selenium (89%). Whereas non-ferrous metal production was the largest source of atmospheric emissions of cadmium (73%), arsenic (69%), copper (70%), and zinc (72%).

Table 3.3. *Global emissions of heavy metals from major anthropogenic source categories in 1995 (in t/y; adapted from Pacyna and Pacyna, 2001)*

Source category	**Cd**	**As**	**Cr**	**Cu**	**Ni**	**Se**	**Zn**
Stationary fossil fuel combustion	691	809	10 145	7 081	86 110	4 101	9 417
Non-ferrous metal production	2 171	3 457	–	18 071	8 878	466	40 872
Iron and steel production	64	353	2 825	142	36	7	2 118
Cement production	17	268	1 335	–	134	3	2 670
Waste disposal	40	1 124	425	621	129	24	1 933
Total	**2 983**	**5 011**	**14 730**	**25 915**	**95 287**	**4601**	**57 010**

Figure 3.6 shows contribution of different continents to global anthropogenic emissions of heavy metals (*Pacyna and Pacyna,* 2001). As seen Asia was the largest contributor to global heavy

metal emissions made up 40-60% of total value. Contributions of other continents vary for different metals. Europe contributed 23% of global emissions of chromium and 21% of nickel; North America made up 24% of selenium and 22% of chromium emissions; South America was responsible for 21% of global copper emissions. The lowest contributions were made by Australia.

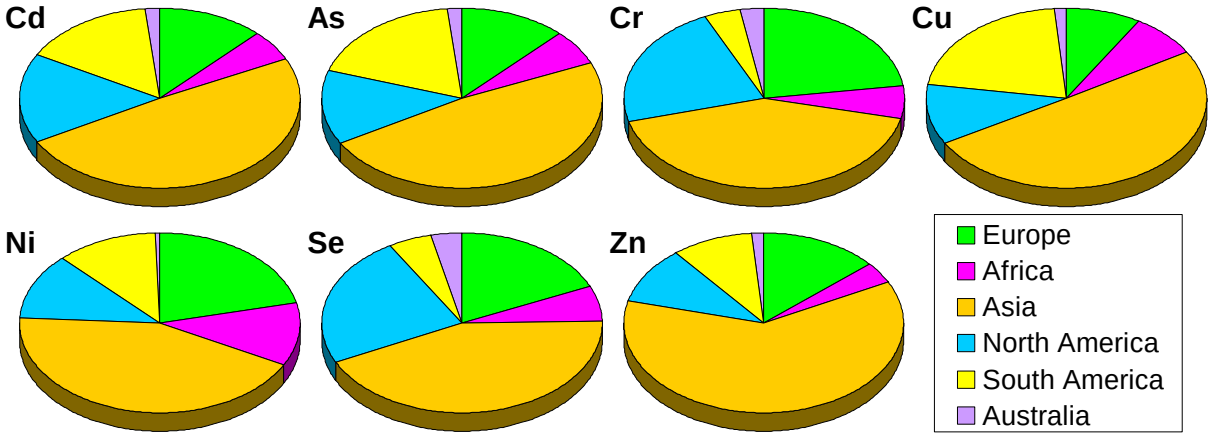


Figure 3.6. *Contribution of different continents to global anthropogenic emissions of heavy metals in 1995*

3.4 Available emissions of Persistent Organic Pollutants

A number of inventories of the global production, usage and emission for selected POPs were developed recently. The information on available inventories including the estimates of POP usage, production and emissions is presented in Table 3.4. Some of the inventories contain the data on total emissions for particular countries, their temporal variations, and gridded emission data. In particular, PCB emission inventory presents a full set of information on spatial variations of emissions on global scale and their temporal variations with annual resolution in period 1930-2100. Gridded emissions data are also available for α-HCH and β-HCH for 1980, 1990, and 2000. For other POPs the estimates of total emissions of particular countries are available. Nevertheless this information can also be used for modeling purposes. Below a description of available POP emission inventories is presented.

3.4.1 PCBs

A global historical emission inventory for selected PCB congeners (22 PCB congeners) was prepared by *Breivik et al.* (2002a). The available data of the historical production of PCBs and the chemical composition of various technical mixtures have been compiled from the literature. Information of imports, exports and consumption, as well as restrictions on production and import has further been analyzed for individual countries. The estimates account for a reported historical global usage of approximately 1300 kt of PCBs. The results suggest that almost 97% of the global historical use of PCBs have occurred in the Northern Hemisphere (*Breivik et al.*, 2002a).

In the second part of this study, an attempt has been made to estimate the historical anthropogenic emissions as a direct result of the widespread usage of PCBs for a period of 70 years. The historical emission of ΣPCB_{22} is estimated to be between 440 and 91722 t, with 7709 as the mid value (*Breivik et al.*, 2002b). In accordance with this three emission scenario were developed, in particular, low emission, average emission, and high emission scenario. In spite of

the major uncertainties, the study is considered as a first important step towards the establishment of a global PCB emission inventory with congener resolution. The importance of temperature as a key parameter in controlling and affecting both the absolute value of PCB emissions as well as the PCB emission profile is emphasized.

Table 3.4. *Global inventories of POP emissions*

Chemical	Use	Period of time	Global use/ production, kt	Global emissions, kt	Reference
PCDD/Fs (as kg I-TEQ)	By-products	1990		50	*Brzuzy and Hites*, 1996
		1995		10.5 8.3 - 36	*UNEP*, 1999
HCB	Pesticide By-product	Mid 1990		0.023 0.012 – 0.092	*Bailey*, 2001
Technical HCH	Insecticide	1948-1997	10000		*Li*, 1999
Lindane	Insecticide	1948-1993	720		*Voldner and Li*, 1995
			6000		*Sang et al.*, 1999
		1950-2000	600		*Li and Vijgen*, 2005
α-HCH	Insecticide	1980	290	184	*Li et al*., 2000
		1990	59	44	
β-HCH	Insecticide	1945-2000	850	230	*Li et al*., 2003
		1980	36	9.8	
		1990	7.4	2.4	
		2000	-	0.066	
PCBs	Various	1930-2000 and up to 2100	1320	7.71 0.44 – 91.7	*Breivik et al.*, 2002a *Breivik et al.*, 2002b *Breivik et al.*, 2007
PAHs	By-products	2004		520	*Zhang and Tao*, 2008

The third part of this study presents an update of preceding PCB emission database. This work takes into account updated information on PCB production, as well as new data on the chemical composition of various technical mixtures for which less information had been available earlier. The methodology to estimate temporal trends of PCB usage is improved. The authors have also included projected emissions up to year 2100, which can facilitate predictions of future environmental exposure. The national emission data for each of the 114 countries considered is spatially resolved on a 1º×1º grid for each congener and year, using population density as a surrogate (*Breivik et al.*, 2007), as indicated for PCB-153 for 2000 in Figure 3.7.

3.4.2 HCHs

HCH is an organochlorine insecticide employed throughout the world. It has also been used for seed protection, poultry and livestock treatment, household vector control, lumber protection, and even for rodent baits. The HCH has two main compositions: technical HCH and lindane. Lindane consists almost entirely of γ-HCH, the insecticidal form, technical HCH contains a total of eight HCH isomers, among which only the α, β, γ, δ and ε isomers are stable and commonly identified. Around 10,000 kt of technical HCH has been released to the environment between 1948 and 1997 (*Li*, 1999). Global gridded emission data for HCH isomers are available for α-HCH and β-HCH.

<u>α-HCH</u>

Estimations of global gridded α-HCH emissions in 1980 and 1990 resulted from the application of technical HCH were presented *Li et al*. (2000). The total global usage of α-HCH was estimated to be around 290 kt in 1980 and 59 kt in 1990. Total global α-HCH emissions in 1980

were 184 kt and 44 kt in 1990. In 1980 approximately 74% of the global α-HCH emissions originated in the Northern Hemisphere mid latitudes, followed by around 19% in the Northern Hemisphere tropics, while in 1990 approximately 54% of the global α-HCH emissions originated in the Northern Hemisphere tropics, followed by around 37% in the Northern Hemisphere mid latitudes. Among total global α-HCH emissions, 85% and 75% originated in Asia in 1980 and 1990, respectively (*Li et al.*, 2000). The spatial distribution of annual emission of α-HCH for 2000 is presented in Figure 3.8.

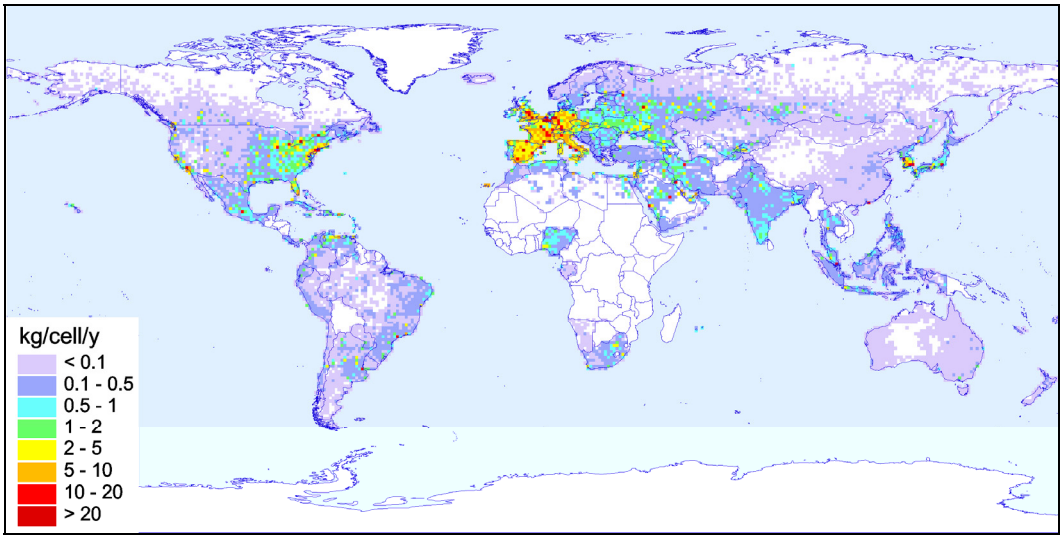


Figure 3.7. *Spatial distribution of annual emission of PCB-153 emission for 2000 (Breivik et al., 2007), kg/cell/y*

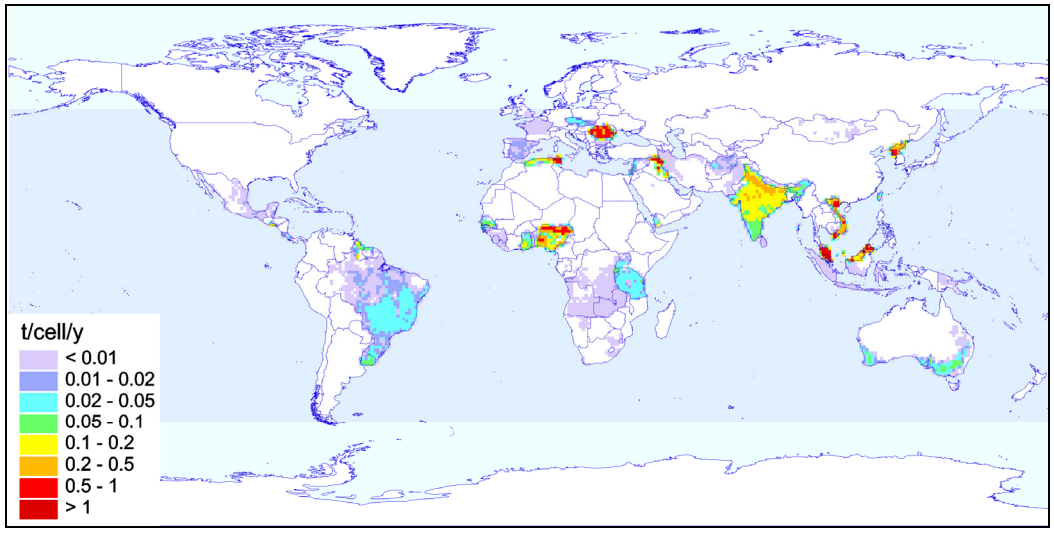


Figure 3.8. *Spatial distribution of annual emission of α-HCH emission for 2000 (Li et al., 2000), t/cell/y*

β-HCH

Global gridded β-HCH emission inventory was prepared by *Li et al.* (2003). The total global usage of β-HCH between 1945 and 2000 is estimated as 850 kt, 230 kt of which was emitted to the atmosphere over the considered period of time. Usage of β-HCH was estimated to be around 36 kt in 1980 and 7.4 kt in 1990. Total β-HCH emissions in 1980 were 9.8 kt and 2.4 kt in 1990. While it is assumed that no usage of technical HCH occurred in 2000, the global β-HCH emissions during this year due to soil residues were estimated at 66 t. It has shown that the global

β-HCH emissions have undergone a "southward tilt" over the time period studied as more northern countries have banned the use of technical HCH (*Li et al*., 2003).

γ-HCH

The majority of the developed countries prohibited the application of HCH in the 1970s. Lindane replaced technical HCH. The amount of global lindane usage has been estimated by a few scientists. *Voldner* and *Li* (1995) gave worldwide lindane usage between 1948 and 1993 as 720 kt , and *Sang et al.* (1999) suggested the global consumption of lindane should be 6,000 kt, more than 8 times as that given by *Voldner and Li* (1995). Global lindane use for agricultural purpose between 1950 and 2000 is estimated to be 450 kt, among which 280 kt was used in Europe, 73 kt in Asian, 64 kt in America, 29 kt in Africa, and 1 kt in Oceania countries (*Li and Vijgen*, 2005). Although lindane use in most European countries has been banned, the historical lindane use from 1950 to 2000 in Europe reached approximately 63% of the total global use. Global annual lindane usage was highest in the 1960s and the beginning of 1970, and has since been declining. Global use in 2000 was approximately 2 kt (2204 tons). The total global lindane usage for all purposes could be approximately 600 kt (*Li and Vijgen*, 2005; *IHPA*, 2006). The last estimate of global lindane use is less than those given by both *Voldner and Li* (1995) and *Sang et al.* (1999).

3.4.3 HCB

HCB had been actively used in industry and for agricultural purposes. HCB was first introduced in 1933 as a fungicide on the seeds of onions, sorghum and crops such as wheat, barley, oats and rye. It is believed that agricultural use of HCB dominated its emissions during the 1950s and 1960s. The peak of HCB production was the late 1970s and early 1980s worldwide. Production of HCB declined as a result of restrictions on its use. The banning of HCB for agricultural use in the 1970s removed the largest single primary source of HCB in the environment (*Barber et al.*, 2005). Although HCB production has ceased in most countries, it is still being generated inadvertently as a by-product and/or impurity in several chemical processes, such as the manufacture of chlorinated solvents, chlorinated aromatics and pesticides (*Jacoff et al.*, 1986). HCB can be also released to the environment by incomplete combustion and from old dumpsites.

Global emission estimate for HCB was presented in (*Bailey*, 2001). This study presents the information from a variety of sources of HCB emissions and gives a quantitative estimate of the global HCB emissions in the mid 1990s. The best estimates of global HCB emissions from different categories of sources are as follows: pesticides application – 6500 kg/y; manufacturing – 9500 kg/y; combustion – 7000 kg/y, including 500 kg from biomass burning. This adds up to total current HCB emissions of approximately 23,000 kg/y with an estimated range 12,000–92,000 kg/y (*Bailey*, 2001). The author suggests that a substantial portion of HCB measured in the atmosphere is thought to come from volatilization of “old” HCB on the soil from past contamination along with unidentified sources. Globally, the magnitude of these secondary emissions is currently not known, but this is a potentially very important aspect of the current global HCB cycle (*Barber et al.*, 2005).

3.4.4 PCDD/Fs

PCDD/Fs enter the environment as by-products of industrial processes. The most significant type of sources of PCDD/F emission to the atmosphere is low-temperature, incomplete incineration of chlorine-containing materials such as plastics. Other major sources include thermal processes, such as motor vehicle fuel combustion in countries where leaded fuel containing chlorine scavengers is still used, and metallurgical industries. PCDD/Fs are also trace contaminants in

chlorophenoxy herbicides, PCB formulations, and chlorophenol wood preservatives (*AMAP*, 2004).

Rough estimates of the global emissions of PCDD/Fs are available from *Brzuzy and Hites* (1996). The authors calculated annual global emission to approximately 50 kg I-TEQ for 1990. They based on assumptions that the global PCDD/F emissions from different categories of sources are as follows: waste incineration – 20 kg I-TEQ/y; cement kilns – 17 kg I-TEQ/y; manufacturing – 7 kg I-TEQ/y, and combustion – 6 kg I-TEQ/y.

In addition, an overview of available emission inventories for PCDD/Fs on global level can be found in (*UNEP*, 1999). This report includes estimates of PCDD/F emissions for Western Europe, Northern America (Canada and the United States). For Asian region there is only one inventory for Japan, which covers a few source categories, and the estimates for emissions from waste incinerators from South Korea. From the southern hemisphere, only Australia has estimated annual emissions. Japan and the United States were considered the largest emission sources of PCDD/Fs (4 kg I-TEQ/y and 2.7 kg I-TEQ/y, respectively). The average estimate of annual "global" PCDD/F emission is amounted to approximately 10.5 kg I-TEQ for 1995. The lower estimate is about 8.3 kg I-TEQ/y and the upper estimate is approximately 36 kg I-TEQ/y. The waste incineration is still the major emitter of dioxins and furans to the atmosphere with contribution to the total emission reaching almost 70 %. The majority of PCDD/PCDF emissions from this sector are due to the municipal (and to a lesser extent to the hazardous) waste incinerators located in Japan. This source sector alone is responsible for almost 34 % of the total PCDD/PCDF inventory to air from identified sources so far (*UNEP*, 1999).

3.4.5 PAHs

Evaluation of global emission of PAHs from about 200 countries to the atmosphere for 2004 was performed in (*Zhang and Tao*, 2008). According to their estimates the total global atmospheric emission of the 16 PAHs listed as the US EPA priority pollutants is accounted for 520 kt/y. Authors supposed that biomass burning including both biofuel and wildfire are dominating emission sources of PAHs with relative contributions of 57% and 17% to the total PAH emission, respectively. At the same time source profiles of individual countries varied remarkably depending on their status of development, population density and vegetation cover. The maximum annual emission of PAHs was estimated for Asia accounting for 290 kt/y (55% of the global PAH emission). The emissions of Africa, Europe, North America, South America and Oceania contributed 19%, 10%, 8%, 6% and 1.5% to the global PAH emission, respectively (*Zhang and Tao*, 2008).

3.5 References

Andreae, M.O. and Merlet, (2001), Emissions of trace gases and aerosols from biomass burning, *Glob. Biogeochem. Cycles*, 15, pp 955--966

AMAP, 2004 AMAP Assessment 2002: Persistent Organic Pollutants in the Arctic. Arctic monitoring and Assessment Programme (AMAP), Oslo, Norway. xvi+310 pp.

Bailey R.E., 2001 Global hexachlorobenzene emissions. *Chemosphere 43*, pp.82-167.

Barber J.L., Sweetman A.J., Wijk D. and K.C.Jones 2005 Hexachlorobenzene in the global environment: Emissions, levels, distribution, trends and processes. *Sci, Total Environ*., 349, pp.1-44.

Berglen, T.F., T.K. Berntsen, I.S.A. Isaksen and J.K. Sundet (2004) A global model of the coupled sulphur/oxidant chemistry in the troposphere: the sulphur cycle*, J. Geophys. Res.*, 109, D19310, doi:10.1029/2003JD003948

Bouwman, A.F., D.S. Lee, W.A. Asman, F.J. Dentener, K.W. Van Der Hoek and J.G.J. Olivier (1997), A global high-resolution emission inventory for ammonia, *Glob. Biogeochem. Cycles,* 11, No. 4, pp 561--587

Breivik K., Sweetman A., Pacyna J.M. and K.C.Jones, 2002a Towards a global historical emission inventory for selected PCB congeners – a mass balance approach. 1. Global production and consumption. *Sci.Total Environ.,* 290, pp.181-198.

Breivik K., Sweetman A., Pacyna J.M. and K.C.Jones, 2002b Towards a global historical emission inventory for selected PCB congeners – a mass balance approach. 1. Emissions. *Sci.Total Environ*., 290, pp.199-224.

Breivik K., Sweetman A., Pacyna J.M. and K.C.Jones, 2007 Towards a global historical emission inventory for selected PCB congeners – a mass balance approach. 3. An update. *Sci.Total Environ*., 377, pp.296-307.

Brzuzy L.P. and R.A. Hites, 1996 Global Mass Balance for Polychlorinated Dibenzo-p-dioxins and Dibenzofurans. *Environ. Sci. Technol.* 30, pp.1791-1804.

Endresen, O., E. Sørgaard, G. Gravir, J.K. Sundet, S. Dalsøren, T.F. Berglen and I.S.A. Isaksen (2003) Emission from international sea transportation and environmental impact, *J. Geophys. Res.*, 108, D17, doi:10.1029/2002JD002898

IHPA, 2006 The legacy of lindane HCH isomer production. Annexes. A global overview of residue management, formulation and disposal by John Vijgen. http://www.ihpa.info/library_access.php.

IPCC, 2006 IPCC guidelines for National Greenhouse Gas Inventories, IPCC National Greenhouse Gas Inventory Programme, Hayama, Japan

Jacoff F.S., Scarberry R. and D.Rosa 1986 Source assessment of hexachlorobenzene from the organic chemical manufacturing industry. In: Morris C.R., Cabral J.R.P, editors. Hexachlorobenzene: Proceedings of an international symposium, vol.77. IARC Sci Publ; pp.7-31.

Jonson, J.E., Tarrasón, L., Wind, P., Gauss, M., Valiyaveetil, S. S., Tsyro, S, Klein, H., Isaksen, I.S.A. and A. Benedictow (2007) *First evaluation of the global EMEP model and comparison with the global OsloCTM2 model*. EMEP-MSC-W Technical report 1/2007, Norwegian Meteorological Institute, Oslo, Norway. Available from www.emep.int

Müller, J.-F., (1992), Geographical distribution and seasonal variation of surface emissions and deposition velocities of atmospheric trace gases, J. Geophys. Res., 97, D4, pp 3787-3804

Li Y.F. and J.Vijgen, 2005 Draft. Global Lindane Usage. Report to ECE.

Li Y.F., 1999 Global Technical Hexachlorocyclohexane usage and its contamination consequences in environment: from 1948 to 1997, Sci. Total. Environ. 232, pp.123-160.

Li Y.F., Scholtz M.T. and B.J.van Heyst, 2000 Global gridded emission inventories of α-hexachlorocyclohexane. Journal of geophysical research, ISSN 0148-0227, v.105, n D5, pp.6621-6632.

Li Y.F., Scholtz M.T. and B.J.van Heyst, 2003 Global gridded emission inventories of β-hexachlorocyclohexane. Environ. Sci. Technol., v.37, pp.3493-3498.

NILU, 2002 Proceedings of AMAP Workshop on Sources, Emissions and Discharges, Kjeller, 23-24 August 2001. NILU OR 3/2002. JP/BMo/O-100079.

Olivier, J.G.J. and J.J.M. Berdowski, (2001), Global emission sources and sinks, in Berdowski, J.J.M., A. Guicherit and P.J. Heij (Eds.)The climate system, pp 33—78, A.A. Balkena Publishers/ Swets & Zeihtlinger Publishers, Lisse, The Netherlands, ISBN 90 5809 255 0

Pacyna J. M., Scholtz M. T. and Y.-F. Li (1995) Global Budgets of Trace Metal Sources, *Environmental Reviews*, 3, 145-159

Pacyna, J.M. and Pacyna, E.G. (2001): An assessment of global and regional emissions of trace metals to the atmosphere from anthropogenic sources worldwide. Environmental Reviews, 9: 269-298

Pacyna J, Pacyna E, Steenhuisen F, Wilson S. (2003) Mapping 1995 global anthropogenic emissions of mercury. *Atmos. Environ*. 37(S1), S109-S117

Pacyna E. G., Pacyna J. M., Steenhuisen F. and Wilson S. (2006) Global anthropogenic mercury emission inventory for 2000. *Atmos. Environ*. 40(22), 4048-4063

Sang S., Petrovic S. and V.Cuddeford, 1999 Lindane – A Review of Toxicity and Environmental Fate, prepared by World Wildlife Fund Canada.

D.G. Streets, T.C. Bond, G. R. Carmichael, S. D. Fernandes, Q. Fu, D. He, Z. Klimont S. M. Nelson, N. Y. Tsai, M. Q.Wang, J.-H.Woo, and K. F. Yarber. An inventory of gaseous and primary aerosol emissions in Asia in the year 2000. *J. Geophys. Res*.,108, 2003. doi:10.1029/2002JD003093.

UNECE, 2007, Hemispheric Transport of Air Pollution, Air pollution studies no. 16, ECE/EB.AIR/94, Geneva, Switzerland.

UNEP, 1999 Dioxin and Furan Inventories. National and Regional Emissions of PCDD/PCDF. Prepared by UNEP Chemicals. Geneva, Switzerland.

Voldner E.C. and Y.F.Li, 1995, Global usage of selected persistent organochlorines, *Science of the Total Environment,* v. 160/161, pp.201-210.

Yienger, J.J., and H. Jevy II (1995), Empirical model of global soil-biogenic NOx emissions, *J. Geophys. Res.,* 100, pp 11447—11464

Zhang Y. and S.Tao, 2008 Global atmospheric emission inventory of Polycyclic Aromatic Hydrocarbons (PAHs) for 2004. *Atmos. Environ.,* in press.

4. Geophysical input data: Land cover and LAI

Alexey Gusev and David Simpson

The geophysical information is important part of model input data and is required in many model parameterizations. Geophysical information include: land use and land cover information (LU/LC), leaf area index (LAI), surface temperature and wetness, and soil properties. Surface characteristics and vegetation cover significantly affect the air-surface interaction of air pollutants and their atmospheric transport. Thus the use of accurate and detailed information on underlying surface properties is of importance for modeling of air pollution transport on global and regional scales.

As part of the joint effort to develop a common EMEP global modeling framework MSC-E and MSC-W have started an evaluation of the geophysical data necessary for modeling on the global scale. As a first step land cover and leaf area index information have been analyzed. It is planned that the common EMEP global modeling framework will be flexible enough to permit the use of several LC and LAI datasets depending on the particular application. In this chapter the progress in collection and preliminary comparison of various land cover and leaf area index data is described.

4.1 Description of available land cover data

The evaluation of geophysical data began with the compilation and comparison of four different sets of land cover data. The selected data sets were all originating from remote sense data and included: UMD, USGS GLCC, GLC2000, and MODIS Land Cover products. A brief description of these datasets is presented below.

4.1.1 UMD land cover dataset

This global land cover dataset was prepared by the University of Maryland Department of Geography in 1998 (Hansen et al., 2000). Images acquired from the Advanced Very High Resolution Radiometer (AVHRR) satellites between 1981 and 1994 were analyzed to distinguish fourteen land cover classes. This product is available at three spatial scales: 1 degree, 8 km and 1 km pixel resolutions. The data on land cover classifications are represented as dominant types and fractional cover.

The objective was to create a land cover map derived from 1 km AVHRR data using all available bands and derived Normalized Difference Vegetation Index (NDVI) for a full year of data (April 1992-March 1993). This thematic map was re-sampled to 0.25, 0.5 and 1.0 degree grids for the International Satellite Land Surface Climatology Project (ISLSCP) data initiative II. The data set describes the geographic distributions of 13 classes of vegetation cover (plus water and unclassified classes) based on a modified International Geosphere-Biosphere Programme (IGBP)

legend (Table 4.1). The water class is derived from a land/water mask. Note that the Urban and built-up class was not mapped from the AVHRR imagery but was developed from the populated places layer of the Digital Chart of the World (DCW). The data set also provides the fraction of each of the 15 classes within the coarser resolution cells.

***Table 4.1**. UMD land cover classification scheme.*

Class No	Description	Class No	Description
0	Water bodies	8	Closed Shrubland
1	Evergreen Needleleaf Forests	9	Open Shrubland
2	Evergreen Broadleaf Forests	10	Grassland
3	Deciduous Needleleaf Forest	11	Cropland
4	Deciduous Broadleaf Forests	12	Bare ground
5	Mixed Forests	13	Urban and Built-up
6	Woodlands	14	Unclassified
7	Wooded Grassland		

4.1.2 USGS Global Land Cover Characteristics database (GLCC)

This global land cover data set with 1-km resolution was prepared by the U.S. Geological Survey's (USGS) National Center for Earth Resources Observation and Science (EROS), the University of Nebraska-Lincoln (UNL) and the Joint Research Centre of the European Commission (Loveland et al., 2000). This data set was intended for use in a wide range of environmental research and modeling applications.

The land cover characteristics were obtained from the IGBP 1-km AVHRR 10-day composites for the period April 1992 - March 1993 (Eidenshink and Faundeen, 1994). AVHRR data quality evaluation was conducted and is reported by Zhu and Yang (1996). This set of AVHRR NDVI data was used to divide the landscape into land cover regions, based on seasonality. The regions are composed of relatively homogeneous land cover associations which exhibit distinctive phenology (that is, onset, peak, and seasonal duration of greenness), and have common levels of primary production. While the primary AVHRR data used in the classification were NDVI, the individual channel data sets were used for post-classification characterization of certain landscape properties. The methods used can be described as a multitemporal unsupervised classification of NDVI data with post-classification refinement using multi-source earth science data (Brown et al., 1993). Using these methods three datasets were derived: Global Ecosystems, IGBP Land Cover, and the USGS Land Use/Land Cover system. The USGS Land Use/Cover classification is given in the Table 4.2. It should be noted that urban areas were extracted from the Digital Chart of the World (Defense Mapping Agency, 1992).

The accuracy of the Global Land Cover Characteristics database (GLCC) was evaluated for the IGBP classification layer. While the results of this validation exercise do not provide conclusive evidence of the accuracy of the other land cover classifications included in GLCC database, they provide a general indication of data quality. According to the evaluation study (Scepan, 1999) the accuracy of sample point is about 60% and area-weighted accuracy is about 67%. A comprehensive presentation of all aspects of the validation, including per class accuracy tables, can be found in the September 1999 volume of Photogrammetric Engineering and Remote Sensing: Special Issue on Global Land Cover Validation.

The first version of the global land cover database GLCC was completed in 1997. The second version of the database contains updated land cover and water classes. However is still based on the 1992-1993 AVHRR time series, and therefore, represents the land cover patterns for that period.

***Table 4.2**. Land cover classification of USGS GLCC database.*

Value	Description	Value	Description
1	Urban and Built-Up Land	13	Evergreen Broadleaf Forest
2	Dryland Cropland and Pasture	14	Evergreen Broadleaf Forest
3	Irrigated Cropland and Pasture	15	Mixed Forest
4	Mixed Dryland/Irrigated Cropland and Pasture	16	Water Bodies
5	Cropland/Grassland Mosaic	17	Herbaceous Wetland
6	Cropland/Woodland Mosaic	18	Wooded Wetland
7	Grassland	19	Barren or Sparsely Vegetated
8	Shrubland	20	Herbaceous Tundra
9	Mixed Shrubland/Grassland	21	Wooded Tundra
10	Savanna	22	Mixed Tundra
11	Deciduous Broadleaf Forest	23	Bare Ground Tundra
12	Deciduous Needleleaf Forest	24	Snow or Ice

4.1.3 Global Land Cover 2000 database (GLC 2000)

The Global Land Cover 2000 database (GLC 2000) was developed by the European Commission's Joint Research Centre (JRC) in collaboration with over 30 research teams around the world, including FAO, UNEP/GRID and CIESIN. The general objective was to provide a harmonized land cover database over the whole globe for the year 2000 which was considered as a reference year for environmental assessment in relation to various activities, in particular, International Conventions on Climate Change, the Convention to Combat Desertification, the Ramsar Convention and the Kyoto Protocol. To achieve this objective GLC 2000 made use of the VEGA 2000 dataset of 14 months of pre-processed daily global data acquired by the VEGETATION instrument on board the SPOT 4 satellite available through a sponsorship from members of the VEGETATION Programme, including JRC.

The GLC 2000 approach is based on satellite images (1 km resolution). A bottom-up approach is used to perform global mapping where each defined region has been mapped by local experts that guaranteed an accurate classification, based on local knowledge. Each partner used the Land Cover Classification System (LCCS) developed by FAO. The LCCS is a comprehensive, standardized a priori classification system which describes land cover according to a hierarchical series of classifiers and attributes. These separate vegetated or non-vegetated surfaces, terrestrial or aquatic/flooded, cultivated and managed, natural and semi-natural, phenology, etc.. The hierarchical structure makes it possible to translate the regional classifications into a more general global legend (Table 4.3).

The GLC 2000 database provides data for various area types including forests, cultivated and managed areas, bare areas, water bodies and urban and built-up areas. The original data was produced in the geographic projection lat/long (http://www-gvm.jrc.it/glc2000). At the same

time the areas in the northern hemisphere can be displayed in the Interrupted Goode's Homolosine projection to maintain the actual size and good visual appearance.

This database has been visually validated by a number of experts and the overall response has been very positive.

***Table 4.3**: Land Cover Classification System (LCCS) with its vegetation classes*

Nr.	Global Land Cover Class
1	Tree Cover, broadleaved, evergreen
2	Tree Cover, broadleaved, deciduous, closed
3	Tree Cover, broadleaved, deciduous, open
4	Tree Cover, needle-leaved, evergreen
5	Tree Cover, needle-leaved, deciduous
6	Tree Cover, mixed leaf type
7	Tree Cover, regularly flooded, fresh
8	Tree Cover, regularly flooded, saline
9	Mosaic: Tree cover / Other natural vegetation
10	Tree Cover, burnt
11	Shrub Cover, evergreen
12	Shrub Cover, deciduous
13	Herbaceous Cover
14	Sparse Herbaceous or sparse shrub cover
15	Regularly flooded shrub and/or herbaceous cover
16	Cultivated and managed areas
17	Mosaic: Cropland / Tree Cover / Other Natural Vegetation
18	Mosaic: Cropland / Shrub and/or Herbaceous cover
19	Bare Areas
20	Water Bodies (natural and artificial)
21	Snow and Ice (natural and artificial)
22	Artificial surfaces and associated areas
23	No data

4.1.4 MODIS Land Cover products

The MODerate-resolution Imaging Spectroradiometer (MODIS) land cover and land cover change products were developed in the University of Maryland on the basis of data obtained by Moderate Resolution Imaging Spectroradiometer instrument onboard of the Terra and Aqua platforms. The satellite data acquisition period is January 1, 2001 - December 31, 2004. The distinct features of MODIS with respect to other instruments are different spectral observations, radiometric fidelity, precise geolocation, narrower bands, and depth in the shortwave infra-red which makes it possible to characterize land cover conditions not previously possible on a global basis.

The MODIS Land Cover product consists of two suites of scientific datasets. The Land Cover product includes 5 classifications of globally distributed land cover types. These are: 1) the 17 classes of land cover in the International Geosphere-Biosphere Programme (IGBP) global vegetation classification scheme, 2) the University of Maryland modification of the IGBP scheme, 3) the MODIS LAI/fPAR scheme, 4) the MODIS Net Primary Production scheme, and 5) the Plant Functional Types provided in connection with the climate modeling. Table 4.4 illustrates the first four classifications of the MODIS Land Cover data (http://edcdaac.usgs.gov/mois/mod12q1v4.asp).

The primary land cover scheme identifies 17 classes of land cover defined according to IGBP which includes 11 natural vegetation classes, 3 developed land classes, one of which is a mosaic with natural vegetation, permanent snow or ice, barren or sparsely vegetated, and water. The data are available in HDF-EOS file format at 1-km spatial resolution for the global extent of land area in 317 sinusoidally projected 10° tiles.

Table 4.4. *The MODIS Land Cover classifications.*

Class	IGBP	UMD	LAI / FPAR	NPP
0	Water	Water	water	water
1	evergreen needle leaf forest	evergreen needle leaf forest	grasses/cereal crops	evergreen needle leaf vegetation
2	evergreen broadleaf forest	evergreen broadleaf forest	Shrubs	evergreen broadleaf vegetation
3	deciduous needle leaf forest	deciduous needle leaf forest	broadleaf crops	deciduous needle leaf vegetation
4	deciduous broadleaf forest	deciduous broadleaf forest	Savanna	deciduous broadleaf vegetation
5	mixed forests	mixed forests	broadleaf forest	annual broadleaf vegetation
6	closed shrublands	closed shrublands	Needle leaf forest	annual grass vegetation
7	open shrublands	open shrublands	Unvegetated	non-vegetated land
8	woody savannas	woody savannas	Urban	urban
9	Savannas	Savannas		
10	Grasslands	Grasslands		
11	permanent wetlands	Croplands		
12	Croplands	urban and built-up		
13	urban and built-up	barren or sparsely vegetated		
14	cropland/natural vegetation mosaic			
15	permanent snow and ice			
16	barren or sparsely vegetated			

Additionally, the MODIS Land Cover product is a set of data describing seasonal time scale dynamics in ecosystems. This product provides estimates of the timing of vegetation phenology at global scales including the timing of vegetation growth, maturity, and senescence that mark the seasonal cycles.

The MODIS Land Cover products have been validated using several techniques to obtain information on uncertainty. The validation includes comparisons with in situ data collected over distributed set of validation test sites, the comparison with aircraft data and products of other sensors (e.g., AVHRR, MISR, etc.), and inter-comparison of trends derived from independently-obtained reference data. Results of all validation activities are available through both published literature and the internet. It is interesting to note that the accuracy of the IGBP classification layer was estimated to be 75-80 percent globally, 70-85 percent by continental regions, and 60-90 percent for individual classes (Zhang et al., 2006).

4.2 Comparison of land cover datasets

An initial comparison of the four datasets mentioned above is presented here.. The UMD and USGS GLCC land cover data are based on the satellite measurements obtained using AVHRR instruments. MODIS land cover products and GLC 2000 were prepared on the basis of information from more sofisticated instruments. The actual measurements were performed in the different time periods, so that can account for some of the differences in the land cover datasets. Another reason for differences in the datasets can be connected with the use of different methods for the processing and classification of satellite data.

To compare the selected datasets their land cover classes were aggregated into three groups, in particular, forests of all types, cropland areas, and urban areas in 1°x1° grid resolution. This aggregated information was then compared in global maps and we also studied the correlation between the classification of the different datasets.

Figure 4.1 illustrates the spatial distribution of the fraction of 1°x1° grid cells covered by forests of all types from the four selected datasets.

In general, the distribution of forests in the four datasets looks rather similar. There is significant correlation between the data on forests of the four datasets (correlation coefficient for each pair is about 0.85). At the same time some differences can be seen in north-eastern part of Russia, in India, and in central Africa, especially between the MODIS land cover and other datasets.

In Figure 4.2 the spatial distribution of the fraction of grid cells covered by cropland areas of four selected datasets is presented. The general distribution of cropland areas in these four datasets looks rather similar. The highest correlation, about 0.9, is obtained between the MODIS land cover and the USGS GLCC dataset, while the lowest correlation, about 0.7, is found between the UMD land cover and the GLC 2000. Some differences can be seen in India and in central Africa.

More essential differences in land cover data of the four datasets are observed for the urban areas (Figure 4.3). The lowest correlation, about 0.6, was found between the MODIS land cover and other datasets. The correlation between the USGS GLCC dataset and the UMD land cover is rather high (0.97), most likely due to the use of DCW data for the urban areas. Some differences can be seen also in Europe, eastern Asia, and in Africa.

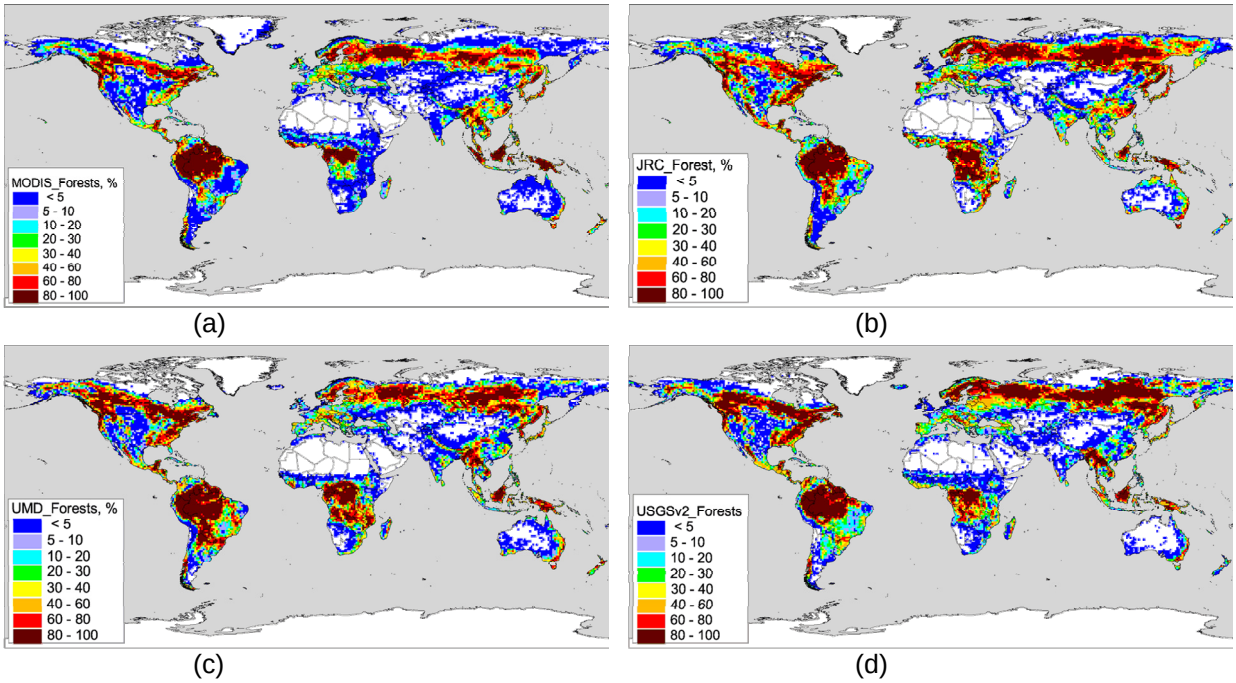


Figure 4.1. *Spatial distribution of area fraction covered by forests of all types in 1°x1° grid according to the selected LC datasets: MODIS (a), GLC 2000 (b), UMD (c), and USGS GLCC(d).*

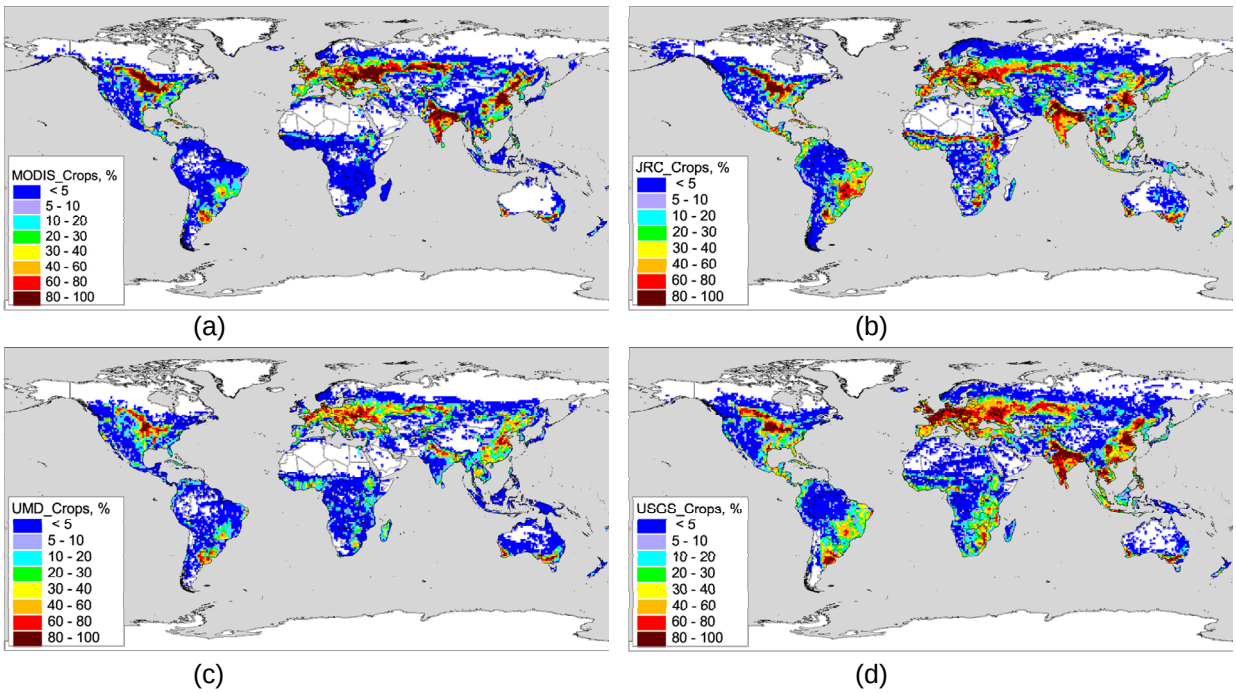


Figure 4.2. *Spatial distribution of area fraction covered by croplands in 1°x1° grid according to the selected LC datasets: MODIS (a), GLC 2000 (b), UMD (c), and USGS GLCC(d).*

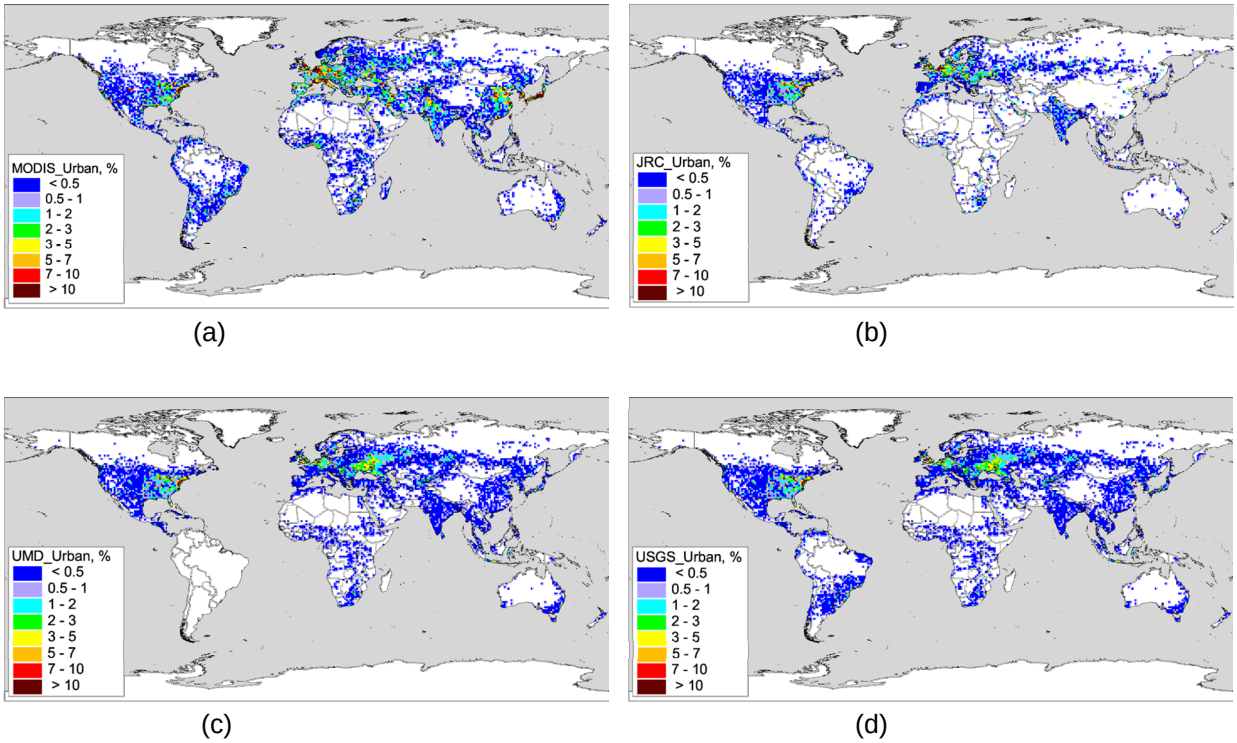


Figure 4.3. *Spatial distribution of area fraction covered by urban areas in 1°x1° grid according to the selected LC datasets: MODIS (a), GLC 2000 (b), UMD (c), and USGS GLCC(d).*

In addition to global maps, we compared the area covered by three aggregated groups of land cover types (forests, croplands, and urban and built-up areas) as derived by the four different data sets. Results are shown in Figure 4.4. It can be seen that the area covered by forests is more or less similar in the four selected datasets. However, the area covered by the croplands and urban categories is essentially different. The global area covered by croplands is twice as high in the GLC 2000 and USGS GLCC datasets than in the UMD land cover. In case of the urban land

class, the area covered by this category in the MODIS dataset exceeds more than twice the area values in the other datasets. The explanation of these differences requires further analysis but still provides a first estimate of the uncertainties related to the description of land use cover data.

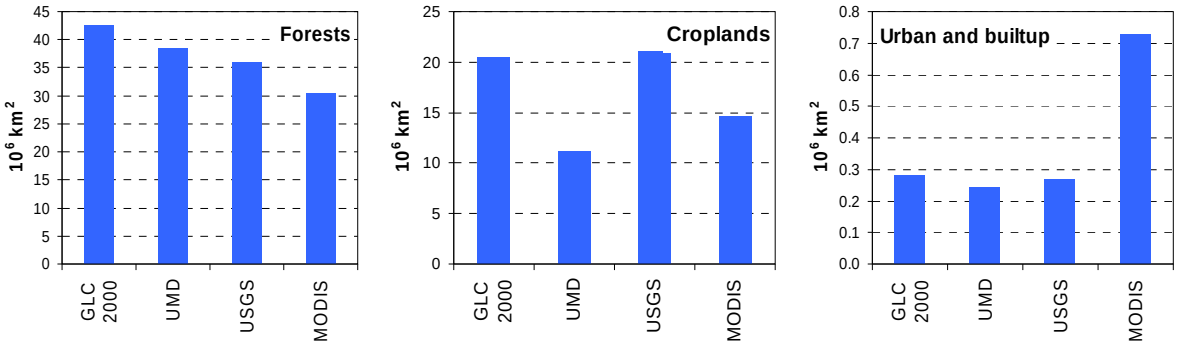


***Figure 4.4**. The area covered by aggregated categories in the selected aggregated LC datasets.*

In addition to the evaluation of differences in the aggregated data classes, we have carried out an overview comparison of the characteristics of the different datasets. We have ranged the different datasets in terms of the time periods covered by the data and in terms of the level of detail of the available information. The summary of the properties and peculiarities of the considered four global land cover datasets are collected in the Table 4.5. All of considered datasets have their advantages and disadvantages which are reflected in the table.

***Table 4.5.** Summary of information on the considered four global land cover datasets (advantages and disadvantages of particular LC datasets are marked by plus and minus signs).*

LC dataset	Time period	LAI information included	Flexibility of land use classifications	Other advantages
MODIS	+	+	+	LC changes in time
GLC 2000	+	-	-	
UMD	-	-	-	
USGS GLCC	-	+	+	Is used in WRF

MODIS Land Cover and GLC 2000 are the most recent datasets produced on the basis of more efficient instruments and using more sophisticated methods of the analysis and classification of satellite measurements. MODIS Land Cover and USGS GLCC provide the possibility to use land cover information in several alternative classifications and contain also information on the spatial distribution of LAI. In addition, some of the datasets provide sub-grid information on different vegetation classes (mosaics) that can share a single grid cell. The last column of the table presents the peculiarities of considered LC datasets. For instance, an important positive feature of the MODIS Land Cover product is that it provides information from 2001 to 2004, and thus it may describe the land cover change in time. The USGS GLCC data are used in the WRF modeling system and thus is widely used by the modeling community, although the review and feedback mechanisms need yet to be considered by MSC-E and MSC-W. For the GLC 2000 and UMD land cover datasets there are substantial differences in distribution of urban areas in comparison with the MODIS and USGS land cover data.

4.3 Available Leaf Area Index information

In addition to land cover, the LAI data is an important piece of information for the description of gaseous species interaction between the atmosphere and vegetation cover. The leaf area index represents the ratio between the area of plant leaves (one-sided or projected) and the area of the underlying surface. For modeling purposes it is of importance to use the geographically resolved LAI data. In addition the seasonal variations of vegetation cover should also be taken into account.

Gridded LAI information is available from several sources: mainly remote sensing sources and from model results. Satellite based LAI spatial information is available from the MODIS and USGS GLCC datasets as described above. These two datasets contain spatially resolved LAI data with 1 km resolution and include information also on the seasonal variations of LAI. In addition, the model of emission of biogenic volatile organic compounds MEGAN (Guenther et al., 2006) and the dynamic global vegetation model of NCAR Community Land Model (Lawrence and Chase, 2006) provide LAI estimates.

It is important to note though that satellite-based LAI values have a potentially-important limitation. The relationship between satellite-measured Normalized Difference Vegetation Index (NDVI) index and LAI shows a "saturation", in which LAI values greater than 3-4 are hard to distinguish from much higher LAI values. Thus, NDVI-based maps of LAI tend to show lower LAI values (typically < 6, often 2-3) than those known to occur in reality in many areas. It can be argued that many characteristics of a canopy also show saturation (e.g. deeper levels in a canopy emit smaller amounts of biogenic VOC, or are less important for deposition), but previous work with the EMEP deposition model has suggested that deposition is sensitive to LAI even for LAI > 3 (Simpson et al., 2003, Tuovinen et al., 2001, 2004), so the low LAI from NDVI-based maps are of concern.

Alternatively, leaf area indexes and their temporal variations can be defined for particular land-use classes which describe vegetation cover on the basis of ground based measurements. For instance, this information can be obtained from the LAI ground based measurements database (Asner et al., 2003) containing the synthesis of global observations on global scale. The Figure 4.5 illustrates the distribution of mean LAI values by the land cover classes and their standard deviations.

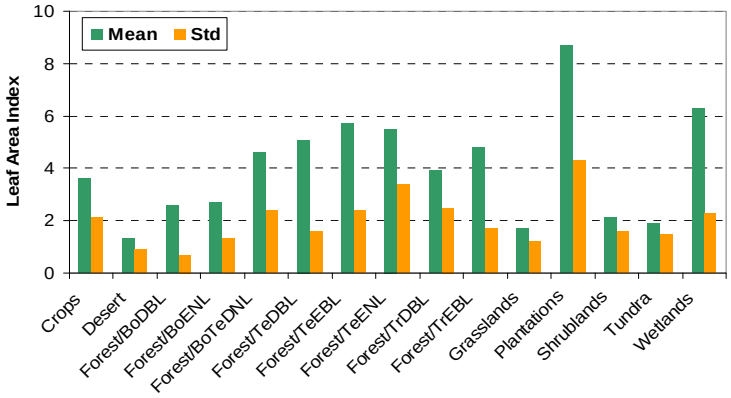


Figure 4.5. *Distribution of mean leaf area index values by land cover classes obtained on the basis of measurements (Asner et al., 2003). The meaning of the acronyms of forest land cover classes used in the Figure is as follows: Forest/BoBDL - Boreal deciduous broadleaf forests; Forest/BoENL - Boreal evergreen needle leaf forests; Forest/BoTeDNL - Boreal/ temperate needle leaf forests; Forest/TeDBL - Temperate deciduous broadleaf forests; Forest/TeEBL - Temperate evergreen broadleaf forests; Forest/TeENL - Temperate evergreen needle leaf forests; Forest/TrDBL - Tropical deciduous broadleaf forests; Forest/TrEBL - Tropical evergreen broadleaf forests.*

4.4 Concluding remarks on geophysical data

Information on land cover from several remote sensing sources has been collected and an initial comparison of MODIS LC, GLC 2000, UMD, and USGS GLCC has been carried out. In general, the distribution of forests and cropland areas in four datasets looks rather similar in the four datasets when compared in aggregated levels at 1°x1° grid resolution. At the same time substantial differences between the datasets were identified for the urban category. The reasons behind these differences require further analysis but this initial comparison provides a first estimate of the extent of the uncertainties related to the description of land use cover data.

None of the land-use data sets presented here contains information on the species-composition of forests. The general classifications of coniferous, deciduous, mixed forest etc. are generally sufficient for deposition modeling, but species-level information can provide much better estimates of biogenic VOC emissions when available (Simpson et al., 1995, 1999; Koeble and Seufert, 2005; NatAir, 2007). It is not clear if such species-level data are appropriate for global modeling studies (the complexity of ecosystems and lack of data probably preclude this), this issue requires further investigation.

The work on collection and comparison of LAI information is ongoing. Several sources of LAI information are considered, in particular, MODIS Land Cover, USGS GLCC, MEGAN, and NCAR CLM. The geographically resolved LAI information along with the description LAI seasonal variations from these sources can be used. On further stages of the work the collection of LAI information will be continued and a detailed comparison of the data will be performed.

4.5 References

Asner G.P., J.M.O. Scurlock, and J.A. Hicke, (2003) Global synthesis of leaf area index observations: implications for ecological and remote sensing studies, Global Ecology & Biogeography, 12, 191-205.

Brown J.F., Loveland T.R., Merchant J.W., Reed B.C., and Ohlen D.O., (1993), Using multisource data in global land cover characterization: concepts, requirements and methods: Photogrammetric Engineering and Remote Sensing , v. 59, p. 977-987.

Eidenshink J.C. and Faundeen J.L., (1994), The 1 km AVHRR global land data set-first stages in implementation: International Journal of Remote Sensing, v. 15, no. 17, p. 3443-3462.

Guenther A., T. Karl, P. Harley, C. Wiedinmyer, P.I. Palmer, and C. Geron, (2006) Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181-3210.

Hansen, M., R. DeFries, J.R.G. Townshend, and R. Sohlberg, (2000), Global land cover classification at 1km resolution using a decision tree classifier, International Journal of Remote Sensing. 21: 1331-1365.

Köble, R. & Seufert, G. (2005), 'Novel Maps for Forest Tree Species in Europe.', In Proceedings, 'A Changing Atmosphere, 8th European Symposium on the Physico-Chemical Behaviour of Atmospheric Pollutants, Torino, Italy, 17-20 Sept., 2001'.

Lawrence P.J. and T.N. Chase (2007) Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0), Journal of Geophysical Research, Vol. 112, G01023, doi:10.1029/2006JG000168.

Loveland T.R., Reed B.C., Brown J.F., Ohlen D.O., Zhu J., Yang L., and Merchant J.W., (2000), Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1-km AVHRR Data: International Journal of Remote Sensing, v. 21, no. 6/7, p. 1303-1330.

NATAIR (2008), 'Improving and applying methods for the calculation of natural and biogenic emissions and assessment of impacts to the air quality' EU 6th Framework Proposal No. 5136999, R. Friedrich (coordinator), Final report, University of Stuttgart, Germany.

Scepan J., (1999), Thematic Validation of High-Resolution Global Land-Cover Data Sets: Photogrammetric Engineering and Remote Sensing, v. 65, no. 9, p. 1051-1060.

Simpson, D., Tuovinen, J., Emberson, L. and Ashmore, M. (2003), Characteristics of an ozone deposition module II: sensitivity analysis, Water, Air and Soil Pollution 143, 123-137.

Simpson, D.; Guenther, A.; Hewitt, C. & Steinbrecher, R. (1995), 'Biogenic emissions in Europe 1. Estimates and uncertainties', J. Geophys. Res. 100(D11), 22875-22890.

Simpson, D.; Winiwarter, W.; Börjesson, G.; Cinderby, S.; Ferreiro, A.; Guenther, A.; Hewitt, C. N.; Janson, R.; Khalil, M. A. K.; Owen, S.; Pierce, T. E.; Puxbaum, H.; Shearer, M.; Skiba, U.; Steinbrecher, R.; Tarrasón, L. & Öquist, M. G. (1999), 'Inventorying emissions from Nature in Europe ', J. Geophys. Res. 104(D7), 8113-8152.

Tuovinen, J., Ashmore, M., Emberson, L. and Simpson, D. (2004), Testing and improving the EMEP ozone deposition module, Atmos. Environ. 38, 2373-2385.

Tuovinen, J., Simpson, D., Mikkelsen, T., Emberson, L., M., M. R. A., Aurela, Cambridge, H. M., Hovmand, M. F., Jensen, N. O., Laurila, T., Pilegaard, K. and Ro-Poulsen, H. (2001), Comparisons of measured and modelled ozone deposition to forests in Northern Europe, Water, Air and Soil Pollution: Focus 1, 263-274.

Zhang X., M. A. Friedl C. B. Schaaf (2006) Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements, Journal of Geophysical Research, Vol. 111, G04017, doi:10.1029/2006JG000217.

Zhu Z., and Yang L., (1996), Characteristics of the 1-km AVHRR data set for North America: International Journal of Remote Sensing v. 17, p. 1915-1924.

5. Comparison of meteorological input data

Alexey Gusev, Jan Eiof Jonson, Peter Wind, Michael Gauss,
Valery Sokovych and Leonor Tarrasón

The preparation of the meteorological input data necessary for the evaluation of pollution levels at global scale is an essential part of the development of the EMEP global modeling framework. It is planned to perform the testing of several currently available global scale meteorological models, in particular, GEM of Environment Canada, PUM of UK Met Office, ECMWF IFS, and WRF of US EPA. For the driving meteorological input, meteorological analyses of ECMWF will be used for the comparison of all these different meteorological drivers.

This chapter presents the initial evaluation and testing of two meteorological drivers, namely, Canadian forecasting and assimilation system GEM and ECMWF IFS. The MSC-E has also started to work with the recently developed global version of the meteorological model WRF. However testing of this model is in a very preliminary stage, so its results and comparison with other selected models will hopefully be presented in the final report on testing of meteorological drivers planned for 2009.

5.1 Description of the GEM and the IFS meteorological drivers

This section presents a brief description of two meteorological drivers under evaluation. The meteorological drivers compared throughout this chapter are: the Global Environmental Multiscale Model (GEM) developed at Canadian Meteorological Center (CMC) and the Integrated Forecast System(IFS) developed at the European Centre for Medium Range Weather Forecasts (ECMWF).

5.1.1 Short description of the IFS at ECMWF

The IFS (Integrated Forecast System) is a comprehensive earth-system model system developed and used at the European Centre for Medium Range Weather Forecasts (ECMWF, http://www.ecmwf.int) and forms the basis for all data assimilation, forecasting and reanalysis activities at ECMWF. It is comprehensively documented on the web pages of ECMWF: http://www.ecmwf.int/research/ifsdocs/ .

In 2008 the numerical scheme has been run on Gaussian T799L91 resolution (799 waves around a great circle on the globe, 91 levels between 0-80 km) with a semi-Lagrangian formulation and a time step of 12 minutes. Prognostic variables include wind, temperature, humidity, cloud fraction and water/ice content, pressure at surface grid-points, and ozone.

The IFS includes a 4D VAR data assimilation system (since 1999), a numerical weather prediction model, and an ocean model. Physical aspects included in IFS are:

- Orography (terrain height and sub-grid-scale characteristics)

- Four surface and sub-surface levels (allowing for vegetation cover, gravitational drainage, capillarity exchange, surface / sub-surface runoff)
- Deep and shallow convection, stratiform and convective precipitation
- Diffusion
- Ground & sea roughness
- Ground and sea-surface temperature
- Ground humidity
- Snow-fall, snow-cover and snow melt
- Radiation (incoming short-wave and out-going long-wave)
- Friction (at surface and in free atmosphere)
- Sub-grid-scale orographic drag
- Gravity waves and blocking effects
- Evaporation, sensible and latent heat flux
- Carbon dioxide (345 ppmv fixed), aerosol, ozone, methane oxidation

The IFS is under continuous development, the different versions being called cycles. The version we have been using for generating our meteorological data for the year 2001 is referred to as cycle '29r2'. It is based on cycle 28r1 (March 2004), which is fully documented at:http://www.ecmwf.int/research/ifsdocs/CY28r1/index.html

The subsequent steps that finally resulted in cycle 29r2 (June 2005) are 28r2, 28r3, 29r1, and 29r2. All changes made in these steps are listed at:http://www.ecmwf.int/products/data/technical/model_id/index.html

The intoduction of Cycle 29r2 included the following changes and improvements (as detailed in http://www.ecmwf.int/products/data/operational_system/evolution/evolution_2005.html#28June2005)

- Assimilation of rain-affected SSM/I radiances
- Set of changes to AIRS assimilation
- Use of Meteosat-8 (MSG) winds
- Jb statistics from latest ensemble data assimilation
- Refinement of surface-pressure bias correction
- Humidity analysis changes affecting spin-up & stratosphere
- Use of SMHI Baltic sea-ice analysis
- Convection changes
- Revised Gaussian sampling for EPS (ensemble prediction system) perturbations.

The IFS presently provides input meteorological data for the Unified EMEP model global applications. To generate meteorological data for the Unified EMEP model, the IFS model system was run by the University of Oslo through remote login at the ECMWF supercomputer facility. The IFS was run in 40 layers between the surface and 10 hPa where the uppermost layer is centered. The horizontal resolution was spectral T319, which corresponds to about 0.5 degrees horizontal resolution. The model was initialized from assimilated data every 24 hours and run for 36 hours. The first 12 hours were used as spin-up and the following 24 hours were archived. The IFS model generates data on the so-called 'grib' format, a highly compressed, binary, data format. From the grib files, EMEP/MSC-W have extracted the variables used in the Unified EMEP model and converted the data into EMEP-readable netCDF or FELT format. This step included the interpolation from the Gaussian spectral T319 grid into the EMEP 1ºx1º regular grid and EMEP's 20 layers. A mass filter was applied to ensure mass balance after the interpolation. Data are available for the years 2001 up to the present, but in this report we focus only on the

2001 data set. In this report, the meteorological input data compared to GEM is the actual input used in the Unified EMEP model. This means that the IFS fields are already interpolated to the geographical grid and that the vertical extent of the fields analysed reaches only up to 100hPa and not 10hPa.

5.1.2 Short description of the GEM model at CMC

The Global Environmental Multiscale Model (GEM) (Côté et. al., 1998a, Côté et. al., 1998b, Yeh et. al., 2002) is an integrated forecasting and data assimilation system developed in the Canadian Meteorological Center (CMC) to meet all of the current and foreseeable operational weather-forecasting needs of Canada. The export version of the model is available for use and can be downloaded from the site http://collaboration.cmc.ec.gc.ca/science/rpn.comm/.

The GEM model dynamics is formulated in terms of the non-hydrostatic equations with a terrain following pressure vertical coordinate. For time integration an implicit two-time-level semi-Lagrangian scheme is applied. For the spatial discretization in horizontal direction the Galerkin grid-point formulation on an Arakawa C-grid is used. The horizontal grid can be global or limited area and can have uniform or variable resolution (Figure 5.1). Furthermore the horizontal grid system can be arbitrarily rotated. Along the vertical direction hybrid coordinates are applied with variable mesh sizes.

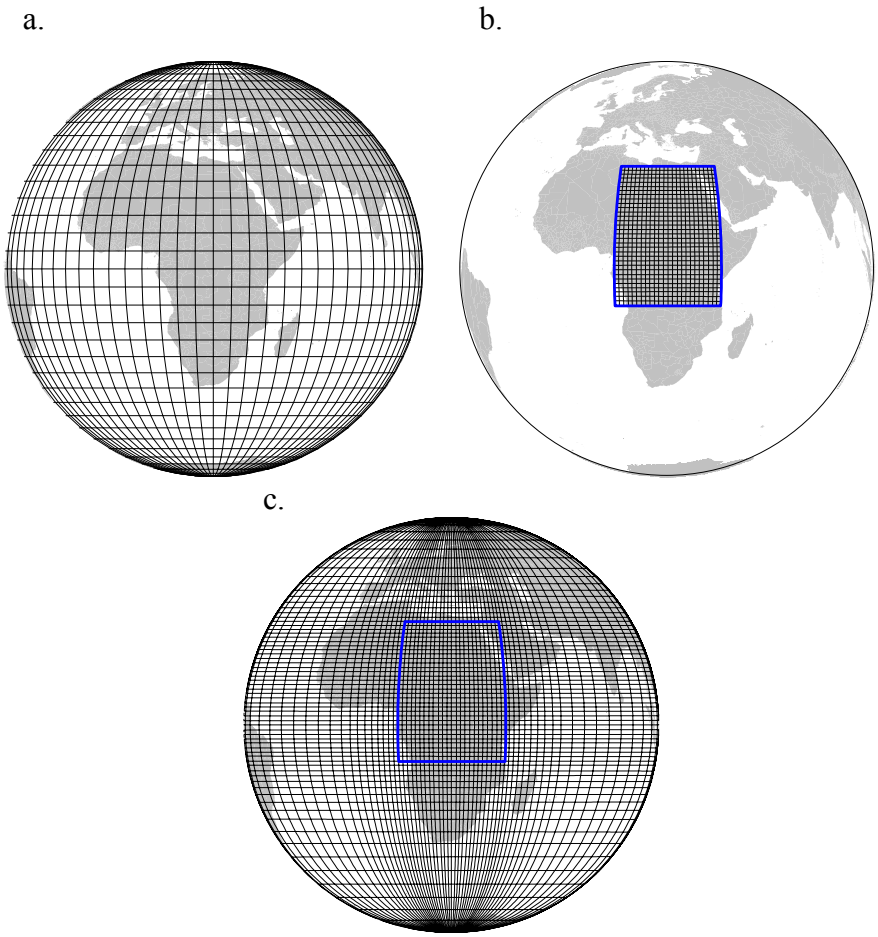


Figure 5.1. *Examples of GEM modelling domains: a – global domain with uniform spatial resolution, b – limited area domain with uniform spatial resolution, c – global domain with variable spatial resolution and limited area domain with uniform resolution.*

The operational GEM model is interfaced with a full complement of physical parameterizations (Mailhot et al., 1997). Currently they include:

- turbulent vertical diffusion;
- land and water surface processes;
- gravity wave drag;
- deep convective processes and shallow convection;
- condensation processes at resolved scales;
- cloud parameterization;
- infrared and solar radiation.

GEM input data

To run the GEM model on the global scale there is a need to supply it with gridded geophysical and climatological data. These input data are distributed as part of the GEM model system. In addition, the preparation of fields of meteorological parameters for the initial conditions is necessary. For this purpose the ECMWF meteorological analyses for 2001were adapted for the use as initial data for GEM. The list of meteorological parameters required to run the GEM model is presented in Table 5.1. The preparation of necessary driving meteorological input for GEM includes the transformation of ECMWF data in GRIB format into the format required by GEM.

Table 5.1*. The list of meteorological parameters of ECMWF reanalysis required to run GEM model*

Parameter	**Dimension**
2m air temperature	2-D
Air temperature	3-D
Albedo	2-D
Geopotential	3-D
Sea-ice cover	2-D
Sea surface temperature	2-D
Skin temperature	2-D
Snow density	2-D
Snow depth	2-D
Soil temperature	2-D
Specific humidity	3-D
Surface pressure	2-D
U wind velocity	3-D
V wind velocity	3-D
Volumetric soil water	2-D

Selection of physical parameterizations and calculation cycle settings

The GEM system is flexible with respect to the choice of physical parameterizations.The following modules are available for different set ups of the GEM system:

convective scheme *kfc* (Kain and Fritsch, 1993),
kuosym (Wagneur, 1991),
ras (Moorthi and Suarez, 1992),
oldkuo (Geleyn, 1985)

shallow convection	conres, conres-ktrsnt (Mailhot et. al., 1988)
type of condensation	consun, newsund (Sundqvist et al., 1989)
mixing length formulation	blackadar (Blackadar, 1962), boujo (Bougeault and Lacarrère, 1989)
land surface processes	fcrest (Mailhot et al., 1997)
radiation scheme	newrad (Garand and Mailhot, 1990)

Four differenr configurations of the system were tested at MSC-E and were evaluated against observations in order to decide an optimal configuration of the GEM system for MSC-E GLEMOS model applications. The different configurations are summarized in Table 5.2. The choices of GEM parametrizations took into account the experience of Canadian ensemble forecasts (www.weatheroffice.gc.ca/ensemble/index_e.html)

Alternative parameterization of land surface processes, in particular, ISBA (Noilhan and Planton, 1989), often used in various applications of GEM, was not included in testing. The reason of this is connected with the absence of necessary data on ice content in the soil, liquid water retained on vegetation and liquid water retained in the snow pack in the meteorological analyses of ECMWF.

Table 5.2. *Different configurations of the GEM system tested at MSC-E*

№	Convective Scheme	Shallow Convection	Type of Condensation	Mixing Length Formulation	Land Surface Processes	Radiation Scheme
1	Kfc	conres-ktrsnt	Consun	Blackadar	Fcrest	newrad
2	Kuosym	conres-ktrsnt	Consun	Boujo	Fcrest	newrad
3	Ras	conres-ktrsnt	Consun	Blackadar	Fcrest	newrad
4	oldkuo	conres	Newsund	Boujo	Fcrest	newrad

The process of preparation of meteorological data for a long period of time (one or several years) is organized as the sequence of short period meteorological model runs. Each of these runs consists of the two parts: spin-up and forecast. A typical period for a spin-up is 12 h. The length of the forecast period can be different. To filter out the high-frequency oscillations each short period model run begins with the 6-h digital filtering (Fillion at. al.,1995). Filtered modeling results of the forecast period are stored and further used as the meteorological input for the chemical transport models.

Eight different set-ups were tested at MSC-E to select the GEM configuration of physical parameterizations and the optimal parameters of the calculation sequence. The tests were performed for January and July of 2001.. Test simulations used different GEM parameterizations, duration of spin-up cycle, duration of forecast cycle, and the number of vertical layers, as summarized in Table 5.3. All GEM model runs were performed using the lat-lon grid with horizontal resolution 1°x1° and model domain upper boundary at 10hPa.

The different configurations were tested against surface observations of precipitation and also air temperature, U and V wind speed components, dew point temperature, and geopotential height at different pressure levels. The different configurations of GEM system in Table 5.3 were also compared with regard to the required computer time. It was concluded that computational time is not very sensitive to the variation of model parameterizations (at least for the selected parameterizations) but that it significantly depends on the number of vertical layers and the duration of the spin-up and forecast cycles.

***Table 5.3**. List of test configurations with different parameterizations of GEM and different settings of calculation sequence, carried out at MSC-E.*

№	Parameterization Number (Table 2)	Spinup Cycle, hours	Forecast Cycle, hours	Levels
1	1	12	24	21
2	2	12	24	21
3	3	12	24	21
4	4	12	24	21
5	4	12	24	28
6	4	12	12	21
7	4	12	48	21
8	4	24	24	21

The evaluation of these eight different configurations is not shown here, but it resulted in the following recommendation for the use of the GEM system at MSC-E

- The choice of physical parametrizations giving the better statistical performance of the model in the areas of interest for modeling at MSC-E is as follows:

  - **convective scheme** *oldkuo*
  - **shallow convection** *conres*
  - **type of condensation** *newsund*
  - **mixing length formulation** *boujo*
  - **land surface processes** *fcrest*
  - **radiation scheme** *newrad*

- The optimal durations of the spin-up and forecast cycles is chosen to be 12 hours and 24 hours, correspondingly.

- The increase of the number of vertical layers improves the agreement between the modeling results and measurements at the levels higher than 300 hPa. Below this level the influence of the increasing is quite small. It is believed that at current stage of the work it is possible to obtain reasonable modeling results using 21 vertical layers in GEM simulations at MSC-E.

5.2 Selected sites and fields used for meteorological driver evaluation

Two kinds of measurement data were used in the evaluation of GEM and IFS modeling results: a) surface measurements and b) vertical profile observations.

In this initial evaluation of the meteorological drivers, the following meteorological variables were selected for the comparison with the surface measurements: 1) air temperature at 2m, 2) wind speed at 10m, and 3) precipitation amount. For the comparison of the vertical profile data on pressure levels, air temperature and wind speed components U and V were selected.

Concerning the temporal resolution of the meteorological fields used in this comparison, the initial tests presented in this chapter, deal primarily with, annual and monthly mean values. Further analysis of hourly data is to be carried out in the following stages on this meteorological driver inter-comparison.

Surface observations of daily precipitation amount from GCOS Surface Network (GSN) were used to compare the precipitation fields calculated by GEM and IFS models with measurements. The Global Climate Observing System (GCOS) Surface Network (GSN) data are available online from the National Climatic Data Center (NCDC) in Asheville, North Carolina (www.gosic.org/gcos/GSN-data-access.htm). Daily precipitation sums for 2001 were available from more than 1000 sites over the globe, as shown in Figure. 5.2.

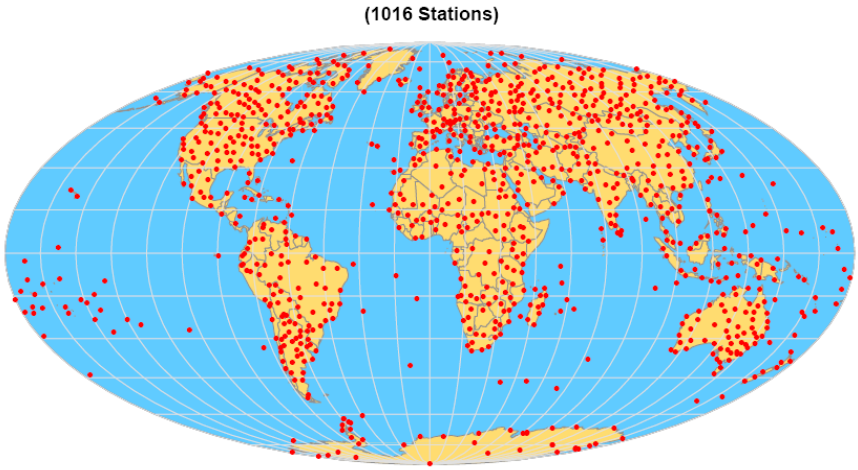


Figure 5.2. *Location of meteorological measurement sites of GSN*

Vertical profile observations from Radiosonde Database (RAOB) were used for the comparison of computed and observed data on pressure levels. The RAOB database is developed jointly by National Climatic Data Centre (NCDC) and Forecast Systems Laboratory (FSL) and is freely available at the site (http://raob.fsl.noaa.gov). It is based on reports available through GTS. The information for 2001 is presented from 1902 sites over the globe (Figure. 5.3). It can be seen that most of measurement sites belong to the North America, Europe, and Southeast Asia.

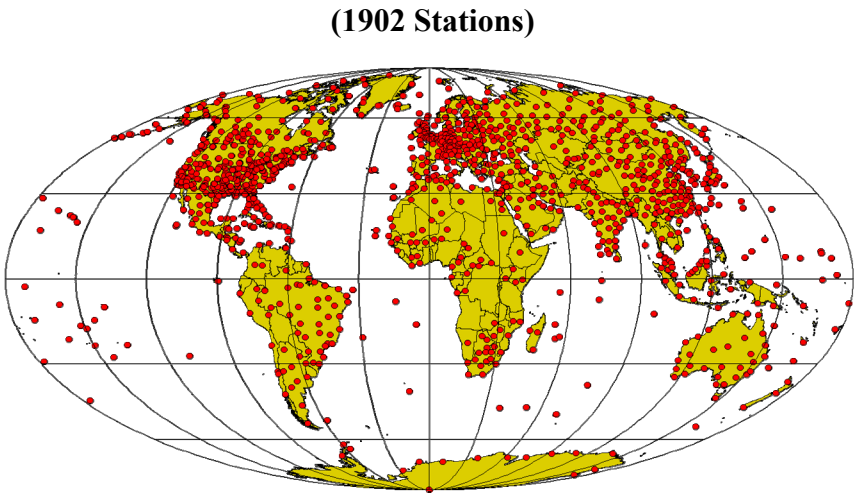


Figure 5.3. *Location of meteorological measurement sites of RAOB*

Precipitation amounts calculated by the GEM and IFS models was also compared with the precipitation dataset for 2001 prepared by the GPCP project (Rudolf et al., 1994) which can be

found in the internet (http://cics.umd.edu/~yin/GPCP). This is a combined dataset based on a blending of rain gauge observations and satellite data. Over land areas with dense monitoring network, the GPCP data are mainly represented by ground level measurements. Over areas with the scarce network, like over the oceans, the data are mostly based on the information obtained from satellites. The spatial resolution of these data is 1°x1°, and temporal resolution is one day (Huffman et al., 2001).

In addition to the comparison with measurements from GSN and RAOB, the spatial distribution of selected meteorological parameters calculated by the GEM and IFS models for 2001 was compared with the data of two meteorological re-analysis products ECMWF ERA-40 and NCEP/DOE [*Kanamitsu et al.*, 2002].

5.3 Statistics used

To characterize the level of agreement between the computed and observed values of selected meteorological variables the following statistical parameters were used: BIAS, root mean square error (RMSE), and correlation coefficient (R_{corr}). The BIAS parameter describes the absolute difference between the calculated and measured data. The correlation coefficients show similarity of fields or time series of observed and calculated values of selected meteorological parameters.

$$1.\ BIAS = \frac{1}{N_t}\sum_{i=1}^{N_t}\left(\overline{M^i} - \overline{O^i}\right) = \frac{1}{N_s}\sum_{j=1}^{N_s}\left(\overline{M_j} - \overline{O_j}\right),$$

$$2.\ RMSE_{time} = \frac{1}{N_s}\sum_{j=1}^{N_s}\sqrt{\frac{1}{N_t}\sum_{i=1}^{N_t}\left(M_j^i - O_j^i\right)^2}\ ,\quad RMSE_{space} = \frac{1}{N_t}\sum_{i=1}^{N_t}\sqrt{\frac{1}{N_s}\sum_{j=1}^{N_s}\left(M_j^i - O_j^i\right)^2}\ ,$$

$$3.\ R_{corr-time} = \frac{1}{N_s}\sum_{j=1}^{N_s}\frac{\sum_{i=1}^{N_t}\left(M_j^i - \overline{M_j}\right)\left(O_j^i - \overline{O_j}\right)}{\sqrt{\sum_{i=1}^{N_t}\left(M_j^i - \overline{M_j}\right)^2}\sqrt{\sum_{i=1}^{N_t}\left(O_j^i - \overline{O_j}\right)^2}},\quad R_{corr-space} = \frac{1}{N_t}\sum_{i=1}^{N_t}\frac{\sum_{j=1}^{N_s}\left(M_j^i - \overline{M^i}\right)\left(O_j^i - \overline{O^i}\right)}{\sqrt{\sum_{j=1}^{N_s}\left(M_j^i - \overline{M^i}\right)^2}\sqrt{\sum_{j=1}^{N_s}\left(O_j^i - \overline{O^i}\right)^2}},$$

where $\overline{M^i} = \frac{1}{N_s}\sum_{j=1}^{N_s} M_j^i$ $\overline{O^i} = \frac{1}{N_s}\sum_{j=1}^{N_s} O_j^i$ $\overline{M_j} = \frac{1}{N_t}\sum_{i=1}^{N_t} M_j^i$ $\overline{O_j} = \frac{1}{N_t}\sum_{i=1}^{N_t} O_j^i$,

M_j^i, O_j^i are the computed and observed values in point j and at the time i, $i = \overline{1, N_t}$, $j = \overline{1, N_s}$.

For the mean monthly values of surface air temperature and precipitation amount the statistical parameters mentioned above were calculated in standard way. For the vertical profiles and upper-air measurements (RAOB), correlation coefficients and root mean square errors for the spatial distribution and temporal variations of meteorological variables were considered separately.

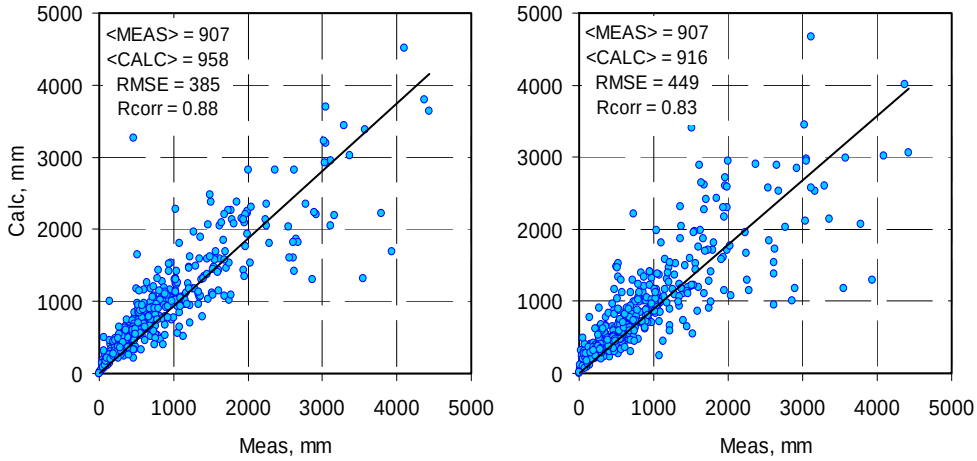
5.4 IFS and GEM output data comparison for 2001

The initial evaluation of these two meteorological drivers was carried out for precipitation fields, air temperature and wind fields. In the following, the main conclusions of the inter-comparison for monthly and annual fields in 2001 are presented. A more comprehensive evaluation of the results is planned to be included in the final joint report on the selection of meteorological drivers which is expected to be completed in 2009.

5.4.1 Precipitation

Figure 5.4 presents the scatter plots of annual precipitation amount calculated by IFS and GEM models versus the measurements of GSN network for 2001. Both models describe reasonably well global scale distribution of annual precipitation amount, both with a positive bias, indicating a tendency to overestimate precipitation amounts. On average the IFS model provided slightly higher annual precipitation in comparison to the results of GEM The spatial correlation of observed and calculated precipitation is estimated to 0.88 for IFS and 0.83 for GEM.. At the same time, precipitation amount computed by IFS has lower RMSE parameter comparing to the data of GEM model.

Annual precipitation amount simulated by the GEM and IFS models for 2001 is presented in Figure 5.5. Re-analyses ERA-40 and NCEP/DOE, and from GPCP product are also given in the figure. The general pattern of annual precipitation amount is rather similar but there are significant differences in equatorial regions where re-analyses provide higher precipitation amount comparing to GEM and IFS model results and GPCP.



a) IFS model b) GEM model

Figure 5.4. *Scatter plots of measured annual precipitation amount at monitoring sites of GSN network versus calculated by IFS (a) and GEM (b) models for 2001*

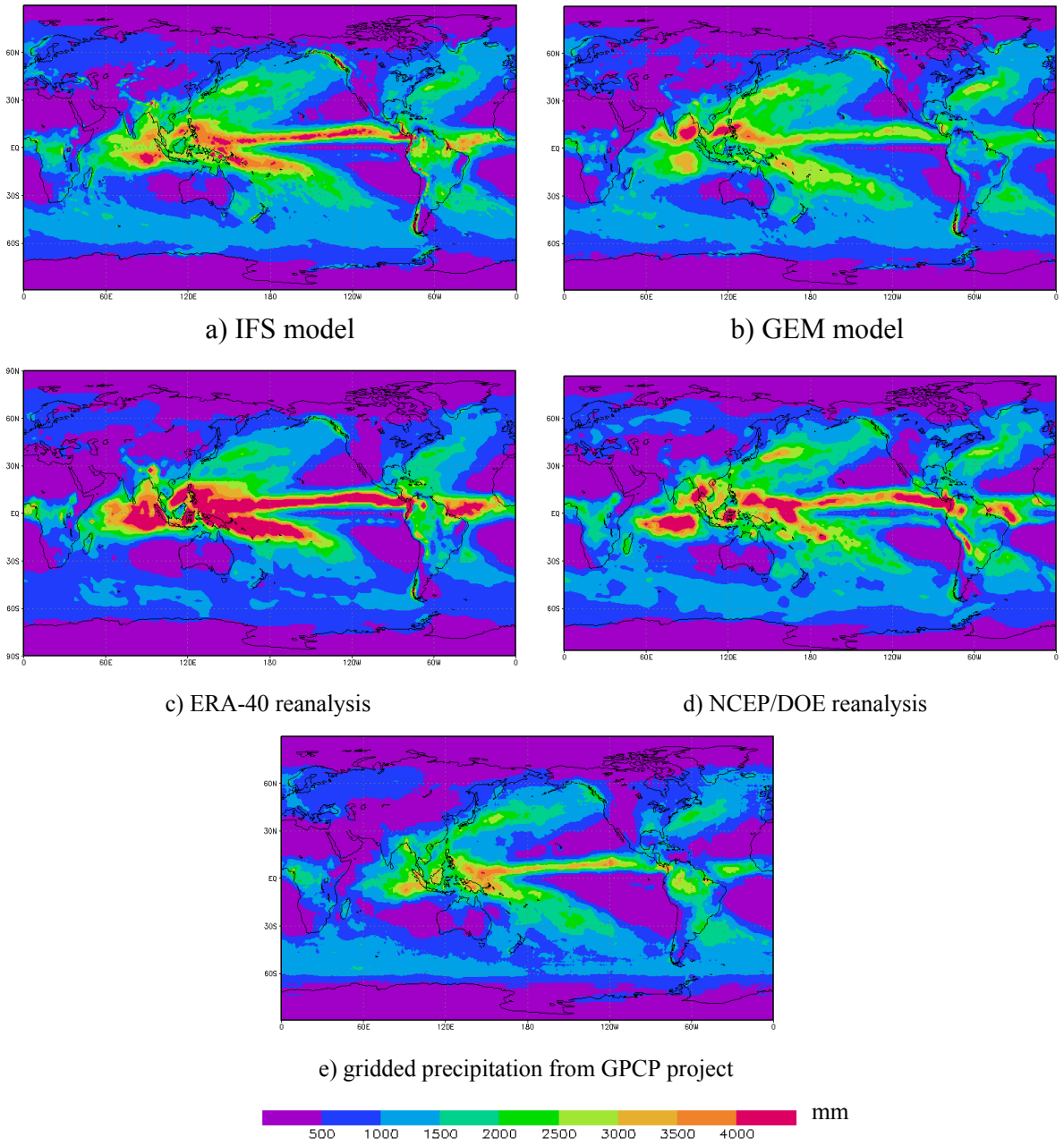
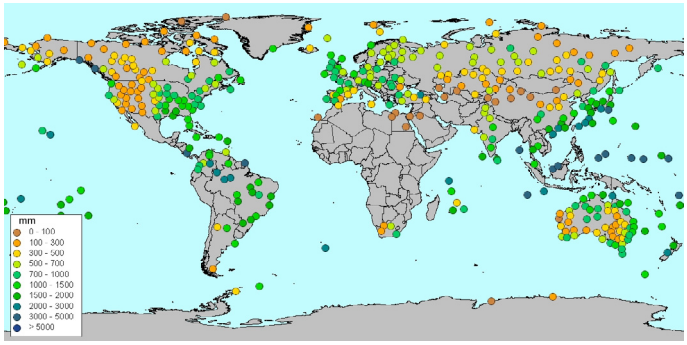
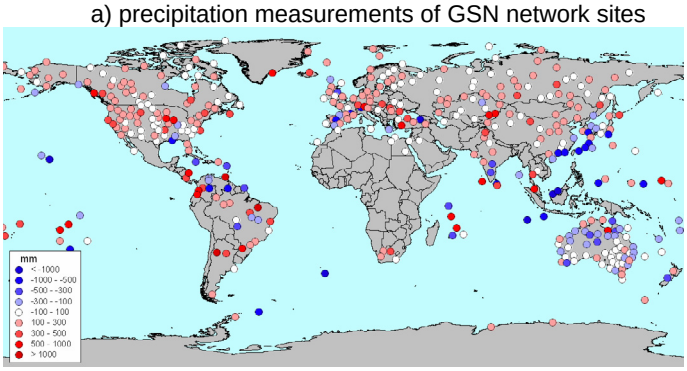


Figure 5.5. *Spatial distribution of annual precipitation amount for 2001 calculated by ECMWF IFS and GEM models, obtained from re-analyses ERA-40 and NCEP/DOE, and according to the GPCP project which is based on ground measurements and satellite data.*

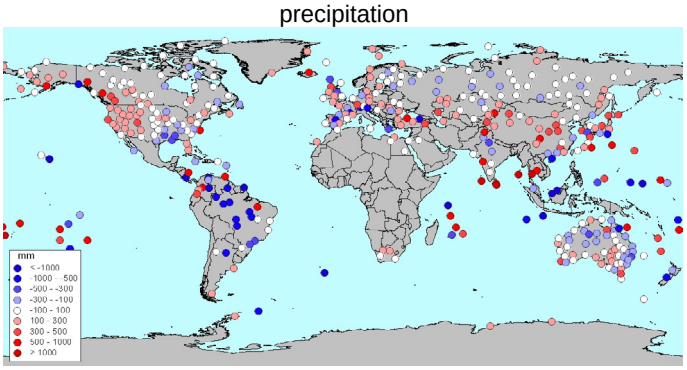
Figure 5.6a presents the annual precipitation amount measured at the sites of GSN network in 2001. It should be noted that the map shows only the sites for which the measurements were available for each month of the year 2001. Figs. 5.6b and 5.6c present the difference between the calculated and measured annual precipitation amount for IFS and GEM models, respectively. In general, the pattern of deviations between the model results and measurements is quite similar. It can be seen that both models reasonably reproduce low levels of precipitation. At the same time the GEM model tends to underestimate precipitation in Russia and Canada whereas the IFS provided some overestimation for these regions. Maximum absolute differences (more than 500 mm) can be noted for the Southeast Asia, Oceania, and South America, characterized by significant annual precipitation. Differences exceeding a factor of two can be noted for a number of sites located at high altitudes in China and the western part of the USA.



a) precipitation measurements of GSN network sites



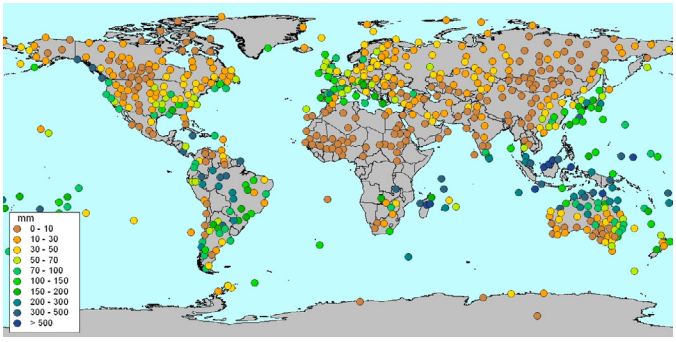
b) difference between IFS calculations and measurements of precipitation



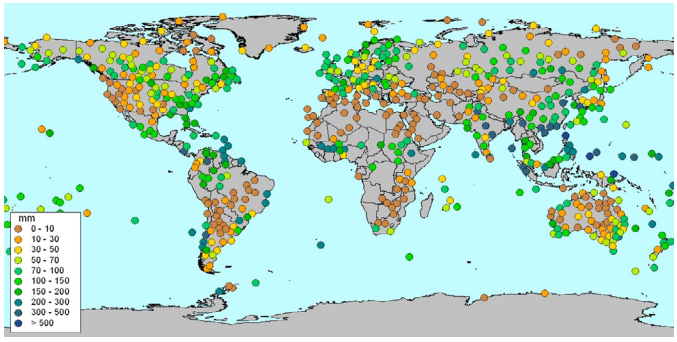
c) difference between GEM calculations and measurements of precipitation

Figure 5.6 *Comparison of annual precipitation amount calculated by GEM and IFS for 2001 with measurements of GSN network.*

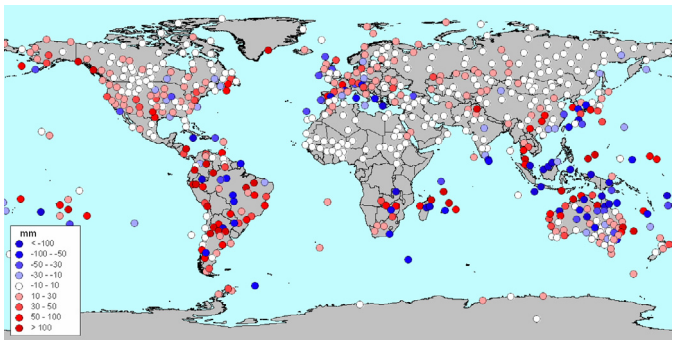
The number of sites presented in Figure 5.6 is somewhat lower in comparison to the whole amount of sites shown in Figure 5.2 due to exclusion of the sites which measurements did not cover the whole year 2001. This exclusion criteria affects mostly African measurements. Taking this into account more detailed information on calculated and measured precipitation amount is shown in Figure 5.7. In particular, the figure presents measured monthly precipitation amount sums for January and July in comparison with the modeling results of GEM and IFS. On the whole better agreement can be seen between the calculated and measured values for the January than for the July. Both models provided good agreement with winter time measurements of significant number of sites in North America, Asia, and Central Africa. Somewhat worse agreement can be seen for the summer period especially in the Northern Hemisphere and the tropical regions.



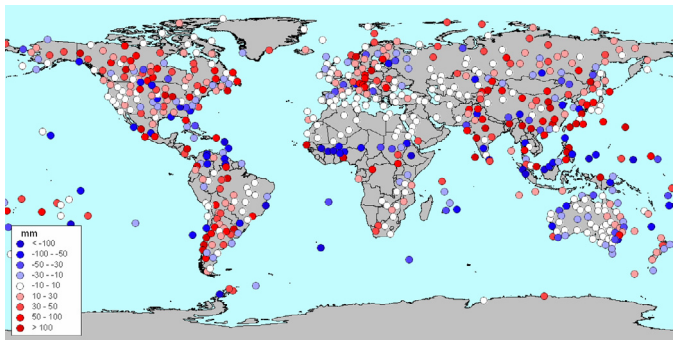
a) measured precipitation amount for January (GSN)



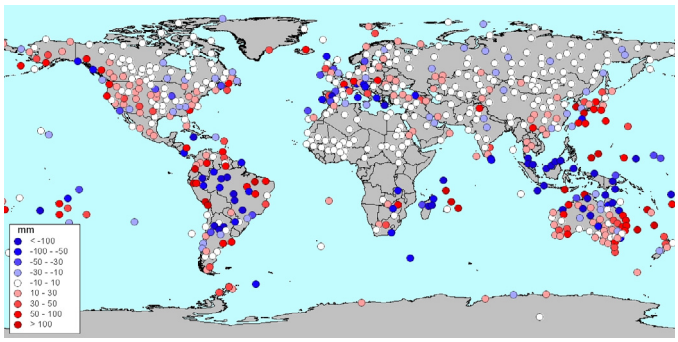
b) measured precipitation amount for July (GSN)



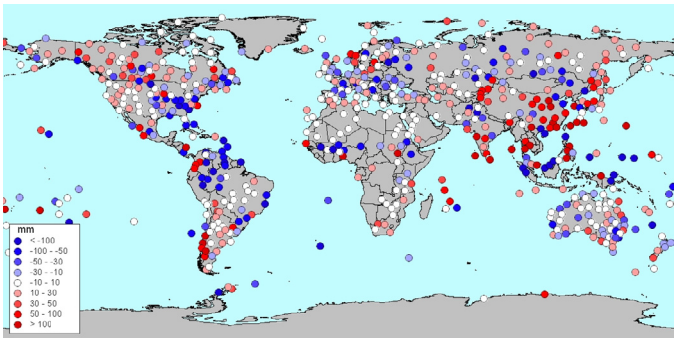
c) difference between IFS calculations and measurements for January



d) difference between IFS calculations and measurements for July



e) difference between GEM calculations and measurements for January



f) difference between GEM calculations and measurements for July

***Figure 5.7**. Comparison of monthly precipitation amount calculated by GEM and IFS for January and July of 2001 with measurements of GSN network*

5.4.2 Air Temperature

Air temperature calculated by the meteorological models IFS and GEM was compared with the data of re-analyses and measurements. The comparison of spatial distribution of annual mean air temperature for the pressure level of 925 hPa is shown in Figure 5.8. The distribution of air temperature is rather similar for all estimates. The GEM model provided slightly lower values in areas with close to maximum temperatures and somewhat overestimated minimal values of temperatures. GEM temperatures over the Arctic and Antarctic regions were also higher than those given by re-analyses. The spatial distribution of annual mean temperature obtained by IFS

model is rather close to the data of ERA-40 re-analysis, which is not surprising, as ERA-40 uses an old version of the IFS as basis for the re-analysis..

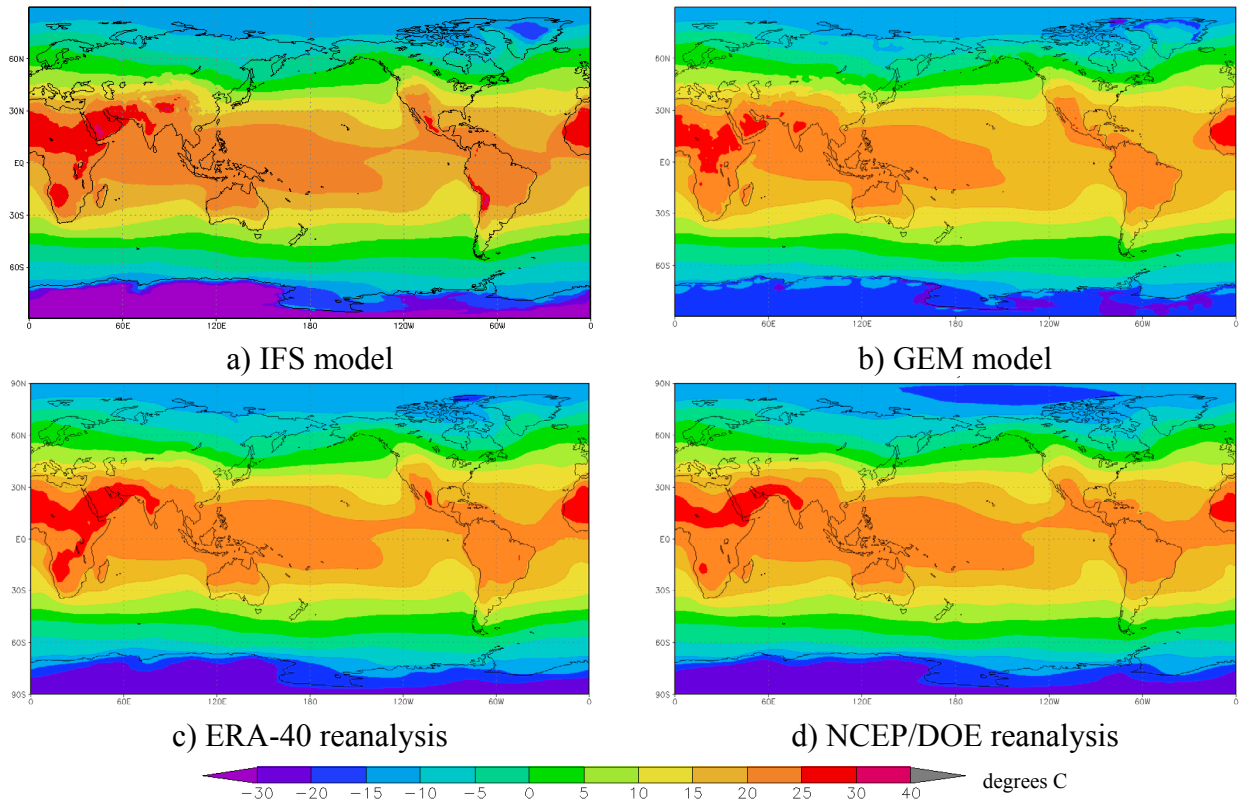
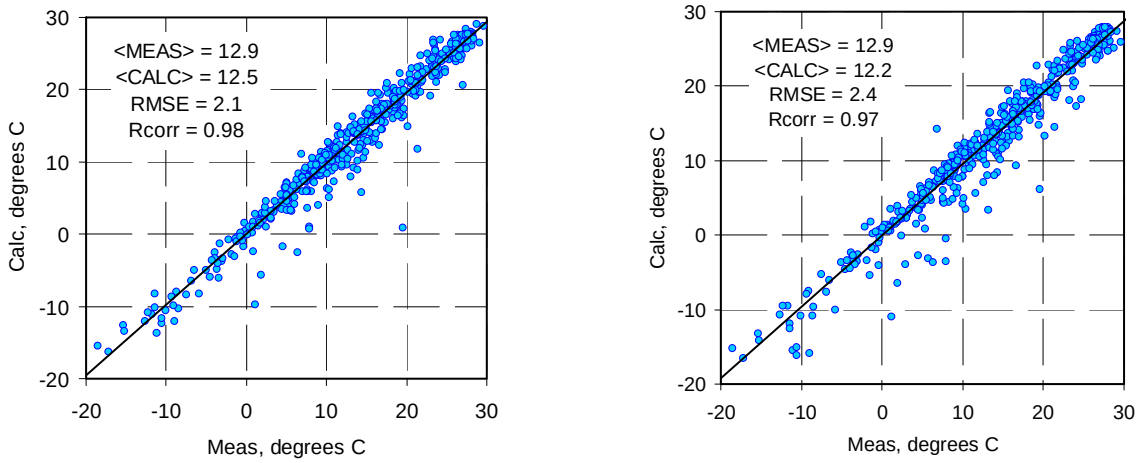


Figure 5.8. *Spatial distribution of annual mean air temperature at 925 hPa for 2001 calculated by ECMWF IFS (a) and GEM (b) models, and obtained from re-analyses ERA-40 (c) and NCEP/DOE (d).*

Calculated surface air temperature values were compared with measurements of RAOB database for 2001. The scatter plots in Figure 5.9 present the annual mean air temperature at 2 m calculated by IFS and GEM models versus measurements of RAOB database. It can be seen that modeling results well agree with observed values of surface air temperature. The IFS and GEM models provided slightly lower annual mean temperature compared to the observed one. At the same time for some of the sites the deviation of calculated air temperature from measured exceeds 10° C.

The spatial correlation between computed and measured annual surface air temperatures is slightly higher for IFS (R_{corr} is equal to 0.98 for IFS) than for GEM (R_{corr} is 0.97 for the GEM).

Results of the comparison of calculated annual mean air temperature vertical profiles for 2001 with measurements of RAOB sites are given in Figure 5.10. Upper left diagram presents the averaged vertical distribution of temperature measured at the RAOB sites and calculated by GEM and IFS models. Additional diagrams in this figure contain the vertical distribution of statistical parameters BIAS, RMSE, and R_{corr} calculated on the basis of model results and measurements. It can be seen that the deviations of mean over time between calculated and measured values of temperature for most of layers do not exceed 1°C. IFS modeling are characterized by somewhat higher positive bias in comparison to the GEM model results. Other statistical parameters like RMSE and R_{corr} obtained on the basis of the GEM and IFS model results are rather close, slightly better for IFS. The values of $RMSE_{time}$ and $RMSE_{space}$ are in the range of 1°-3 C and correlation coefficients $R_{corr\text{-}time}$ and $R_{corr\text{-}space}$ are higher than 0.9 for both of the models.



a) IFS model b) GEM model

Figure 5.9. *Scatter plots of measured annual mean surface air temperature at 2m height obtained from RAOB database versus calculated by IFS (a) and GEM (b) models for 2001*

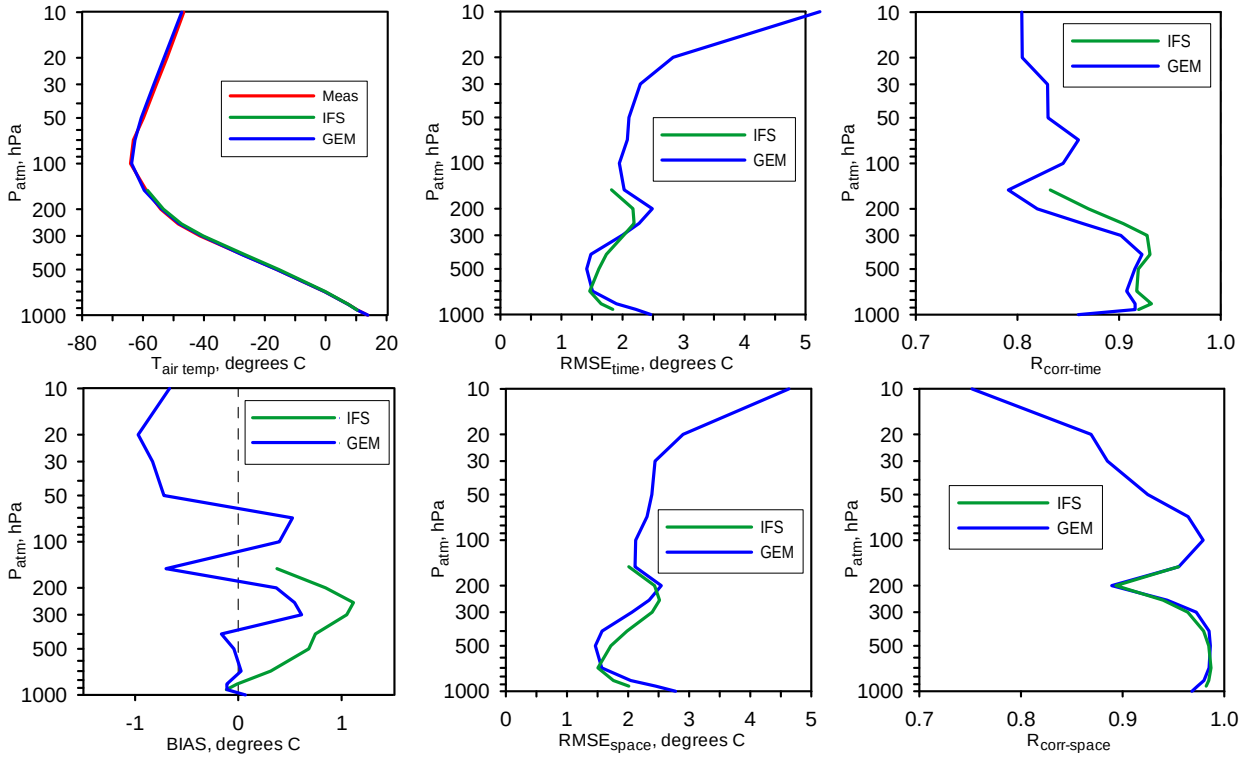
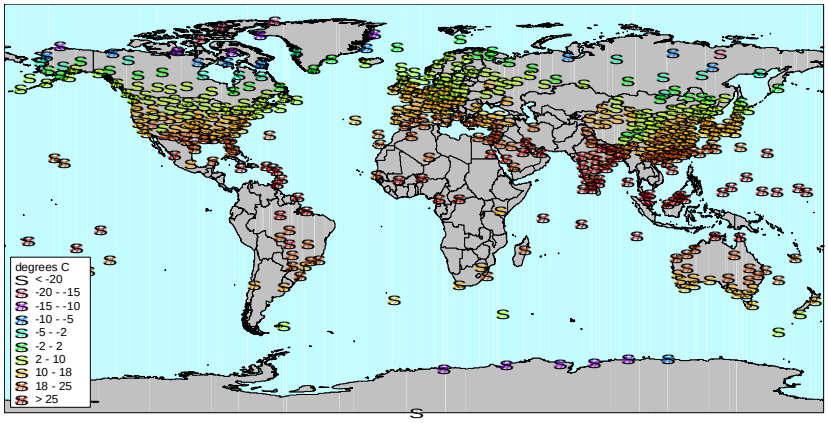


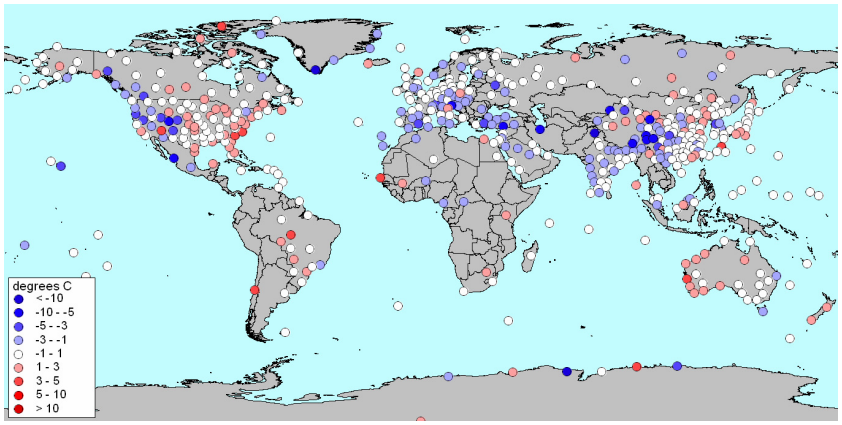
Figure 5.10. *Comparison of annual mean air temperature profiles for 2001 calculated by GEM and IFS model with measurements of RAOB database.*

Results of the comparison of calculated and measured annual mean surface air temperature are also shown in Figure 5.11. The figure presents the geographical distribution of measured air temperature values (Figure 5.11a) along with the differences between the computed and observed temperatures for IFS (Figure 5.11b) and GEM (Figure 5.11c) models. There is a reasonable agreement of modeling results of both models with measurements for the European

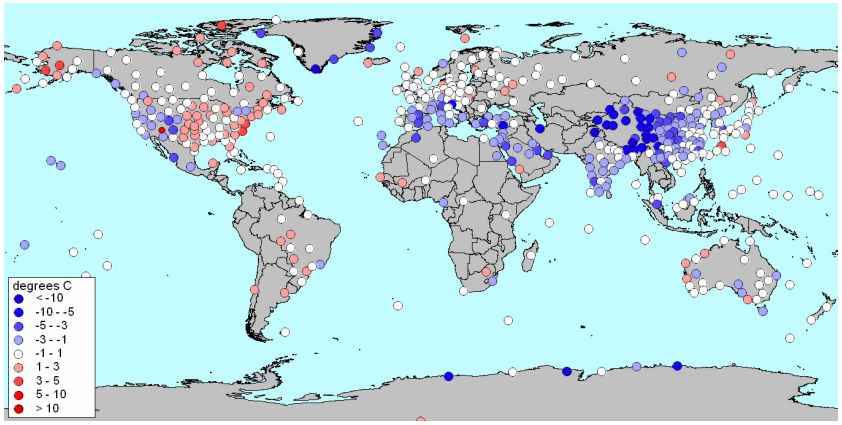
and North American regions. At high altitude sites, both models tend to provide lower surface air temperatures (by -10° - -5° C) compared to measurements. Some underestimation of observed temperature by the models is also noted for the sites in coastal areas of the Mediterranean Sea and Arabian Peninsula (by -5° - -1° C). This underestimation is more significant for the GEM. For the eastern part of the USA, GEM overestimates (by 1°-5° C) of measured surface air temperature.



a) surface air temperature measurements of RAOB sites



b) difference between IFS calculations and measurements of temperature



c) difference between GEM calculations and measurements of temperature

Figure 5.11. *Comparison of annual mean surface air temperatures calculated by the GEM and IFS models for 2001 with measurements of RAOB sites*

5.4.3 Wind Speed

For wind speed, the comparison so far includes only an evaluation of annual mean vertical profiles produced by GEM and IFS models with measurements of RAOB sites for 2001. Results are shown in Figure 5.12. Similar to the comparison of temperature profiles, the first diagram presents the averaged vertical distribution of wind speed measured at the RAOB sites and calculated by GEM and IFS models. Additional diagrams of the figure contain the vertical distributions of statistical parameters BIAS, RMSE, and R_{corr} calculated on the basis of model results and measurements.

Significant correlation between observed and calculated values of wind speed can be noted for the both models. It can be seen that IFS and GEM model results are rather close to measurements up to the height of 300 hPa. Some underestimation of the wind speed at the heights 200-300 hPa and above 50 hPa can be noted for the GEM model results. It is also can be seen in the BIAS diagram where the maximum deviation for the GEM is about 1.5 m/s at 200 hPa and for the IFS is about 1 m/s at 300 hPa. The correlation coefficients for the both models have minimum values at the lowest level (0.6-0.7). For other levels the correlation between observed and calculated wind speed is higher (0.7-0.9).

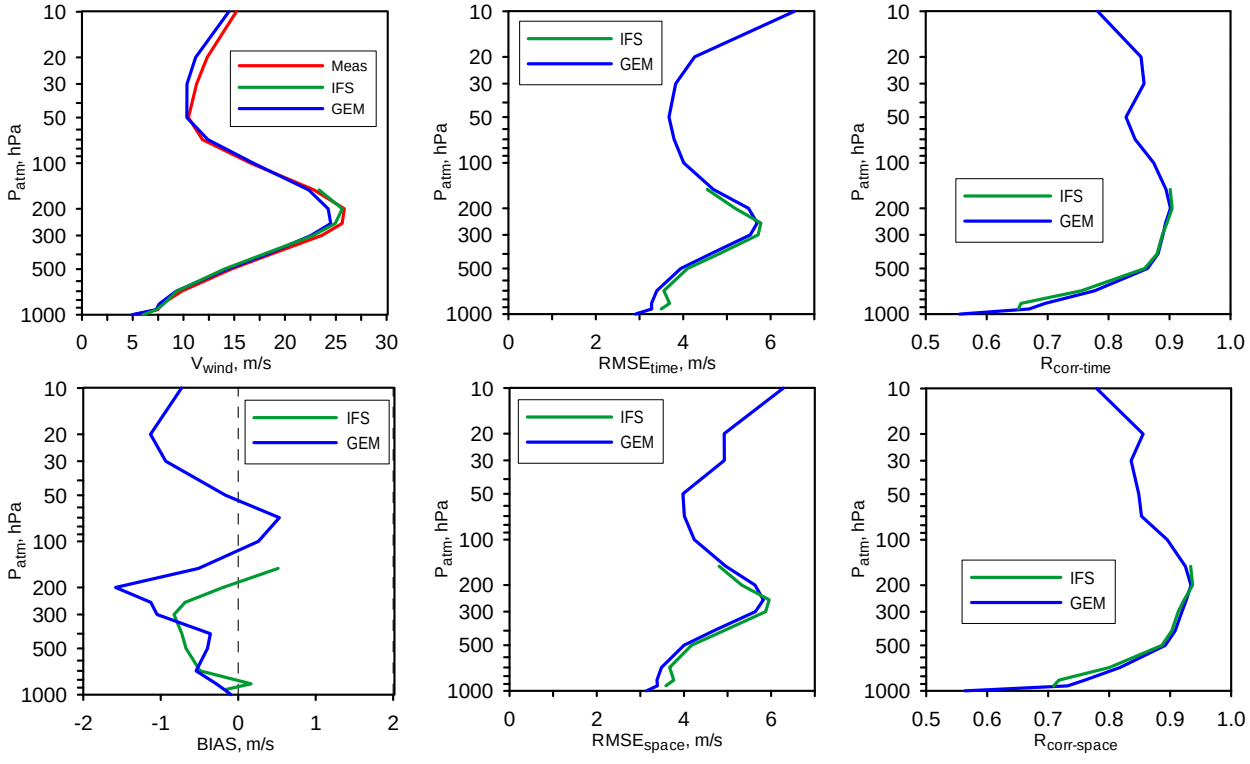


***Figure 5.18**. Comparison of annual mean wind speed profiles for 2001 calculated by GEM and IFS model with measurements of RAOB database.*

5.5 Spatial variability of validated vertical profiles

Along with the analysis of measurements and modeling results for the whole set of monitoring sites an attempt to perform regional comparison was carried out. This was to characterize the performance of the IFS and GEM models in different regions of the world. For this purpose the vertical profiles of selected meteorological parameters calculated by GEM and IFS were aggregated and compared with upper-air measurements of RAOB database for the seven conventionally defined regions, in particular, North America, South America, Europe, Russia, Africa, Asia, and Australia.

For each of these regions vertical profiles of annual mean air temperature and wind speed magnitude were prepared. Averaging was performed for the sites located in the particular region. Along with that the statistical parameters describing deviations (BIAS), relative mean square error (RMSE), and temporal correlation (R_{corr}) between time-series of selected parameters have been calculated. Results of the comparison are shown in Figure 5.13 for annual mean air temperature and in Figure 5.14 for the annual mean wind speed magnitude.

It can be seen that both models reasonably reproduce vertical profiles of air temperature for 2001 for all of the selected regions. Absolute difference between measured and calculated annual mean air temperature described by the BIAS parameter does not exceed 2°C excluding the level 10 hPa. The value of the root mean square error ($RMSE_{time}$) is in the range 1°-3°C. For both models the highest correlation with observations among the selected regions takes place for the North America and Europe ($R_{corr\text{-}time} > 0.9$). The lowest correlation can be noted for the South America region.

The IFS model in most of the cases slightly overestimates air temperature observations, especially at higher levels. At the same time, near the surface the differences between calculated and observed values are higher for the GEM model, in particular, for the North America and Africa.

Comparison of vertical profiles of annual mean wind speed for 2001 show that the models generally reproduce well the observed values. There is some tendency to underestimate measurements which is more significant at higher levels. The most significant values of BIAS (about -4 m/s) were obtained for the GEM model in case of South America and Australia at the levels 300 and 200 hPa. On average GEM model results are characterized by more essential negative bias than the results of the IFS model.

The values of RMSE are in the range 2-6 m/s and tend to increase at higher levels. For both models the highest correlation with observations among the selected regions takes place in free troposphere for the North America and Europe ($R_{corr\text{-}time} > 0.8$). At lower levels the values of correlation decrease to approximately 0.6 with the lowest correlation in case of the South America region (about 0.4).

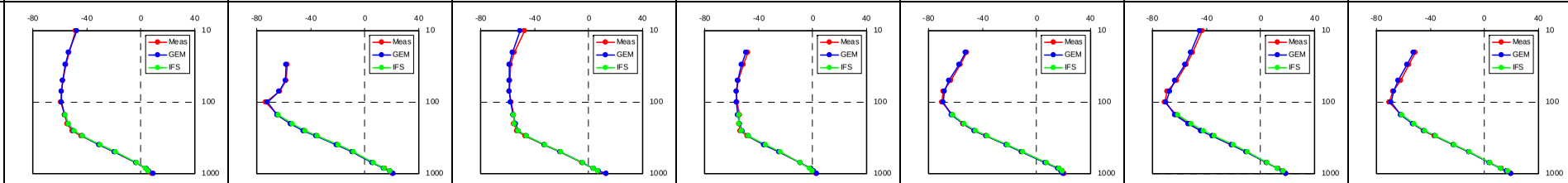
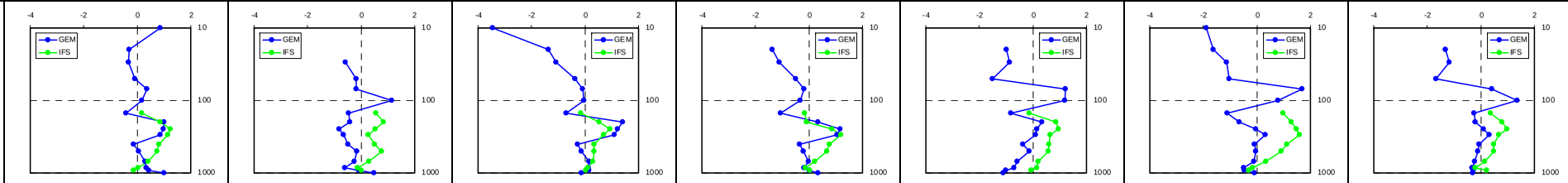
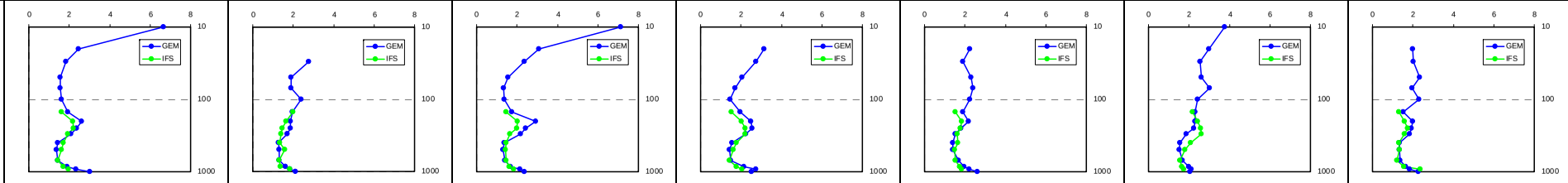
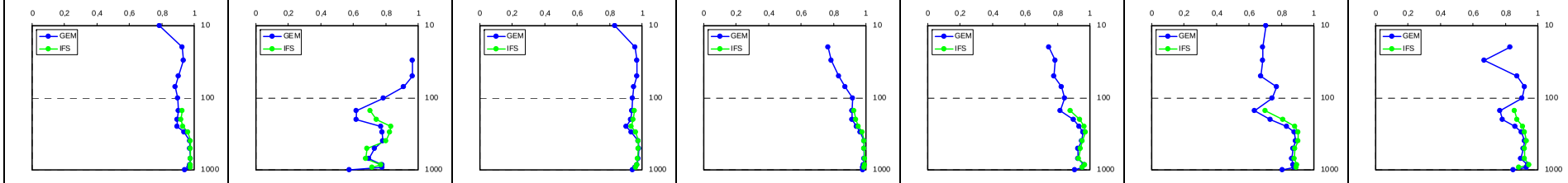
	North America	South America	Europe	Russia	Africa	Asia	Australia
Measured and calculated values, degrees C							
BIAS, degrees C							
$RMSE_{time}$, degrees C							
$R_{corr\text{-}time}$							

Figure 5.13. *Comparison of annual mean air temperature (°C) calculated by GEM and IFS models for 2001 with measurements of RAOB database. Measured and calculated values as long as statistical parameters BIAS, $RMSE_{time}$, $R_{corr\text{-}time}$ are presented for the selected geographical regions. The vertical axis of the diagrams represents the atmospheric pressure in hPa (logarithmic scale).*

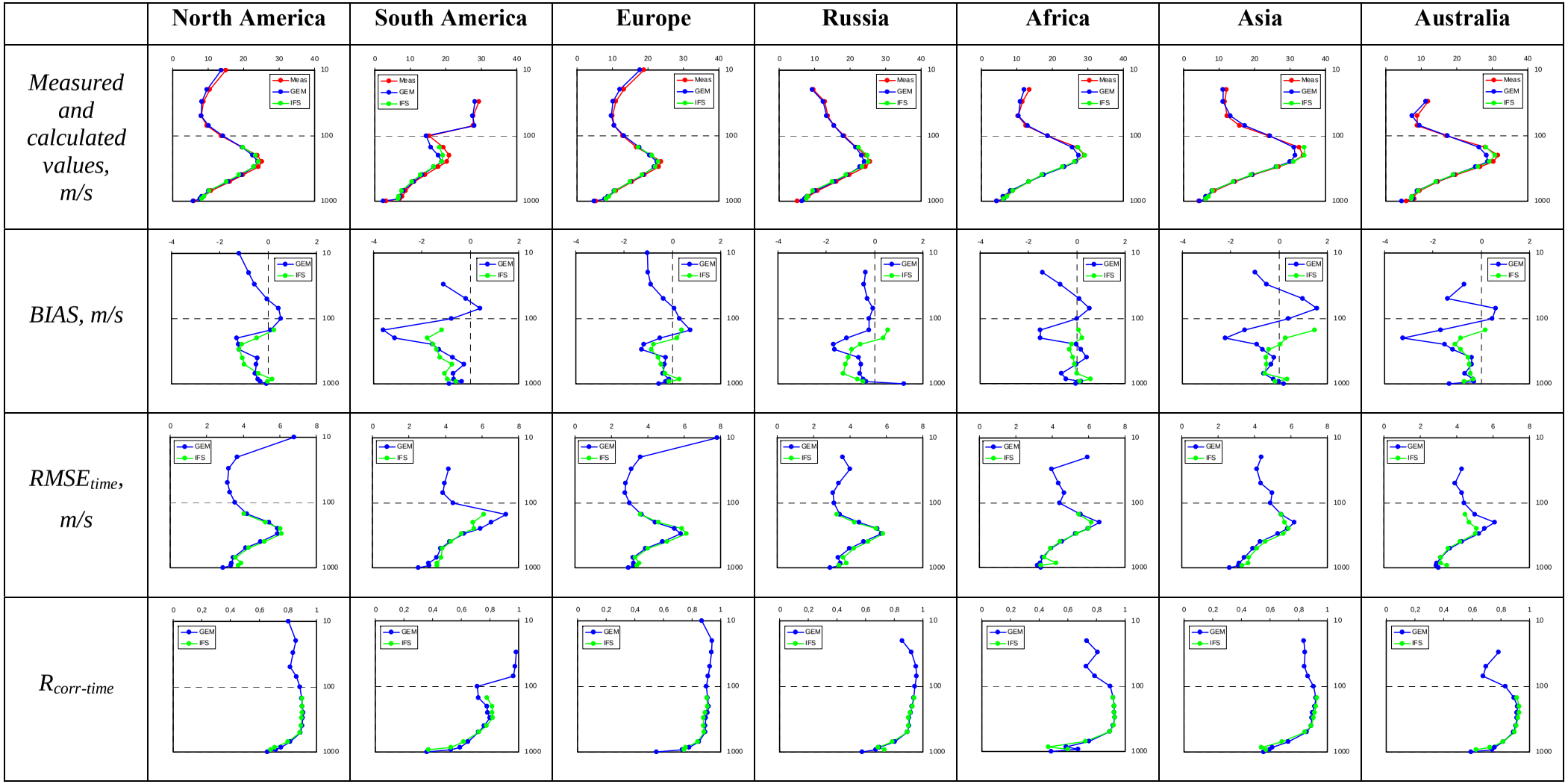


Figure 5.14. *Comparison of annual mean magnitude of wind speed (m/s) calculated by GEM and IFS models for 2001 with measurements of RAOB database. Measured and calculated values as long as statistical parameters BIAS, $RMSE_{time}$, $R_{corr\text{-}time}$ are presented for the selected geographical regions. The vertical axis of the diagrams represents the atmospheric pressure in hPa (logarithmic scale).*

5.6 Concluding remarks

An initial comparison of results from two meteorological drivers, the IFS from ECMWF and the GEM from the Canadian Meteorological Center was presented here. The version used for the IFS was based on cycle 29r2. For the GEM model, different configurations of the system were evaluated and the final one used was the one providing better agreement with meteorological measurements.

Test simulations were carried out for 2001 and the results were compared with meteorological measurements (from ground stations, radiosondes and satellites) and with data from re-analyses, namely ERA-40 and NCEP/DOE. At this stage, a limited number of meteorological fields were used for the comparison. These were: air temperature, wind speed components, and precipitation amounts. For air temperature and wind speed both surface values and vertical profile data on pressure levels were compared. The level of agreement between the measurements and modeling results was analyzed for annual and monthly data and averaged over different regions. The characterization was done based on statistical parameters.

For precipitation amounts, the comparison of GEM and IFS simulation results with the measurements of GSN network for 2001 revealed reasonable agreement between the calculated and observed values and high correlation between them. Spatial distributions of annual precipitation amount calculated by GEM and IFS have rather similar pattern and agree well with the precipitation amounts derived from the GPCP project, that combine ground level data with satellite information. In comparison with re-analyses data, the results from GEM and IFS provided somewhat lower precipitation amounts in equatorial regions, especially over sea areas.

The comparison of spatial distribution of GEM and IFS annual mean air temperature and data from re-analyses was carried out for the pressure level of 925 hPa. GEM model provided slightly lower values in areas with close to maximum temperatures and somewhat overestimated minimum temperatures. Comparison with surface air temperature measurements has shown that modeling results of GEM and IFS agree well with the observed values and that there is high spatial correlation between computed and measured temperatures. At the same time, the models tend to provide lower surface air temperatures compared to measurements at high altitude areas of China and North America. Differences between the calculated vertical profiles of air temperature and upper-air measurements of vertical profiles do not exceed 1°C. The IFS model results are characterized by somewhat higher positive bias in comparison to the GEM model, while RMS errors are generally lower for the IFS model results.

The comparison of vertical profiles of calculated annual mean magnitudes of wind speed produced by the GEM and IFS models with measurements at RAOB sites has demonstrated significant correlation between the observed and calculated values of both models. The IFS and GEM model results are rather close to measurements up to the height of 300 hPa. Some underestimation of the wind speed at the heights 200-300 hPa and above 50 hPa can be noted for the GEM model results.

Evaluation of the agreement between observed and calculated vertical profiles of air temperature and wind speed was carried out for the seven conventionally defined regions, in particular, North America, South America, Europe, Russia, Africa, Asia, and Australia. Both models reasonably reproduced vertical profiles of air temperature and wind speed for 2001.

The differences for air temperature do not exceed 2°C in average for most of the regions. In case of wind speed, there is some tendency to underestimate measurements which is more significant at higher levels. The highest level of agreement with measurements was obtained for the North American and European regions while the lowest was obtained for the South American region.

This comparison of the meteorological input data to be used in the common EMEP modeling framework will continue in 2009 and further results will be presented then. We intend to analyze in more detail the temporal accuracy of the data and extend the comparison to other parameters of relevance for atmospheric dispersion modeling, such as cloud cover and turbulent surface fluxes. It is also planned to include in the comparison two other available global scale meteorological drivers, namely, the PUM from the UK Met Office, and WRF developed by US EPA.

5.7 References

Blackadar A. K. (1962) The Vertical Distribution of Wind and Turbulent Exchange in a Neutral Atmosphere, Journal of Geophysical Research, Vol. 67, No. 8, 1962, pp. 3095-3102.

Bougeault P. and Lacarrère P. (1989) Parameterization of orography-induced turbulence in a meso-beta-scale model. Mon. Wea. Rev., 117, pp. 1872-1890.

Côté J., Desmarais J.-G., Gravel S., Méthot A., Patoine A., Roch M., and Staniforth A. (1998b) The operational CMC-MRB Global Environmental Multiscale (GEM) model: Part II - Results, Mon. Wea. Rev. 126, 1397-1418.

Côté J., Gravel S., Méthot A., Patoine A., Roch M., and Staniforth A. (1998a) The operational CMC-MRB Global Environmental Multiscale (GEM) model: Part I - Design considerations and formulation, Mon. Wea. Rev. 126, 1373-1395.

Fillion L., Mitchell H. L., Ritchie H., and Staniforth A. (1995): The impact of a digital filter finalization technique in a global data assimilation system. *Tellus,* **47A,** 304–323.

Garand L. and Mailhot J. (1990) The influence of infrared radiation on numerical weather forecasts. Proceedings of the Seventh Conference on Atmospheric Radiation, July 23-271990, San Francisco, U.S.A., *Amer. Meteor. Soc.*, pp. J146-J151.

Geleyn J.-F. (1985) On a Simple, Parameter-Free Partition between Moistening and Precipitation in the Kuo Scheme., Mon. Wea. Rev., 113, pp. 405-407.

Huffman G.J., Adler R.F., Morrissey M.M., Bolvin D.T., Curtis S., Joyce R., McGavock B. and Susskind J. (2001). Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations. J. of Hydrometeorology, vol 2., pp. 36 – 50.

Kain J.S. and Fritsch J.M. (1993) Convective parameterization for mesoscale models: The Kain- Fritsch scheme. The representation of cumulus convection in numerical models. Meteor. Monogr., No. 24, Amer. Meteor. Soc., pp. 165-170.

Kanamitsu M., Ebisuzaki W., Woolen J., Yang S-K., Hnilo J.J., Fiorino M. and Potter G.L. [2002]. NCEP-DOE AMIP-II reanalysis (R-2). Bulletin of American Meteorological Society, 11, pp. 1631 – 1643

Mailhot J., Bélair S., Benoit R., Bilodeau B., Delage Y, Fillion L., Garand L., Girard C.and Tremblay A. (1998) Scientific Description of RPN Physics Library. Version 3.6. Recherche en Prévision Numérique,188 pp.

Mailhot J., Sarrazin R., Bilodeau B., Brunet N. and Pellerin G. (1997) Development of the 35-km version of the Canadian regional forecast system. Atmos-ocean, 35, pp. 1-28

McFarlane N.A. (1987) The effect of orographically excited gravity-wave drag on the general circulation of the lower stratosphere and troposphere. J. Atmos. Sci., 44, pp. 1775-1800.

Moorthi S. and Suarez M. (1992) Relaxed Arakawa-Schubert: A Parameterization of Moist Convection for General Circulation Models. Mon. Wea. Rev., 120, pp. 978-1002.

Noilhan J. and Planton A. S. (1989) A simple parameterization of land surface processes for meteorological models, Mon. Wea. Rev., 117, pp. 536-549.

Rudolf B., Hauschild H., Rueth W. and Schneider U., 1994. Terrestrial precipitation analysis: operational method and required density of point measurements. NATO ASI Series, Vol. I 26, Global Precipitation and Climate Change. Ed. by F. Desalmand. Springer-Verlag Berlin Heidelberg, pp. 173 – 186.

Sundqvist H., Berge E. and Kristjansson J.E. (1989) Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. Mon Wea Rev, 117, pp. 1641--1657.

Wagneur N. (1991) Une évaluation des schémas de type Kuo pour le paramétrage de la convection, Msc Thesis, UQAM, 76 pp.

Yeh K.-S., Côté J., Gravel S., Méthot A., Patoine A., Roch M., and Staniforth A. (2002) The CMC–MRB Global Environmental Multiscale (GEM) Model. Part III: Nonhydrostatic Formulation. Monthly Weather Review: Vol. 130, No. 2, pp. 339–356.

6. Comparison of atmospheric dispersion by the EMEP global models

Jan Eiof Jonson, Oleg Travnikov, and Peter Wind

Within the framework of the Task Force on Hemispheric Transport of Air Pollution, a simple tracer experiment has been set-up in order to investigate differences between models caused by their description of atmospheric dispersion, involving both advection and diffusion. The recently developed global models at MSC-E and MSC-W have participated in the TFHTAP tracer experiments, a two of many models.

Since the focus of the two global MSC models is on different pollutants, a good way to compare them is through the study of their respective transport characteristics. This chapter presents an initial comparison of the results from both models, based on their results from the tracer experiment at TFHTAP.

6.1 Global models at MSC-W and MSC-E

Only a short description of the models is included here. More detailed model descriptions of the models are found in the references cited in the model descriptions below.

6.1.1 The Unified global model at MSC-W

The application and validation with measurements of the Unified EMEP model in global scale developed at MSC-W was already described in Jonson et al. (2007). The EMEP Unified model is routinely tested. The first in-depth description of the EMEP Unified model was provided in Simpson et al. (2003). The model has been coded flexible so that the domain can be specified in scripts along with the links to the input data files required to run the model. Thus the same model version can be used in local, regional and global applications. In the model experiment reported here, we have used the rv3_0 release, now available as open source through http://www.emep.int/OpenSource. A detailed description of the model is also available from this site. In the current model version the horizontal grid is regularly spaced latitude/longitude with 1x1 degree resolution. In Jonson et al. (2007) model results from the global application were validated against measurements. It was shown that the model performed well for most species and regions considered. However, the model underestimated ozone in the free troposphere, in particular in the summer months. This is most likely caused by lack of convection in the model. The inclusion of a convection scheme in the EMEP Unified model is currently under progress.

6.1.2 The GLEMOS global model at MSC-E

The new global model developed at MSC-E is named the Global EMEP Multi-Media Modelling System (GLEMOS). In spite of the fact that the new global model is largely based on parameterisations of the well developed and tested hemispheric model (MSCE-HM-Hem), a number of principal modifications have been done to follow the general requirements of a global modelling system. First of all, the vertical model domain has been extended up to 10 hPa (ca. 30 km). Extension of the vertical coverage is required for modelling of atmospheric dispersion of long-lived substances on a global scale in order to avoid the need of setting boundary conditions at the upper boundary as well as to take into account possible stratosphere-troposphere exchange. The current vertical structure of the model domain is shown in Figure 6.1. It consists of 20 irregular terrain-following sigma layers. Among them 10 layers cover the lowest 5 km of the troposphere and height of the lowest layer is about 75 m.

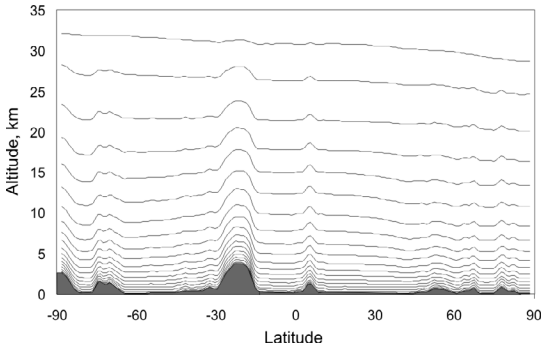
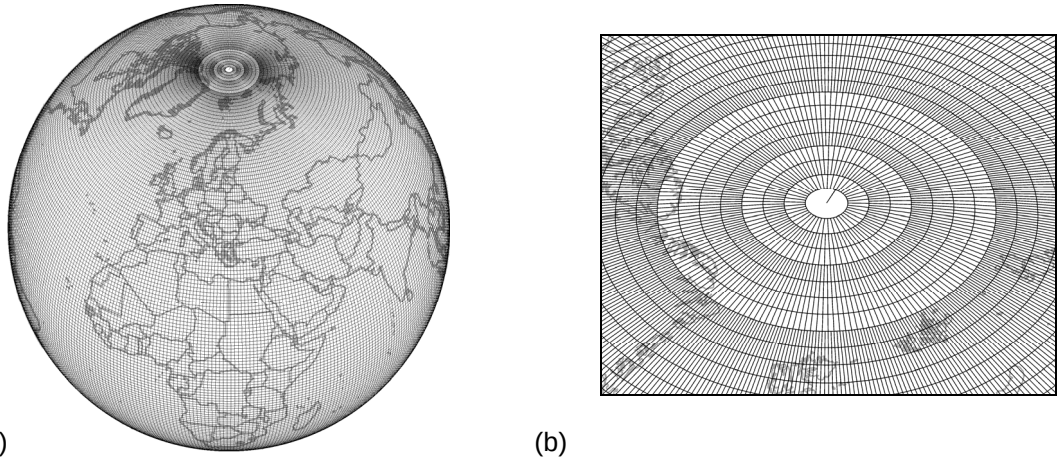


Figure 6.1 *Vertical structure of the model grid*

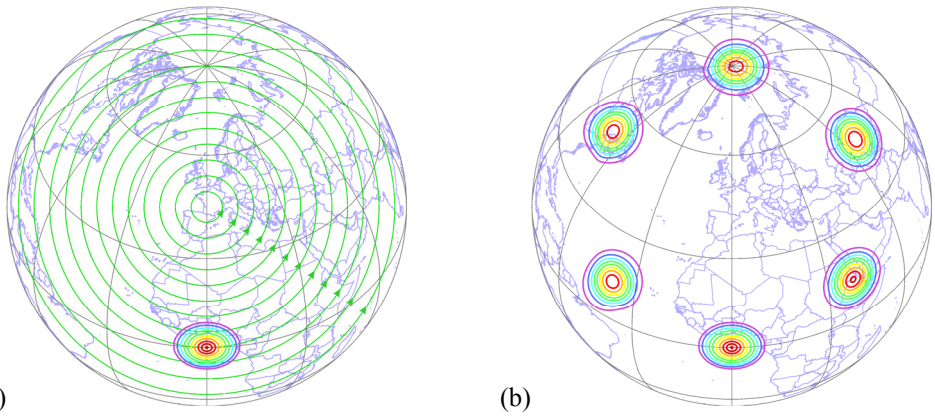
Secondly, the spatial resolution of the model grid was refined down to 1°×1°. The horizontal configuration of the global model grid is shown in Figure 6.2a. It should be noted that refining the grid in angular co-ordinates resulted in extremely small spatial steps in the zonal direction near the poles (down to a few kilometers) which, in turn, significantly complicated the time integration procedure with an explicit scheme. To solve the problem the model grid was modified in the vicinity of the poles. In particular, starting from the latitude 82° northward and southward, respectively, the gridcells were aggregated in couples to keep the zonal step of the grid not smaller than 15 km. Besides, the areas in the immediate vicinity of the poles (within the radius 1 km) are presented by single circle gridcells (Figure 6.2b). This allowed avoiding extra small integration time steps and the necessity to involve meso-scale physics for the description of atmospheric processes at so fine resolution.



(a) (b)

Figure 6.2 *Configuration of the horizontal 1°×1° model grid for (a) the whole model domain and (b) area around the pole*

This sort of modifications of the model grid requires thorough testing of the model advection scheme performance, particularly, in the vicinity of the poles. To test the advection scheme we performed a numerical experiment simulating transport of a tracer with a cone-shaped initial distribution in an artificial rotational wind flow. Conditions and results of the experiment are shown in Figure 6.3. As seen, a tracer cone-shaped plume initially located at the equator moves over the circle trajectory coming through the pole. No significant distortions of the initial distribution were observed during the tracer movement through the irregular grid near the pole. Besides, the level of numerical diffusion of the advection scheme observed after the whole cycle of rotation was not significant. Thus, the advection scheme demonstrated acceptable performance for the irregular grid applied in the model.



(a) (b)

Figure 6.3 *Rotational flow numerical experiment of the model advection scheme: (a) streamlines of the artificial wind and the initial cone-shaped distribution of a tracer and (b)results of a tracer transport in the rotational flow).*

Parameterisations of major physical and chemical processes were inherited form the hemispheric MSCE-HM-Hem and MSCE-POP-Hem models. A list of major atmospheric parameterisations applied in the model is presented below:

- ***Advection:*** Second-order Bott scheme modified for spherical geometry
- ***Diffusion:*** Second-order implicit finite-difference scheme
- ***Mass conservation:*** Calculation of vertical velocities using analytical inversion of the Bott scheme
- ***Dry deposition:*** Resistance analogy approach, size-segregated deposition velocities of particle
- ***Wet deposition:*** In-cloud and below-cloud scavenging, empirically derived approach
- ***Mercury atmospheric chemistry***
- ***POPs gas-particles partitioning, gaseous exchange with the surface etc.***

Detailed description of the applied parameterisations can be found in (Gusev et al., 2005). At the moment only atmospheric transport module was adapted for the global model. Inclusion of other media description (soil, seawater, vegetation) is planed for near future.

In the tracer experiment the MSC-E global model was driven by meteorological data pre-processed by the GEM meteorological model. Main characteristics of the driver and evaluation of the generated data are presented in Chapter 5 of the current report. Besides, the MODIS dataset was used for the land cover information.

6.2 The TP1x experiment under TFHTAP

In order to contribute to the revision of the Gothenburg protocol a model experiment (below denoted the SR experiment) was set up under the TFHTAP to give a first assessment of the source receptor relationships between the main source regions in the Northern Hemisphere, roughly representing North America (NA, 125W to 60W and 15N to55N), Europe (EU, 10W to 50E and 25N to 65N), East Asia (EA, 95E to 160E and 15N-50N) and South Asia (SA, 50E to 95E and 5N to 35N). A description of the SR experiment and key findings are published in the interim report from TFHTAP (UNECE, 2007). Furthermore several papers based on the comparison of these source receptor calculations are in preparation (Shindell et al. 2008; Sanderson et al. 2008; Fiore et al. 2008). In the SR experiment substantial differences between the models were identified. The differences were difficult to explain because of the complexity of the models. In order to highlight differences between the models caused purely by advection, a simple tracer experiment (TP1) was set up. This tracer experiment has later been expanded with more species. This new tracer experiment is denoted TP1x.

The global models at MSC-W and MSC-E are both taking part in this model inter-comparison. The global models at MSC-W and MSC-E have different scopes and thus the parameterizations and chemistry differ between the two models. As no chemistry is involved, the TP1x experiment allows a comparison of the advection in the models and thus the two models can be compared on equal terms.

All tracers included in TP1x are listed in Table 6.1, but not all models have included all the tracers in their model setup. The first set of tracers are similar to CO and are designated as COfromEA, COfromEU, COfromNA, COfromSA, depending on their area of origin. All models taking part in the TP1x experiment use the same emissions. The emitted species are then removed assuming simple exponential decay rates listed in Table 6.1. Emissions are obtained through: ftp://ftp.retro.enes.org/pub/other/htap. Starting with zero concentration fields, the models are required to run for one year to spin up the models. The experiment is described in more detail in the TFHTAP wiki pages available at: http://aqm.jrc.it/HTAP/ (open, but require registration).

In the analysis of the TFHTAP SR data described above it was seen that differences between the models are very large when it comes to wet deposition, and this is also the main reason for the large differences between models in the advection of soluble species to the Arctic (Shindell et al. 2008). In TP1x, extra tracers with wet deposition (sink added) were therefore included in the inter-comparison study. These tracers are denoted SAdirect, SAfromNA, SAfromEU, SAfromEA, SAfromSA in Table 6.1. These tracers are similar to the CO like tracer, but with wet deposition included as an additional sink. It is believed that the effects of the different parameterizations used for wet deposition in the models will be easier to identify with a simple tracer experiment, rather than in a full chemistry model. These tracers are not included in the analysis here as the SA experiments so far only been run with the MSC-W global model and not with the MSC-E model.

Table 6.1 *List of tracers that should be submitted to TFHTAP for the TP1x tracer experiment. The species from EA,EU,NA,SA refers to the regions described in the text. The sink column refers to the constant uniform exponential decay in days supplemented by wet deposition for the SA species.*

Tracer	**Emission/Source**	**Sink**	**Tracer**	**Emission/source**	**Sink**
CO direct	1019.65 Tg CO	50 days	AVOC	79.2 TG as C	7 days
COfromEA	137.07 TG CO	As COdirect	BVOC	725.7 Tg as C	1 day
COfromEU	81.07 Tg CO	As COdirect	COfromAVOC	τ AVOC*0.7	As COdirect
COfromNA	94.17 Tg CO	As COdirect	COfromBVOC	τ AVOC*0.4	As COdirect
COfromSA	121.33 Tg CO	As COdirect	COfromCH_4 (CH_4 1760ppb)	τ_{CH4} 8.5 yr^{-1} * 0.86	As COdirect
SA direct	As CO direct	50 days + wet dep. As SO_4^{2-}	SVOC	As AVOC	5.6 days
SAfromEA	As CO fromEA	As SAdirect	MVOC	As AVOC	13 days
SAfromEU	As CO fromEU	Aa SAdirect	LVOC	As AVOC	64 days
SAfromNA	As COfromNA	As SAdirect			
SAfromSA	As COfromSA	As SAdirect			

6.3 Comparison of tracer dispersion by the EMEP/MSC global models

The global EMEP models from MSC-E and MSC-W have submitted data for the TP1x tracer experiment. Several of the species in TP1x are of particular interest for the models having submitted results for ozone, nitrogen and depositions of acidifying species in the SR calculations, and less useful for models focusing on Hg and POPs. For that natural reason only the MSC-W global model has performed the calculations for all the species listed in Table 6.1., and the results are not available from GLEMOS.

Here we will concentrate on model results for the CO-like tracer, called "COdirect" and the CO-like tracer from the 4 regions (CO from NA, CO from EU, CO from EA and CO from SA). Results from both models are available for these tracers. Both models have also submitted results for LVOC, MVOC and SVOC. An analysis of these tracers is not included here as it brings little added information to the comparison of the two models.

Including the models from MSC-W and MSC-E, altogether 19 models have submitted results for "COdirect". The focus below is on the MSC-W and MSC-E models. However, it is also useful to see how different models perform. Thus model results are also shown for other models and/or for a model ensemble.

Figure 6.4 shows the annually averaged contour plots of "COdirect" for the lowest model layer calculated with the both MSC models and also the ensemble mean for all 19 participating models. The results calculated with the MSC-Wand the MSC-E models are both very similar to the ensemble mean. High levels of "COdirect" are calculated in particular for parts of India, China and equatorial Africa. Elevated levels are also calculated for North America, Europe and South America. High levels of "COdirect" over Africa and South America are related to biomass burning.

Figure 6.5 shows calculated "COdirect" fields in the surface layer from all the 19 models as bar plots for the globe, the Northern and Southern hemispheres and also for the 4 selected source receptor regions described above. With a few exceptions, results for all the models are relatively similar. Again, both the MSC-W and the MSC-E models are close to the ensemble mean.

Figure 6.6 shows "COdirect" results zonally averaged for the MSC-W and MSC-E models. At two different level heights : near the surface and also at about 500 hPa. In the figure, results from the Geos-Chem model and the MozartGFDL models are also included. The models show similar patterns, but the MSC-E model is higher over the tropics. The Geos-Chem model is lower than the other models at northern mid latitudes.

At 500hPa, the different models generally show a bimodal distribution with high levels over the northern mid latitudes and above the equator. However, the MSC-W model is clearly lower over the equatorial region and higher over the northern mid latitudes as already pointed out above. In the MSC-W model convective processes are not yet included. As a result, the

lifting of mass in particular in the inter-tropical convergence zone is underestimated in the MSC-W model. It remains to be seen, when the parameterization of convection is implemented, in what way this may be linked to the high levels over the northern mid latitudes.

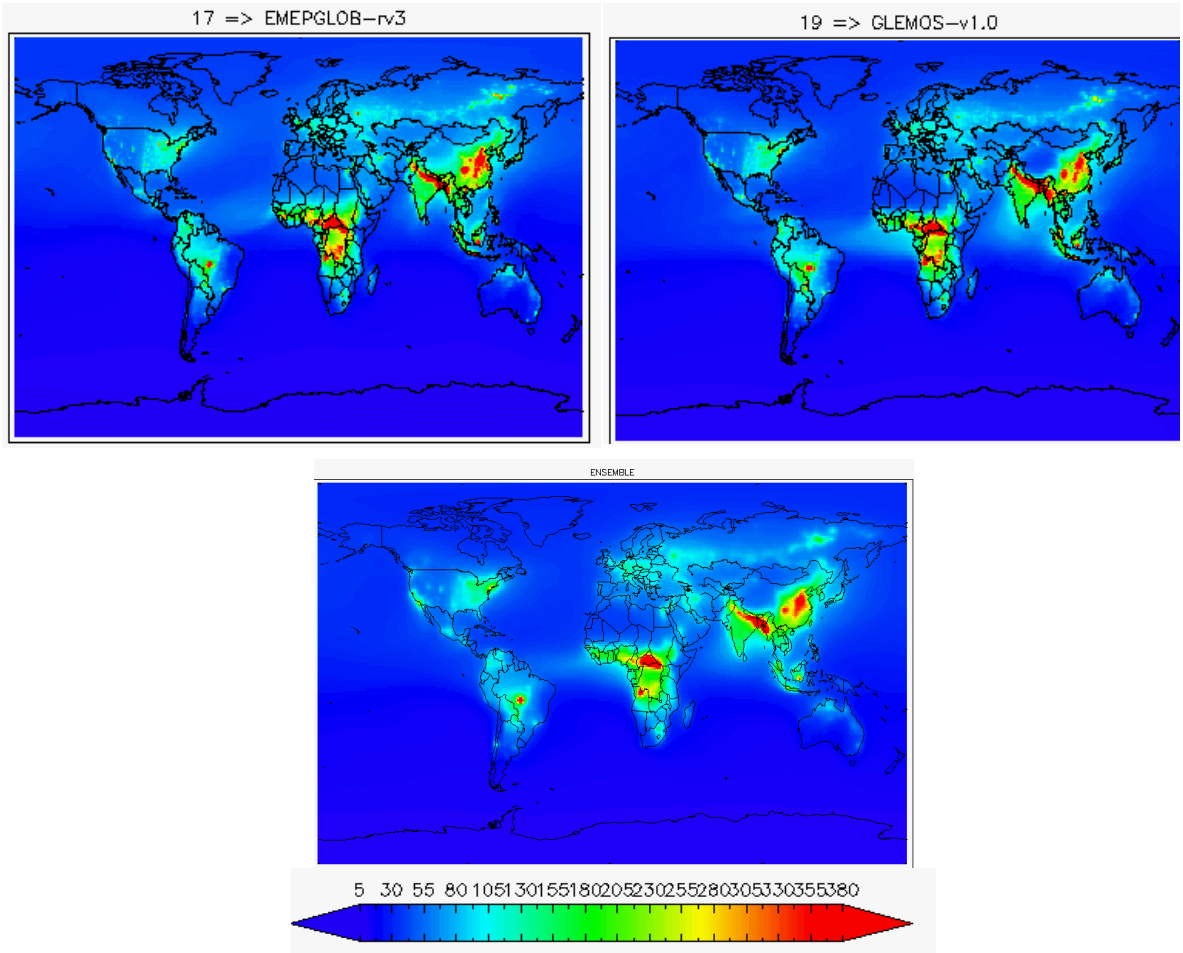


Figure 6.4 *The annual average of the tracer experiment "COdirect" in ppb in the lowest model layer. Upper left: MSC-W model (EMEPGLOB-rv3),Upper right: MSC-E model (GLEMOS-v1.0), C enter Bottom: ensemble middle for all 19 models.*

6.4 Comparisons in the free troposphere, including vertical profiles

Figure 6.7 shows the annually averaged concentrations of COdirect in the middle free troposphere. At this level concentrations are higher in the MSC-W model compared to the MSC-E model in the northern mid latitudes, and in particular so over the southwest corner of China. Over equatorial Africa, emissions are lifted higher in the MSC-E model and pollution

is subsequently advected eastwards over the Atlantic. In the southern hemisphere, concentrations are higher in the MSC-E model.

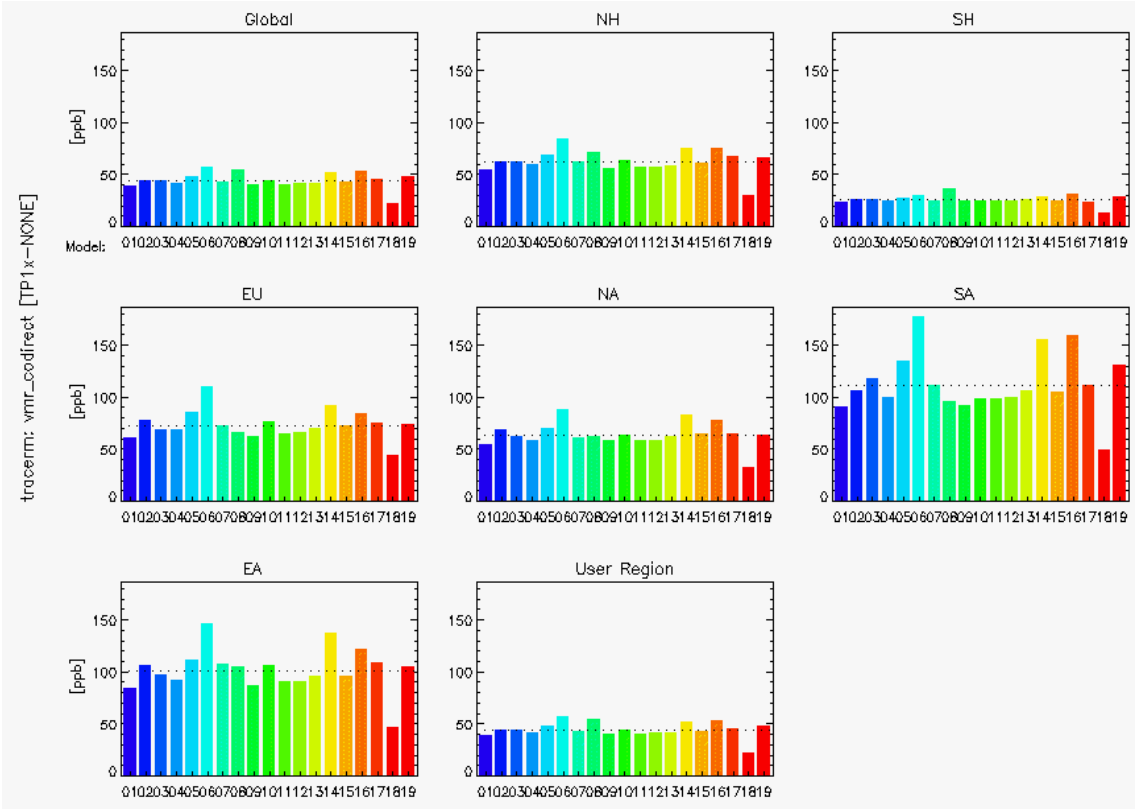


Figure 6.5 *Annually averaged COdirect in the lowest model layer in ppb averaged for the globe, for the Northern and Southern hemispheres and for the 4 source receptor regions described in the text. The MSC_W model is nr 17 (3'rd from the right). The MSC-E model is nr. 19 (far right).*

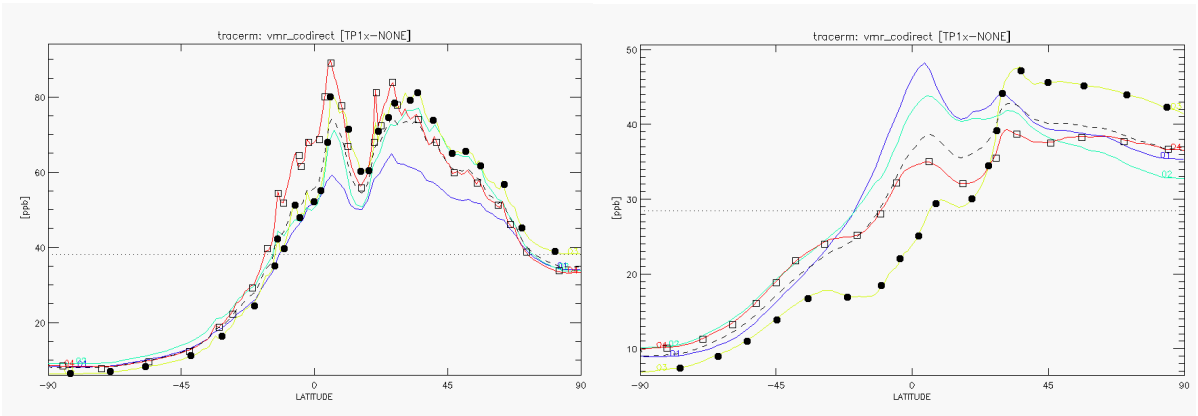


Figure 6.6 *COdirect in the lowest model layer (left) and at 509 hPa (right) zonally averaged for the Geos-Chem. Model (blue), the MozartGFDL model (turquoise), the MSC-W model (yellow with filled black circles), the MSC-E model (red with open squares).*

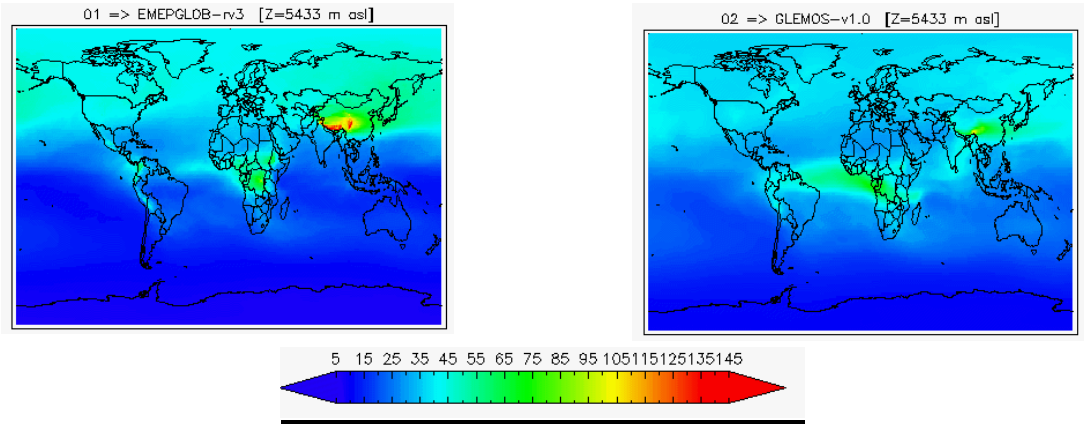


Figure 6.7 *Annually averaged COdirect in ppb at 509 hPa for the MSC_w (left) and MSC-E (right) models.*

In Figure 6.8 and 6.9 vertical profiles from the "COdirect" experiment are shown for January 30 and July 15, for four different locations: Pico and Vienna (in Figure 6.8) and Cheju and Ascension Island (in Figure 6.9). It should be stressed that the figures are snapshots at given times. We have chosen to show profiles for the same dates for all the 4 selected sites. The profiles should be interpreted with caution as there is substantial variability between individual days. For Pico (in the mid North Atlantic) and Vienna (in Austria) most of the models, including the MSC-W and the MSC-E models, are close to the ensemble mean for the two selected days. In winter there is a tendency for concentrations to be higher in the free troposphere in the MSC-W model compared most other models, confirming the results from Figure 6.7 where a comparison with the concentrations with the MSC-E model was presented. For Cheju (South Korea) and Ascension Island (South Atlantic) there are substantial differences between the models, in particular in the winter. Above Cheju the largest spread in model results are seen in the lower troposphere. Cheju is close to the source regions in East Asia. We believe that the spread in model results is caused by differences in model resolution and thus emission hotspots may be included differently in the different models. Above Ascension Island the largest spread in model results are located in the middle and upper troposphere in winter. We believe that the model spread here is caused by differences in the lifting of CO over the African continent and the subsequent advection in the middle and upper troposphere. In Cheju, as for Pico and Vienna, there is a tendency for the MSC-W global model to have higher concentrations in the free troposphere compared to most other models, whereas the opposite is the case for Ascension Island in the southern hemisphere west of equatorial Africa. Again, this is in agreement with Figure 6.7 showing higher outflow from Africa in the free troposphere in the MSC-E model compared to the MSC-W model.

These results are consistent with what was shown in Figure 6.6 (right) concerning results at about 500hPa.

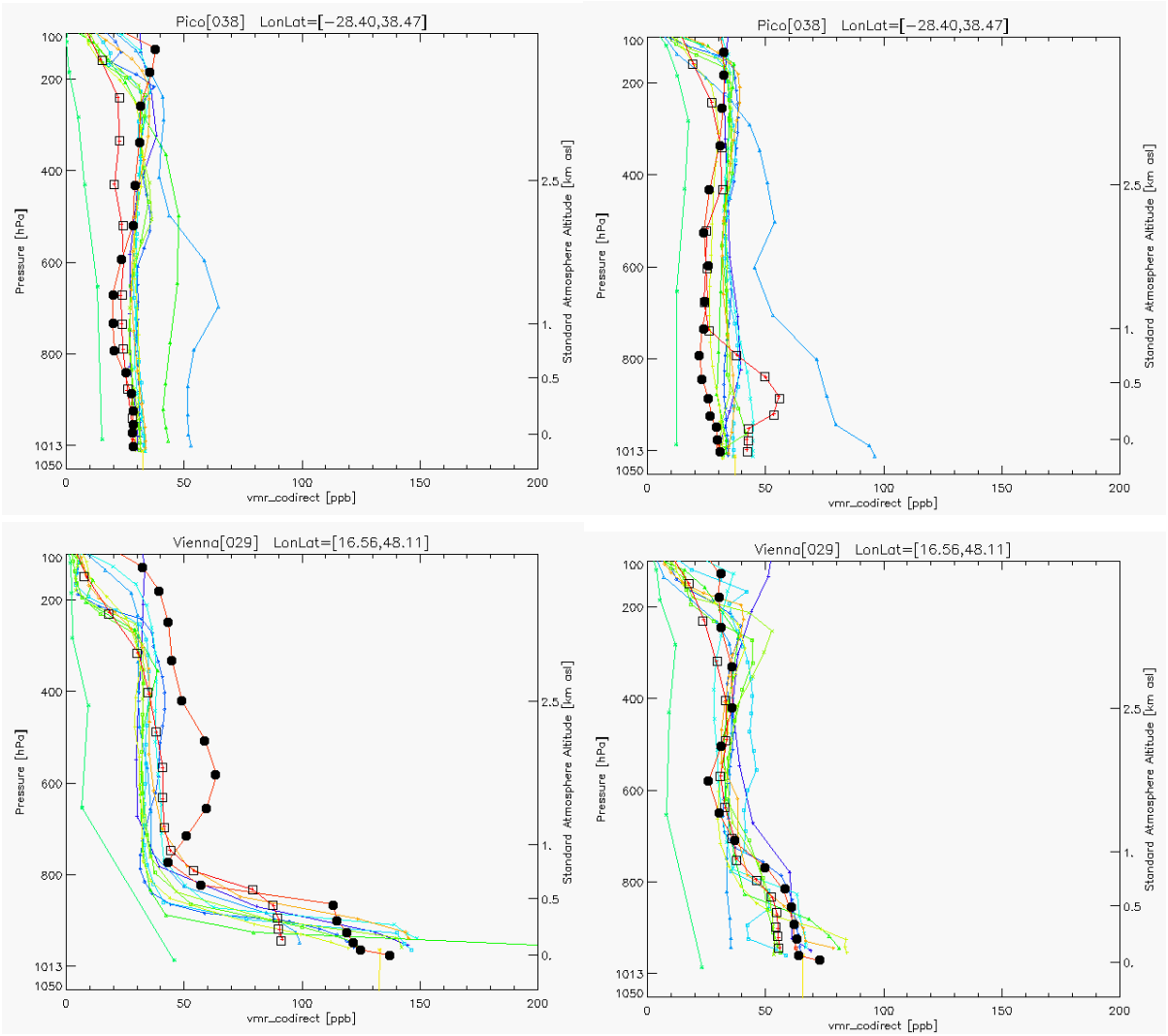


Figure 6.8 *Vertical profiles of the CO tracer for all 19 models in ppb at Pico, the Azores (top) and Vienna (bottom). Figures to the left January 30 12 UTC, to the right July 15 12 UTC. (MSC-W model – filled circles, MSC-E open squares).*

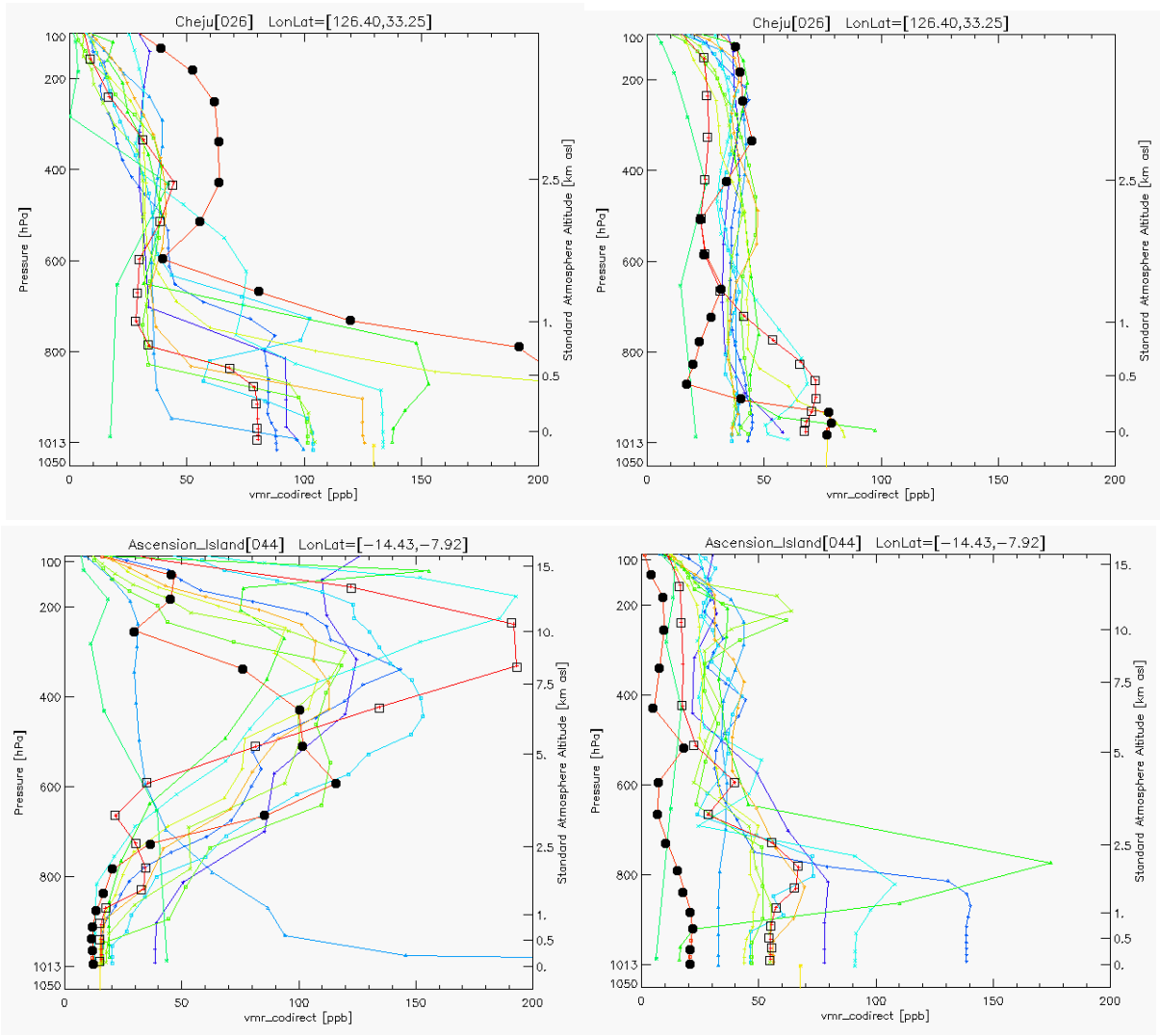


Figure 6.9 *Vertical profiles of the CO tracer for all 19 models in ppb at Cheju, tSouth Korea (top) and Ascension Island (bottom). Figures to the left January 30 12 UTC, to the right July 15 12 UTC. (MSC-W model – filled circles, MSC-E open squares).*

6.5 Source receptor relationships

Figure 6.10 shows the source receptor relationships for the 4 selected regions calculated with the MSC-W and the MSC-E global models. The ensemble mean from all 19 models is also shown. Both the MSC-W and the MSC-E models are close to the 19 models ensemble mean. For all regions there are considerable contributions from "foreign" regions. However, COdirect has a relatively long lifetime of 50 days. From the ensemble means, the foreign/intercontinental contribution to COdirect in the surface layer are about 40% for Europe and North America, somewhat less for East Asia, and about 32% for South Asia. In all 4 regions "other¨ (meaning emissions from outside the 4 selected regions) is the largest foreign contributor to COdirect. For Europe the largest identified contributor is North America. For North America the largest identified contributor is East Asia. East Asia and South Asia are the largest identifiable contributors to each other.

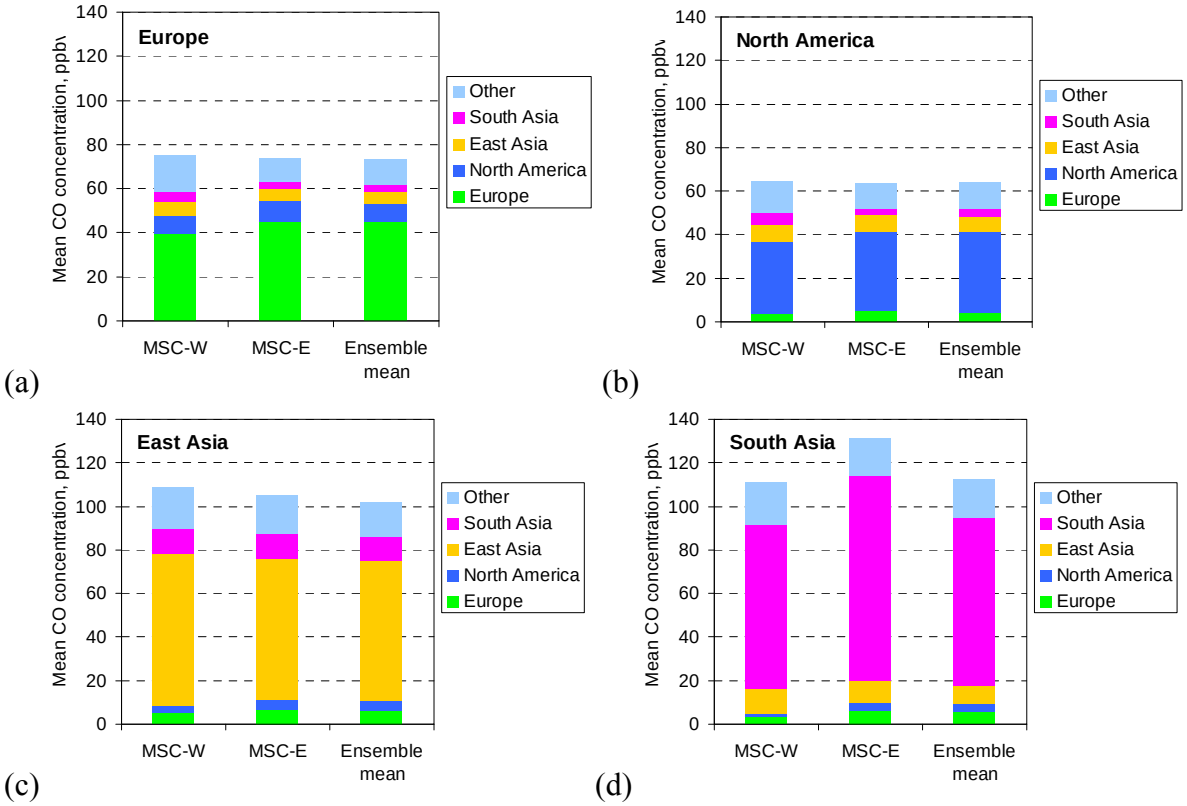


Figure 6.10 *Contribution of different source regions to mean surface concentration of COdirect in various parts of the globe: (a) – Europe; (b) – North America; (c) – East Asia; (d) – South America*

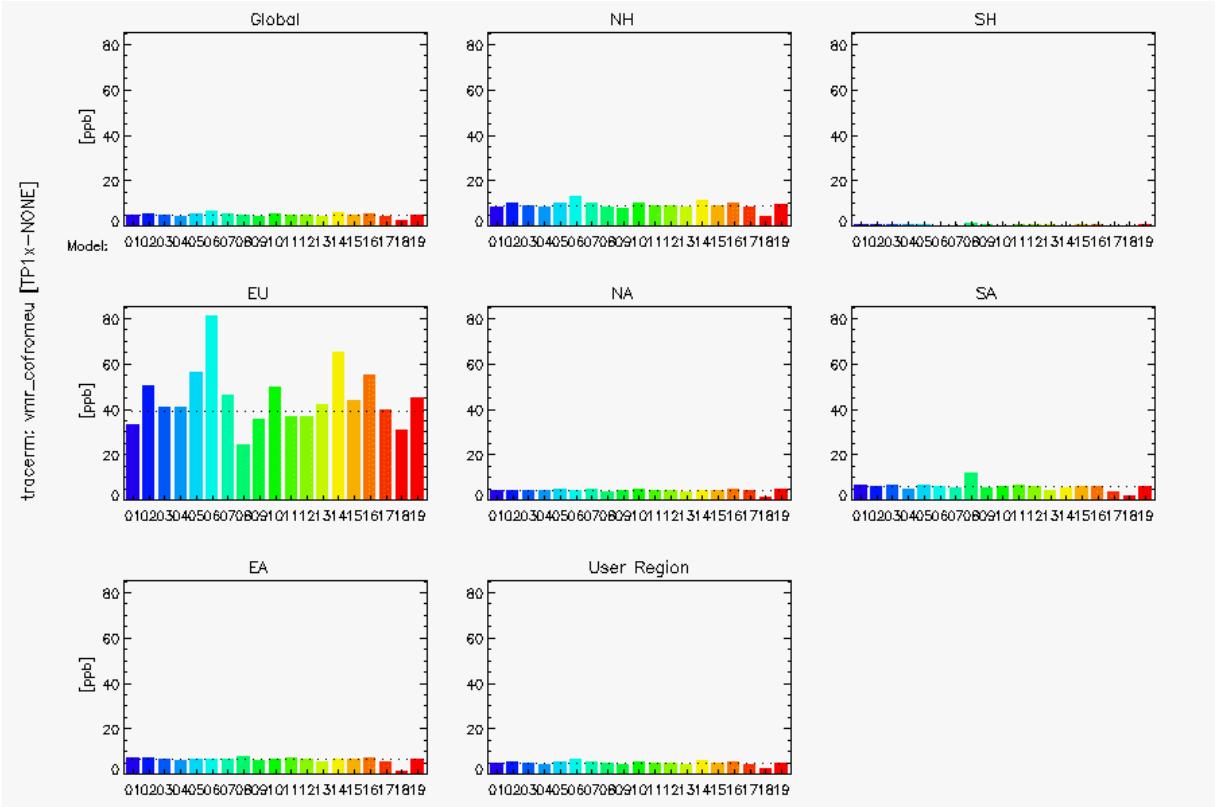


Figure 6.11 *Annually averaged contribution from emissions in the EU region to the CO like tracer in the different regions. The MSC-W model is nr 17 (3'rd from the right).*
The MSC-E model is nr. 19 (far right).

The calculated effects of emissions from the region EU are also illustrated in Figure 6.11, showing the contributions to concentrations averaged globally, for the two hemispheres and for the selected regions. The differences between the models are larger than for "COdirect" (Figure 6.5). Both the MSC-W and the MSC-E models are close to the ensemble mean

6.6 Conclusions from the comparison of tracer dispersion

Above we have used the TP1x experiment as a test case for the MSC-W and the MSC-E models. COdirect is emitted as CO with a decay similar to that of CO in the real atmosphere. As such, the COdirect distribution should resemble that of CO in the atmosphere. However, we have not attempted to compare with CO measurements here. Such a comparison will be facilitated with the implementation of a hemispheric measurement database as described in Chapter 2. Concerning intercontinental source-receptor calculations, in the selected regions the domestic contribution is roughly ranging from 60% to 70% for COdirect. However, most air pollutants will have a much shorter residence time in the atmosphere and as a result the percentage contribution from outside will be markedly lower.

Based on the model results from the participating models in the TFHTAP model inter-comparison we can in principle not identify any model as being better than the others. However, if a model is an outlier in any way the reason for the difference(s) should be investigated. The MSC-W model is low in the mid and upper troposphere in the tropics and in general in the southern hemisphere. We believe that this is caused by lack of convection in the model. This is in agreement with the results in Jonson et al. (2007). Here an underestimation of ozone in the free troposphere, in particular in the summer months, was attributed to insufficient parameterization of convective processes in the MSC-W model. At MSC-W there is ongoing work implementing a parameterization of convective processes in the model.

The MSC-W and the MSC-E models have similar horizontal resolution. Both models use the Bott advection scheme. They use different meteorological drivers. However, as shown in Chapter 5, the meteorological input data are very similar. They do however have a different vertical resolution The number of layers are the same in both models (20), but whereas the model domain only extent to 100hPa in the MSC-W global model, it extends to 10 hPa in the MSC-E global model. On the other hand, near the surface and in the free troposphere, the vertical resolution is coarser in the MSC-E model. Even so model results calculated with the MSC-W and MSC-E models are altogether similar. The difference in model results will increase as more processes (chemistry, deposition processes etc) are added. For the TF HTAP source receptor calculations for ozone the differences between the models are considerably larger (Fiore et al. 2008).

6.7 References

A.M. Fiore, A.M., F.J. Dentener, O. Wild, C. Cuvelier, M.G. Schultz, C. Textor, M. Schulz, C. Atherton, D. Bergmann, I. Bey, G. Carmichael, R. Doherty, B.N. Duncan, G. Faluvegi, G. Folberth, M. Garcia Vivanco, M. Gauss, S. Gong, D. Hauglustaine, P. Hess, T. Holloway, L.W. Horowitz, I.S.A. Isaksen, D.J. Jacob, J.E. Jonson, J.W. Kaminski, T.J. keating, A. Lupu, I. MacKenzie, E. Marmer, V. Montanaro, R. Park, K.J. Pringle, J.A. Pyle, M.G. Sanderson, S. Schroeder, D.T. Shindell, D. Stevenson, S. Szopa, R. Van Dingenen, P. Wind, G. Wojcik, S. Wu, G. Zeng, and A. Zuber.(2008)Multi-model estimates of intercontinental source-receptor relationships for ozone pollution. J. Geophys. Res., in preparation.

Gusev A., Mantseva E., Shatalov V., Travnikov O. (2005) Hemispheric model (MSCE-Hem) of persistent toxic substances dispersion in the environment. MSC-E Technical Report 11/2005, Meteorological Synthesizing Centre East of EMEP, Moscow. Available from www.msceast.org.

Jonson, J.E., Tarrasón, L., Wind, P., Gauss, M., Valiyaveetil, S. S., Tsyro, S, Klein, H., Isaksen, I.S.A. and A. Benedictow (2007) First evaluation of the global EMEP model and comparison with the global OsloCTM2 model. EMEP-MSC-W Technical report 1/2007, Norwegian Meteorological Institute, Oslo, Norway, available from www.emep.int.

Sanderson, M.G., F.J. Dentener, A.M. Fiore, K. Cuvelier, T.J. Keating, A. Zuber, C.S. Atherton, D.J. Bergmann, T. Diehl, R.M. Doherty, B.N. Duncan, P. Hess, L.W. Horowitz, R. Hu, D. Jacob, J.E. Jonson, J.W. Kaminski, A. Lupu, I.A. Mackenzie, E. Marmer, V. Montanaro, G. Pitari, M.J. Prather, K.J. Pringle, S. Schroeder, M.G. Schultz, D.T. Shindell, D. Simpson, S. Szopa, and O. Wild. Oxidised Nitrogen Source-Receptor Relationships. J. Geophys. Res., in preparation, 2008.

Shindell, D.T, 09 H. Teich1, M. Chin, F. Dentener, R.M. Doherty, G. Faluvegi, A.M. Fiore, P. Hess, I.A. MacKenzie, M.G. Sanderson, M.G. Schultz, M. Schulz, D.S. Stevenson, C. Textor, O. Wild, D.J. Bergmann, H. Bian1, C. Cuvelier, B.N. Duncan1, G. Folberth, L.W. Horowitz, J.E. Jonson, J.W. Kaminski1, E.Marmer, R. Park, K.J. Pringle, S. Schroeder, S. Szopa, T. Takemura, G. Zeng, T.J. Keating, and A. Zuber.(2008) A multi-model assessment of pollution transport to the Arctic. J. Geophys. Res., submitted March 2008.

Simpson, D., H. Fagerli, J.E. Jonson, S. Tsyro, P. Wind J-P Tuovinen (2003) Transboundary acidification, eutrophication and ground level ozone in Europe, PART 1. EMEP status report 1/2003, The Norwegian Meteorological Institute, Oslo Norway, 2004, available from www.emep.int.

UNECE, 2007, Hemispheric Transport of Air Pollution, Air pollution studies no. 16, ECE/EB.AIR/94, Geneva, Switzerland.