EMEP/MSC-W Model Unofficial User's Guide

Release rv4_15

https://github.com/metno/emep-ctm

Oct 16, 2017

CONTENTS:

CHAPTER

ONE

WELCOME TO EMEP

This guide gives a brief documentation of the EMEP/MSC-W model version rv4.15. It is intended primarily as a guide on how to run the model, and to help users wishing to understand or change the model in terms of domains, outputs, chemistry, etc.

The main documentation for the EMEP/MSC-W model is an article published in Atmospheric Chemistry and Physics in 2012. This article will be referred to as Simpson et al. (2012) in this manual.

- Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L.D., Fagerli, H., Flechard, C.R., Hayman, G.D., Gauss, M., Jonson, J.E., Jenkin, M.W., Nyíri, Á, Richter, C., Semeena, V.S, Tsyro, S., Tuovinen, J.-P., Valdebenito, Á., and Wind, P.: The EMEP MSC-W chemical transport model – technical description. Atmospheric Chemistry and Physics, 12, 7825-7865, 2012.

  http://www.atmos-chem-phys.net/12/7825/2012/acp-12-7825-2012.html

The model source code is available from the Open Source EMEP/MSC-W model github page:

https://github.com/metno/emep-ctm

1.1 Licenses and Caveats

The EMEP code is provided under the GNU General Public License version 3 (http://fsf.org and/or http://www.gnu.org/copyleft/gpl.html).

Each code module is prefaced with something like:

```
! <EXAMPLE_CODE.f90 - A component of the EMEP MSC-W  Eulerian
!          Chemical transport Model>
!*****************************************************************!
!*
!*  Copyright (C) 2007-2016 met.no
!*
!*  Contact information:
!*  Norwegian Meteorological Institute
!*  Box 43 Blindern
!*  0313 OSLO
!*  NORWAY
!*  email: emep.mscw@met.no
!*
!*    This program is free software: you can redistribute it and/or modify
!*    it under the terms of the GNU General Public License as published by
!*    the Free Software Foundation, either version 3 of the License, or
!*    (at your option) any later version.
!*
!*    This program is distributed in the hope that it will be useful,
```

```
!*    but WITHOUT ANY WARRANTY; without even the implied warranty of
!*    MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the
!*    GNU General Public License for more details.
!*
!*    You should have received a copy of the GNU General Public License
!*    along with this program.  If not, see <http://www.gnu.org/licenses/>.
!*****************************************************************!
```

And a copy of the license file, **gpl.txt**, is provided with the model code source files.

It is important to note that the code is provided "as it is", and EMEP/MSC-W has very limited resources with which to support usage of the code.

1.2 Computer Information

To compile the EMEP/MSC-W model you need:

- Fortran 95 compiler
- NetCDF Library (> 4.1.3)
- MPI Library (> 1.0)

It is necessary to compile with double precision real (8 bytes real). The program has been used on computers ranging from a Linux laptop to supercomputers (Itanium2 cluster, Intel Xeon cluster, Cray XT4, IBM power5+). It is compatible with all compilers tested so far: Intel, PGI, gfortran, XL Fortran. A Makefile is included, the path to NetCDF (`INCL` and `LLIB`) have to be adapted to your machine, and the Fortran compiler (`F90`) and flags (`F90FLAGS`) to the compiler you are using.

The code has been tested with 1 to 1024 CPUs, and scales well (for large grids). If only one CPU is used 1-2 GB memory is required. If more than one, for example 64 CPUs are used, 200 MB of memory per CPU is enough (in the case of a 132 X 159 grid size). For runs on more than 32 CPUs, a fast interconnect is recommended (infiniband for example), for smaller runs, gigabit Ethernet is sufficient. It takes ~5 hours on 64*Xeon X5355 (2.66GHz) for 1 year simulation.

When downloading input data in order to do a "base run" please make sure that there are 35 Gb disc space available, especially due to large meteorology input files. The model can be run for shorter periods, users can download meteorology for only the period they are interested in, plus one day.

1.3 Getting Started

It is recommended to read all the chapters of this EMEP/MSC-W model User Guide before you start downloading anything from the EMEP/MSC-W Open Source website.

This is what you need to do before you can do a "base run" with the EMEP/MSC-W model:

- Read the EMEP/MSC-W model User Guide
- Download input data, description and downloading instructions in Section 2.
- Download the EMEP/MSC-W model source code, description and downloading instructions Section 1.4.
- Follow the instructions for "Submitting a Run" description in Section 4.
- Download some model results for comparison, description in and downloading instructions Section 3.

1.4 Model code

The latest release Open Source of the EMEP/MSC-W model is version rv4_15. This and previous releases can be found on the releases section of the EMEP/MSC-W Open Source github page.

Although the source code is available for download as a single compressed file on the release page, the preferred retrieval method is via the catalog tool as follows:

```
# download the catalog tool
wget https://raw.githubusercontent.com/metno/emep-ctm/master/tools/catalog.py

# make it executable and run it
chmod +x catalog.py

# download the source code for rv4_15 release
catalog.py -R rv4_15 --source
```

The model source code, makefiles, and a copy of the license file will be placed under the directory `EMEP_MSC-W_model.rv4.15.OpenSource/source/`. An overview of the files is given in Table 1.1.

Table 1.1: Model source files

Type	Filename
modules files	`*.f90`
include files	`*.inc`
namelist	`config_emep.nml`
makefiles	`Makefile` and `Makefile.SRCS`
dependency file	`dependencies`
a copy of the license	`gpl.txt`

In addition there is a run script called `modrun.sh`, which will be placed in the `EMEP_MSC-W_model.rv4.15.OpenSource` directory. The run script, `modrun.sh`, can easily be modified to work on your computer system. This script is described in detail in Section 4.

1.5 Model grid

The current EMEP model version, and the provided gridded input data, have a horizontal resolution of $50 \times 50 km^2$ (at $60^\circ N$) and are defined on a polar stereographic projection with 20 sigma levels vertically. The model is very flexible with regard to the horizontal resolution, in that it readily makes use of meteorological data provided with the model. The vertical resolution is currently still restricted to the fixed 20 layer system. The physical description is given in detail in Chapter 2 of the EMEP Status Report 1/2003 Part I (Simpson *et al.*, 2003).

In 2008 the EMEP domain was extended eastwards in order to include the EECCA countries in the EMEP model grid, see Figure 1.1. To distinguish the new grid from the old EMEP grid, the new grid is called EECCA in this text and in the `config_emep.nml`.

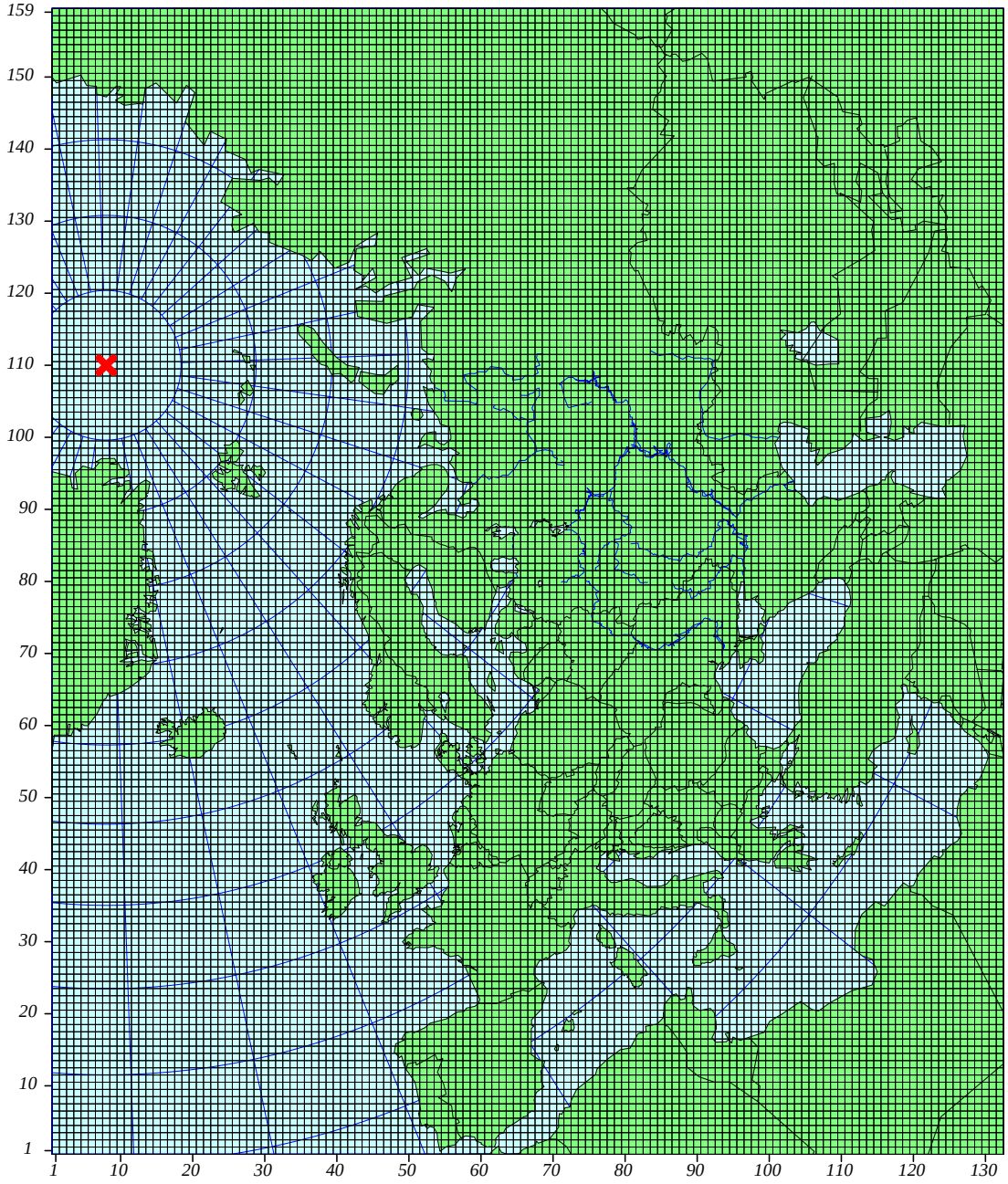

Figure 1.1: The extended EMEP grid covering EECCA area with 132×159 gridpoints on $50 \times 50 km^2$ resolution defined on a polar stereographic projection.

CHAPTER

TWO

INPUT FILES

This chapter provides an overview on the necessary input files to run the EMEP/MSC-W model. A complete set of input files is provided as part of the EMEP/MSC-W Open Source release to allow model runs for the meteorological year 2015. Table 2.1 lists the input files.

In the latest release, meteorology is provided for 2 different model domains and resolutions:

- *EECCA* domain with a horizontal resolution of 50x50 km2 (at 60°N), on polar stereographic projection, and 20 vertical levels;
- *EMEP01* domain with a 0.1x0.1 degrees on long-lat projection, and 34 vertical levels.

Download the input via the catalog tool (Section 1.4) as follows:

```
# download 2015 meteorology for the EECCA domain
catalog.py -R rv4_15 --meteo --met-domain EECCA

# download 2015 meteorology for the EMEP01 domain
catalog.py -R rv4_15 --meteo --met-domain EMEP01

# download other input files
catalog.py -R rv4_15 --input
```

The meteorology files will be placed under `EMEP_MSC-W_model.rv4.15.OpenSource/meteo2015/`, and the remaining input files will be placed under `EMEP_MSC-W_model.rv4.15.OpenSource/input/`

This are all input files needed to run the EMEP/MSC-W model, except the aircraft emissions (`AircraftEmis_FL.nc`), and forest fire emissions (`FINN_ForestFireEmis_2015.nc`). See sections Section 2.1.9 and Section 2.1.11 for details about these emissions data.

IMPORTANT: The input data available in the EMEP/MSC-W Open Source Web site should be appropriately acknowledged when used for model runs. If nothing else is specified according to references further in this chapter, please acknowledge EMEP/MSC-W in any use of these data.

Table 2.1: List of input data files

Data	Name	Format
Meteorology data	`met/`	
Meteorology	`meteoYYYYMMDD.nc` (365+1 files)	netCDF[1]
Other Input files	`input/`	
Global Ozone	`GLOBAL_O3.nc`	netCDF
New Global Ozone	`Logan_P.nc`	netCDF[2]
BVOC emissions	`EMEP_EuroBVOC.nc`	netCDF
Landuse	`LanduseGLC.nc` and `Landuse_PS_5km_LC.nc`	netCDF
Degree-day factor	`DegreeDayFactors.nc`	netCDF

Continued on next page

Table 2.1 – continued from previous page

Data	Name	Format
N depositions	`annualNdep.nc`	netCDF
Road dust	`RoadMap.nc` and `AVG_SMI_2005_2010.nc`	netCDF[3]
Aircraft emissions	`AircraftEmis_FL.nc`	netCDF[3]
Surface Pressure	`SurfacePressure.nc`	netCDF[3]
Forest Fire	`FINN_ForestFireEmis_YYYY.nc`	netCDF[3]
Dust files	`Soil_Tegen.nc`	netCDF[3]
	`SoilTypes_IFS.nc`	netCDF[3]
Emissions	`emislist.POLL` (7 files, EMEP 50km PS grid)	ASCII[4]
	`Emis_TNO7.nc` (regional, 0.125×0.0625 lon-lat)	netCDF[3]
	`Emis_GLOB_05.nc` (global, 0.5×0.5 lon-lat)	netCDF[3]
Vertical level distribution	`Vertical_levels.txt`	ASCII
Time factors for monthly emissions	`MonthlyFac.POLL` (7 files)	ASCII[4]
Time factors for daily emissions	`DailyFac.POLL` (7 files)	ASCII[4]
Time factors for hourly emissions	`HOURLY-FACS`	ASCII
Emission heights	`EmisHeights.txt`	ASCII
Natural SO_2	`natso2MM.dat` (12 files)	ASCII[1]
Volcanoes	`columnsource_emission.csv`	ASCII
	`columnsource_location.csv`	ASCII
Lightning emissions	`lightningMM.dat` (12 files)	ASCII[1]
Emissions speciation	`emissplit.defaults.POLL`	ASCII[4]
	`emissplit.specials.POLL`	ASCII[43]
Emission factors for scenario runs	`femis.dat`	ASCII
Photo-dissociation rates	`jclearSS.dat` (4 files)	ASCII[5]
	`jcl1kmSS.dat` (4 files) and `jcl1.jun`	ASCII[5]
	`jcl3kmSS.dat` (4 files) and `jcl3.jun`	ASCII[5]
Landuse definitions	`Inputs_LandDefs.csv`	ASCII
Stomatal conductance	`Inputs_DO3SE.csv`	ASCII
Sites locations for surface output	`sites.dat`	ASCII
Sondes locations for vertical output	`sondes.dat`	ASCII

2.1 NetCDF files

2.1.1 Meteorology

The daily meteorological input data (`meteoYYYYMMDD.nc`, where `YYYY` is year, `MM` is month and `DD` is day) used for the EMEP/MSC-W Model are based on forecast experiment runs with the Integrated Forecast System (IFS), a global operational forecasting model from the European Centre for Medium-Range Weather Forecasts (ECMWF).

The IFS forecasts has been run by MSC-W as independent experiments on the HPCs at ECMWF with special requests on some output parameters. The meteorological fields are retrieved on a $0.1° \times 0.1°$ longitude latitude coordinates and interpolated to $50 \times 50km^2$ polar-stereographic grid projection. Vertically, the fields on 60 eta (η) levels from the IFS model are interpolated onto the 37 EMEP sigma (σ) levels. The meteorology is prepared into 37 sigma levels since the model is under test for a finer vertical resolution.

[1] `YYYY`: year, `MM`: month, `DD`: day.
[2] New O_3 boundary condition data in 30 levels. Can be used with `NewLogan=.true.` in `BoundaryConditions_ml.f90`.
[3] Optional, in most cases.
[4] `POLL`: pollutant type (NH_3, CO, NO_x, SO_x, NMVOC, $PM_{2.5}$ and PM_{co}).
[5] `SS`: seasons.

The open source code is released with 20 sigma levels and to make the model read the meteorology properly, a description of the 20 vertical sigma levels is needed. This is provided in an ASCII file called `Vertical_levels.txt` together with the other input data (Table 2.1). The version of the IFS model used for preparing these fields, Cycle 38r2, is documented in http://www.ecmwf.int/research/ifsdocs/index.html. Previous years are based on Cycle 36r1 with a resolution of $0.2^\circ \times 0.2^\circ$ on a spherical grid. Meteorological fields currently used for EMEP/MSC-W Model runs are given in Table 2.2. Some verification and description of these meteorological fields are given in Chapter 2 of the EMEP Status Report 1/2016.

Acknowledgement: ECMWF, met.no

Table 2.2: Input meteorological data used in the EMEP/MSC-W Model

Parameter	Unit	Description
3D fields		for 37 σ
u, v	m/s	Horizontal wind velocity components
q	kg/kg	Specific humidity
θ	K	Potential temperature
CW	kg/kg	Cloud water
CL	$\%$	3D Cloud cover
$cnvuf$	kg/sm^2	Convective updraft flux
$cnvdf$	kg/sm^2	Convective downdraft flux
PR	mm	Precipitation
2D fields		for surface
PS	hPa	Surface pressure
$T2$	K	Temperature at $2m$ height
$Rh2$	$\%$	Relative humidity at $2m$ height
SH	W/m^2	Surface flux of sensible heat
LH	W/m^2	Surface flux of latent heat
τ	N/m^2	Surface stress
SST	K	Sea surface temperature
SWC	m^3/m^3	Soil water content
$lspr$	m	Large scale precipitation
cpr	m	Convective precipitation
$sdepth$	m	Snow depth
ice	$\%$	Fraction of ice
$SMI1$		Soil moisture index level 1
$SMI3$		Soil moisture index level 3
$u10, v10$	m/s	Wind at $10m$ height

2.1.2 Gridded emissions

Since 2015 different formats of gridded emissions can be used and mixed (with some restrictions) under one common framework. The different formats that are presently supported are:

"Old style" ASCII emissions format: Total yearly emissions.

The gridded emission files contain 16 columns where the first column represents the country code (http://www.emep.int/grid/country_numbers.txt), the second and the third columns are the i and j indices of the EMEP grid, the fourth and fifth columns are the total emissions from low and high sources, and the last 11 columns contain emissions from 10 anthropogenic SNAP sectors.

The advantage of the ASCII emissions format, is that they are easy to modify, and the interpretation of the numbers is straightforward. The main disadvantage of the ASCII emissions format, is that they are only valid for one specific grid projection. Visualization of these emissions, needs also some more efforts.

Countrywise NetCDF emissions: Yearly totals.

Each country and sector has its own NetCDF field.

The main advantage of NetCDF emissions is that all the information about the data (projection, units) is given in the same file. This allows the code to reproject the emissions to any grid projection on the fly. It is easy to visualize the emissions of one country with simple tools, like ncview. The data is simple to interpret and it is possible to add new countries to an existing file (with appropriate tools).

The disadvantage of countrywise NetCDF emissions, is that there are quite a large number of fields, with most of the data being zero. NetCDF will compress the data, but it will still take some time for the model to read all the data.

"Fraction type" NetCDF emissions: Yearly totals.

The total emissions are stored in one gridded map, and in addition information about which country the emission belongs to.

The main advantage of "fraction type" NetCDF emissions, is that they will keep the grid flexibility, have a more compact form and be faster to read in.

The disadvantage is that the interpretation of the content of the fields is more difficult and it is hard, for instance, to add a new country to the file. Total emissions and coverage of countries can easily be visualized, but not emissions from one single country.

Description of main fields for "fraction type" NetCDF Emissions Table 2.3

Table 2.3: Description of main fields for "fraction type" NetCDF Emissions

Variable name	Description
`Ncodes`	Number of countries sharing the same grid cell
`poll_secNN`	Pollutant from each sector
`Codes`	Country code number
`fractions_poll_secNN`	Fraction of emissions to assign to one country

Monthly "fraction type" NetCDF emissions.

This is similar to the yearly "fraction type" NetCDF emissions, but there are 12 monthly values for each field. This format cannot be combined with other formats.

Using and combining gridded emissions

These gridded emission files are controlled via the `config_emep.nml` file. Each file is assigned as one set of values for `emis_inputlist`. Listing 2.1 line 1 includes an ASCII emission file, where the keyword `POLL` will be replaced by the model by all the emitted pollutants (according to the names defined in `CM_EmisFiles.inc`). An additional NetCDF emission is included in line 2.

Now all emissions from both ASCII file and NetCDF file will be used. In practice some countries might be counted twice. Therefore some new data can be included in the `emis_inputlist`, to specify which countries to keep or to avoid. Listing 2.1 lines 3–4 will include only 'NO', 'SE' and 'FI' from the first file (ASCII), and take all countries except 'NO', 'SE' and 'FI' from the second file (NetCDF).

Sets of countries can in principle be defined; for now only the set 'EUMACC2' is defined.

Listing 2.1: Mixed emission configuration example.

```
1 emis_inputlist(1)%name = '/MyPathToEmissions/emislist.POLL',
2 emis_inputlist(2)%name = '/MyPathToEmissions/Emis_GLOB_05.nc',
```

```
emis_inputlist(1)%incl(1:) = 'NO','SE','FI',
emis_inputlist(2)%excl(1:) = 'NO','SE','FI',
emis_inputlist(1)%PollName(1:2) = 'voc','sox',
```

It is also possible to restrict the number of pollutants from each of the files. If not all pollutants from `CM_EmisFiles.inc` are to be read, one can specify a list of pollutants to be included using "PollName". For instance in the example above , the last line specifies that emissions will include only VOC and SOx from the file defined by emis_inputlist(1)%name. If PollName is not specified at all, all pollutants are included (therefore all pollutants from emis_inputlist(2)%name will be included). The specified pollutants must already be defined in `CM_EmisFiles.inc`

2.1.3 Global Ozone

Initial concentration of ozone are required in order to initialize the model runs. Boundary conditions along the sides of the model domain and at the top of the domain are then required as the model is running.

The `Logan_P.nc` file contains monthly averaged fields in NetCDF format. The initial and background concentrations are based on the Logan (1998) climatology. The Logan climatology is scaled on run time according to the Mace Head measurements as described in Simpson *et al.* (2003). For a number of other species, background/initial conditions are set within the model using functions based on observations (Simpson *et al.*, 2003 and Fagerli *et al.*, 2004).

2.1.4 BVOC emissions

Biogenic emissions of isoprene and monoterpene are calculated in the model as a function of temperature and solar radiation, using the landuse datasets. The light and temperature dependencies are similar to those used in the original open source model, see Chapter 4.2 of the EMEP Status Report 1/2003 Part I (Simpson *et al.*, 2003).

Biogenic VOC emission potentials (i.e. rates at $30°C$ and full sunlight) are included for four different forest types in the NetCDF file `EMEP_EuroBVOC.nc`. These emission potentials have unit $\mu g/m^2/h$, and refer to emissions per area of the appropriate forest category. In addition, default emission potentials are given for other land-cover categories in the file `Inputs_LandDefs.csv`. The underlying emission potentials, land-cover data bases, and model coding have however changed substantially since model version v.2011-06. The new approach is documented in Simpson *et al.*, 2012.

2.1.5 Landuse

Landuse data are required for modelling boundary layer processes (i.e. dry deposition, turbulent diffusion). The EMEP/MSC-W model can accept landuse data from any data set covering the whole of the domain, providing reasonable resolution of the vegetation categories. Gridded data sets providing these landuse categories across the EMEP domain have been created based on the data from the Stockholm Environment Institute at York (SEI-Y) and from the Coordinating Center for Effects (CCE). 16 basic landuse classes have been identified for use in the deposition module in the model, and three additional "fake" landuse classes are used for providing results for integrated assessment modeling and effects work.

There are two NetCDF files included, one file `Landuse_PS_5km_LC.nc` on 5 km resolution over the EMEP domain, and a global `LanduseGLC.nc`. The different landuse types are desribed in Simpson et al (2012).

2.1.6 Degree-day factor

Domestic combustion which contribute to a large part of SNAP 2, varies on the daily mean temperature. The variation is based on the heating degree-day concept. These degree days are pre-calculated for each day and stored in the file `DegreeDayFactors.nc`. See Simpson et al. (2012) section 6.1.2.

2.1.7 NO_x depositions

Areas with high NO deposition loads have greater soil-NO emissions. To include this in the model, a NetCDF file where pre-calculated N-depositions are included. The file made by the results from the EMEP/MSC-W model runs over a 5-year period.

2.1.8 Road Dust

Road traffic produces dust. These emissions are handled in the EMEP/MSC-W model in the `Emissions_ml.f90` module. To include road dust, set `USE_ROADDUST=.true.` in `config_emep.nml`. There are two files included in input data, `RoadMap.nc` and `AVG_SMI_2005-2010.nc`. `RoadMap.nc` include gridded roads and PM emissions over Europe, `AVG_SMI_2005-2010.nc` are global.

2.1.9 Aircraft emissions

In the EMEP/MSC-W model aircraft emissions are 'OFF' by default. They can be switched 'ON' by setting `USE_AIRCRAFT_EMIS=.true.` in `config_emep.nml` and download the data from http://www.pa.op.dlr.de/quantify. The EMEP model uses data provided by the EU-Framework Programme 6 Integrated Project QUANTIFY (http://www.pa.op.dlr.de/quantify). However, before using these data a protocol has to be signed, which is why the data file can not be provided directly on the EMEP/MSC-W Open Source website. If you want to use aircraft emissions go to http://www.pa.op.dlr.de/quantify, click on 'QUANTIFY emission inventories and scenarios', and then click on 'Register'. That page will provide information about the registration process and the protocol that has to be signed. Once you are registered, click 'Login' and provide user name and password. On the new page, search for 'Emissions for EMEP', which links directly to the `Readme` file and the emission data file in NetCDF format. Download the emission data file and place it in the input folder.

2.1.10 Surface Pressure

If `USE_AIRCRAFT_EMIS=.true.` in `config_emep.nml`, then in addition to the Aircraft Emission file, there will be need for a `SurfacePressure.nc` file, which is already in the `/input` folder. The NetCDF file consists of surface pressure fields for each of the months in 2008 called `surface_pressure`, and one field for the whole year called `surface_pressure_year`. All fields are given in Pa.

2.1.11 Forest Fire

Since model version rv3.9 (November 2011), daily emissions from forest and vegetation fires are taken from the "Fire INventory from NCAR version 1.0" (FINNv1, Wiedinmyer et al. 2011). Data are available from 2005, with daily resolution, on a fine $1km \times 1km$ grid. We store these data on a slightly coarser grid ($0.2° \times 0.2°$) globally for access by the EMEP/MSC-W model. To include forest fire emissions set `USE_FOREST_FIRES=.true.` in `config_emep.nml` and download the 2012 GEOS-chem daily data http://bai.acom.ucar.edu/Data/fire/. The data needs to be stored with units mole/day in a NetCDF file called `FINN_ForestFireEmis_2015.nc` compatible with the `ForestFire_ml.f90` module.

2.1.12 Dust files

The annual ASCII data for sand and clay fractions as well as the monthly data for boundary and initial conditions for dust from Sahara are replaced with a single NetCDF file `Soil_Tegen.nc` since 2013. This covers data for a global domain in 0.5×0.5 degree resolution.

The variables 'sand' and 'clay' gives the fraction (in %) of sand an clay in the soil for each grid cell over land.

The files are used by the module `DustProd_ml.f90`, which calculates windblown dust emissions from soil erosion. Note that the parametrization is still in the development and testing phase, and is by default 'turned off'. To include it in the model calculations, set `USE_DUST=.true.` in `config_emep.nml`. The user is recommended to read carefully documentation and comments in the module `DustProd_ml.f90`.

There is also a possibility to include boundary and initial conditions for dust from Sahara. The input file gives monthly dust mixing ratios (MM - month, e.g. 01, 02, 03,...) for fine and coarse dust from Sahara. The files are based on calculations from a global CTM at the University of Oslo for 2000. To include Saharan dust, set `USE_SAHARA=.true.` in `config_emep.nml`.

Another source for dust is an arid surface. This is accounted for by soilmosture calculations in `DustProd_ml.f90`. Together with Soil Water Index from the meteorology files and permanent wilting point (pwp) from `SoilTypes_IFS.nc`. This file is global and NetCDF. See Simpson et al. (2012) section 6.10.

2.2 ASCII files

2.2.1 Natural SO_2

Natural SO_2 emissions (dimethylsulfide (DMS) from sea) are provided as monthly gridded files. The values are given at the surface in $\mu g/m^2$ for each grid cell in the domain.

2.2.2 Volcanoes

Emissions from volcanic passive degassing of SO_2 are included for the active Italian volcanoes, Etna, Vulcano and Stromboli, and based upon the officially submitted data. To consider these volcanic emissions, we need to feed the locations and heights of volcanoes into the model. The input file `columnsource_location.csv` contains the geographical coordinates (latitudes and longitudes) and the heights (in meters) of the included volcanoes, while `columnsource_emission.csv` contains the emission parameters.

Since 2010 the EMEP/MSC-W model has also been used to model the transport of ash and SO_2 from volcanic eruptions. In addition to data for passive degassing of SO_2, the above two input files also contain locations and emission parameters for two recent eruptions of Icelandic volcanoes (Eyjafjallajökull in 2010 and Grimsvötn in 2011). In order to include emissions from these eruptions one needs to set `USE_ASH=.true.` in `config_emep.nml`.

2.2.3 Gridded emissions

The official emission input for the EMEP/MSC-W model consists of gridded annual national emissions based on emission data reported every year to EMEP/MSC-W (until 2005) and to CEIP (from 2006) by each participating country. More details about the emission input with references can be found in Chapter 4 of the EMEP Status Report 1/2003 Part I (Simpson et al., 2003).

Since 2015 different formats of gridded emissions can be used and mixed (with some restrictions) in the EMEP model under one common framework. The new emission system is described in Section 2.1.2. Here we focus only on the "old style" ASCII emission format.

Seven gridded emission input files (`emislist.poll`) are available in ASCII format for the following compounds: CO, NH_3, NO_x, $PM_{2.5}$, PM_{co}, SO_x and VOC.

The gridded ASCII emission files contain 16 columns where the first column represents the country code (http://www.emep.int/grid/country_numbers.txt), the second and the third columns are the i and j indices of the EMEP grid, the fourth and fifth columns are the total emissions from low and high sources, and the last 11 columns contain emissions from 10 anthropogenic SNAP sectors (http://reports.eea.eu.int/technical_report_2001_3/en) and 1 source-sector called "Other sources and sinks", which include natural and biogenic emission sources. The data are given with the Mg.

Acknowledgement: EMEP

2.2.4 Time factors for emissions

Monthly and daily time factors for emission of the 7 compounds (CO, NH_3, NO_x, $PM_{2.5}$, PM_{co}, SO_x and VOC). There is one file available per compound in ASCII format.

The first two columns in the files represent the country code (http://www.emep.int/grid/country_numbers.txt), the second column represents the sector (http://webdab.emep.int/sectors.html). In the monthly files, the 12 consecutive columns represent the time factors corresponding to the months of the year. In the daily files there are 7 consecutive columns representing the time factor for each day of the week.

The file `HOURLY-FACS` includes factors for each of the eleven SNAP sectors for every hour (the columns) for each day of the week, see Simpson et al. (2012) section 6.1.2

2.2.5 Emission heights

A vertical distribution for the eleven SNAP sectors are given in the file `EmisHeights.txt`. The file has seven vertical levels, over the columns and the SNAP sectors given in the first row. Read more in Simpson et al. (2012) section 6.1.1.

2.2.6 Emission factor for scenario runs

Scenario run in the case of the EMEP/MSC-W model means a run to test the impact of one or more pollutants from a particular country.

Emission factors are applied to specified countries and emission sectors and can be set by changing the ASCII file `femis.dat`. This file can be changed by the users according to their needs.

The file contains several columns (the number is flexible). The first column represents the country code (http://www.emep.int/grid/country_numbers.txt), the second represents the sector (http://reports.eea.eu.int/technical_report_2001_3/en) where '0' means all sectors, and then in the remaining columns one can specify which emissions to reduce. Here 1.0 means no reduction of the given pollutant (SO_x, NO_x, VOC, NH_3, CO, $PM_{2.5}$and PM_{co}) from sectors of specified country. The number following the first text ("Name") in line 1 (number 5 in the downloaded file) gives the number of pollutants treated in the file.

2.2.7 Chemical speciation of emissions

Many of the emission files give emissions of a group of compounds, e.g. NO_x includes NO+|NO2|, and VOC can include many compounds. The information needed to retrieve emissions of individual compounds from these the gridded files is given in files labelled `emissplit.defaults.POLL` or `emissplit.specials.POLL`, where `POLL` can be NO_x, VOC, etc.

The defaults file give the emission split for each SNAP sector (one per row, with second index being the SNAP sector), which is applied to all countries by default. For VOC this split was derived from the UK inventory of Passant (2002), as part of the chemical comparison project of Hayman *et al.* (2011).

The specials files are in general optional, and can be used to specify speciation for particular countries or SNAP sectors. The 1st column specifies the country code of interest, the second the SNAP sector.

If forest fires are used, then the file `emissplit.specials.voc` is required (not optional), and the country-code 101 used to specify the VOC speciation of forest fires in this file.

2.2.8 Lightning emissions

Emissions of NO_x from lightning are included in the model as monthly averages on T21 ($5.65° \times 5.65°$) resolution (Køhler *et al.*, 1995). The lightning emissions are defined on a 64×32 grid with 17 vertical levels, with global coverage, and are provided as 12 ASCII files `lightningMM.dat`.

2.2.9 Landuse definitions

For the vegetative landuse categories where stomatal modelling is undertaken, the start and end of the growing season (SGS, EGS, in days) must be specified. The calculation of SGS and EGS with respect to latitude is done in the module `LandDefs_ml.f90`. The parameters needed to specify the development of the leaf area index (LAI) within the growing season are given in the ASCII file `Inputs_LandDefs.csv`. For more information, see chapter 5 of the EMEP Status Report 1/2003 Part I (Simpson *et al.*, 2003).

The file, designed to be opened with excel or gnumeric, contains a header briefly explaining the contents of the 14 columns. The first three columns are representing the landuse name, code (which are consistent with those in `Landuse.Input` file) and type (grouping of the landuse classes). The fourth column (PFT) gives a plant-functional type code (for future use), the fifth gives the maximum height of vegetation (m), the sixth indicates albedo (%) and the seventh indicates possible source of NH_x (0 off/1 on, curently not used). Columns 8 to 11 define the growing season (day number), column 12 and 13 lists the LAI minimum and maximum (m^2/m^2) and columns 14 and 15 defines the length of the LAI increase and decline periods (no. of days). Finally, the last four columns give default values of foliar biomass and biogenic VOC emission potentials. See Simpson et al., (2012) for details.

2.2.10 Stomatal conductance

Parameters for the stomatal conductance model, deposition of O_3 and stomatal exchange (DO3SE) must be specified. That are based upon the ideas in Emberson *et al.*, 2000, and are discussed in Simpson and Emberson, 2006 and Tuovinen et al. 2004.

The ASCII file `Inputs_DO3SE.csv` provides land-phenology data of each landuse type for stomatal conductance calculations. The data are summarised in Table 5.1 in Chapter 5 of the EMEP Status Report 1/2003 Part I (Simpson *et al.*, 2003).

The file contains a **header** with the contents of the file, with different factors needed for each of the landuse classes used in the EMEP/MSC-W model. The first two columns represent the landuse code (which are consistent with those in `Landuse.Input` file) and name. The next 22 values are different phenology factors.

2.2.11 Photo-dissociation rates

The photo-dissociation rates (J-values) are provided as lookup tables. The method is previously described in Jonson *et al.*, (2001). J-values are provided as clear sky, light cloud and dense cloud conditions, and the model interpolates between these according to cloudiness from the meteorological input data. In the lookup tables data are listed for every 10 degree latitude at an interval of 1 degree zenith angle at every model height.

For the two types of cloud conditions there are one ASCII file averaged for each season (`SS`); 01, 02, 03 and 04. For light cloud the four seasonal files are called `jcl1kmSS.dat`, for dense cloud conditions the four seasonal files are called `jcl3kmSS.dat`, and then for clear sky four files called `jclearSS.dat`. In addittion there are two files for June called `jcl1.jun` and `jcl3.jun`.

Each file contains 18 columns. The first column is latitude of zenith angle and then the next 17 are the values for the model levels with the 1/s. For more details about these rates, please read Chapter 7.2 of the EMEP Status Report 1/2003 Part I (Simpson *et al.*, 2003).

2.2.12 Site and Sonde locations for output

The model provides a possibility for extra output data of surface concentration for a set of specified measurement site locations and concentrations for the vertical column above a set of specified locations. These site and sonde locations are listed in the ASCII files `sites.dat` and `sondes.dat` files. These files can be changed by the user, this is described in Section 3.2.

CHAPTER

THREE

OUTPUT FILES

Output files from a model run are written out in either ASCII, or (for most data outputs) in NetCDF format. The different NetCDF files are named after the `runlabel1` parameter set in `modrun.sh`. The model output is written to the same directory as where the runscript where submitted, as described in Section 4.

To check your model run, already prepared model result files can be downloaded using the catalog tool (Section 1.4) as follows:

```
# download the output
catalog.py -R rv4_15 --output
```

Unpacked files are placed in an output directory with model run results for a whole year, and sometimes with a smaller test run for e.g. April.

Table 3.1: List of model output files

Output data files	Short description	Format
`Base_hour.nc`	Gridded hourly values of a	NetCDF
	selection of compounds	
`Base_day.nc`	Gridded daily values of a	NetCDF
	selection of compounds	
`Base_month.nc`	Gridded monthly values of a	NetCDF
	selection of compounds	
`Base_fullrun.nc`	Gridded yearly values of a	NetCDF
	selection of compounds	
`sites_YYYY.nc`	Surface daily values of a	NetCDF[1]
	selection of stations and compounds	
`sondes_YYYY.nc`	Vertical daily values of a	NetCDF[1]
	selection of stations and compounds	
`sites_YYYY.cvs`	ASCII version of `sites_YYYY.nc`	ASCII[2]
`sondes_YYYY.csv`	ASCII version of `sondes_YYYY.nc`	ASCII[2]
Additional files		
`RunLog.out`	Summary log of runs, including total emissions	ASCII
	of different air pollutants per country	
`Timing.out`	Timing log file	ASCII

[1] `YYYY`: year.
[2] Deprecated output.

3.1 Output parameters NetCDF files

Parameters to be written out `Base_day.nc`, `Base_month.nc` and `Base_year.nc` are defined in `My_Derived_ml.f90` and `Derived_ml.f90`. In `My_Derived_ml.f90`, the use can specify the output species (air concentrations, depositions, column values), units and temporal resolution of the outputs (daily, monthly, yearly).

The name of output parameter provides some information about data. The names start with TYPE of the parameter, namely SURF (surface air concentrations), DDEP (Dry deposition), WDEP (Wet deposition), COLUMN (Vertically integrated parameters), Area (Surface area) etc.

For surface air concentrations, the general name pattern is `SURF_UNITS_COMPONENT`. Here, `UNITS` can e.g. be "ug" ($\mu g/m^3$), "ugS" ($\mu g(S)/m^3$), "ugN" ($\mu g(N)/m^3$), or "ppb". Note that the components are classified either as "SPEC" (species) or "GROUP". The content of complex GROUP components can be found in `CM_ChemGroups_ml.f90`.

For column integrated parameters, the names are *COLUMN_COMPONENT_kNLAYERS*, where *NLAYERS* is the number of layers from model top included in the integration. The units for column outputs are "ugm2" ($\mu g/m^2$), "mcm2" ($molec/m^2$) or "e15mcm2" ($10^{15} molec/m^2$).

For dry depositions, given per $1m^2$ of specified landuse, the names look like `DDEP_COMPONENT_m2LANDUSE`, where `LANDUSE` can be either a specific landuse type or a cell average. For wet depositions, the names are `WDEP_COMPONENT`. The units for dry and wet depositions are mg/m^2, $mg(S)/m^2$ or $mg(N)/m^2$.

Surface concentrations, column integrated, wet and dry deposition outputs are defined by the user in `config_emep.nml` file. Surface concentrations and column integrated outputs are described in `OutputConcs_config` namelist, Dry and wet deposition outputs are described in `OutputDep_config` namelist.

`VG_COMPONENT_LANDUSE` are the dry deposition velocities on various landuse types, typically in cm/s.

Table 3.2 lists most of output parameters, providing additional explanation to the complex components. For a complete suit of currently selected output parameters, see provided output NetCDF files, or `My_Derived_ml.f90` module.

Table 3.2: List of output parameters

Parameter name	Short description	Comments
Surface Concentrations		
`SURF_ppb_O3`	O_3 [ppb]	
`SURF_ugN_NO`	NO [$\mu g(N)/m^3$]	Available also in ppb
`SURF_ugN_NO2`	NO_2 [$\mu g(N)/m^3$]	Available also in ppb
`SURF_ugN_HNO3`	HNO_3 [$\mu g(N)/m^3$]	Available also in ppb
`SURF_ugN_NH3`	NH_3 [$\mu g(N)/m^3$]	Available also in ppb
`SURF_ugS_SO2`	SO_2 [$\mu g(S)/m^3$]	Available also in ppb
`SURF_ug_SO4`	SO_4^{2-} [$\mu g/m^3$]	
`SURF_ug_NO3_F`	NO_3^- fine aerosol [$\mu g/m^3$]	As ammonium nitrate
`SURF_ug_NO3_C`	NO_3^- coarse aerosol [$\mu g/m^3$]	Associated with sea salt and mineral dust
`SURF_ug_TNO3`	NO_3^- total [$\mu g/m^3$]	Sum of fine and coarse nitrate
`SURF_ug_NH4_F`	NH_4^+ fine aerosol [$\mu g/m^3$]	As ammonium sulphate and ammonium nitrate
`SURF_ug_SIA`	SIA [$\mu g/m^3$]	Secondary Inorganic Aerosol
`SURF_ug_SIA`	SIA [$\mu g/m^3$]	Secondary Inorganic Aerosol
`SURF_ug_ECFINE`	EC fine [$\mu g/m^3$]	Elemental carbon
`SURF_ug_ECCOARSE`	EC coarse [$\mu g/m^3$]	Elemental carbon
`SURF_ug_PM_OM25`	OM fine [$\mu g/m^3$]	Organic Matter fine aerosol
`SURF_ug_PM_OMCOARSE`	OM coarse [$\mu g/m^3$]	Organic Matter coarse aerosol
`SURF_ug_SEASALT_F`	Sea salt fine aerosol [$\mu g/m^3$]	

Continued on next page

Table 3.2 – continued from previous page

Parameter name	Short description	Comments
SURF_ug_SEASALT_C	Sea salt coarse aerosol [$\mu g/m^3$]	
SURF_ug_SEASALT	Sea salt [$\mu g/m^3$]	Sum of fine and coarse sea salt
SURF_ug_DUST_ROAD_F	Road dust fine aerosol [$\mu g/m^3$]	
SURF_ug_DUST_ROAD_C	Road dust coarse aerosol [$\mu g/m^3$]	
SURF_ug_DUST_WB_F	Windblown dust fine [$\mu g/m^3$]	
SURF_ug_DUST_WB_C	Windblown dust coarse [$\mu g/m^3$]	
SURF_ug_DUST_SAH_F	Saharan dust fine [$\mu g/m^3$]	From Boundary conditions
SURF_ug_DUST_SAH_C	Saharan dust coarse [$\mu g/m^3$]	From Boundary conditions
SURF_ug_DUST_NAT_F	Natural dust fine [$\mu g/m^3$]	Windblown and Saharan
SURF_ug_DUST_NAT_C	Natural dust coarse [$\mu g/m^3$]	Windblown and Saharan
SURF_ug_DUST	Mineral dust [$\mu g/m^3$]	From all sources
SURF_ug_PM10	PM_{10} dry [$\mu g/m^3$]	
SURF_ug_PM10_rh50	PM_{10} wet [$\mu g/m^3$]	PM_{10} particle water at 50 %rh
SURF_ug_PM25	$PM_{2.5}$ dry [$\mu g/m^3$]	Includes fine PM and 27% of coarse NO_3^-
SURF_ug_PM25_rh50	$PM_{2.5}$ wet [$\mu g/m^3$]	$PM_{2.5}$ particle water at 50 %rh
SURF_ug_PM25X	$PM_{2.5}$ dry [$\mu g/m^3$]	Includes fine PM and 27% of coarse NO_3^-, EC and OM
SURF_ug_PM25X_rh50	$PM_{2.5}$ [$\mu g/m^3$]	As PM25X + particle water at 50 %rh
SURF_ug_PMFINE	Fine PM [$\mu g/m^3$]	Sum of all fine aerosols
SURF_ug_PPM25	Primary P|PM25| [$\mu g/m^3$]	Anthropogenic emissions
SURF_ug_PPM_C	Primary coarse PM [$\mu g/m^3$]	Anthropogenic emissions
SURF_ug_PM25_FIRE	$PM_{2.5}$ from forest fires [$\mu g/m^3$]	Sum of BC, OC and rest $PM_{2.5}$
Dry Depositions		
DDEP_SOX_m2Grid	Oxidized sulphur [$mg(S)/m^2$]	For a grid cell landuse area weighted
DDEP_SOX_m2Conif	Oxidized sulphur [$mg(S)/m^2$]	To coniferous forest
DDEP_NOX_m2Grid	Oxidized nitrogen [$mg(N)/m^2$]	For a grid cell landuse area weighted
DDEP_NOX_m2Decid	Oxidized nitrogen [$mg(N)/m^2$]	To decideous forest
DDEP_RDN_m2Grid	Reduced nitrogen [$mg(N)/m^2$]	For a grid cell landuse area weighted
DDEP_RDN_m2Seminat	Reduced nitrogen [$mg(N)/m^2$]	To semi-natural
Wet Depositions		
WDEP_PREC	Precipitation [mm]	
WDEP_SOX	Oxidized sulphur [$mg(S)/m^2$]	
WDEP_SS	Sea salt [mg/m^2]	
Others		
AOD	Aerosol Optical Depth at 550nm	Experimental
Area_Crops_Frac	Area fraction of crops	Available for several landuses
VG_NO3_F_Grid	Dry deposition velocity of fine NO_3^-	Grid cell average
Meteorological parameters		
USTAR_GRID	U^* grid averaged	Available for several landuses
T2m	Temperature at 2m [$°C$]	
rh2m	Fractional relative humidity at 2m	

3.2 Add your own fields

Most standard output can be outputted by adding lines and modifying the parameters in the config_emep.nml file.

The meteorological fields defined in the met array in the MetFields_ml.f90 file, can be retrieved by using the 'MET2D' or 'MET3D' keywords. If a 3D array is requested with the 'MET2D' keyword, only the lowest level is written out.

If you want an array that does not fit in any category, or even make your own special field, you can get it in the output

using the procedure shown below; this will however require that you write in the code and recompile. For 2D fields:

- set a value for `Nspecial2d` in `MetFields_ml.f90` according to the number of new outputs
- write in the code the values of the array special2d(i,j,N), where N is the output number (i.e. 1,2,... Nspecial2d). For instance `special2d(i,j,1)=Grid%invL` in `CellMet_ml.f90`
- include `special2d` array in the same routine by adding the line `use MetFields_ml , only : special2d`
- In `config_emep.nml` include the corresponding line, for instance:

```
'MyinvL','MET2D','special2d1', '-','MyUnit' ,-99,-99,F,1.0,T,'YMD',
```

(and 'special2d2', 'special2d3', ... for additional outputs) For 3D fields replace all the "2d" by "3d".

Detailed emissions by sectors can be obtained by adding the keyword `SecEmisOuPoll`, and specify the pollutants required. For example adding the line:

```
SecEmisOuPoll(1:)  = 'pm25', 'nox',
```

will give you the value for all the 11 SNAP sectors for PM25 and NOx.

To get emissions partitioned into splitted compounds (up to 18), the value `EmisSplit_OUT=.true.` must be set in `My_Derived_ml.f90`, and the code recompiled. (This parameter cannot be set in `config_emep.nml` for now)

3.3 ASCII outputs: sites and sondes

Two main options are available for the output of ASCII files for comparison with measurements or detailed model analysis. These are

sites output of surface concentrations for a set of specified measurement site locations.

sondes output of concentrations for the vertical column above a set of specified locations.

Both sites and sondes are specified and handled in similar ways, in the module `Sites_ml.f90`, so we treat them both together below. Locations are specified in the input files `sites.dat` and `sondes.dat`. The files start with a description of its content followed by a list of the stations. For example, a sondes.dat input file may look like this:

Listing 3.1: Minimum `modrun.sh` example.

```
# "Sondes: names, locations, elevations"
# "Area: EMEP-Europe"
# "ix: x coordinate"
# "iy: y coordinate"
# "lev: vertical coordinate (20=ground)"
: Units index
: Coords LatLong
: VertCoords EMEPsigma
: DomainName NA
#
name lat long lev #HEADERS
-    deg deg level #SKIP
#DATA:
Uccle              50.80    4.35  20  ! comment
Lerwick            60.13   -1.18  20  ! comment
Sodankyla          67.39   26.65  20  ! comment
Ny_Alesund         78.93   11.88  20  ! comment
Hohenpeissenberg   47.80   11.02  20  ! comment
```

The first line in each file is a header with file content. Then, the contents are described in more detail. Text strings after # are just clarifying comments. 'Area', e.g., is the domain to which the stations belong, e.g. 'Northern Hemisphere'.

Text after : is read in by the model:

Units Either 'deg' (degrees) or 'index' (model grid indices).

Coords Either 'LatLong' (latitudes/longitudes) or 'ModelCoords' (indices of the grid box in which the station is located).

VertCoords Vertical coordinate system that is used in the model (usually 'EMEPsigma').

Both `sites.dat` and `sondes.dat` files are optional, but recommended. The species and meteorological data requested for site and sonde output are specified in `My_Outputs.f90` by the use of arrays. Only a few met fields are defined so far but more can be added into `Sites_ml.f90` as required.

The output files `sites_2015.csv` and `sondes_2015.csv` are comma separated files that can be read by excel. If you include the whole year, or the 31st December, `sites_2016.csv` and `sondes_2016.csv` are also included in the output.

CHAPTER

FOUR

SUBMITTING A RUN

In this chapter we provide detailed information on how to run the regional EMEP/MSC-W model for two different types of simulations, namely:

Base run This is the default set up for yearly transport model calculations in $50 \times 50 km^2$ grid.

Scenario run A run with reduced emissions from a particular country or several countries is called a "Scenario run". It is the basic type of run for the source-receptor calculations.

Details about the submission of these different types of runs are given below. We suggest that users test the "Base run" first, which can be done without significant changes in the code itself. One can also use the outputs of such a run in the future as a reference run for the other simulations.

4.1 Base run

This is an example of a minimum `modrun.sh` script to run the model.

Listing 4.1: Minimum `modrun.sh` example.

```
#!/bin/bash

# Minimalistic script for run the Unified EMEP model
GRID=EECCA
NLEV=20lev

# Link the input data
inputdir=../
ln -s $inputdir/input/* .         # input files except meteorology
ln -s $inputdir/input/$NLEV/* .   # num-level dependant input
ln -s $inputdir/input/$GRID/* .   # grid dependant input

# Run the model
mpiexec $inputdir/code/Unimod

# Clean the links to the input data
find -maxdepth 1 -type l -delete
```

This bash shell script is designed so that users can easily adapt it to fit their needs. It contain the minimum information required to run the EMEP/MSC-W model. The script should be self-explanatory. It assumes one directory for input data other than meteorology data. The metdata for the year, and for January 1^{st} the following year (365 +1 files) is linked directly in the `config_emep.nml` file. You need to set the right paths for the input directories. All the input files in the input directories are linked to the directory you are working from.

4.1.1 config_emep.nml

The model has a namelist system. It is possible to set different constants and flags for running the model. The constants and flags itself is defined in `ModelConstants_ml.f90`, while they are set in the namelist file under `ModelConstants_config` parameter. Some of these are briefly explained in Section 2. Model gets information about running for special cases from this file. The datasets provided are for the EMEP grid EECCA.

The different parameters for the model run are set in the `config_emep.nml` file. In the very beginning of this, the section `INPUT_PARA` has all these variables including the link to the meteorology data. The trendyear can be set to change the boundary emissions for earlier and future years, see the modules `BoundaryConditions_ml.f90` and `GlobalBCs_ml.f90` to understand better what the trendyear setting does. The default setting is the meteorological year you are running for, in this case 2015. The runlabel1 option sets the name of the different output NetCDF files, see Section 3. The `startdate` and `enddate` parameters are set for the time period you want the model to run (YYYY,MM,DD), and you need meteorology data for the period, as shown in Listing 4.2.

Listing 4.2: Basic namelist example; `config_emep.nml` extract.

```
&INPUT_PARA
  GRID      = 'EECCA',
  iyr_trend = 2015,
  runlabel1 = 'Base',
  runlabel2 = 'Opensource_Setup_2017',
  startdate = 2015,01,01,00,
  enddate   = 2015,12,31,24,
&end
&Machine_config
  DataPath(1) = '../input',           ! define 'DataDir' keyword
&end
&ModelConstants_config
  meteo                 = '../meteoYYYY/GRID/meteoYYYYMMDD.nc',
  DegreeDayFactorsFile  = 'MetDir/DegreeDayFactors.nc',
!------------------------------
  EmisDir               = 'DataDir/EECCA',
  emis_inputlist(1)%name= 'EmisDir/gridPOLL', !example of ASCII type
!--------Sub domain x0, x1, y0, y1
  RUNDOMAIN =  36, 100, 50, 150,      ! EECCA sub-domain
&end
```

In Listing 4.2, the model is run for the period 1 January to 10 Januray 2014 and the trend year used is 2014. Output files will be stored with the name 'Base' and the meteorological correspond to the 'EECCA' grid.

It is possible to run the model on a smaller domain than the full regional model domain, as defined by indexes x and y. For the 'EECCA' gtid $x = 1, \ldots, 132; y = 1, \ldots, 159$. To set a smaller domain, use `RUNDOMAIN` variable in the `ModelConstants_config` namelist to idicate the sub-domain indexes. In Listing 4.2, `RUNDOMAIN` defines a subdomain with $x = 36, \ldots, 100; y = 50, \ldots, 150$.

To run the model, the correct path to the EMEP/MSC-W model code has to be set (`mpiexec path_to_the_modelcode/Unimod`).

It is recommended to submit the script as a batch job. Please check the submission routine on the computer system you are running on. In the newer model versions (since 4.0) the number of nodes is set automatically from what is asked for when submitting a job. The approximate time and CPU usage is described in Section 1.2.

When the job is no longer running or in the queue, it is either finished or has crashed for some reason. If the model run crashed, an error message will give information on what was missing or wrong in the routine. If the run was successful, a message

```
++++++++++++++++++++++++++++++++++++++++++++++++
programme is finished
```

will be stated at the end of the log file, before printing the `Timing.out` file.

The model results will be written to this same directory. Please make sure there is enough disk place for the model results. For more information about the model result output files, see Section 3.

If for some reason the model crashed, please check both the log and the error file for any clue of the crash. After fixing the problem the job can be submitted again. If the model has crashed, then the links to the input data are not removed.

The script can also be submitted interactively, and either have the output written to the screen or to named error and output log files. The variables wanted in the output are specified in the `OutputConcs_config`, `OutputDep_config` and in the `OutputMisc_config` parameters respectively for surface concentrations, depositions and some miscellaneous outputs.

4.2 Nesting

The boundary conditions needed for EMEP MSC-W model is provided with the input data. The model can read Boundary conditions data from other models as well. These data has to be in NetCDF format. The boundary conditions needed for EMEP MSC-W model is provided with the input data. The model can read Boundary conditions data from other models as well. These data has to be in NetCDF format.

Different Nesting modes are:

- read the external BC data only,
- produce EMEP BC data from the simulation,
- read the external BC data and produce EMEP BC data,
- using the default EMEP BC data from the input data directory and write out EMEP BC at the end of the simulation,
- read the external BC data only in the beginning of the simulation,
- read external BC at the beginning of the simulation and write out EMEP BC at the end of the simulation.

These options are controlled by the `MODE_READ` and `MODE_SAVE` variables in `Nest_config` namelist, in `config_emep.nml` file. The mode options are:

`MODE_READ`

'NONE' do nothing (default).

'START' read at the start of run.

'FORECAST' read at the start of run, if the files are found.

'NHOUR' read at given `NHOURREAD` hourly intervals. `NHOURREAD` is set in `Nest_config` and should be an integer fraction of 24.

`MODE_SAVE`

'NONE' do nothing (default).

'END' write at end of run.

'FORECAST' write every `OUTDATE(1:FORECAST_NDUMP).OUTDATE` and `FORECAST_NDUMP` are set in `Nest_config`.

'NHOUR' write at given *NHOURSAVE* hourly intervals. `NHOURSAVE` is set in `Nest_config` and should be an integer fraction of 24.

If BC data are read at NHOURREAD intervals from the file defined by template_read_BC in Nest_config. The IC data (entire 3D domain) will be set at start of run from the file defined by template_read_3D in Nest_config.

4.2.1 Write BCs from EMEP MSC-W model

Listing 4.3 shows an example to write every 3 hours into daily Nest/BC files. Output file name is defined by template_write ('BC_YYYYMMDD.nc'), where 'YYYYMMDD' is replaced by corresponding date.

All advected model variables will be written out for the sub-domain defined by out_DOMAIN ($x = 60, \ldots, 107; y = 11, \ldots, 58$). If no out_DOMAIN is given, the model inner domain will be written out.

Listing 4.3: Write BCs configuration example.

```
&Nest_config
  MODE_READ        = 'NONE',          ! do not read external BC
  MODE_SAVE        = 'NHOUR',         ! write BCs
  NHOURSAVE        = 3,               !  every 3 hours
  template_write   = 'BC_YYYYMMDD.nc' !  to your (daily) BC output file
                                      !  (8 records per file)
  !-------- Sub domain for write modes
  out_DOMAIN       = 60,107,11,58,    ! istart,iend,jstart,jend
&end
```

Reduce the size of BC files

The size of the files obtained in a nesting configuration can be very large if the out_DOMAIN is large. If the inner domain is known in advance, only the part matching exactly the part needed to construct the BC can be stored. To achieve this, two parameters have to be passed in &Nest_config:

1. RUNDOMAIN_inner which must match exactly the RUNDOMAIN used in for the inner run
2. MET_inner which should be a link to any metdata of the inner grid, in order to define the projection parameters of the inner grid.

Listing 4.4: Inner domain options for nested BC output example.

```
&Nest_config
[...]
  RUNDOMAIN_inner = 30,70,30,70,
  MET_inner = 'inner_domain/wrfout_d03_YYYY-MM-DD_00:00:00'
&end
```

Note that the file will have the same dimensions, but zeros are put into the unused parts. The NetCDF internal compression will take care of reducing the actual size, as measured by used disc space.

4.2.2 Read BCs from EMEP MSC-W model

Listing 4.5 shows an example to read every 3 hours from the Nest/BC files created previously by running Listing 4.3. Please note that the model sub-domain for a nested run is set by RUNDOMAIN, as shown in Listing 4.2.

Listing 4.5: Read BCs configuration example.

```
&Nest_config
  MODE_READ        = 'NHOUR',         ! read external BC
```

```
  NHOURREAD        = 3,                !   every 3 hours
  template_read_BC = 'BC_YYYYMMDD.nc' !   your (daily) BC input file
  MODE_SAVE        = 'NONE',          ! do not write BCs
&end
```

4.2.3 Read external BCs

Reading BCs from a different model is more involved than the previous example. The vertical axis and variables in the file need to be mapped to the corresponding model variables.

Listing 4.6 shows an example to read every 3 hours from an external BC file. The model will read 3 variables from `MyBC.nc`: O_3, NO, and NO_2. The maping between the `MyBC.nc` variables and the corresponding model variables is defined in the `ExternalBICs_bc` namelist.

Listing 4.6: External BCs configuration example.

```
&Nest_config
  MODE_READ        = 'NHOUR',         ! read external BC
  NHOURREAD        = 3,               !   every 3 hours
  template_read_BC = 'MyBC.nc'        !   from your BC input file
  MODE_SAVE        = 'NONE',          ! do not write BCs
&end
&ExternalBICs_config
  USE_EXTERNAL_BIC = T,
  EXTERNAL_BIC_NAME= 'MyBICScenario', ! variable mapping, see ExternalBICs_bc
  TOP_BC           = T,               ! use top BC also from your BC file
  filename_eta     = 'filename.zaxis',! text file containing the
                                      !   description of your BC file
&end
&ExternalBICs_bc
  description='MyBICScenario','Version name',3,  ! name,version,size
  map_bc=! emep,external,frac,wanted,found,IXADV,
    'O3' ,'O3_VMR_inst' ,1.0,T,F,-1,
    'NO' ,'NO_VMR_inst' ,1.0,T,F,-1,
    'NO2','NO2_VMR_inst',1.0,T,F,-1,
&end
```

4.2.4 Vertical coordinate

In ordet the determine the vertical levels on the external BC file ('MyBC.nc' in Listing 4.6), the following checks will take place in the following order:

1. η **(eta) coordinate:** If the variable $hyam$ (hybrid `a` coefficient at layer midpoint) is found, η is calculated from $hyam$ and $hyam$ variables on the file.
2. σ **(sigma) coordinate:** If the vertical level is indexed by variable `k`.
3. η **coordinate:** If the file defined by variable `filename_eta` exist ('filename.zaxis' in Listing 4.6), η is derived from `vct` and `vctsize` variables on the file.
4. **pressure coordinate:** If the vertical level is indexed by variable `lev`.

Independent of the coordinates of the BC file, the BC levels will be interpolated into EMEP model levels. If the BC file level structure is not recognized, and there is no `filename_eta` provided, the model will write an error message and stop execution.

An example of the `filename_eta` for EMEP model levels is given below. Here the `vct` variable describes the model level boundaries in hybrid eta coordinate:

Listing 4.7: `filename_eta` for EMEP model levels.

```
zaxistype = hybrid
size      : 20
name      : k
longname  : vertical sigma coordinates
units     : sigma_level
levels    : 0.0200 0.0600 0.1000 0.1425 0.1950 0.2635 0.3470
        0.4365 0.5215 0.5990 0.6695 0.7330 0.7895 0.8390
        0.8815 0.9170 0.9455 0.9670 0.9820 0.9940
vctsize   = 42
vct       =
        10000. 09600. 09200. 08800. 08350. 07750. 06980.
        06080. 05190. 04380. 03640. 02970. 02370. 01840.
        01380. 00990. 00670. 00420. 00240. 00120. 00000.
        0.0000 0.0400 0.0800 0.1200 0.1650 0.2250 0.3020
        0.3920 0.4810 0.5620 0.6360 0.7030 0.7630 0.8160
        0.8620 0.9010 0.9330 0.9580 0.9760 0.9880 1.0000
```

`vct` is the vertical coordinate table describing the hybrid `a` and `b` paramters at the layer interfaces in η coordinate system ($hyai$ and $hybi$). They must respect the following constraints:

- $hyai_1 = 0$ and $hybi_1 = 1$;
- $hyai_0 = P_t$ and $hybi_0 = 0$.
- P_t is the pressure at top of the model domain.

In this file, the first 21 values in `vct` represent $hyai$ and the remaining 21 represent $hybi$ values in hPa.

4.2.5 Variable mapping

The variables to be used from the external boundary condition data are given in the `ExternalBICS_bc` namelist in the `config_emep.nml` file. In Listing 4.6, *map_bc* maps 3 variables in the 'MyBC.nc' file to 3 model variables (O_3, NO, and NO_2). On each line of *map_bc* contains the 6 elements:

1. Variable name in EMEP MSC-W model, e.g. *'O3'*.
2. Variable name in the External BC data file, e.g. *'O3_VMR_inst'*.
3. External BC component to EMEP component fraction, e.g. *1.0*.
4. Is this component wanted or not, set to *T* or *.true.* to read the variable.
5. Was the BC variable found on the file, will be set by the model on run time.
6. Index of the advected model variable, will be set by the model on run time.

The fraction is helpful, when one has to map a variable that is explicitly not in the model but a fraction of that variable can be mapped to a matching variable in the model.

Unit conversions are delegated to the `Units_ml.f90` module. The supported units are: $\mu g/m^3$, $\mu g(S)/m^3$, $\mu g(N)/m^3$, $\mu g(C)/m^3$, ppb, mixing ratio (mol/mol) and mass mixing ratio (kg/kg).

If the BC data has different units, either convert them into one of the above mentioned units in pre-processing or add the respective conversion factor in the module `Units_ml.f90`.

4.3 Source Receptor (SR) Runs

The EMEP/MSC-W model can be used to test the impact of reduced emission of one or more pollutants from a particular country or a number of countries. Such runs are called "Scenario runs". They are the basic runs for source-receptor calculations.

Emission factors for reduced emissions of pollutants from different sectors and countries can be defined in the input file called `femis.dat`, which can be found in the downloaded input data directory, see section Section 2.2.6.

Listing 4.8: `femis.dat` for a base run.

```
Name  7  sox  nox  co   voc  nh3  pm25   pmco
27    0  1.0  1.0  1.0  1.0  1.0  1.0    1.0
```

This base run example means that there are (1.0), no emission reductions of SO_x, NO_x, CO, VOC, NH_3, $PM_{2.5}$ and PM_{co} from all sectors in the UK.

- The first column of the second line represents the country code. (27 is the code for UK.) The codes for all countries can be found in Fortran module `Country_ml.f90`. Please note that the country code must be the same as in the emission files for the given country. Some countries and areas are divided into sub-areas in the emission files. In this case, one line for each sub-area has to be included into the `femis.dat` file. Countries and areas where emissions are given for sub-areas include the Russian Federation, Germany and all sea areas.
- The second column of the second line represents the sector and "0" means all sectors. Here one can write the appropriate sector code if the emission is reduced only from a specific sector. The description of each sector can also be found in the Fortran module `EmisDef_ml.f90`.
- The columns under the pollutant names show the emission factors for the given pollutants. For example, 0.7 would mean 70% of the original emission, thus 30% reduction.
- The number ("7") following the first text ("Name") in the first line gives the number of pollutants treated in the file.

An example of `femis.dat` file describing 50% reduced emission of SO_x from sector 10 (the emission from agriculture) in the UK is shown in Listing 4.9.

Listing 4.9: `femis.dat` for 50% SO_x reduction from sector 10 over UK.

```
Name  7  sox  nox  co   voc  nh3  pm25   pmco
27   10  0.5  1.0  1.0  1.0  1.0  1.0    1.0
```

For a scenario run `femis.dat` file should be edited manually depending on the level of reduction one would like to test with any pollutant from any sector and/or any country. Several lines can be written in the file.

4.4 Other less used options

There are many internal settings that are set to a default value by the model. These default values can often be overridden by setting specific values for keywords in `config_emep.nml`.

The `RUNDOMAIN` is divided by the model into subdomains which are assigned to each processor. Normally this partioning is done such that the X and Y direction are divided into approximately equal number of parts. For runs in lat lon projection containing poles the Y division is done into one or two parts, so that each processor has the same share of pole regions. The default partitioning can be overrided using the `DOMAIN_DECOM_MODE` parameter in `config_emep.nml`. Recognized values are: *'X*Y'*, *'XY'*, *'X*1'*, *'Y=1'*, *'X'*, *'1*Y'*, *'X=1'*, *'Y'*, *'2*Y'*, *'X=2'*, *'X*2'*, *'Y=2'*. See also in `Par_ml.f90` for details.

The main timestep parameter `dt_advec` can be set manually in config_emep.nml (in seconds, real number). It must be a factor of 3600.

`JUMPOVER29FEB` : if set to T , will not treat the 29th of February even for leap years. (Useful if that date is missing, for instance in Climate runs).

`NETCDF_DEFLATE_LEVEL`: The level of compression used in the NetCDF output files (integer). Negative values means netcdf3 format.

CHAPTER

FIVE

REFERENCES

Andersson-Sköld and Simpson, 1997 Y. Andersson-Sköld and D. Simpson. Comparison of the chemical schemes of the EMEP MSC-W and the IVL photochemical trajectory models *Atmos. Environ.*, 1999, 33, 1111-1129

Emberson *et al.*, 2000 L. Emberson, D. Simpson, J.-P. Tuovinen, M.R. Ashmore and H.M. Cambridge. Towards a model of ozone deposition and stomatal uptake over Europe. EMEP/MSC-W Note 6/2000.

Fagerli *et al.*, 2004 H. Fagerli, D. Simpson and S. Tsyro. Unified EMEP model: Updates. In *EMEP Status Report 1/2004, Transboundary acidification, eutrophication and ground level ozone in Europe. Status Report 1/2004*, pages 11–18. The Norwegian Meteorological Institute, Oslo, Norway, 2004.

Gardner *et al.*, 1997 R.M. Gardner, K. Adams, T. Cook, F. Deidewig, S. Ernedal, R. Falk, E. Fleuti, E. Herms, C.E. Johnson, M. Lecht, D.S. Lee, M. Leech, D. Lister, B. Masse, M. Metcalfe, P. Newton, A. Schmitt, C. Vandenbergh and R. Van Drimmelen. The ANCAT/EC global inventory of NO:math:_*x* emissions from aircraft. *Atmos. Environ.*, 31 (12): 1751-1766 , 1997.

Jonson *et al.*, 2001 J.E.Jonson, J.K.Sundet and L. Tarrasón. Model calculations of present and future levels of ozone and ozone precursors with a global and a regional model. *Atmos. Environ.*, 35: 525-537, 2001.

Köhler *et al.*, 1995 I. Köhler, R. Sausen and G. Klenner. NO:math:_*x* production from lightning. The impact of NO:math:_*x* emissions from aircraft upon the atmosphere at flight altitudes 8-15 km (AERONOX), edited by U. Schumann, final report to the Commission of the European Communities, Deutch Luft und Raumfart, Oberpfaffenhofen, Germany, 343–345, 1995.

Logan, 1998 J. A. Logan. An analysis of ozonesonde data for the troposhere: Recommendations for testing 3-D models and development of a gridded climatology for troposheric ozone. *J. Geophys. Res.*, 10 (D13):16, 115–116, 149, 1998.

Passant, 2002 Passant NR., Speciation of UK emissions of non-methane volatile organic compounds. AEA Technology Report ENV-0545, Culham, Abingdon, United Kingdom, 2002.

Simpson *et al.*, 2003 D. Simpson, H. Fagerli, J.E. Jonson, S. Tsyro, P. Wind and J.-P. Tuovinen. *The EMEP Unified Eulerian Model. Model Description. EMEP MSC-W Report 1/2003*. The Norwegian Meteorological Institute, Oslo, Norway, 2003.

Simpson and Emberson, 2006 D. Simpson and L. Emberson. Ozone fluxes – Updates. In *Transboundary Acidification, Eutrophication and Ground Level Ozone in Europe from 1990 to 2004 in support for the review of the Gothenburg Protocol. EMEP Status Report 1/2006*, pages 63–79. The Norwegian Meteorological Institute, Oslo, Norway, 2006.

Simpson *et al.*, 2012 Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, Á., and Wind, P. The EMEP MSC-W chemical transport model – technical description *Atmos. Chem. Physics*, 2012, 12, 7825-7865

Tuovinen *et al.*, 2004 Tuovinen, J.-P.; Ashmore, M.; Emberson, L. & Simpson, D. Testing and improving the EMEP ozone deposition module Atmos. Environ., 2004, 38, 2373-2385

Wiedinmyer *et al.*, 2011 Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., and Soja, A. J.: The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning, Geosci. Model Dev., 4, 625–641,doi:10.5194/gmd-4-625-2011, 2011