Documentation of the chemistry-transport model

Chimere

[version V200310+]

Last update: December 16, 2003

— Preliminary version —

Introduction

The CHIMERE multi-scale model is primarily designed to produce daily forecasts of ozone, aerosols and other pollutants and make long-term simulations for emission control scenarios. CHIMERE runs over a range of spatial scale from the regional scale (several thousand kilometers) to the urban scale (100-200 Km) with resolutions from 1-2 Km to 100 Km.
CHIMERE proposes many different options for simulations which make it also a powerful research tool for testing parameterizations, hypotheses. Its use is relatively simple provided input data is correctly provided. It can run with several vertical resolutions, and with a wide range of complexity. It can run with several chemical mechanisms, simplified or more complete, with or without aerosols.

- This documentation is freely available at the following internet adresse:

`http://euler.lmd.polytechnique.fr/chimere/V200310+/`

- CHIMERE is distributed under the GNU General Public License
- Copyright (C) 2003 Institut Pierre-Simon Laplace (C.N.R.S.), INERIS, LISA (C.N.R.S.)

- CHIMERE authors and contacts:

  - ☐ Bernard AUMONT, Matthias BEEKMANN, Bertrand BESSAGNET, Nadège BLOND, Claude DEROGNAT, Alma HODZIC, Cécile HONORÉ, Laurent MENUT, Jean ROUX, Hauke SCHMIDT, Robert VAUTARD
  - ☐ chimere@merlin.lmd.polytechnique.fr

Contents

1 Short Description of the model **5**
- 1.1 Purpose . . . 5
- 1.2 Input data required . . . 5
- 1.3 OS, performance and software . . . 5
- 1.4 The physics and numerics in short . . . 6
- 1.5 What can be done with . . . 6
- 1.6 Current and previous applications . . . 6

2 Model Documentation **7**
- 2.1 Main options . . . 7
- 2.2 Model geometry . . . 7
  - 2.2.1 Horizontal grid . . . 7
  - 2.2.2 Vertical grid . . . 7
  - 2.2.3 Grid nesting . . . 8
- 2.3 Transport and mixing . . . 9
  - 2.3.1 Horizontal transport . . . 9
  - 2.3.2 Vertical transport . . . 9
  - 2.3.3 Turbulent mixing . . . 9
- 2.4 Aerosol processes . . . 10
- 2.5 Land use and related parameters . . . 11
- 2.6 Meteorological input data . . . 11
- 2.7 Boundary concentrations . . . 11
- 2.8 Initial concentrations . . . 12
- 2.9 Chemistry . . . 13
  - 2.9.1 Chemical mechanisms (gas phase) . . . 13
  - 2.9.2 Species list of the reduced MELCHIOR2 gas-phase chemical mechanism . . . 13
  - 2.9.3 Aqueous-phase chemistry . . . 17
  - 2.9.4 Heterogeneous chemistry . . . 18
  - 2.9.5 Secondary organic aerosol chemistry . . . 18
- 2.10 Clouds and impact on photolysis . . . 18
- 2.11 Emissions . . . 19
  - 2.11.1 Anthropic emissions . . . 19
  - 2.11.2 Biogenic emissions . . . 20
- 2.12 Dry deposition . . . 20
  - 2.12.1 For gases: . . . 20
  - 2.12.2 For aerosols: . . . 21
- 2.13 Wet scavenging . . . 21
- 2.14 Time numerical solver . . . 22

3 Technical Documentation **25**
- 3.1 General structure . . . 25
  - 3.1.1 Installation . . . 25
  - 3.1.2 Software-languages-systems . . . 25
  - 3.1.3 General software structure . . . 25
  - 3.1.4 Directories . . . 27
  - 3.1.5 Model Files . . . 27
- 3.2 Steps to perform a simulation . . . 29
- 3.3 CHIMERE code structure . . . 30
  - 3.3.1 The top calling script . . . 30
  - 3.3.2 The temporary directory . . . 30
  - 3.3.3 Files and programs . . . 30
  - 3.3.4 Routine calls sequence for the gas-phase version . . . 30
  - 3.3.5 Routine calls sequence for the aerosol version . . . 31

3.3.6 Input files to chimere . . . 31
3.3.7 Output files . . . 33
3.4 The chemical preprocessor: CHEMPREP . . . 33
3.4.1 Conditions on the mechanism . . . 33
3.4.2 Syntax of the REACTIONS file . . . 35
3.4.3 Syntax of the PHOTORATES file . . . 35
3.4.4 Files produced . . . 35
3.5 Meteorological interface . . . 35
3.5.1 First stage . . . 36
3.5.2 Second stage: the diagmet diagnostic model . . . 39
3.6 Land-use interface . . . 40
3.6.1 Model domains . . . 40
3.6.2 Output files . . . 41
3.6.3 Land-use types . . . 41
3.6.4 The prepdomain interface for Europe . . . 41
3.6.5 Extension to other areas . . . 42
3.6.6 Steps for the definition of a new CHIMERE domain . . . 42
3.7 Anthropic emission interface . . . 43
3.8 Boundary conditions interface . . . 44
3.9 Initial Conditions . . . 45
3.10 Running the model . . . 46
3.10.1 Running the model directly without the top calling script . . . 46
3.10.2 Running the model with the top calling script . . . 46
3.11 Visualization and post-processing . . . 46
3.11.1 The output files . . . 46
3.11.2 The mapit visualization example program . . . 47

4 References using CHIMERE **49**
4.1 Peer-Reviewed Articles . . . 49
4.2 Submitted manuscripts . . . 50
4.3 Other documents . . . 50

Chapter 1

Short Description of the model

1.1 Purpose

The CHIMERE multi-scale model is primarily designed to produce daily forecasts of ozone, aerosols and other pollutants and make long-term simulations for emission control scenarios. CHIMERE runs over a range of spatial scale from the regional scale (several thousand kilometers) to the urban scale (100-200 Km) with resolutions from 1-2 Km to 100 Km. On this server, documentation and source codes are proposed for the complete multi-scale model. However most data are valid only for Europe and should be revisited for applications on other continents.
CHIMERE proposes many different options for simulations which make it also a powerful research tool for testing parameterizations, hypotheses. Its use is relatively simple provided input data is correctly provided. It can run with several vertical resolutions, and with a wide range of complexity. It can run with several chemical mechanisms, simplified or more complete, with or without aerosols.
The Figure below shows the example of the ozone simulation for August 8th, 2003 at 14h UT, during the exceptional heat wave that occured over Western Europe. The gray dots show monitoring stations where the 180 ug/m3 exceedance is reached at this exact time.

1.2 Input data required

CHIMERE is a portable model which has been adapted to various types of input data. It requires meteorological data, boundary conditions, land-use information, emissions. In all cases but emissions, a default solution is proposed for the user. Refinements on input data have to be made by the user.
Meteorology: Meteo data is not provided on the server, but a CHIMERE interface for ECMWF and for the MM5 mesoscale is proposed. MM5 (free software [http://box.mmm.ucar.edu/mm5]) can be used combined with AVN/NCEP (as forcing) public weather forecasts. ECMWF data or forecasts are not public.
Boundary conditions: A set of boundary conditions from the MOZART II model is proposed as a default solution. Data is graciously provided by MPI, Hamburg, thanks to M. Schultz, G. Brasseur, C. Granier and D. Niehl. This allows tropospheric simulations below 200 hPa.
Land-use: CHIMERE needs input landuse information, as well as biogenic emission potenials, based on land cover. The default proposed land use is the GLCF data base [http://glcf.umiacs.umd.edu]). The data for deriving biogenic emission potentials only cover Europe. It needs to be completed for other continents.
Emissions: Emissions are usually coming from local, nonpublic databases. Here, for model testing, emissions are proposed for Western Europe at a resolution of 1/2 degree, and for the region (150Kmx150Km) around Paris.

1.3 OS, performance and software

The model has only been tested on PCs under LINUX. It should also easily work on SUN and other UNIX systems with minor changes. Most changes should be due to the binary data interfaces. Please report to us all problems. A typical time for simulation on a PC, using the European version with a 65x33x8 grid, is 5 minutes for one simulated day with the gas-phase version, and 30 minutes with the aerosol version. It can then easily simulate entire seasons or years. A minimal setting of 1 Gb RAM is necessary. Software required is Fortran compiler (g77 is really great

and free!) for the gas-phase version. A F90 compiler is needed for running the aerosol version, in this release. In future releases, the aerosol version will be written in F77. Asimple graphical interface is available using the GMT freeware (to be installed).

1.4 The physics and numerics in short

- The chemical mechanism (MELCHIOR) is adapted from the original EMEP mechanism.
- Photolytic rates are attenuated using liquid water or relative humidity
- Boundary layer turbulence is represented as a diffusion (Troen and Mahrt, 1986, BLM)
- Vertical wind is diagnosed through a bottom-up mass balance scheme.
- Dry deposition is as in Wesely (1989). Wet deposition is included
- Six aerosol sizes represented as "bins" in the model.
- Aerosol thermodynamic equilibrium is achieved using the ISORROPIA model.
- Several aqueous-phase reactions considered
- Secondary organic aerosols formation considered
- Advection is performed by the PPM (Piecewise Parabolic Method) 3d order scheme for slow species.
- The numerical time solver is the TWOSTEP method.

1.5 What can be done with

- Use the model for simulating episodes or long-term periods.
- Use the model for operational forecasting.
- Use the model for studying emission scenarios.
- sensitivity tests: CHIMERE is an easy-to-modify code which allows testing of various processes and making sensitivity tests. For instance the influence of new chemical reactions can be tested.
- Model development: It is the wish of the authors that CHIMERE be an open model, hence all model modifications in the direction of larger realism are welcome. There are a number of such improvements still to be done.

1.6 Current and previous applications

- Operational air quality forecasts over Western Europe .....[http://prevair.ineris.fr]
- Operational Forecasts over:
  - ☐ The Paris Region .....[http://www.airparif.asso.fr]
  - ☐ The Alsace Region .....[http://www.atmo-alsace.net]
  - ☐ The Lyon Region .....[http://www.atmo-rhonealpes.org]
  - ☐ The Toulouse Region .....[http://www.oramip.org]
- European model intercomparisons
- European projects on model comparisons of emission scenarios impact: CITYDELTA, EURODELTA
- Simulations and forecasts over other countries: Thailand, China and Portugal.
- Ozone representation using data assimilation.
- Sensitivity studies using the adjoint model.
- Many other research projects.

Chapter 2

Model Documentation

2.1 Main options

CHIMERE is an air-quality simulation model with several options making it a flexible tool for detailed episode analysis as well as forecasting or long-term simulations. It can be adapted to many meteorological input data. CHIMERE is invoked by a **top calling script** named chimere-[...] where the options for the simulations/forecasts are defined. This section lists these main options:

- The user must choose between the **gas-phase** or the **aerosol** version. This is done in the top calling script by setting the value of parameter "model" [=gas or =aero]. The gas-phase version is much lighter in terms of computational cost (about 4-5 times faster) than the aerosol version. The aerosol version requires at present a F90 compiler, not the gas-phase version. The source codes are placed in different directories: src_gas for gas-phase and src_aero for aerosols.
- The chemical mechanism is (in the very near future) optional: it will contain mechanisms MELCHIOR1 (complete MELCHIOR mechanism, Lattuati (1997), MELCHIOR2 (reduced MELCHIOR mechanism, Derognat, (1999) and MoCA Aumont et al. (19XX) and RACM in future versions. The mechanism option is given in the top calling script by the "meca" parameter (=MELCHIOR2, =MELCHIOR1)
- The meteorological data to be used is a main option. In the present version we propose an interface for the MM5 model or for data from the ECMWF model. The top script parameter is "meteo". If the user's data come from an other model or are given in another format, user should build his own interface and name it meteo=[new_interface_name].

2.2 Model geometry

2.2.1 Horizontal grid

The model uses any regular rectangular horizontal grid system as presneted in the Figure 2.1 . The grid is entirely defined by a user-provided coordinate file named COORD_[domain] containing the geographical coordinates of the grid cell centres. From that file the cell edge lengthes are calculated and used in the calculation of transport fluxes. The cell surface is also calculated. All these calculations take into account sphericity to first order. One should not use grid cells with too large dimensions (less than, say, 2 degrees or 200 Km). All geometrical definitions and calculations are set up at once in the initialization routine **Binigeom.f** in the src_[model] directories.
The horizontal grid is generally derived from the grid on which the emission inventory is given, in order not to interpolate (and then smooth) emissions. As input to CHIMERE, emissions must be given on the model grid. There is no particular need for the grid to coincide with the meteorological or land-use grid. Interpolation programs are provided (in particular for MM5 outputs).

2.2.2 Vertical grid

The model uses any number of vertical layers, described in hybrid sigma-p coordinates. The pressure in hPa at the top of each layer k is given by the following formula:

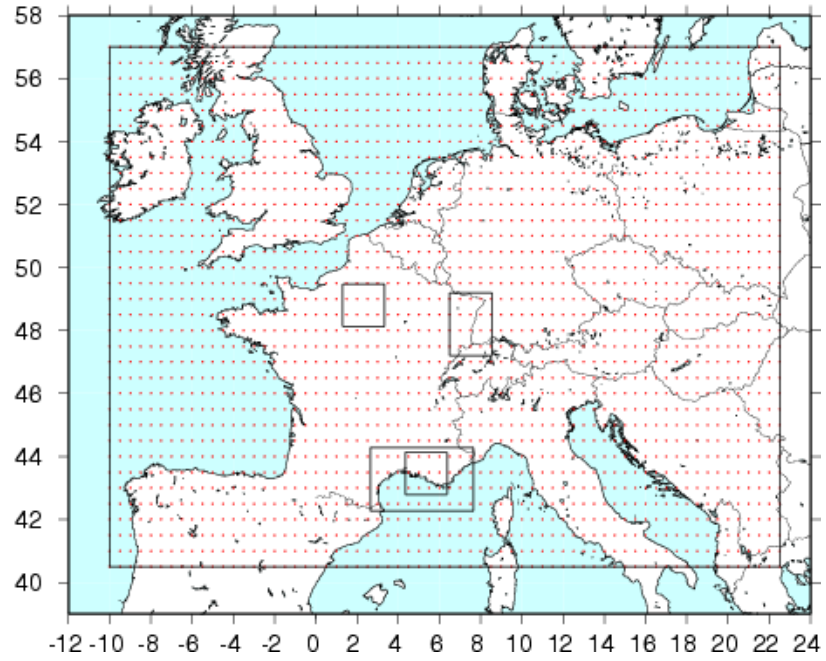


Figure 2.1: *CHIMERE domains*

$$P_k = a_k 10^5 + b_k P_{surf}$$

where P_{surf} is the surface pressure and the coefficients a_k and b_k are given.
The hybrid coefficients a_k and b_k are provided by the user as an input file, VCOORD.[vgname], in the data_[model]_[meca] directory. [vgname] is a user-defined name. As an option, the vertical grid can be defined automatically given the top pressure and sigma at the top of the surface layer. Then the resolution varies upward as a geometric progression. These options are defined in the top calling script.

Recommandations

- When defining manually the hybrid coefficients the user must verify that under all possible meteorological conditions these layers do not cross one another or get mixed up.
- It is highly recommended that top level be a pressure level (b_k = 0), since it has been found to minimize mass loss through top boundary.
- Using the MOZART II boundary conditions, one should not use a model top above 200 hPa
- In the present formulation for cloud/radiation/photolysis, it is assumed that the model domain is "below the clouds", i.e. no cloud albedo is taken into account. This puts constraints on model top to be not too high. This will be changed in future releases.
- Deep convection mass fluxes are not taken into account in the present formulation. Again, this will be changed in future releases, but puts also constraints on model top.
- For these reasons model ozone and other gases and particles should not be considered too seriously far above the boundary layer. In the present version the model top should not exceed 500 hPa.

2.2.3 Grid nesting

Model simulations can be nested in a one-way mode. That is, a coarse CHIMERE simulation can give boundary conditions to a high-resolution simulation. Nested simulations must be performed one after the other. There is no need for any correspondence or relation between the coarse and the fine horizontal grids. However the vertical grids must correspond. In fact, a nested grid can have a higher vertical resolution, by vertical grid subdivisions defined in the top calling script. These subdivided layers are equal-depth layers within coarse-grid layers. This allows to

perform a coarse simulation wiith moderate vertical resolution and a high-resolution simulation with accuracy in all three directions.

2.3 Transport and mixing

2.3.1 Horizontal transport

The horizontal transport is taken into account by calculating pollutant fluxes at the boundaries of each model cell. For a given pollutant, let us denote by C_{ij} the concentration at the center of each model cell, and U_{ij} and V_{ij} the associated wind components (zonal+meridian). First, wind is calculated at each of the 4 cell boundaries, taking into account only the normal component, by averaging the values of this wind component between the cell and its adjacent one. Let us assume that we are dealing with the interface between a cell i,j and its western one i-1,j. The normal component of the wind is thus **U**=1/2(U_{ij}+U_{i-1j}), and the thickness of the interface is **H**=1/2(H_{ij}+H_{i-1j}). The flux **F** used for changing concentration in cell i,j is then calculated as:

$$F = U \frac{H}{H_{ij}} \frac{C}{L_{ij}}$$

which has the dimension of a concentration time derivative, and where **C** is the pollutant concentration at the interface, and L_{ij} the zonal extent of the cell.
The concentration **C** is calculated either by the **UPWIND** method or by the **PPM** (Parabolic Piecewise Method, Collela and Woodward,198?), depending on the pollutant. In order to reduce computer time, PPM is used only for a reduced set of "slow" species. These species are listed in the **PPM_SPECIES** input file. All the other ones are transported using the **UPWIND** method (that is, using **C**=C_{i-1j} when **U**>0 and **C**=C_{ij} when **U**<0).
Since most large-scale weather models do not contain any "urban parametrization", the possibility of correcting the wind speed in the surface layer (due to increased roughness) in cities is offered. This will automatically be balanced by a vertical wind component calculated in the mass balance (see vertical transport below). This correction has however no effect at continental scale where anyway the fraction of urban area in model grid cells remains very small. It has effects on urban versions of the model, and mostly on primary pollutants. The urban parameterization option switch lies in the METINFOS file to be given as input to diagmet.f. It is advised to use the switch for ECMWF data but not for MM5 data.
Horizontal mass fluxes (independent of species, first order) are calculated first at each coarse time step in routine **Bhtransport.f**, then the complete calculation using the species concentration and PPM is performed at each fine time step in routine **Btransmix.f**.

2.3.2 Vertical transport

Vertical transport is assumed to balance horizontal mass divergence/convergence. The scheme used is the first-order UPWIND scheme. More precisely, the algorithm is as follows: In the lowest layer, the mass imbalance is calculated by summing up all incoming mass fluxes and substracting all outgoing mass fluxes, and a vertical wind through the interface between the lowest and the second layer is calculated to compensate this imbalance. Then the same operation is performed on the second layer, including now in the mass imbalance the fluxes at the cell bottom interface. Then all layers are scanned upward to the top of the model. Vertical transport mass fluxes (independent of species, first order) are calculated first at each coarse time step in routine **Bvtransport.f**, then the complete calculation using the species concentration is performed at each fine time step in routine **Btransmix.f**.

2.3.3 Turbulent mixing

Horizontal turbulent fluxes are not considered. Vertical turbulent mixing takes place only in the boundary-layer. The formulation uses K-diffusion following the parameterization of Troen and Mahrt (1986), without counter-gradient term. In each model column, diffusivity K(z) is calculated as:

$$K_z = k w_s z \left(1 - \frac{z}{h}\right)^{1/3} \tag{2.1}$$

where w_s is a vertical scale given by similarity formulae.

- In the stable case (surface sensible heat flux < 0): $w_s = u_*/(1 + 4.7z/L)$
- in the unstable case: $w_s = (u_*^3 + 2.8ew_*^3)^{1/3}$

where e = max(0.1,z/h), **L** is the Obukhov Length, w_* is the convective velocity scale, u_* the friction velocity and h the boundary layer height. A minimal K(z) is assumed, with a value of 0.01 m^2/s.
The boundary layer height h can be either directly provided by the meteorological model or recalculated in the **diagmet.f** meteo diagnostic model. In the latter case the boundary layer height is considered as the maximum of the Troen and Mahrt (1986) boundary layer height calculaated from the Richardson number profile, as the lowest altitude where R_i = 0.5, and a more convectively-based boundary layer height calculation. The latter is based on a simplified and diagnostic version of the approach of Cheinet (2002). It consists in the resolution of the (dry) thermal plume equation with diffusion. The in-plume vertical velocity and buoyancy equations are solved and the boundary layer is taken as the height where vertical velocity stops. Thermals are initiated with a non-vanishing vertical velocity and potential temperature departure, depending on the turbulence similarity parameters in the surface layer.
At each interface between layers k andk+1, one calculates an equivalent turbulent vertical velocity V_k and vertical mixing is expressed in layer k by the mass flux:
Net Flux = $\mathbf{V}_k(\mathbf{C}_{k+1}.\mathbf{D}_k/\mathbf{D}_{k+1} - \mathbf{C}_k)/\mathbf{H}_k$,
where **C** denotes concentrations, **D** densities and **H** layer thicknesses. Density ratios are applied since turbulent mixing must conserve mass in each model cell. An equivalent formula is used for the upper layer. The turbulent velocity V_k is deduced from K(z) using:

$$V_k = \frac{K_z}{(\frac{1}{2}(H_k + H_{k+1}))} \tag{2.2}$$

When the top of the boundary layer is above the model top, H_{k+1} is replaced by h - <model-top> in the previous equation.
The K(z) profile is calculated in routine Bmixing.f, as well as the turbulent velocity profile. Vertical mixing mass fluxes (independent of species, first order) are calculated first at each coarse time step in routine **Bvtransport.f**, then the complete calculation using the species concentration is performed at each fine time step in routine **Btransmix.f**.

2.4 Aerosol processes

CHIMERE has an "aerosol" version consisting in the gas-phase version with MELCHIOR2 chemistry, together with a sectional aerosol module. This module accounts for 7 species (primary particle material, nitrate, sulfate, ammonium, biogenic secondary organic aerosol (SOA), anthropogenic SOA and water). Potentially, Chloride et Sodium can be included (high computing time). In its initial version the module uses 6 bins from 10 nm to 40 Âţm. Physical processes into account are described below. See also sections on chemistry (aqueous-phase chemistry, heterogenous chemistry and SOA formation) below. Dry and wet scavenging/deposition are also described in separate sections.
The other physical processes accounted for are:

- **Coagulation** : Classical theory already used by Gelbard and Seinfeld (1980).
- **Absorption** : Absorption is taken into account for both inorganic and organic species but their processing is different. For inorganic species, equilibrium concentrations are computed with the thermodynamic module ISORROPIA[1] (or a tabulated version) presented in Nenes et al. (1998). For secondary organic species, equilibrium concentrations are calculated through a temperature dependent partitioning coefficient (Pankow, 1994). Please note that the ISORROPIA code is included in CHIMERE, please check the ISORROPIA web site for downloading the latest version.
- **Nucleation** : The parameterization of Kulmala et al. (1998) for sulfuric acid nucleation is used. This process, favoured by cold humid atmospheric conditions, affects the number of ultrafine particles. The nucleated flux is added to the smallest bin in the sectional distribution. Although nucleation of condensable organic species has been clearly identified in many experimental studies (Kavouras et al., 1998), there is no available parameterization. Since the sulfuric acid nucleation process competes with absorption processes, it is expected to occur in weakly particle polluted conditions.

[1] http://nenes.eas.gatech.edu/ISORROPIA

2.5 Land use and related parameters

2.6 Meteorological input data

The CHIMERE model requires meteorological input variables; these are:

- **3D data:**
  - ☐ Horizontal wind (transport)
  - ☐ Temperature (chemistry,thermodynamics)
  - ☐ Density (chemistry and transport)
  - ☐ Specific humidity (chemistry/thermodynamics)
  - ☐ Height of model layers (model geometry)
  - ☐ Cloud liquid water content (for the aerosol version, aqueous-phase chemistry, scavenging)
- **2D data:**
  - ☐ Temperature at 2m (deposition, biogenic emissions)
  - ☐ Photolysis attenuation in the column due to clouds (chemistry/photolysis)
  - ☐ Boundary layer height for mixing
  - ☐ Friction velocity u^* for mixing and deposition
  - ☐ Obukov length for mixing
  - ☐ Convective velocity scale w^* for mixing
  - ☐ Aerodynamic resistance for deposition
  - ☐ Total Precipitation rate for aerosol scavenging

These data are obtained either directly or through the proposed 2-step chain of preprocessing (see the technical documentation) from meteorological analyses or forecasts from standard variables issued in meteorological simulations. An interface is proposed for MM5 and ECMWF models. This interface contains horizontal and time interpolation, and a diagnostic meteo model (the **src_[model]/diagmet.f** program), allowing diagnostic calculation and vertical interpolation of meteo variables. The model has now been tested using meteorological data from AVN/NCEP, MM5, ECMWF, ARPEGE and ALADIN models.
Among the variables that may be calculated from the interface are:

- Radiation attenuation
- Boundary layer height
- Friction velocity U* (From the Louis 1982 parameterization)
- Aerodynamic resistance (Seinfeld and Pandis, 1998)
- Sensible heat flux (read or recalculated using the Priestley, 1948 formula.
- Obukov length
- Convective velocity W*

The other variables (3D) are vertically averaged into the CHIMERE layers.

2.7 Boundary concentrations

Boundary conditions can be either "external" or given by a coarse-resolution CHIMERE simulation. In the first case and in the present release, only monthly climatologies from the MOZART second-generation global chemistry-transport model (MOZART2) are used Horowitz et al. (2003). The concentration of a few species are considered as constant (but depending on month) throughout the simulation. The MOZART climatologies have been calculated and are provided here graciously by the Max-Planck Institut, Hamburg, by M. Schultz, G. Brasseur, C. Granier and D. Niehl. In the second case, a first coarse run must have been carried out, and the vertical resolution cannot be changed relative to the coarse run. Supplementary layers can be defined by subdividing the coarse-run layers.
The user can also provide a different, independent set of boundary conditions to the model, which can be time-varying or constant during the simulation. A new interface to build data for CHIMERE has to be constructed.

The current interfaces are **preplat_MOZART2.f** and **preplat_CHIMERE.f** for lateral boundaries, and **preptop_MOZART2.f** and **preptop_CHIMERE.f** for top boundaries. See technical documentation for the format of boundary conditions files.
Several other files are needed, for describing the species considered at the boundaries. For MOZART2 lateral or top boundaries, only the following species are present:

- O3: Ozone
- NO2: Nitrogen Dioxide
- CO: Carbon Monoxide
- PAN: Peroxy-Acetyl-Nitrate
- CH4: Methane
- C2H6: Ethane
- HCHO: Formaldehyde
- HNO3: Nitric Acid

The file containing the MOZART2 "lateral" species names is **LAT_SPEC.MOZART2**. The file containing the "top" species names is **TOP_SPEC.MOZART2**. If other interfaces had to be built, two other files like these ones would have to be present in the data_[model]_[meca] directories.
There is no yet coupling to any "global" climatology for aerosols. Therefore the boundary aerosol concentrations are set to 0.
When performing a nest run (option latboun=CHIMERE in the top calling script), the order and the number of species names in file LAT_SPEC.CHIMERE **must be exactly the same as the order of the output species of the coarse run** (file **OUTPUT_SPECIES.[your-choice]**). It is therefore advised to output a certain number of species in the coarse simulation. The recommended minimal set of species for mechanism **meca=**MELCHIOR2 is given in the file **OUTPUT_SPECIES.complete**. Here is the list of compounds passed from the coarse run to the high-resolution run at the boundaries. The list is only valid for the gas-phase model. Check the boundary species for the aerosol version equivalent file in the data_aero_[meca] directory.

- O3 ..... Ozone
- NO2 ..... Nitrogen Dioxide
- NO ..... Nitrogen Monoxide
- PAN ..... Peroxy-Acetyl-Nitrate
- HNO3 ..... Nitric Acid
- SO2 ..... Sulphur Dioxide
- CO ..... Carbon Monoxide
- CH4 ..... Methane
- C2H6 ..... Ethane
- NC4H10 ..... Butane
- C2H4 ..... Ethene
- C3H6 ..... Propene
- OXYL ..... o-Xylene
- C5H8 ..... Isoprene
- APINEN ..... Alpha-Pinene
- HCHO ..... Formaldehyde
- CH3CHO ..... Acetaldehyde
- GLYOX ..... Glyoxal
- MGLYOX ..... Methyl-Glyoxal
- CH3COE .. Methyl-Ethyl Ketone

A high-resolution simulation cannot yet be "vertically nested". Therefore at the top boundary only MOZART2 monthly climatologies can be applied in the present version.
The boundary species names and species are read in routine **Biniboun.f**. Then in practice the boundary concentrations values are put in the edges of the **conc** array containing all concentrations in routine **Blocvalues.f**

2.8 Initial concentrations

There are two possible initialization modes:

- Initialization by reading initial concentrations in a restart file. This file must contain the "end" concentrations of a previous simulation, as provided in the end.[simul-name]. In this case the order of species in the file is as in the **ACTIVE_SPECIES** file. The two consecutive runs must therefore be made **with the same chemical mechanism**.
- Initialisation by interpolating boundary conditions. In this case Initialization by interpolating boundary conditions, if no initial file available.

In practice, the top calling script checks the existence and nonemptiness of a file called **ini.[simul-name]** in the simulation directory, which can be a copy of an "end" file. If it exists and is nonempty, then the initial conditions are used. In the other case, initial conditions are interpolated from boundary conditions. Initial conditions are read and put into arrays **conc** and **conco** in routine **Biniconc.f**.

2.9 Chemistry

2.9.1 Chemical mechanisms (gas phase)

CHIMERE offers the option to include different gas phase chemical mechanisms. The original, complete scheme (Lattuati (1997)), hereafter called MELCHIOR1, describes more than 300 reactions of 80 gaseous species. The hydrocarbon degradation is fairly similar to the EMEP gas phase mechanism (Simpson, 1992). Adaptations are made in particular for low NO_x conditions and NO_x-nitrate chemistry. All rate constants are updated according to Atkinson et al. (1997) and DeMore et al. (1997). Heterogeneous formation of HONO from deposition of NO2 on wet surfaces is now considered, using the formulation of Aumont (2003). In order to reduce the computing time a reduced mechanism with 44 species and about 120 reactions is derived from MELCHIOR (Derognat, 1998), following the concept of "chemical operators" (Carter, 1990, Aumont et al., 1997). This reduced mechanism is called MELCHIOR2 hereafter. The MoCA mechanism (Aumont et al., 1997) has also been implemented into CHIMERE. It contains also about 80 species and it has about 200 reactions.
The chemical mechanisms files are placed in the data_[model]_[meca]. New mechanisms can be introduced. In this case, copy for instance the directory data_gas_MELCHIOR2 onto data_gas_XXXX, and change the **REACTIONS** file for a new set of reactions. Not all forms of reactions are allowed. Arrhenius, Troe, photolysis and few other classical reaction rate types are only allowed. Introducing other types implies some recoding. See the technical documentation (the chemprep preprocessor). A file with altitude-dependent photolysis rates must also be provided, named **PHOTO_PARAMETERS**. Then run the **chemprep** chemical preprocessor to make FORTRAN-readable files.
Photolysis rates are calculated under clear sky conditions as a function of height using the TUV model (Madronich et al. (1998). Then clouds are taken into account in a highly parameterized fashion (see next Section).
Gas-phase chemistry constants is read in initialization routine **Binichem.f**. Then rates are calculated at the coarse time step in routine **Brates.f**, but chemical fluxes are calculated for all species at the fine time step in routine **Bchemistry.f**.

2.9.2 Species list of the reduced MELCHIOR2 gas-phase chemical mechanism

Species	Definition
Inorganic compounds	
O3	ozone
H2O2	hydrogen peroxide
OH	hydroxy radical
HO2	hydroperoxy radical
NO	nitric oxide
NO2	nitrogen dioxide
NO3	nitrogen trioxide
N2O5	dinitrogen pentoxide
HONO	nitrous acide
HNO3	nitric acide
CO	carbon monoxide
SO2	sulfur dioxide
Hydrocarbon species[2]	
CH4	methane
C2H6	ethane
NC4H10	n-butane
C2H4	ethene
C3H6	propene
OXYL	o-xylene
C5H8	isoprene
APINEN	α-pinene
Carbonyls	
HCHO	formaldehyde
CH3CHO	acetaldehyde
CH3COE	methyl ethyl ketone
GLYOX	glyoxal
MGLYOX	methyl glyoxal
CH3COY	dimethyl glyoxal
MEMALD	unsaturated dicarbonyls, reacting like 4-oxo-2-pentenal
MVK	methyl vinyl ketone
MAC	methacroleine
Organic nitrates	
PAN	peroxyacetyl nitrate
CARNIT	nitrate carbonyl taken as α-nitrooxy acetone
ISNI	unsaturated nitrate from isoprene degradation
Organic peroxides	
CH3O2H	methyl hydroperoxide
PPA	peroxy acetyl acid
(Per)oxy radicals	
CH3O2	methyl peroxy radical
CH3COO	peroxy acetyl radical
Operators[3]	
oRO2	representing peroxy radicals from OH attack to C2H5, NCHH10, C2H4, C3H6, OXYL, CH3COE, MEMALD, and MVK
oROOH	representing organic peroxides from oRO2+HO2 reactions
obio	representing peroxy radicals produced by C5H8 and APINEN + OH reaction
obioH	representing biogenic organic peroxides from obio+HO2 and obio+obio reactions

[3] For the concept of "chemical operators" see Carter (1990) and Aumont et al. (1997)

oPAN ..... representing PAN homologue compounds (except PAN)
PANH ..... representing results from oPAN+HO2 reaction
toPAN ..... representing results from oPAN+NO2 reaction
oRN1 . representing organic nitrate peroxy radicals from NO3 attack to C2H4, C3H6, C5H8, APINEN and OH attack to ISNI

Reaction list of the reduced MELCHIOR2 chemical mechanism

Reactions	Kinetic constants	Ref.
Inorganic chemistry		
O_3+NO → NO_2	$Ae^{-B/T}$,A=1.8e-12,B=1370	[1]
O_3+NO_2 → NO_3	$Ae^{-B/T}$,A=1.2e-13,B=2450	[1]
O_3+OH → HO_2	$Ae^{-B/T}$,A=1.9e-12,B=1000	[1]
O_3+HO_2 → OH	$Ae^{-B/T}$,A=1.4e-14,B=600	[1]
NO+HO_2 → OH+NO_2	$Ae^{-B/T}$,A=3.7e-12,B=-240	[1]
NO_2+OH+M→ HNO_3	troe(3.4e-30,0,3.2,4.77e-11,0,1.4,0.30)	[2], n9
HO_2+OH→ H_2O	$Ae^{-B/T}$,A=4.8e-11,B=-250	[1]
H_2O_2+OH→ HO_2	$Ae^{-B/T}$,A=2.9e-12,B=160	[1]
HNO_3+OH→ NO_3	$Ae^{-B/T}$,A=5.5e-15,B=-985	[3], n1
CO+OH→ HO_2+CO_2	$Ae^{-B/T}(300/T)^N$, A=2e-13, B=0, N=1	[3], n1
HO_2+HO_2 → H_2O_2	$Ae^{-B/T}$,A=2.2e-13,B=-740	[1], n1
HO_2+HO_2+H_2O→ H_2O_2	$Ae^{-B/T}$,A=4.52e-34,B=-2827	[1], n1
NO_3+HO_2→ NO_2+OH	4e-12	[1]
NO_3+H_2O_2 → HNO_3+HO_2	2e-15	[3]
NO_3+NO→ 2*NO_2	$Ae^{-B/T}$,A=1.8e-11,B=-110	[1]
NO_2+NO_3→ NO+NO_2	$Ae^{-B/T}$,A=4.5e-14,B=1260	[3]
NO_2+NO_3+M→ N_2O_5	troe(2.7e-30,0,3.4,2e-12,0,-0.2,0.33)	[1], n9
N_2O_5+M→ NO_3+NO_2	troe(1e-3,11000,3.5,9.7e14,11080,-0.1,0.33)	[1], n9
N_2O_5+H_2O→ 2*HNO_3	2.6e-22	[4]
N_2O_5+H_2O+H_2O→ 2*HNO_3	2e-39	[4]
NO+OH+M→ HONO	troe(7.e-31,0,2.6,1.5e-11,0,0.5,0.6)	[1], n9
HONO+OH→ NO_2	$Ae^{-B/T}$,A=1.8e-11,B=390	[1]
NO+NO+O_2→ 2*NO_2	$Ae^{-B/T}$,A=3.30E-39,B=-530.0	[1]
NO_2→ HONO	0.5*depo(NO_2)	[18]
SOx chemistry		
SO_2+CH_3O_2→ H_2SO_4+HCHO+HO_2	4e-17	[3]
SO_2+OH+M→ H_2SO_4+HO_2	troe(4e-31,0,3.3,2e-12,0,0,0.45)	[3], n9
OH attack to hydrocarbons		
CH_4+OH→ CH_3O_2	$Ae^{-B/T}$,A=2.3e-12,B=1765	[1]
C_2H_6+OH→ CH_3CHO+oRO_2	$Ae^{-B/T}$,A=7.9e-12,B=1030	[1]
NC_4H_{10}+OH→ 0.9*CH_3COE +0.1*CH_3CHO +0.1*CH_3COO +0.9*oRO_2	$Ae^{-B/T}(300/T)^N$,A=1.36e-12, B=-190, N=-2	[5]
C_2H_4+OH+M→ 2*HCHO+oRO_2	troe(7e-29,0,3.1,9e-12,0,0,0.7)	[1], n9
C_3H_6+OH+M→ HCHO+CH_3CHO+oRO_2	troe(8e-27,0,3.5,3e-11,0,0,0.5)	[1], n9
OXYL+OH→ MEMALD+MGLYOX+oRO_2	1.37e-11	[5]
$C5H_8$+OH→ 0.32*MAC +0.42*MVK +0.74*HCHO +obio	$Ae^{-B/T}$, A=2.55e-11, B=-410	[1]
APINEN+OH→ 0.8*CH_3CHO+0.8*CH_3COE+obio	$Ae^{-B/T}$, A=1.21e-11, B=-444	[17]
OH attack to carbonyls and peroxides		
HCHO+OH → CO+HO_2	$Ae^{-B/T}$,A=8.6e-12,B=-20	[1]
CH_3CHO+OH →CH_3COO	$Ae^{-B/T}$,A=5.6e-12,B=-310	[1]
MEMALD+OH →GLYOX+MGLYOX+oRO_2	5.6e-11	[1]
CH_3COE+OH →CH_3COY+oRO_2	$Ae^{-B/T}(300/T)^N$, A=2.92e-13, B=-414, N=-2	[1]
GLYOX+OH → 2*CO+HO_2	1.1e-11	[1]
MGLYOX+OH → CH_3COO+CO	1.5e-11	[1]

Reactions	Kinetic constants	Ref.
MVK+OH → 0.266*MGLYOX +0.266*HCHO +0.684*CH_3CHO +0.684*CH_3COO +0.05*ISNI +0.95*oRO_2	$Ae^{-B/T}$,A=4.1e-12,B=-453	[6]
MAC+OH → 0.5*CH_3COE+0.5*CO_2+0.5*oPAN	$Ae^{-B/T}$,A=1.86e-11,B=-175	[6]
CH_3O_2H+OH → CH_3O_2	$Ae^{-B/T}$,A=1.9e-12,B=-190	[1]
PPA+OH → CH_3COO	$Ae^{-B/T}$,A=1.9e-12,B=-190	n2
CH_3O_2H+OH → HCHO+OH	$Ae^{-B/T}$,A=1.e-12,B=-190	[1]
oROOH+OH → 0.8*OH+0.2*oRO_2	$Ae^{-B/T}$,A=4.35e-12,B=-455	n2
obioH+OH → OH	8e-11	n2
PANH+OH → 0.2*oPAN	1.64e-11	n2
CARNIT+OH → CH_3CHO+CO+NO_2	k(T)=$Ae^{-B/T}$,A=5.6e-12,B=-310	
NO_3 attack to hydrocarbons and aldehydes		
C_2H_4+NO_3 → 0.5*CARNIT+HCHO+oRN1	2e-16 [1]	
C_3H_6+NO_3 → 0.5*CARNIT +1.5*HCHO +0.5*CH_3CHO +0.5*HO_2 +oRN1	9.45e-15	[1]
APINEN+NO_3 → CH_3CHO+CH_3COE+oRN1	$Ae^{-B/T}$,A=1.19e-12,B=490	[8]
C5H_8+NO_3 → 0.85*ISNI +0.1*MAC +0.05*MVK +0.15*HCHO +0.8*HO_2 +oRN1	7.8e-13	[6]
HCHO+NO_3 → CO+HNO_3+HO_2	5.8e-16	[1]
CH_3CHO+NO_3 → CH_3COO+HNO_3	2.8e-15	[9]
CH_3O_2+NO_3 → HCHO+HO_2+NO_2	1.2e-12	[10]
CH_3COO+NO_3 → CH_3O_2+NO_2+CO_2	4e-12	[10]
oRO_2+NO_3 → NO_2+HO_2	1.2e-12	[10]
obio+NO_3 → NO_2+HO_2	1.2e-12	[10]
oPAN+NO_3 → NO_2+HO_2	4e-12	[10]
oRN1+NO_3→ 1.5*NO_2	1.2e-12	[10]
O_3 attack to unsaturated hydrocarbons		
C_2H_4+O_3 → HCHO+0.12*HO_2+0.13*H_2+0.44*CO	$Ae^{-B/T}$,A=9.1e-15,B=2580	[1]
C_3H_6+O_3 → 0.53*HCHO +0.5*CH_3CHO +0.31*CH_3O_2 +0.28*HO_2 +0.15*OH +0.065*H_2 +0.4*CO +0.7*CH_4	$Ae^{-B/T}$,A=5.5e-15,B=1880	[1]
C5H_8+O_3 → 0.67*MAC +0.26*MVK +0.55*OH +0.07*C_3H_6 +0.8*HCHO +0.06*HO_2 +0.05*CO +0.3*O_3	$Ae^{-B/T}$,A=1.2e-14,B=2013	[6]
MAC+O_3 → 0.8*MGLYOX +0.7*HCHO +0.215*OH +0.275*HO_2 +0.2*CO +0.2*O_3	$Ae^{-B/T}$,A=5.3e-15,B=2520	[6]
MVK+O_3 → 0.82*MGLYOX +0.8*HCHO +0.04*CH_3CHO +0.08*OH +0.06*HO_2 +0.05*CO +0.2*O_3	$Ae^{-B/T}$,A=4.3e-15,B=2016	[6]
APINEN+O_3 → 1.27*CH_3CHO +0.53*CH_3COE +0.14*CO +0.62*oRO_2 +0.42*HCHO +0.85*OH +0.1*HO_2	$Ae^{-B/T}$,A=1.0e-15,B=736	[8]
organic radical conversion		
CH_3O_2+NO → HCHO+NO_2+HO_2	$Ae^{-B/T}$,A=4.2e-12,B=-180	[1]
CH_3COO+NO → CH_3O_2+NO_2+CO_2	2e-11	[1]
oRO_2+NO → NO_2+HO_2	4e-12	n3
obio+NO → 0.86*NO_2+0.78*HO_2+0.14*ISNI	$Ae^{-B/T}$,A=1.4e-11,B=180	[6]
oPAN+NO → NO_2+HO_2	1.4e-11	n3
oRN1+NO → 1.5*NO_2	4e-11	n3
organic radical recombination		
CH_3O_2+HO_2→ CH_3O_2H	$Ae^{-B/T}$,A=4.1e-13,B=-790	[1]
CH_3COO+HO_2→ 0.67*PPA+0.33*O_3	$Ae^{-B/T}$,A=4.3e-13,B=-1040	[1]
oRO_2+HO_2→ oROOH	$Ae^{-B/T}$,A=2.7e-13,B=-1000	n4
obio+HO_2→ obioH	$Ae^{-B/T}$,A=2.7e-13,B=-1000	n4
oPAN+HO_2→ PANH	$Ae^{-B/T}$,A=2.7e-13,B=-1000	n4
oRN1+HO_2→ X	$Ae^{-B/T}$,A=2.7e-13,B=-1000	n4
CH_3O_2+CH_3O_2→ 1.35*HCHO+0.7*HO_2	$Ae^{-B/T}$,A=1.13e-13,B=-356	[1], n5
CH_3COO+CH_3O_2→ 0.5*CH_3O_2+0.5*CO_2+HCHO+0.5*HO_2	$Ae^{-B/T}$,A=3.34e-12,B=-400	[1], n5

Reactions	Kinetic constants	Ref.
$oRO_2+CH_3O_2 \rightarrow$ 0.65*HCHO+0.8*HO_2+0.35*CH_3OH	$Ae^{-B/T}$,A=1.5e-13,B=-220	n5
obio+$CH_3O_2 \rightarrow$ 0.8*HO_2+0.5*HCHO	$Ae^{-B/T}$,A=2.44e-11,B=223	n5
oPAN+$CH_3O_2 \rightarrow$ HCHO+0.5*HO_2	$Ae^{-B/T}$,A=7.9e-12,B=-140	n5
$CH_3COO+CH_3COO \rightarrow$ 2.*CH_3O_2+2.*CO_2	$Ae^{-B/T}$,A=2.8e-12,B=-530	[1], n5
$oRO_2+CH_3COO \rightarrow$ 0.8*CH_3O_2+0.8*CO_2+0.8*HO_2	$Ae^{-B/T}$,A=8.6e-13,B=-260	n5
obio+$CH_3COO \rightarrow$ 0.5*HCHO+1.5*HO_2+0.7*CO_2	$Ae^{-B/T}$,A=1.18e-11,B=127	n5
oPAN+$CH_3COO \rightarrow$ CH_3O_2+CO_2+HO_2	$Ae^{-B/T}$,A=3.34e-12,B=-400	n5
$oRO_2+oRO_2 \rightarrow$ 1.3*HO_2	6.4e-14	n5
organic nitrates and pernitrates		
$CH_3COO+NO_2+M \rightarrow$ PAN	troe(2.7e-28,0,7.1,1.2e-11,0,0.9,0.3)	[1], n9
oPAN+NO_2+M $\rightarrow$ toPAN	troe(2.7e-28,0,7.1,1.2e-11,0,0.9,0.3)	n6 , n9
PAN+M $\rightarrow$ $CH_3COO+NO_2$	troe(4.9e-3,12100,0,5.4e16,13830,0,0.3)	[1], n9
toPAN+M $\rightarrow$ oPAN+NO_2	troe(4.9e-3,12100,0,5.4e16,13830,0,0.3)	n6 , n9
PAN+OH $\rightarrow$ HCHO+NO_3+CO_2	$Ae^{-B/T}$,A=9.5e-13,B=650	[1]
toPAN+OH $\rightarrow$ NO_3+CO_2	$Ae^{-B/T}$,A=3.25e-13,B=-500	n6
ISNI+OH $\rightarrow$ 0.95*CH_3CHO +0.475*CH_3COE +0.475*MGLYOX +0.05*ISNI +0.05*HO_2 + oRN1	3.4e-11	[6]
photolytic reactions		
$O_3 \rightarrow$ 2*OH	photorate	[3], n7
$NO_2 \rightarrow$ NO+O_3	photorate	[3]
$NO_3 \rightarrow$ NO_2+O_3	photorate	[3], [11]
$NO_3 \rightarrow$ NO	photorate	[3], [11]
$N_2O_5 \rightarrow$ NO_2+NO_3	photorate	[3]
$H_2O_2 \rightarrow$ 2*OH	photorate	[3]
$HNO_3 \rightarrow$ NO_2+OH	photorate	[3]
HONO $\rightarrow$ NO+OH	photorate	[3]
HCHO $\rightarrow$ CO+2*HO_2	photorate	[3]
HCHO $\rightarrow$ CO+H_2	photorate	[3]
$CH_3CHO \rightarrow$ CH_3O_2+HO_2+CO	photorate	[1]
$CH_3COE \rightarrow$ $CH_3COO+CH_3CHO+oRO_2$	photorate	[11], [12]
$CH_3COY \rightarrow$ 2*CH_3COO	photorate	[13], [14]
MGLYOX $\rightarrow$ $CH_3COO+HO_2$+CO	photorate	[1]
GLYOX $\rightarrow$ 0.6*HO_2+2*CO+0.7*H_2	photorate	[13], [14]
MEMALD $\rightarrow$ 0.5*MVK +0.5*MALEIC +0.5*oPAN +0.5*HCHO +0.5*HO_2	photorate	[15]
$CH_3O_2H \rightarrow$ HCHO+OH+HO_2	photorate	[3]
PPA $\rightarrow$ CH_3O_2+CO_2+OH	photorate	n8
oROOH $\rightarrow$ OH+HO_2	photorate	n8
obioH $\rightarrow$ OH+HO_2	photorate	n8
PANH $\rightarrow$ OH+HO_2	photorate	n8
PAN $\rightarrow$ $CH_3COO+NO_2$	photorate	[3]

- [1] Atkinson et al., 1997
- [2] Donahue et al., 1997
- [3] DeMore et al., 1997
- [4] Mentel et al., 1996
- [5] Atkinson, 1990
- [6] Paulson and Seinfeld, 1992
- [7] Le Bras, 1997
- [8] Atkinson et al., 1994
- [9] DeMore et al., 1990
- [10] Canosa-Mas et al., 1996
- [11] Johnston et al., 1996
- [12] Martinez et al., 1992
- [13] Jenkins et al., 1997
- [14] Plum et al., 1983
- [15] Bierbach et al., 1994
- [16] Lightfoot et al., 1996
- [17] Kirchner and Stockwell, 1996
- [18] Aumont, 2002

- n1 : Rate constants containing a density dependent term in a minor reaction channel have been simplified by

putting an average boundary layer molecular density of 2.5*10^{19} molecules/cm^3.

- n2 : For operator reactions with OH, oROOH is interpreted as 2-butyl hydrogen peroxide, obioH as a peroxide from isoprene degradation, PANH represents unsaturated organic peroxides; due to lack of data, the rate constants of these compounds and of peroxy acetyl acid have been estimated using structure reactivity relationships in Kwok and Atkinson (1995).
- n3 : The rate constant of the operator reaction oRO_2+NO has been set to that of the i-C_3H7O_2+NO reaction given in [16], that of oPAN+NO to that of CH_3COO_2+NO [1], the rate constant for oRO_2NO+NO is taken from that of ISNIR (nitrate peroxy radicals from isoprene degradation) + NO [6].
- n4 : The rate constants for the operator reactions oRO_2, obio, RO_2NI+HO_2 have been set to that of the reaction C_2H5O_2+HO_2 [1], that of oPAN+HO_2 has been taken from CH_3COO_2+HO_2 [1].
- n5 : Radical conserving and terminating pathways are combined to single reactions; rate constants in operator reactions are affected by interpreting oRO_2 as 2-butyl peroxy radical [16], obio as a peroxy radical from isoprene degradation [6] and oPAN as CH_3COO_2.
- n6 : Reaction rates of the operators oPAN and OPAN have been set to those of the peroxy acetyl radical and PAN, respectively.
- n7 : The reactions O_3+hν $\rightarrow$ O(1D), O(1D)+H_2O $\rightarrow$ 2OH, O(1D)+M $\rightarrow$ O(3P)+M and O(3P)+O_2+M $\rightarrow$ O_3+M have been combined to reaction O_3+h$\nu\rightarrow$ O(1D) by taking into account H_2O and M concentrations in order to adjust the photolysis frequency.
- n8 : Due to lack of data, photolysis frequencies of higher organic peroxides are taken as that of CH_3OOH.
- n9 : Three body Troe reactions [1], given in the form:

$$k = \frac{k_0[M]}{1+\frac{k_0[M]}{k_\infty}} f^p$$

with

$$p = (1+(log_{10}(k_0[M]/k_\infty))^2)^{-1},$$

$$k_0 = A_0 e^{\left(-\frac{B_0}{T}\right)} \left(\frac{T}{300}\right)^{-n}, \text{ and } k_\infty = A_\infty e^{\left(-\frac{B_\infty}{T}\right)} \left(\frac{T}{300}\right)^{-n},$$

The parameters are given in the order $A_0, B_0, n, A_\infty, B_\infty, m, f$

2.9.3 Aqueous-phase chemistry

In the aerosol version of the model only, Aqueous sulphate chemistry is considered; Sulfate is produced in the following aqueous reactions (Berge, 1993; Hoffman and Calvert, 1985; Lee and Schwartz, 1983):

- $SO_2^{aq} + O_3^{aq} \rightarrow SO_4^{2-}$
- $HSO_3^- + O_3^{aq} \rightarrow SO_4^{2-}$
- $SO_3^{2-} + O_3^{aq} \rightarrow SO_4^{2-}$
- $SO_2^{aq} + H_2O_2^{aq} \rightarrow SO_4^{2-}$
- $SO_2^{aq} + NO_2^{aq} \rightarrow SO_4^{2-}$
- $SO_3^{2-}(Fe^{3+}) \rightarrow SO_4^{2-}$
- $HSO_3^-(Mn^{2+}) \rightarrow SO_4^{2-}$

SO_2, H_2O_2 and O_3 in the aqueous phase are in equilibrium with the concentrations in the gas phase. Moreover, aqueous SO2 is dissociated into HSO_3^- and SO_3^{2-}. Catalyzed oxidation reactions of sulfur dioxide in aqueous droplets with iron and manganese are considered, following (Martin et al. (1991); Martin and Hill (1987) ; Hoffman and Calvert (1985)). Henry's law coefficient and other aqueous equilibrium constants are used (Seinfeld and Pandis, 1998). Sulfur chemistry is very pH sensitive. pH is kept bound between 4.5 and 6.0.

2.9.4 Heterogeneous chemistry

In the aerosol version of the model only, a few heterogeneous reactions are also considered. Nitric acid is produced onto existing particles and fog droplets. Although aerosol particles and cloud droplets occupy a very small fraction of the atmosphere, it is now well established that reactions involving gas species onto their surfaces may significantly contribute to atmospheric chemistry cycles. Jacob (2000) recommends for ozone model to include a minimal set of reactions with associated uptake coefficients given by Harrison and Collins (1998), and other references in Jacob (2000):

- $HO_2 \rightarrow 0.5H_2O_2$ ..... γ = 0.2
- $NO_3 \rightarrow HNO_3$ ..... γ = 0.001
- $NO_2 \rightarrow 0.5HNO_3$ + 0.5HONO ..... γ = 0.0001
- $N_2O_5 \rightarrow 2HNO_3$ ..... γ = 0.01 - 1

The first-order rate constant k for gas heterogeneous loss onto particles is given by :

$$k = \sum_l \left(\frac{d_l}{2D_g} + \frac{4}{V\gamma} \right)^{-1} A_l \tag{2.3}$$

with d_l, the particle diameter (m), D_g, the reacting gas molecular diffusivity (m2 s^{-1}), n, the mean molecular velocity (m s-1), A_l, the total surface area in the particle bin l and gamma the uptake coefficient of reactive species. The uptake coefficient for N_2O_5 is assumed to be temperature-dependent in the range 0.01 - 1 (DeMore et al., 1997) with increasing values for decreasing temperatures. A recent study (Aumont et al., 2003) suggests that NO2 reactions on wet surfaces could be an important source for HONO production during wintertime smog episodes, so, a new reaction is added (also present in the gas-phase mechanism, see above).

2.9.5 Secondary organic aerosol chemistry

Secondary organic aerosols (SOA), produced through photochemical reactions in the atmosphere contain only oxidized species, carbonyls, carboxylic acid, etc.. Pun et al. (2002) and Griffin et al. (2002a, 2002b) propose a very complete modeling system for secondary organic formation and partitioning between gas and particle phases.
Up to now, a very simplified scheme for SOA formation has been implemented in the chemical module MELCHIOR. Anthropogenic aerosol yields (ASOA) come from Grosjean and Seinfeld (1989), Moucheron and Milford (1996), Odum et al. (1996, 1997) and Schell et al. (2001). For biogenic secondary organic aerosols (BSOA), Pankow et al. (2001) have proposed aerosol yields for terpene oxidation. In the model, precursor volatile organic compounds able to form secondary aerosol species are high chain alkanes, aromatics and monoterpenes. ASOA and BSOA are partitioned between gas and aerosol phases. Mass transfer is not only driven by the gas phase diffusion but also by the thermodynamic equilibrium through a temperature dependent partition coefficient (Pankow, 1994).

2.10 Clouds and impact on photolysis

In this version clouds are assumed to lie above the model top, so that there is no cloud albedo effect within the model domain depth. For all photolysed species, clear sky photolysis rates $J_c(z)$ are multiplied throughout model columns by an **attenuation coefficient** A(d) depending on the **total cloud optical depth (COD)** d. Using the TUV model, and a large set of CODs for clouds at various altitudes, the attenuation relative to the clear-sky case has been fitted as a function of COD with the formula:

$$A(d) = e^{-0.11d^{2/3}}$$

Several options are offered in order to calculate the COD. Total COD, d, is the sum of partial CODs from 3 cloud layers, low clouds d_l, medium clouds d_m and high clouds d_h. The limits between these cloud is user-chosen, but depend on the meteorological model. For MM5, limits of 2000m and 6000m are proposed. For each cloud layer three options are possible for the calculation of the partial cloud optical depth:

- Calculation as a function of Liquid/Ice Water Content in the column; using an assumption of sphericity and equivalent droplet size of 6 microns, an assumption of hexagonal shape for ice particles, the formula for cloud optical depth is 180.C_w + 67.C_i, where C_w and C_i are respectively the liquid water column (in Kg/m^2) and ice column for this cloud layer. This option is chosen by setting the cloud option to "1" in the **METINFOS** file.
- Liquid/Ice water are generally unverified and unstable parameters in atmospheric models. Basing a cloud parameterization on such parameters can be risky. A more robust parameterization consists in using relative humidity only. The proposed parameterization (option "2" in **METINFOS** file) consists in parameterizing the COD as a function of the integral **R**, over the cloud depth, of the relative humidity above 75%. It is assumed that small cloud formation (in particular cumulus clouds) starts at 75% relative humidity. Normalization of **R** leads to the formulation (for instance for low clouds): d_l = aR/dz, where a=0.02 is chosen such that a 1000m-thick layer has an optical cloud depth of 20.
- An even simpler parameterization (Option "0" in METINFOS) can be achieved by making the COD simply proportional to the cloud fraction (if available) for each cloud layer. Coefficient tuning led to proportionality coefficients of: 50 for low-clouds, 10 for medium clouds and 2 for high clouds. This means, for instance, that a sky covered with 100% of high clouds has an optical depth of 2. Tuning was performed with ECMWF cloud fraction data and should change with the meteorological model.

Recommandations The first option using liquid/ice water (whenever available),is based on a variable which can be very model-dependent and which is rarely verified against observations. Moreover it does not take into account fractionary cloudiness. However it may be suited for high/medium clouds which are often stratiform clouds in the midlatitudes. The second option uses cloud fractions (whenever available), which is less meaningful to radiation but could be used when cloud/ice water is not available. The third option uses relative (to liquid water) humidity, and has been tuned using J(NO2) measurements taken over a long period near Paris (manuscript in preparation). The formulation, modifying radiation for RH$>$0.75 should only be used for low clouds. It has the advantage of being a more robust parameterization since it depends on a highly verified variable. In summary, our preference is: RH formulation (option 2 in METINFOS) for low clouds, and for liquid/ice water formulation (whenever available) for medium and high clouds, and cloud fractions otherwise.
The attenuation coefficient is calculated in the **diagmet.f** diagnostic interface model. It is read in Biniphys.f, photolysis rates are calculated in routine **Bphotorates.f** every coarse time-step.

2.11 Emissions

2.11.1 Anthropic emissions

CHIMERE requires input **Anthropic** emission data for 14 model species. These model species are:

- NO
- NO2
- SO2
- CO
- CH4[4]
- C2H6 Ethane
- NC4H10 n-Butane
- C2H4 Ethene
- C3H6 Propene
- C5H8 Isoprene
- OXYL o-Xylene
- HCHO Formaldehyde
- CH3CHO Acetaldehyde
- CH3COE Methyl ethyl Ketone

In the present version, emissions of Methanol and Ethanol are also included but are not used. They are only used in the **complete** version of the chemical mechanism (soon available). Annual data of anthropogenic emissions for the four classes NOx, SO2, CO and NMVOC are taken from the EMEP data base for 1998 (Mylona, 1999, see also the EMEP web site[5]. The data are spatially interpolated from the EMEP grid onto the CHIMERE grid.
The NMVOC emissions have to be split into 9 classes represented within the model chemical mechanism. To this end they are first distributed for each country into different activity sectors (traffic, solvents, industrial and residential combustion, others), according to data prepared by IER (Institute for Energy Economics and Rational Use of Energy, University of Stuttgart; GENEMIS, 1994). For each sector, NMVOC emissions are then split into 32 classes with similar structure and reactivity, following a classification of Middleton et al., 1990, and using standard NMVOC

[4] Actually CH4 emissions are zero in this model version
[5] http://www.emep.int

profiles again from IER (genemis, 1994). Third, NMVOCs from these 32 classes are aggregated into the 9 classes represented within CHIMERE, by applying mass and reactivity weighting as proposed by Middleton et al., 1990. It should be noted that the uncertainty in the spatial distribution of individual VOC emissions is large, given that the VOC profiles are assumed to be the same all over Europe, with the exception of traffic, where national differences (e.g. the ratio gasoline / diesel) are taken into account. Monthly, daily and hourly variations of the emissions are modeled by imposing respective variations from the GENEMIS (1994) data base.
In rural areas NO emissions from microbial processes may be an important source. Since these emissions strongly depend on temperature, they are processed in the model as "biogenic" emissions and often called in this manual "biogenic NO emissions". In soils NO is produced in a reaction chain of oxidation and reduction from ammonium which is used in fertilizers (e.g. Williams et al., 1992). CHIMERE uses a European inventory of soil NO emissions from Stohl et al. (1996). This inventory estimates the soil emissions to be during the summer months at the order of about 20differences between the countries. In the model, these NO emissions are only considered during the months of May to August.
Anthropic emissions (except soil NO) are given in 12 files corresponding to each month (**EMISSIONS.**MM), MM being the calendar month. They are given for three typical days: Week day, Saturday, Sunday and for each UT hour. For some days weekday emissions are considered as a sunday (for holidays), but this assumption is often valid only for France (e.g. July 14 is considered as a sunday), although it is extended to Europe!

2.11.2 Biogenic emissions

Biogenic emissions of isoprene and terpenes (affected to alpha-pinene in the chemical mechanism) are parameterized as fluxes $F_i = \epsilon_i D \gamma_i$ (Guenther, 1997), with ϵ_i being the species dependent "emission potential" and D the "foliar density", both values taken from Simpson et al. (1999).
The "environmental correction factor" γ_i accounts for the dependence on temperature and insolation (e.g. Guenther, 1997). The spatial distribution of tree species is established following the methodology outlined in Simpson et al. (1999). Therefore the SEI (Stockholm Environment Institute) land cover data base, which details the fraction of different tree species over Europe, is interpolated on the CHIMERE grid. The 136 SEI land use classes are aggregated into 11 “emitter” classes. Oaks, for example are differentiated in the classes highly isoprene emitting, highly terpene emitting and low biogenic VOC emitting oaks. As this attribution of classes is sometimes ambiguous, national tree species inventories from Simpson et al. (1999) are used to verify the fraction of each emitter class on a country-wise level, and to adjust it, if necessary.

2.12 Dry deposition

2.12.1 For gases:

Dry deposition is considered for model gas species i and is parameterized as a downward flux $F_{d,i} = -v_{d,i} c_i$ out of the lowest model layer with c_i being the concentration of species i. The deposition velocity is, as commonly, described through a resistance analogy (Wesely and Hicks (1977)):

$$v_{d,i} = \frac{1}{R_a + R_{b,i} + R_{c,i}} \tag{2.4}$$

The aerodynamic resistance R_a is calculated as the integral of the inverse of the diffusivity coefficient K(z) up to the middle of the model surface layer, which can be estimated using the analytical formulae of the surface-layer similarity profiles for K (Seinfeld and Pandis (1998)). The formulae read:
Stable case

$$R_a = \frac{1}{k u_*} \left[ln\left(\frac{z}{z_0}\right) + 4.7(\zeta - \zeta_0) \right] \tag{2.5}$$

Neutral case

$$R_a = \frac{1}{k u_*} \left[ln\left(\frac{z}{z_0}\right) \right] \tag{2.6}$$

Unstable case

$$R_a = \frac{1}{ku_*}\left[ln\left(\frac{z}{z_0}\right) + ln\left(\frac{(\eta_0^2+1)(\eta_0+1)^2}{(\eta_r^2+1)(\eta_r+1)^2}\right) + 2\left(tan^{-1}\eta_r - tan^{-1}\eta_0\right)\right] \quad (2.7)$$

where $\eta_0 = (1-15\zeta_0)^{1/4}$, $\eta_r = (1-15\zeta_r)^{1/4}$ and $\zeta_0 = z_0/L$, **L** is the Obukov length, k the von Karman constant and z_0 is the dynamical roughness length. z_0 depends on the fraction of land-use for each category and on the season. See Section on land-use and on meteorology for more details.
According to Hicks et al. (1987) the quasi laminar resistance R_b, i is calculated as:

$$R_{b,i} = \frac{2}{ku_*}\left(\frac{Sc_i}{Pr}\right)^{2/3} \quad (2.8)$$

where Sc_i and P_r are the Schmidt and Prandtl number, respectively, k the von Karman's constant and u_* the friction velocity.
The formulation of the surface resistance R_c follows Erisman et al. (1994). It uses a number of different other resistances accounting mainly for stomatal and surface processes which are again dependent on the land use type and season. Necessary chemical parameters for the calculation of R_c are also taken from Erisman et al. (1994) except for carbonyls (Sander, 1999; Baer and Nester, 1992) and peroxide species (Hall et al., 1999). All species-related parameters are gathered in the DEPO_SPEC file located in all data_[model]_[meca] directories.
The main land/seasonal parameters can be found in the file DEPO_LAND file which can be found in all data_[model]_[meca] directories. Seasonal variations of resistances are taken from 3 Wesely Seasons:

- #1 (summer with full vegetation), from may to october. In August, September and October, however, LAI, stomatal resistance and deposition height are divided by two for the CROPLAND category in order to take into account harvested fields.
- #3 (winter without frost) from november through march.
- #5 (transitional season with short annuals) for April.

Most land parameters are taken from Wesely (1989), but LAI are drawn from the NASA/EOSDIS Oak Ridge National Laboratory[6] using average LAI field measurements, for Summer. Some other data come from the article of Brook (19XX)
Dry deposition (for gases) parameters are initialized in routine **Binidepo.f**. At the coarse time step frequency, deposition velocities are calculated in routine **Bdepvel.f**. Deposition fluxes are calculated at the fine time step frequency in routine **Bdeposition.f**.

2.12.2 For aerosols:

Dry deposition for aerosols also makes use of a resistance scheme. The dry deposition velocity is expressed as:

$$v_d = v_s + \frac{1}{r_a + r_b + r_a r_b v_s} \quad (2.9)$$

where v_s is the sedimentation velocity expressed as in Seinfeld and Pandis (1998), R_a the aerodynamic resistance and R_b the quasi-laminar resistance for particles.

2.13 Wet scavenging

The aerosol version (only) takes into account scavenging for gas/aerosols in clouds or rain droplets. Wet scavenging is formulated as follows:

- For **gases** in **clouds** : Nitric acid, ammonia in the gas phase are scavenged by cloud droplets. This process is assumed to be revertible. Moreover, for in-cloud scavenging, dissolved gases in a non precipitating cloud can reappear in the gas phase due to cloud dissipation. Equilibrium between dissolved gases concentration and gas-phase concentrations follow Seinfeld and Pandis (1998).

[6] http://www.daac.ornl.gov

- For **gases** in **rain droplets below the clouds** : Dissolution of gases in precipitating drops is assumed to be irrevertible, both for HNO_3 and NH_3. The scavenging coefficient is expressed as:

$$\Gamma = \frac{pD_g}{6.10^5 u_g D^2}(2 + 0.6R_e^{1/2}S_c^{1/3}) \tag{2.10}$$

  p being the precipitation rate (mm h^{-1}), D_g the molecular diffusion coefficient ($m^2 s^{-1}$), u_g the raindrop velocity (m s^{-1}), Re and Sc respectively the Reynolds and Schmidt numbers of drops. Mircea and Stefan (1998) and references therein give relationships between u_g and hydrometeor diameter for various types of precipitation. In the model, sulfur dioxide and hydrogen peroxide are also scavenged by precipitation.
- For **particles** in **clouds** : Particles can be scavenged either by coagulation with cloud droplets or by precipitating drops. Particles also act as cloud condensation nuclei to form new droplets. This latter process of nucleation is the most efficient one in clouds. According to Tsyro (2002) and Guelle et al. (1998), the deposition flux is written as:

$$\left[\frac{dQ_l^k}{dt}\right] = -\frac{\epsilon_l P_r}{w_l h} Q_l^k \tag{2.11}$$

  with p_r being the precipitation rate released in the grid cell (g cm^{-2} s^{-1}), w_l the liquid water content (g cm^{-3}), h the cell thickness (cm) and eipsilon an empirical uptake coefficient (in the range 0 - 1) depending on particle composition. l and k are respectively the bin and composition subscripts.
- For **particles** in **rain droplets below the clouds** : Particles are scavenged by raining drops, the deposition flux of particles being:

$$\left[\frac{dQ_l^k}{dt}\right] = -\frac{\alpha p E_l}{u_g} Q_l^k \tag{2.12}$$

  with a being an empirical coefficient, p the precipitation rate in the grid cell (g cm^{-2} s^{-1}), E a collision efficiency coefficient between particles and raining drops (Slinn, 1983) and u_g the falling drop velocity (cm s^{-1}).

2.14 Time numerical solver

The numerical method for the temporal solution of the stiff system of partial differential equations is adapted from the second-order TWOSTEP algorithm originally proposed by Verwer (1994) for gas phase chemistry only. It is based on the application of a Gauss-Seidel iteration scheme to the 2-step implicit backward differentiation (BDF2) formula:

$$c^{n+1} = \frac{4}{3}c^n - \frac{1}{3}c^{n-1} + \frac{2}{3}\Delta t R(c^{n+1}), \tag{2.13}$$

with c^n being the vector of chemical concentrations at time t_n, Δt the time step leading from time t_n to $t_n + 1$ and $R(c) = \dot{c} = P(c) - L(c)c$ and $R(c) = \dot{c} = P(c) - L(c)c$"> the temporal evolution of the concentrations due to chemical production and emissions (P) and chemical loss and deposition (L". Note that L is a diagonal matrix here. After rearranging and introducing the production and loss terms this equation reads

$$c^{n+1} = \left(I + \frac{2}{3}\Delta t L(c^{n+1})\right)^{-1} \left(\frac{4}{3}c^n - \frac{1}{3}c^{n-1} + \frac{2}{3}\Delta t P(c^{n+1})\right). \tag{2.14}$$

The implicit nonlinear system obtained in this scheme can be solved pertinently with a Gauss-Seidel method (Verwer, 1994).
In CHIMERE the production and loss terms P and L in equation 2.14 are replaced by the modified terms $\tilde{P} = P + P_h + P_v$ and $\tilde{L} = L + L_h + L_v$, respectively. P_h and P_v denote the temporal evolution of the concentrations due to horizontal (only advection) and vertical (advection and diffusion) inflow into the concerned grid box,
L_h and L_v the temporal evolution due to the respective outflow divided by the concentration itself. With this integration scheme, at each time step all physical and chemical processes are updated simultaneously. The “operator splitting” technique (see e.g. McRae et al. (1982), which is still a standard way of solving the 3D transport-chemistry-problem, and the associated “splitting error” are therefore avoided. Further advantages of the scheme are

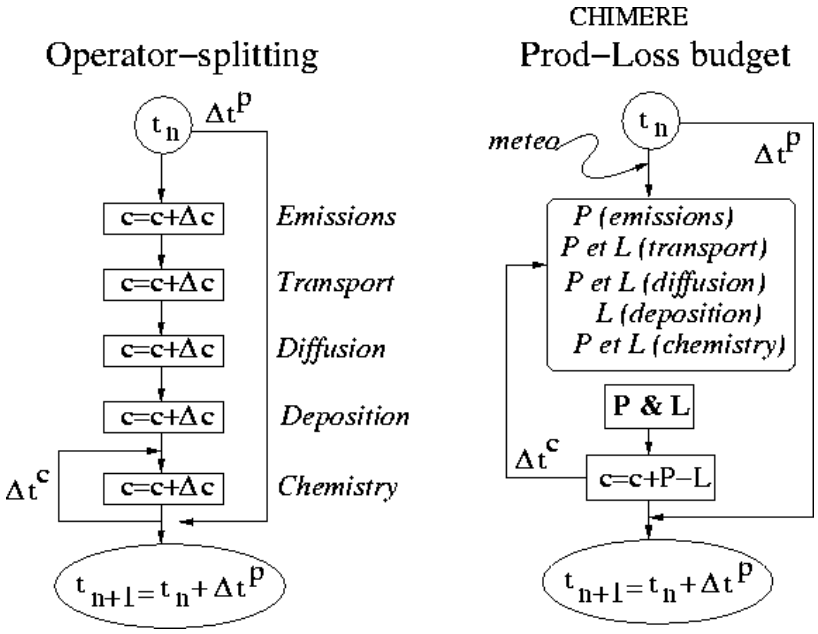


Figure 2.2: *Principle of 'operator-splitting' versus chimere integration*

its stability even for quite long time steps due to the implicitness of the formulation and the simplicity of the code which facilitates the development of secondary models (adjoint, tangent linear) enormously.
In practice, time stepping has two time steps: A **coarse time step**, where all physical variables are calculated, and a **fine step** (refined step in order to gain accuracy on chemistry. During loops over fine steps, physical variables are maintained constant.
In the provided version for the continental models, we chose a "very quick formulation", with a 10-minute physical step, no sub-chemical steps, i.e. all processes are stepped to 10 minutes, with only 1 Gauss-Seidel iteration. **It is STRONGLY recommanded**, in order to have more accurate results, to use 2 Gauss-Seidel iterations, which increases by 2 the computer time. This chang affect mostly concentrations during morning. However, for values of ozone concentration, the proposed version does not appear to degrade simulations, especially during afternoon hours.
The TWOSTEP formulae and loops are performed in routine **Btwostep.f**.

Chapter 3

Technical Documentation

3.1 General structure

3.1.1 Installation

After having downloaded the model directory into tar+gzip form (**V200310F.tar.gz** for instance), gunzip the file and type **tar xvf V200310F.tar**. It will be expanded into a directory containing subdirectories and shell script files. The same operation has to be done with the MOZART2 climatologies (MOZART2-200303.tar.gz). However in this case 600 mbytes of disk space is required to have these on disk. The prepdomains-200310.tar.gz must be processed in the same way, and 1 gbyte of disk space is required. If using the provided emission files, the same processing has also to be applied on the emission files. Some disk space is also required (less than 500 mbytes per emission directory). All these separate directories can be placed anywhere in the file system.
There is no need to compile anything at this stage. Compilation is actually made each time the model is run (it takes a few seconds...). For the **prepdomains** module, compilation is done first, see the README file there. The next stage now is to worry about the model inputs.

3.1.2 Software-languages-systems

The CHIMERE model has only been tested on LINUX systems. However it should be working on most UNIX systems provided the following software is provided. The model should work on other systems with minor changes, mostly due to binary files which may have to be converted.
The model has two separate versions: the gas-phase version and the aerosol version. The **gas-phase** version is written in FORTRAN77 language for which there are free compilers (g77 for instance, the GNU f77 compiler). The **aerosol** version requires an f90 compiler. In future releases the aerosol version will be converted in fortran 77 so that CHIMERE will only require free software to make it work. The scripts piloting simulations are written in Bourne-Shell (sh), and use the free GNU software **gawk** that must be dowloaded prior to the execution of the model. The post-processing stage which is also proposed with the **mapit** graphics program uses the **GMT3.x** free software (separate downloading).

3.1.3 General software structure

The general structure contains a central chemistry-transport calculation unit with interfaces to several databases (emissions, meteorology, boundary conditions, land use, as shown in the figure below. Each interface transforms original or processed data to CHIMERE-input formatted data adapted to the time period and region to simulate. For instance, the emission interface here links a data base containing emissions for 3 typical days per month to CHIMERE input, by constructing a hourly emission file adapted to the period to simulate. Other emission interfaces can be defined by the user.
The emission interface here links a data base containing emissions for 3 typical days per month to CHIMERE input, by constructing a hourly emission file adapted to the period to simulate. Other emission interfaces can be defined by the user (see the Section on emission interface)

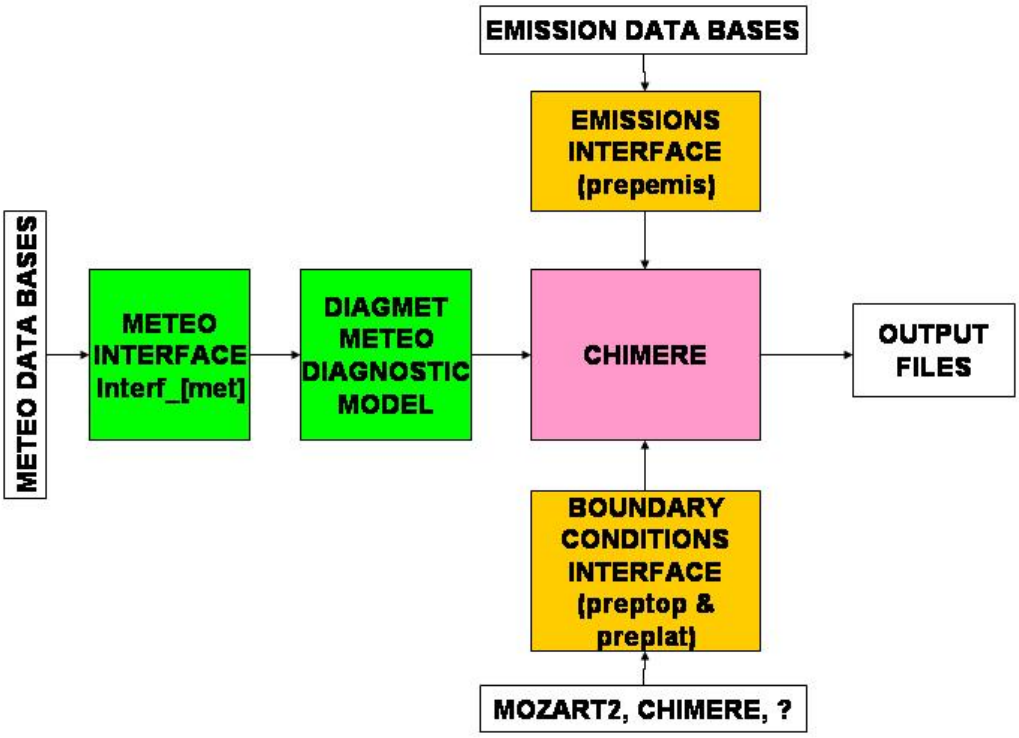


Figure 3.1: *Schematic general structure of CHIMERE*

The most complex interface is the meteorological interface because it uses two steps. The first one (**meteo**) transforms original meteorological data (standard variables) given on the meteo model grid, at a frequency that is not necessarily one hour, to the same variables but on the CHIMERE grid (space interpolation) at the hourly sampling rate (time interpolation) in a specific file format. Two examples (the **interf-mm5** interface and the **interf-ecm** interface) are provided (in the meteo directory), respectively for data coming from MM5 and from the ECMWF model. The mm5 interface can be directly applied to a standard output MM5 file (MMOUT files), while the ecm interface takes ECMWF already put in a binary specific format.

The second step for the meteorological interface is the diagnostic model **diagmet**, used to take standard meteorological variables on the CHIMERE grid and transform them into variables necessary for input to CHIMERE. For instance, the boundary layer height can be diagnosed from the vertical profile of temperature, humidity and wind.

The boundary condition interface takes either a "coarse run" made with chimere and interpolates hourly pollutant concentrations at the domain boundaries, or provide climatological (monthly mean) values at the domain boundaries from a global database consisting of simulation outputs of the MOZART (second generation) model, ran at Max-Planck Institute, Hamburg and provided to the developing team by M Schultz, G. Brasseur, C. Granier and D. Niehl.

The land-use interface (**prepdomains**) is used to construct necessary terrain files and is used only once per domain construction.

The model core organization is described below. Briefly, it consists of the call of a main FORTRAN program (**chimere.f**) which initializes and reads all data, and runs an integration convering the time period specified. The output files are then available for postptrocessing.

The whole system is piloted by a **top calling script** which:

- reads parameters of the simulation
- copies all necessary files into a **temporary directory** where all programs are executed
- compiles all necessary code
- runs the 3 "dynamic" interfaces (meteo, emission and boundary conditions)
- runs the model itself

The top calling scripts are executed to carry out a simulation. They must be edited and the options/directories must be chosen and defined by changing the values of parameters inside the scripts.

3.1.4 Directories

The **model** directory, e.g. **V200310F**, contains individual files, example **top calling scripts** and subdirectories. These subdirectories are:

- **src_gas** : Contains the source routines of CHIMERE, gas-phase version.
- **src_aero**: Contains the source routines of CHIMERE, aerosol version.
- **data_gas_[meca]**: **[meca]** is the name of the chemical mechanism (only MELCHIOR1 and MELCHIOR2 in release V200310F). The directory contains data files for the gas-phase version and the mechanism.
- **data_aero_MELCHIOR2**: The directory contains data files for the gas-phase version and the mechanism MELCHIOR2. In V200310F release, only the MELCHIOR2 gas-phase mechanism can be used for aerosols.
- **util**: The directory contains some utility programs
- **domains**: A subdirectory where coordinate files COORD_[domain], biogenic emission potential files BIOFACS_[domain] and land-use files LANDUSE_[domain] must be placed for each CHIMERE model domain. This directory contains also the file **DOMAINS** to which one line should be added per new model domain.
- **meteo**: This directory contains all elements for the first step of meteo interfaces, i. e. the horizontal/time interpolation and reformatting of data. When the user starts from a new meteorological data base, a shell script named **interf-[name]** should be placed in this directory, which does this task. See the examples **interf-mm5** and **interf-ecm**. [name] is the name of the interface and can be any character string.
- **mapit**: This is the simple graphical interface. See the documentation below.

The **prepdomains** directory contains files and routines for constructing new domains **in Europe**. Follow instruction in the README file, and read the documentation below.
The **emission** directories contain only data files with the MELCHIOR (1 or 2) emissions. There is one file per month.
The **MOZART2** directory contains files with the monthly climatologies of the MOZART2 model, processed from files sent by the Max-Planck Institute, Hamburg.
There are other directories created by running the model itself:

- A *temporary directory*, where all temporary files go. Optionally this directory can be left after a model simulation (especially when debugging); By default it is removed at the end of top calling script execution. The address of this directory is given in the top calling script (environment variable **TMP**. It does not need to be created before the run.
- An *output* directory where simulation output files are placed during the run. The address of this directory is given in the top calling script (variable **simuldir**. It does not need to be created before the run.

For these two directories, choose a disk location where there is a lot of space. Typically these directories need 500 mbytes.

3.1.5 Model Files

Files in the **model** directory are grouped into 5 different categories. Addresses are given here relative to the model directory.
Scripts (sh)

- **Top calling script** examples are found directly in the model directory. Their name starts with **chimere-**. They are Bourne shell scripts. When all data sets are ready for a simulation, a top calling script must be executed for running the simulation. The V200310F version provides the following examples:
  - **chimere-ecm-iop6-cont-gas** Example of Western Europe 5-day simulation from 14 to 18 July 1999 using the model gas-phase version and ECMWF meteorological data up to 700 hPa. Domain is the CONT1 domain that was used for several years in the pioneer experiments (2000-2003). The period correspond to an intensive observation period of the ESQUIF[1] campaign over the Ile-de-France region.

[1] http://esquif.lmd.polytechnique.fr

- **chimere-ecm-iop6-cont-aero** Same as above but with the aerosol model version.
- **chimere-ecm-iop6-idf-gas** Nested model example, for the Ile-de-France area and the gas-phase version. The horizontal resolution is 6 Km.
- **chimere-mm5-hw**: 5-day simulation example using an MM5 simulation over Western Europe. Simulation starts on July 30 2003 and ends August 3 2003. Gas-phase only.
- **chimere-ecm-auto**: Allows automatic long-term simulations with ECMWF data with the IPSL (internal format) data bases.
- **chimere-mm5-auto**: Allows automatic long-term simulations with MM5 data with the IPSL (internal format) data bases.

- **data_gas_[meca]/chemprep**: runs the gas-phase chemical preprocessor. See documentation below
- **data_aero_MELCHIOR2/chemprep.aero**: runs the chemical preprocessor for the aerosol version. It produces input chemistry files for both gases and aerosols. See documentation below
- **meteo/interf-[metname]**: runs the first-step meteorological interface (interpolation + formatting). The script is called by the top calling script.
- **util/define_geom**: called by the top calling script to define the geometric progressing layer vertical grid. It outputs a file with hybrid sigma-p coefficients defining the grid.
- **util/make_mm5_coord**: utility script allowing to extract coordinates and terrain altitude, in longitude/latitude format, of a MM5 **TERRAIN** file. Edit the script itself for giving information. The output files are needed in the **meteo** directory in order to use the MM5 interface **interf-mm5**.
- **vcoord_subdiv**: called by the top calling script, used to refine a vertical grid for nested domains.
- **define_10d**: Utility program for defining time slices of automatic long-term runs.

Main programs

- **src_[model]/chimere.f**: Main program for chimere runs. Compiled to chimere.e and called by the top calling script.
- **src_[model]/diagmet.f**: Second step (diagnostic meteorological model) of the meteorological interface. It is not origin-dependent for meteo data. It ingests the output files of the first meteo interface (files **meteo/interf-[metname]**).
- **src_[model]/prepemis_std.f**: Standard emission interface.
- **src_[model]/preplat_[bounds].f**: lateral boundary conditions interface.
- **src_[model]/preptop_[bounds].f**: top boundary conditions interface.
- **meteo/exdom-[metname].f**: horizontal interpolation of original meteorological grid onto the CHIMERE grid.
- **meteo/prepmm5.f**: Program reading the MM5 standard output files.
- **meteo/altipres-ecm.f**: Transformation of pressure to altitude in ECMWF data
- **meteo/tinter-ecm.f**: Transformation from 3-hourly to hourly ECMWF data
- **meteo/wcal1.f**: Program calculating spatial interpolation weights for a quasi-rectuangular grid.
- **chemprep.awk**: awk program for the chemical preprocessor. It is not called, since it has been used to prepare the chemical input files.

Subroutines

Subroutines are located in the **src_[model]** directories and are named B*.f or B*.f90. Except a for a few files like **Bcalendar.f**, which is a set of utility routines for the manipulation of dates, almost all of the subroutine files contain one subroutine, which has the same name as the file, e.g. **Biniemis.f** contains the subroutine **iniemis**. Many comments can be found within the subroutines.

Headers

Only one header file has to undergo macro substitutions, **src_[model]/chimere.h.sed**. The *sed* command is executed in the top calling script, and transforms this file into a file named **chimere.h**. It contains parameter, arrays, COMMON definitions and are included in most of the chemistry-transport routines.

Input data files

All other provided files are data files. Their name is in capital letters and they lie in the **data_[model]_[meca]** and **domains** directories. A special file is **INFOS.sed**. It is a template which is transformed by the **sed** filter into a file named **INFOS** which is read by **chimere.e**. This transformation is done in top calling script. The same remark holds

for the **meteo/METINFOS-[metname].sed** file which, after sed transformation, becomes an input to the diagnostic meteo interface **diagmet.f**.
Input data files are separated into categories. Their detailed description is given below.
Input to chemistry-transport stage: **Files used for land processes (deposition+emissions)**

- domains/COORD_[domain]
- domains/BIOFACS_[domain]
- domains/LANDUSE_[domain]
- domains/DOMAINS
- data_[model]_[meca]/DEPO_SPEC
- data_[model]_[meca]/DEPO_LAND
- data_[model]_[meca]/BIOGENIC
- data_[model]_[meca]/ANTHTROPIC
- data_[model]_[meca]/LANDPAR

Files for chemistry and model information used for running the model

- data_[model]_[meca]/VCOORD.[vgrid]
- data_[model]_[meca]/LAT_SPEC.[bounds]
- data_[model]_[meca]/TOP_SPEC.[bounds]
- data_[model]_[meca]/PPM_SPECIES
- data_[model]_[meca]/OUTPUT_SPECIES.[outspec]
- data_[model]_[meca]/OUTPUT_PARAMETERS
- data_[model]_[meca]/INFOS.sed
- data_aero_MELCHIOR2/AEROMIN
- data_aero_MELCHIOR2/AEROSOL

Output files from chemprep preprocessor:

- data_[model]_[meca]/ACTIVE_REACTIONS
- data_[model]_[meca]/ACTIVE_SPECIES
- data_[model]_[meca]/CHEMISTRY
- data_[model]_[meca]/STOICHIOMETRY
- data_[model]_[meca]/FAMILIES
- data_[model]_[meca]/REACTION_RATES
- data_aero_MELCHIOR2/AERO_SPECIES **Input files to chemprep preprocessor:**
- data_[model]_[meca]/REACTIONS
- data_[model]_[meca]/PHOTO_PARAMETERS
- data_[model]_[meca]/E-DEPENDENCIES

Files for the meteo interface

- meteo/COORD_MM5[domain]
- meteo/METINFOS.[metname]

3.2 Steps to perform a simulation

Here are the following steps to achieve for running a simulation with the model:

- **(1)** First step is to install all the code. The model directory does not need much space (50 mbytes. The data directories, by contrast, must be installed where much space exist. It is advised in any case to have, for performing several simulations, say about 5-10 gbytes of disk space at hand.
- **(2)** Second step to achieve is to find meteorological data. There is no meteorological model provided with CHIMERE. CHIMERE can use many meteorological models but users should provide their own interface, except for MM5, for which an interface is proposed in the code, which takes as input standard MM5 output files (MMOUT files). Examples of meteorological data are provided on this server, in the **downloading** menu. See the documentation for this interface below. If MM5 (which is a free software, available on the NCAR[2] web site,

[2] http://www.ncar.ucar.edu

is chosen as the meteorological driver, it has to be initialized with other global meteorological data (analyses), which often are not public. However the National Centres for Environmental Prediction (NCEP) and the National Centre for Atmospheric Research provide a wide variety of meteorological data, analyses etc... online. Our experience deals with real-time AVN/NCEP analyses and forecasts that can be obtained on a web server. When setting up a meteorological simulation using MM5, choose a meteorological domain entirely containing the CHIMERE domain, and a time period entirely containing the CHIMERE period to simulate. The PBL option "MRF" (Option #5), based on the Troen and Mahrt (1986) parameterization should be prefered because it is most consistent with the model mixing formulation. The Schultz (Option #8) microphysics parameterization has also been tested with CHIMERE and is recommended. Also in case of the use of MM5, the MM5 coordinate files (**COORD_MM5[dom]** and **COORD_MM5[dom]_DOT** must be constructed from the MM5 TERRAIN file.

- **(3)** Prepare the land-parameters for the CHIMERE domain. There are 3 necessary land files: The coordinate file, the file containing the fractions in all of 9 land-use classes for all grid cells, and the file containing biogenic emission potentials. The directory provided online (**prepdomains**) only works **for Europe**. It works with the Global Land Cover Facility (GLCF) land-use data base (see documentation below). Other methodologies should be found for other parts of the world, in particular for constructing the emission potentials. Perhaps as a very first experiment for a simulation outside Europe it could be advised to remove them by setting a BIOFACS_[domain] file with 5 columns with zeros in them. LANDUSE_[domains] files are easier to construct from the GLCF data base LANDUSE_[domains] files are easier to construct from the GLCF data base.
- **(4)**Emission files have to be constructed and specified for the chosen mechanism. This is often a difficult task. In the emission documentation below, indications are provided on how to produce emission files to be input to CHIMERE. If the standard emission procedure is followed (one emission file per calendar month containing hourly emissions for all MELCHIOR emitted species and for 3 typical days (week day, saturday and sunday), the interface program **prepemis_std.f** proposed here (in the **src_[model]** directories can be used. Otherwise the interface program must be recoded.
- **(5)**The **top calling script** (**chimere-mm5-hw** for instance) must be modified according to the conditions of the simulation. Edit this Bourne-shell script and make changes in the values of the variables all the way to where it is written: "Only experienced users should change this below". Give names to output directories, simulations etc...
- **(6)** Run the model by invoking the top calling script.
- **(7)** Enjoy the results by making programs to read them and run them !

3.3 CHIMERE code structure

3.3.1 The top calling script

The CTM program is executed by the **top calling script** (chimere-* in the top directory), which, after running all the interfaces, and constructing the INFOS file, executes the **chimere.e** compiled version of **chimere.f**.

3.3.2 The temporary directory

All executable and temporary data files are put in a *temporary directory* that can be cleaned or not at the end of execution. This allows in particular several simulations to be performed simultaneously by calls from a unique directory, provided the name of the directory contains the process number (in Bourne Shell). The temporary directory address is defined in the top calling script. There is no need to build the directory before the execution.

3.3.3 Files and programs

The diagram below shows all the input and output files together with the executable programs or script that are used at all stages of the calling script. See the color legend in the right bottom corner. The yellow files are constructed before the run.

3.3.4 Routine calls sequence for the gas-phase version

The general organigram of calls of the CHIMERE model (chimere.f) is given in the figure below. Only major routines are cited, not the utility ones. The main program **chimere.f** first calls all initialization routines (in (**inichimere** and

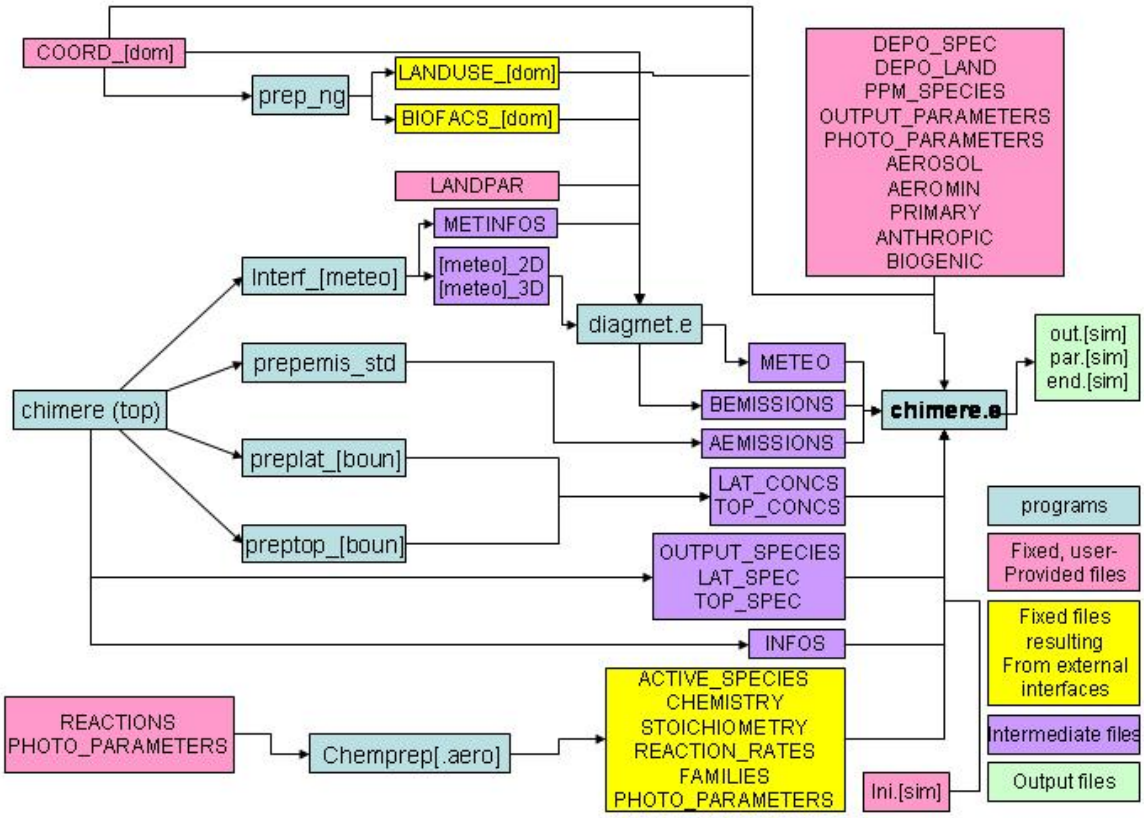


Figure 3.2: *Schematic general structure of CHIMERE*

further calls) where all "static data" data are read and many variables defined. Then it calls the integration routine **integrun** which then (i) reads hourly data every hour (routine **readhour** (ii) loops on the coarse time step, (iii) outputs results on files (routine **outprint**) and (iv) shifts hourly data (routine **renewhour**). At each model step, one needs hourly data for the start and the end of the hour in order to linearly interpolate, for meteo data, boundary conditions and biogenic emissions. The code keeps only these two instants in memory. Within the coarse time step, a call to routine **physics** is made, in order to define all variables that will be taken as constants during the time step (linear interpolation of meteo data, calculation of reaction rates etc...), and a loop on the fine time step is carried out, with a call to the numerical solver **twostep**. This latter routine calls the core routine **prodloss** which calculates, for all processes, the loss and production fluxes and sums them. Thus **prodloss** calls **emissions** (emission fluxes), **transmix** (transport and mixing fluxes), **chemistry** (chemical fluxes) and **deposition** (deposition fluxes). The routine transmix calculate the PPM concentrations at the boundaries with routines contained in the **Bppm.f** package.

3.3.5 Routine calls sequence for the aerosol version

3.3.6 Input files to chimere

As discussed above, a CHIMERE run consists in several actions which are performed in the top calling script: (1) running the interfaces, (2) building the INFOS and (3) executing **chimere.e**. This section describes the necessary input files to the third action. All these files must be present in the temporary directory before running the executable **chimere.e**. The top calling script copies and constructs these files in the temporary directory automatically. A short description is given here, for more details click on the link.

Static files that are just copied from the data_[model]_[meca] directory

- DEPO_SPEC: Dry deposition constants for gases

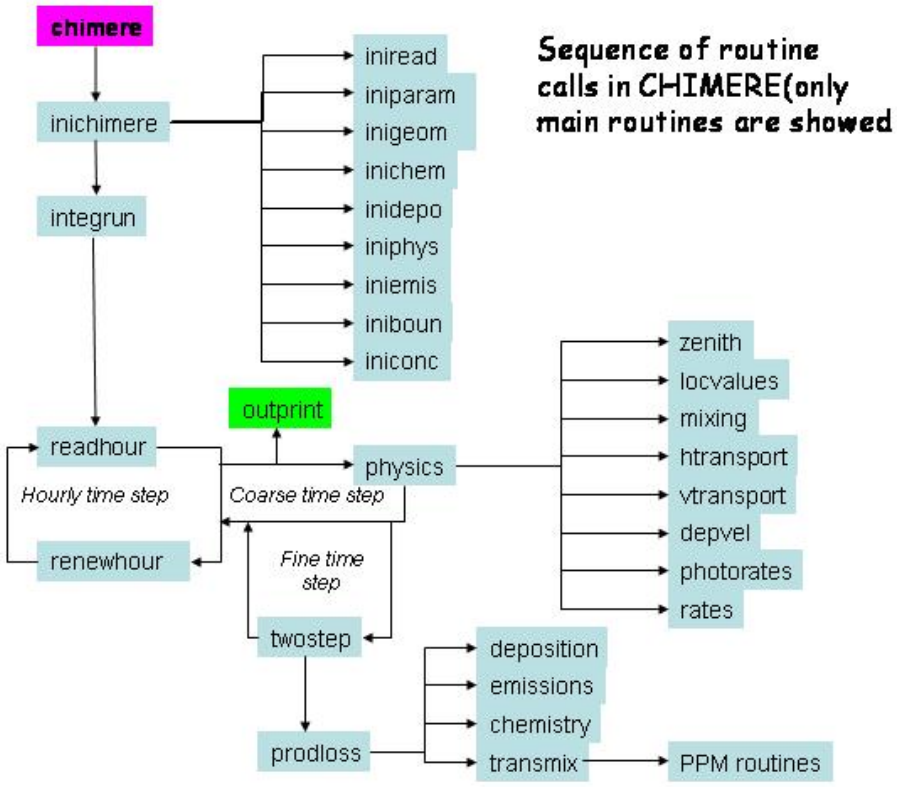


Figure 3.3: *Schematic general structure of CHIMERE*

- DEPO_LAND: Dry deposition constants depending on land-use category
- PPM_SPECIES: List of species transported by the PPM scheme
- OUTPUT_PARAMETERS: Meteo and other parameters to put in the par.[sim] output file.
- PHOTO_PARAMETERS: Tabulated photolysis rates from TUV and related parameters
- AEROSOL:
- AEROMIN:
- PRIMARY:
- ANTHROPIC: List of species emitted in anthropic emissions.
- BIOGENIC: List of species emitted in biogenic emissions.

Terrain files copied from the **domains** *directory*

- COORD_[dom]: Grid coordinates Longitude-Latitude
- LANDUSE_[dom]: Land-use fractions into 9 categories.

Chemistry files copied from the **data_[model]_[meca]** *directory and coming out of the* **chemprep** *preprocessor.*

- ACTIVE_SPECIES: List of all "active" species (in the RHS of reactions) in the mechanism.
- CHEMISTRY: Chemical reactions in CHIMERE format.
- STOICHIOMETRY: Stoichiometric coefficients for 4 different temperatures.
- REACTION_RATES: Reaction rate constants.
- FAMILIES: List and definition of family species (for output).

Intermediate files originating from the interfaces

- METEO: Meteorological 3D and 2D fields in one file. It is an output of **diagmet.f**.

- BEMISSIONS: Biogenic emissions 2D fields. It is an output of **diagmet.f**.
- AEMISSIONS: Anthropic emissions 2D fields. It is an output of the emission interface **prepemis_std.f**.
- LAT_CONCS: Lateral boundary conditions. It is an output of the boundary conditions interface **preplat_[bound].f**.
- TOP_CONCS: Top boudary conditions. It is an output of the boundary conditions interface **preptop_[bound].f**.

Other files

- INFOS: File containing the input file names and some options
- OUTPUT_SPECIES: List of output species names
- LAT_SPEC: List of boundary species
- TOP_SPEC: List of top boundary species

3.3.7 Output files

3.4 The chemical preprocessor: CHEMPREP

chemprep allows to modify reactions, add reactions or even completely change the chemical mechanism. It allows an easy writing of reactions and from two initial input files **REACTIONS** and **PHOTORATES** produces other files which are read by **chimere**.
There are however certain conditions for the chemical set to be fulfilled for **chemprep** to work as it is proposed, otherwise some rewriting has to be done in the pre-processor. Also the two input files must be follow a certain syntax. Once the two input files are prepared by the user, the preprocessor is invoked by running the **chemprep** script.

3.4.1 Conditions on the mechanism

The conditions for the set of reactions to be used are listed below:

- Only reactions of the type A1(+B1(+C1....)->(a*)A2+(b*)B2+... are accepted. Letters with index 1 denote the reactants and those with index 2 the products. a, b, (if any) are stoichiometric coefficients. If omitted they are considered to equal 1. Negative stoichiometric coefficients are accepted, but not recommended since it can lead to negative concentrations. Stoichiometry in the reactants is not allowed.
- Only species which are *consumed* at least once in the mechanism will be considered as **active species** in the chimere model. If the user intends to use in chimere a species which is only results of chemical reactions but which is not reactant in any reaction, she (he) should add an artificial reaction to the mechanism, for instance A->X (A is the concerned product), with a negligible reaction rate, i. e. k=1e-10 (s-1).
- There are special species whose concentration is only prescribed in the model. These are H2O (water vapor), N2 (molecular nitrogen), O2 (molecular oxygen) and M (80% N2 + 20% O2). Their values is given by the meteo data. Reaction with these species is possible with no constraint. Changes of these "prescribed species" imply several changes in the chimere code and is not an easy change.
- Reactions have a maximum of 4 reactant species and as many products as desired (less than 1000!). To increase the number of reactants is easy: change the parameter **nreactamax** in header file **chimere.h**.
- It may be desirable that stoichiometric coefficients vary with some parameters like temperature. This is however not possible here in this version.
- All reactions will be processed by the solver in an equivalent manner. There is no family grouping (like often used for the QSSA solver) possible.
- Clear sky photolysis rates must be precalculated by the user, and depend only on zenith angle Z. No dependence on height is allowed, even in the presence of clouds. The proposed tabulated values correspond to an average value between 0 and 1000m. This should be a good compromise for the boundary layer. Introducing a dependence with height will be considered in a future model version. The 12 Tabulated J values must correspond to zenith angles of (in degree) 90,88,86,83,80,75,70,60,45,30,15,0

- There is a possibility of defining *families*. These are groups of species whose concentrations sum can be output in the model. However these families are only used in outputs, not in the chemistry itself. See syntax in the next section.
- The mathematical expressions for reaction rates (the K's and J's) can take 20 forms, but here only a limited number is available. The table below show the available possibilities, together with associated syntax.

Type number	Name	Math	syntax in REACTIONS file
1	Constant K	K	k=*value*
2	Simplified Arrhenius	K=Aexp(-B/T)	k(T)=Aexp(-B/T), A=*value*, B=*value*
3	Arrhenius	K=Aexp(-B/T)(300/T)**N	k(T)=Aexp(-B/T)(300/T)**N, A=*value*, B=*value*, N=*value*
4	Troe or Falloff	K=K0*f**E/(1+M*(K0/K_∞)) K0=A0exp(-B0/T)(300/T)**N0 K_∞=A_∞exp(-B_∞/T)(300/T)**N_∞ E=1/(1+(Log10(M*K0/K_∞))**2) M=air-density	K(T,M)=troe(A0,B0,N0,A_∞,B_∞,N_∞,f)
5	Photolysis	J=J(Z)	J(Z)=photorate(*label*)[3].
6	Special type	K=K1*K2/(1+K2) K1=A1exp(-B1/T) K2=A2exp(-B2/T)	k(T)=SPECIAL_1(A1,B1,A2,B2)[4]
7	Special type	K=K1/(1+K2) K1=A1exp(-B1/T) K2=A2exp(-B2/T)	k(T)=SPECIAL_2(A1,B1,A2,B2)
8	Special type	K=2*sqrt(f1*f2*f3*f4/((1+f3)*(1+f4))) f1=A1exp(-B1/T) f2=A2exp(-B2/T) f3=A3exp(-B3/T) f4=A4exp(-B4/T)	k=SPECIAL_3(A1,B1,A2,B2,A3,B3,A4,B4)
9	Special type	K=2*sqrt(f1*f2)*(1-sqrt(f3*f4))*(1-f4)/(2-f3-f4) f1=A1exp(-B1/T) f2=A2exp(-B2/T) f3=A3exp(-B3/T) f4=A4exp(-B4/T)	k=SPECIAL_3(A1,B1,A2,B2,A3,B3,A4,B4)
10 to 12	Not used [5]		
13	Aggregated ozone photolysis	J=J0*H2O/(H2O+M*(0.02909*exp(70/T) +0.06545*exp(110d0/T))) H2O=water vapor (molec/cm3) M=air density (molec/cm3)	J(T,Z,H2O)=photorate(O3) [6]
14	Modified Troe [7]	Like Type #4 (Troe) but: E=1/(1+((Log10(M*K0/K_∞)-0.12)/1.2)**2)	K(T,M)=troe(A0,B0,N0,A_∞,B_∞,N_∞,f)
15 to 20	Not used [8]		

[3] *label* is a label standing for recognition of **chemprep** in the **PHOTORATES** file. All photolysis rates must appear tabulated in this latter file, the tabulated values corresponding to specific zenith angles (see above)

[4] * Special form for radical recombination (not used in this mechanism)

[5] User can define a new type. Changes must be done in **chemprep.awk** (recognition of a new syntax), **Binichem.awk** (define the number of constant **ltabrate**, and in **Brates.f** (new math expression)

[6] Aggregation of the ozone photolysis with reactions with atomic oxygen. The photolysis rates must be those of ozone only, hence the budget reaction is O3->2*OH.

[7] Fall-off specific to NO2+OH+M->HNO3

[8] User can define a new type. Changes must be done in **chemprep.awk** (recognition of a new syntax), **Binichem.awk** (define the number of constant **ltabrate**, and in **Brates.f** (new math expression).

3.4.2 Syntax of the REACTIONS file

Following are the rules for writing the REACTIONS file

- Each line is considered as a reaction except blank lines or lines starting with a "#" sign (left for comments).
- Species names must contain at most 15 characters.
- A reaction is written as: **A1(+A2(+A3(+A4)))->(a1*)B1+(a2*)B2+(a3*)B3...** *Rate expression* where **Ai** are the reactants (max=4), **Bi** the products and **ai** are possible stoichiometric (written if different from 1).
- *Rate expression* follows the syntax rules defined in the above table.
- Reactants can also be one (or several) of the prescribed species (H2O, N2, O2 or M).
- A1+A1(+A1)...->... are allowed (repetition of the same reactant, like HO2+HO2...
- Syntax like A1+A2-A3 is NOT allowed
- Negative stoichiometric coefficients are allowed, but not recommended (correctness of results not garanteed because never tested). In that case, however, use the syntax ±a1*B1 in the products.
- Families are defined using the syntax:
  FAM=A1+A2+A3...
  Ai must be active species, and can be repeated as many times as desired. There is no limitation to the number of family elements.

3.4.3 Syntax of the PHOTORATES file

The syntax is very simple: Each line contains information for one photolysis reaction. There can be more photolysis reactions than in the **REACTIONS** file. The unused ones are ignored. There must be an exact correpsondence between the **labels** of the photolysis reactions defined in the **REACTIONS** file. This label is the first column. The other columns contain the 12 tabulated photolysis rate values for the zenith angles from 90 to 0 degrees (see above table).

3.4.4 Files produced

Once the files **REACTIONS** and **PHOTORATES** are prepared, run the **chemprep** script. The produced files are (see description in the INPUT FILES to chimere Section above):

- ACTIVE_SPECIES (the list of "active species")
- CHEMISTRY (the list of reactions as read by subroutine **inichem**
- REACTION_RATES (the list of constants used for calculating rates)
- STOICHIOMETRY (the list of stoichiometric coefficients)
- FAMILIES (the list and definitions of families)

3.5 Meteorological interface

CHIMERE does not contain any meteorological driver. It is only a chemical-transport model, meaning that meteorological data have to be provided in order to perform a simulation. The code provided on this server only contains a meteorological interface which allows to start from standard meteorological variables and transforms them to CHIMERE inputs. In practice the meteorological interface consists in the construction of two files to be made available as input to the executable chimere.e in the temporary directory:

- The METEO file containing all hourly meteorological variables in a single file (click to see the data format).
- The BEMISSIONS file containing all biogenic emissions. In V200310, This file contains 2D hourly files with Isoprene, Terpene and Nitrogen Monoxide emissions.

The construction of these files is made in two stages:

- The first stage is an **interpolation stage**: it takes meteorological variables from a meteorological database, simulation or forecast file and transforms it into standard variables given on the horizontal CHIMERE grid as hourly values. **This stage has to be constructed by the user**. Specific files with specific formats and filenames must be built. Two example, with interface to a ECMWF data and to MM5 output files are given.

- The second stage is a diagnostic meteorological interface (the **src_[model]/diagmet.f**. It is a generic program allowing to transform standard meteo variables into CHIMERE inputs. It also allows construction of possibly non-available variables from standard variables and some parameterizations.

3.5.1 First stage

The top calling script calls a program named **interf-[metname]** in the **meteo** directory. This program (usually a bourne shell script) has to be written by the user. Two examples are given: **interf-ecm**, which is an interface to ECMWF data (there are 5 example days provided separately in the matter to download), and **interf-mm5** which is a generic interface to mm5. If other meteorological data is to be used, the user give it a name, [metname], and write the interf-[metname] bash program. It should also change the **"meteo="** line in the top calling script.
In the end, there are mandatory and optional variables that must be output by the **interf-[metname]** program. The constraints for the user is to provide (mandatory or more) variables into separate fortran-binary files with:

- Hourly values of the 3D and 4D variables starting on the first CHIMERE hour simulation and ending on the last simulation hour.
- Interpolated variables on the chimere horizontal grid. One should, by contrast, leave the data on the original vertical grid.
- A binary format which could be read in fortran as (for a 3D file):

```
do nh=0,nhours_simul
   read(unit)idate,(((var(i,j,k),i=1,nx),j=1,ny),k=1,nz)
enddo
```

idate is an integer giving the date in the form YYYYMMDDHH, which must therefore be incremented at each record, **nhours_simul** is the duration of the simulation (note that there must be 1 additional value relative to the simulation time in hour, **nx** and **ny** are the zonal and meridional dimensions of the CHIMERE grid, and **nz** is the number of meteorological levels used in the original data.

Here follows the lists of mandatory and optional variables which must or can be output by the interpolation stage.

Recommandations

There are a number of recommendations that can be given about the choices of the optional variables. Some are listed below.

- If Boundary layer height is not available from meteo data, it will be calculated using the maximum between a formula for stable air (using a critical Richardson number=0.5), and a parameterization for unstable boundary layer based on a simplification of the Cheinet [2002, PhD thesis] work. These calculations assume a dry boundary layer. The user should use meteorological model's boundary layer height if more elaborate and whenever possible, especially if it accounts for clouds.
- u_* is used for deposition and calculation of diffusivities. it is a particularly sensitive parameter for ozone in Summer through the calculation of aerodynamic resistance Ra. It should strongly depend on land use type, which are critical to deposition. In meteorological models, especially large-scale, roughness lengths are sometimes estimated too roughly for the application to high-resolution deposition. The recalculation of u_* which is proposed in the **diagmet.f** interface uses the Louis [1982] formulation which is particularly robust and gave satisfactory results. We advise to use it anyway, in particular to have deposition consistent with a high-resolution land use.
- Heat fluxes are also used for the calculation of W* and therefore mixing, and the height of the boundary layer. In fact only the virtual heat flux is required, which can be recalculated from an empirical formula [Priestley, 1948] using temperatures in the first meteorological model layers. This formula is not very accurate. It is strongly advised to use heat fluxes from the meteorological model.
- For radiation several calculation options are offered. All are based on a statistical model of the clouds optical thickness. The first option uses liquid/ice water (whenever available), a variable which can be very model-dependent and which is rarely verified against observations. Moreover it does not take into account fractionary cloudiness. However it may be suited for high/medium clouds which are often stratiform clouds in the midlatitudes. The

File Name	Variable	Dimension	Units
Mandatory Variables			
ZWIN_3D	Zonal component of the wind	3D	m/s
MWIN_3D	Meridional component of the wind	3D	m/s
TEMP_3D	Temperature	3D	K
MIXR_3D	Water vapor mixing ration	3D	Kg/Kg
ALTI_3D	Altitude of half layers on which all 3D meteo variables should be given	3D	m
PRES_3D	Pressure at half layers on which all 3D meteo variables should be given	3D	Pa
CLIQ_3D	Cloud liquid water content (not including rain water). Mandatory ONLY for the AEROSOL VERSION	3D	K
TEM2_2D	2m Temperature	2D	K
COPC_2D	Convective Precipitation. Mandatory ONLY for the AEROSOL VERSION	2D	mm/h
LSPC_2D	Large-scale Precipitation. Mandatory ONLY for the AEROSOL VERSION	2D	mm/h
Optional Variables			
RAIN_3D	3D cloud water content for radiation, optional for gas-phase, mandatory for aerosols	3D	Kg/Kg
RAIN_3D	3D rain water (to be added to cloud water for radiation)	3D	Kg/Kg
CICE_3D	3D ice (for radiation)	3D	Kg/Kg
SSHF_2D	Surface sensible heat flux, optionally read or recalculated	2D	W/m2
SLHF_2D	Surface latent heat flux, optionally read	2D	W/m2
USTA_2D	Friction velocity, optionally read or recalculated from Louis [1982]	2D	m/s
PBLH_2D	Boundary layer height, optionally read or recalculated	2D	m
LOWC_2D	Low cloud fraction, optional for radiation	2D	m
MEDC_2D	Medium cloud fraction, optional for radiation	2D	m
HIGC_2D	High cloud fraction, optional for radiation	2D	m

second option uses cloud fractions (whenever available), which is less meaningful to radiation but could be used when cloud/ice water is not available. The third option uses relative (to liquid water) humidity, and has been tuned using J(NO2) measurements taken over a long period near Paris (manuscript in preparation). The formulation, modifying radiation for RH>0.75 should only be used for low clouds. It has the advantage of being a more robust parameterization since it depends on a highly verified variable. In summary, our preference is: RH formulation (option 2 in METINFOS) for low clouds, and for liquid/ice water formulation (whenever available) for medium and high clouds, and cloud fractions otherwise.

The METINFOS file

The **METINFOS** file construction is the last part of the first stage of meteorological interface. It is a file that needs to be given to **diagmet.f**. This file gives informations to **diagmet.f**, in particular about the chosen options. In the two example interfaces it is constructed from a template file **METINFOS-ecm.sed** or **METINFOS-mm5.sed** where macros are replaced by values in the interface. In the METINFOS file should be placed in the temporary directory and look like (example drawn from MM5):

20030801 120 1132 [Starting date] [Lenght (hour) of simulation] [grid cell for screen display]
25 [Number of meteo half-levels]
1 [Recalculate (1) or read (0) u_*]
0 [Recalculate (1) or read (0) surface fluxes]
0 [Recalculate (1) or read (0) boundary layer height]

0 [Use ice water (1) for radiation or not (0)]
0 [Use rain water (1) for radiation or not (0)]
0 [Use urban corrections (1) or not (0) for wind, fluxes and 2m temperature]
2 01 15 [Low cloud radiation option, 0=cloud fraction, 1=liq. water, 2=RH] [Cloud bottom] [cloud top]
1 16 21 [Medium cloud radiation option, 0=cloud fraction, 1=liq. water, 2=RH] [Cloud bottom] [cloud top]
1 22 25 [High cloud radiation option, 0=cloud fraction, 1=liq. water, 2=RH] [Cloud bottom] [cloud top]

The ECMWF example interface

The **interf-ecm** interface starts from a number of ECMWF files in a data base, and constructs inputs to **diagmet.f** as described above. This example is made for ECMWF files with the first 10 (before 2000) or 18 ECMWF operational forecast model levels, 3-hourly sampled (by putting short-term forecasts one after another). It assumes a domain containing Western Europe. A tiny slice of these data, corresponding to the ESQUIF (pollution experiment over Paris) Intensive Observation Period #6, from 14 July 1999 to 18 July 1999, is provided in the download menu for testing the model. Otherwise ECMWF data is not public and can be ordered on the ECMWF[9] web server.
The files are supposed to be in binary format and be named as [variable]_bin.[year]. There are 2D data and 3D data. For instance, 2D data should be read as:

```
do nh=0,N
   read(unit)idate,(((var(i,j,k),i=1,nx),j=1,ny),k=1,nz)
enddo
```

Data should contain at least all hours to be simulated by CHIMERE. Files can contain more (before or after) data. In the example, data are assumed to be gather by yearly slices of time.
Click on the link above to see the interface bash program. The sequence of operations is as follows:

- Definitions of dates and directories variables, level number (which vary from year to year) and flag for availability of boundary layer height.
- Meteo horizontal interpolation: files are read one by one and variables are interpolated to the CHIMERE model grid, through the prgoram **exdom-ecm.f**. This requires the CHIMERE grid coordinates file **COORD_[domain]**.
- Time interpolation: This transforms the new temporary files (tmp/*-ecm) containing 3-hourly fields into 1-hourly fields (*_3D and *_2D) for **diagmet.f** input. This uses the linear interpolator program **tinter-ecm.f**.
- Then, with the **altipres.f** program, missing altitude and pressure mandatory files are constructed from the known hybrid σ-p ECMWF levels and from the surface pressure [SURP_2D].
- Finally the **sed** stream editor replaces macros in the **METINFOS-ecm.sed** file to form the **METINFOS** file.

The MM5 example interface

The **interf-mm5** interface starts from a UNIQUE MM5 file, **MMOUT_[mm5-domain]**, specified **exactly as in the V3.6 MM5 format**, see the NCAR[10] web site for a full documentation of MM5 and output format. The MMOUT file is assumed to be written with word byte order reversed relative to standard Linux words. In this interface it is assumed that the output frequency for MM5 simulations is 1 hour, so there is no time interpolation for input to CHIMERE.
Before using the MM5 interface, two coordinate files for the MM5 grid must be contructed manually. There is a file for the CRS coordinates (used for all variables but wind) and a file for the DOT coordinates (shifted grid for the wind fields). A utility for making these file coordinates is provided in the CHIMERE **util** directory. Edit the **make_mm5_coord** utility script for more help. These files (**COORD_MM5[domain]** and **COORD_MM5[domain]_DOT**) must be placed in the **meteo** directory once they are constructed.
There are no prior assumptions on the choice of physical options in MM5. In our appliccations, however, the following options have been used and can therefore be recommended:

[9] http://www.ecmwf.int
[10] http://ncar.ucar.edu

- Schultz microphysics,
- MRF (Troen-Mahrt) PBL
- Grell cumulus scheme
- RRTF radiation scheme
- In all applications over Europe where AVN/NCEP is used as forcing data, we found that nudging was necessary (grid FDDA option), with nudging coefficients varying from 0.00001 to 0.0001, depending on the spatial scale, but no meteorological observations were used.

Click on the link above to see the interface bash program. The sequence of operations is as follows:

- Preliminary definitions of variables
- Appllication of the **prepmm5.f** generic reading program. This program reads MM5 output data and converts them into separate binary files.
- Construction of weight files for regridding: The horizontal interpolation from the MM5 grid to the CHIMERE grid is bilinear. **It requires that a coordinate file of the MM5 grid be constructed beforehand** (see above).
- Regridding of each meteo variable file from the MM5 grid to the CHIMERE grid.
- Construction of the **METINFOS** file from the **METINFOS.sed** template file, for the diagmet.f options choices.

3.5.2 Second stage: the diagmet diagnostic model

The second stage of the meteorological interface consists in the production of the two input files for CHIMERE, **BEMISSIONS** and **METEO**, from a number of input files described below, among which the output files of the first stage. There is a unique program that achieves this stage: **diagmet.f**. It is a generic program, designed to be independent on the origin of the meteorological data. It is found in the **src_gas** or **src_aero** directories, depending on the use of the aerosol version or the gas-phase version. The aerosol version of **diagmet.f** requires and outputs more variables than the gas-phase version. These variables are the cloud liquid water and the precipitation.
The **diagmet.f** requires the following set of programs:

- ***_2D** files containing the hourly 2D meteorological variable files constructed by the first stage of the interface. They are in binary format and contain variables interpolated to the CHIMERE grid. There are mandatory and optional files (see first-stage section above).
- ***_3D** files containing hourly 3D variables, hourly, horizontally interpolated to the CHIMERE grid. The vertical grid is still that of the meteorological forcing model. There are mandatory and optional files (see first-stage section above).
- **METINFOS**: This is an ASCII file containing the meteorological option flags and parameters.
- **COORD_[domain]**: This is the coordinate file of the CHIMERE domain, which contains two ASCII columns of numbers describing the grid cell centres by longitude and latitude, from West to East and then from South to North. This file must be constructed once for each model domain using the Land-use (or an equivalent) interface.
- **LANDUSE**: File containing the fraction of land use per grid cell. There are 9 land use categories in this model version, see the Land-use interface below for the description of these classes. The file therefore contains 9 columns with numbers between 0 and 1. This file must be constructed once for each model domain using the Land-use (or an equivalent) interface. This file, which must be present in the temporary directory, is a copy of the **domain/COORD_[domain]** file.
- **BIOFACS**: Biogenic emission potentials file (see the Land-use interface for description. This file must be constructed once for each model domain using the Land-use (or an equivalent) interface. It is possible that the user wants to use his own biogenic emissions, calculated in a completely separate way. In this case, it is advised to replace this file by a file containing 5 columns with zeros and the number of lines equal to the the number of the grid cells, and define another interface. This file, which must be present in the temporary directory, is a copy of the **domain/COORD_[domain]** file.
- **VCOORD**: This file contains the vertical grid hybrid sigma-p coefficients for the top of all CHIMERE layers. It is either copied from the data_[model]_[meca]/VCOORD.[vgname] file, or constructed automatically using a geometric progression of layer thicknesses. The parameters of this method are defined in the top calling script. In fact the file describes the layer structure of the *mother domain*, if any, and the **lsubdiv** parameters in the top calling script define possible sublayers in nested domains.

- **LANDPAR**: This file gives some land-use/season dependent surface parameters: roughness height Zo, extra fluxes, wind reduction or extra urban heat-island temperature increment.

The output file **METEO** contains one record per hour and can be read as follows:

```
       read(unit)idate,(winz(1,nb),nb=1,nbox)  Zonal wind (m/s)
$                     ,(winm(1,nb),nb=1,nbox)   Meridional wind (m/s)
$                     ,(temp(1,nb),nb=1,nbox)   Temperature (K)
$                     ,(sphu(1,nb),nb=1,nbox)   Mixing ratio (Kg/Kg)
$                     ,(airm(1,nb),nb=1,nbox)   Air density (molec/cm3)
$                     ,(hlay(1,nb),nb=1,nbox)   Layer top altitudes (m)
$                     ,(tem2(1,nh),nh=1,nhoriz) 2m-temperature
$                     ,(atte(1,nh),nh=1,nhoriz) Cloud photolysis attenuation (0-1)
$                     ,(hght(1,nh),nh=1,nhoriz) Boundary Layer Heiht (m)
$                     ,(usta(1,nh),nh=1,nhoriz) u$_{*}$ (m/s)
$                     ,(aerr(1,nh),nh=1,nhoriz) Aerodynamic resistance (s/m)
$                     ,(obuk(1,nh),nh=1,nhoriz) Obukov length (m)
$                     ,(wsta(1,nh),nh=1,nhoriz) W* convective velocity scale (m/s)
```

idate is the date with YYYYMMDDHH integer format, **nbox** the total number of grid cells (horizontalXvertical), **nhoriz** is the number of horizontal grid cell per layer, and the other variables are described shortly above.
The record format for the aerosol version is slightly different, and contains 2 more variables: cloud liquid water content and precipitation.
The format of the **BEMISSIONS** file is similar, with 2D records containing the biogenic emission rates in molec/cm2/s for all biogenic species (presently isoprene, terpenes and NO). All species emissions are in the same record as above for meteorology.

3.6 Land-use interface

The land-use interface job is to construct files that depends on the CHIMERE model grid and domain only. Therefore these files only need to be constructed once per model domain. The interface is not called in the top calling script, and it is assumed when running the script that the output files are already constructed and placed in the **domains** directory. The user can either make his own interface to produce the files or use the (**prepdomains**) proposed interface which, at present, *works only for model domains over Europe*.

3.6.1 Model domains

In the following, [domain] is a name given by the user to the model horizontal grid/domain. Model domains are entirely defined by their grid cell centres description. It is given in the COORD_[domain] file which must be placed in the **domains** directory. The **COORD_[domain]** file contains lines with a longitude (in decimal degree) and a latitude (in decimal degree). Each pair of coordinates stands for a grid cell centre, described (from top to bottom of file) from West to East then from South to North.
Model grids can be any quasi-rectangular grid, with weakly varying spatial step. It is assumed, in particular for the Parabolic Piecewise Method for tranpsort that locally (over 5 consecutive cells) the grid size is constant in each direction, with sizes equal to that of the locally central cell. Sphericity effects are taken into account, but therefore linearized.
The file **DOMAINS** contained in the **domains** directory must then have a line containing the name of the domain, the number of cells in the x-direction, the number of cells in the y-direction, and a grid point number for outputing values on screen. When the model is run, either **diagmet.f** or **chimere.f** use this grid cell for displaying values on the screen.
There are predefined domains:

- **CONT1** (1/2x1/2 degree resolution over limited Western Europe domain)
- **CONT2** (1/2x1/4 degree resolution over the same domain)
- **CONT3** (extended 1/2x1/2 resolution, domain now including Spain, Portugal and Italy.
- **IDF** A domain of 150kmx150km around Paris with a 6-km resolution

3.6.2 Output files

The output files of the interface are directly ingested by the second-stage meteorological interface **diagmet.f** and by the CHIMERE main program **chimere.f**. They must be placed manually in the **domains** directory. Here is their description.

- The **COORD_[domain]** file is described above
- The **LANDUSE_[domain]** contains for each grid cell (one cell per file line), 9 values describing the fraction of land use covered by each of the 9 land-use types (see below) recognized by the model. The sum over each line must be equal to 1. The ordering of grid cells is as in **COORD_[domain]**
- The **BIOFACS_[domain]** contains, for each grid cell, 5 values describing the biogenic emission potentials to be combined with the meteorological factors (temperature, PAR) in order to construct the hourly biogenic emission file **BEMISSIONS** in the second-stage meteorological interface **diagmet.f**.
  - The first column is for isoprene emission. It is the isoprene landuse-dependent coefficient Eo (in molec/cm2/s), in the cell, which has to be multiplied by the meteorological factors
  - The second and third columns are the coefficients E_1 and E_2 coefficients for terpene emissions.
  - The fourth and fifth columns are the emission potential Eo for biogenic emission for two seasons: growing season (4th column), assumed to be from may to July, and for the other months (5th column).

Biogenic emissions are calculated in the **Bcalbio.f** subroutine which is called by **diagmet.f**.

3.6.3 Land-use types

In the V200310F model version, 9 land-use types are recognized. They are described here, with the correspondance to the GLCF (University of maryland) data base used for the prpoposed **prepdomains** interface:

Category Number	Description	Corresp. GLCF (0-13)
1	Agricultural land / crops	[11]:Cropland
2	Grassland	[10]+1/2[7]:Grassland+1/2Wooded grassland
3	Barren land/bare ground	[12]:Bare ground
4	Inland Water	
4	Urban	[13]:Urban and built-up
6	Shrubs	[8]+[9]: Shrublands
7	Needleaf forest	[1]+[3]+1/2([5]+[6])+1/4[7]
8	Broadleaf forest	[2]+[4]+1/2([5]+[6])+1/4[7]
9	Ocean	[0]: Water

3.6.4 The prepdomain interface for Europe

There is not one way to construct the previous output files. Different methodologies can be used for calculating biogenic emissions and landuse. However we propose here one way to proceed, the **prepdomains** interface. It is based on the Global Land Cover Facility (GLCF[11]) 1kmX1km resolution database from the University of Maryland, following the methodology of Hansen et al. (2000, J. Remote Sensing).
The **prepdomains** interface downloadable for CHIMERE (see download menu) applies only to Europe because (i) Only a subset of the GLCF land cover is used (over Europe) and (ii) The tree distribution used here for biogenic emissions contains only european countries informations. Therefore biogenic emissions procedure should be completely changed for other continents.
By "Europe" it is meant the domain extending from 15W to 30E and 34N to 61N.
Once the tar file has been downloaded and extracted (about 1gbytes !!!), one should find several files:

[11] http://glcf.umiacs.umd.edu/data/landcover

- **LAND_EUR1**: An ASCII piece of the GLCF landuse classes: Format: Longitude (decimal degree), Latitude (dec. degree) and land cover category [0-13] (see the GLCF documentation on the GLCF[12] web site.
- **COUNTRY_EUR1**: An ASCII file giving the country code (see the order in first line of **FACT_FOREST** file).
- **FACT_FOREST**: ASCII file giving the tree distribution by country, and the total amount of Fertilizer used (second line, in Kg/ha/Year of Nitrogen), from Stohl (1996, Atmospheric Environment).After Line 3 follow the tree species distribution.
- **FACT_RATE**: These are emission factors for individual trees and other characteristics (deciduous or permanent leaves).
- **FACT_LAND**: Emission factors from soil (unit ???)
- **LAND_AGGREGATION**: This is the matrix of correspondence between the 14 GLCF categories and the 9 CHIMERE categories.
- **README**: Read it
- **biofacs.f**: Program constructing the biogenic emission potentials, given the land-use and the other distribution (trees, countries, ...).
- **landuse_ng.f**: Calculates the 9-category fractions on the CHIMERE grid from GLCF.
- **countries_ng.f**: Calculates country area fraction on the CHIMERE grid
- **prep_ng**: This is the program to edit, modify and execute for each new domain.

The interface use is very simple: Provide first, in the **prepdomains** directory, the **COORD_[domain]** file containing the centres of the grid cells. Then edit **prep_ng**. Modifiy the first 3 program lines: domain name, nxx= and nyy=. Then simply execute the **prep_ng** program. This can take 5-30 minutes (or more...). At the end, a directory with the domain name is created containing the necessary files for CHIMERE. Move these files to the CHIMERE **domains** directory. Then edit the **domains/DOMAINS** files and add a line describing the new domain. Interface work is done (if things worked correctly...).

3.6.5 Extension to other areas

The extension to other areas than Europe is not straightforward, but can be managed in different manners. One has to differentiate the construction of **LANDUSE_[domain]** file and that of the **BIOFACS_[domain]** file. In the first case, the methodology available here for Europe is fairly general. The solution is to download the original 1KmX1Km GLCF land cover data base, and use an extraction program as the extract_eur.f program (provided separately). In this program it is assumed that the name of the GLCF global file is **latlon**. Take therefore the latitude-longitude version of this file from the GLCF server.
The extension of biogenic emissions to other regions than Europe is much more difficult. While the general methodology remains correct, the tree species list must be changed for other continents, especially for tropical areas. The tree distribution per country has to be redefined with new countries, and the tree emission factors must be redefined. Also the biogenic emission of NO methodology from Stohl (1996) may be difficult to extend to other areas in the world, at least to tropical areas where soil properties may be very different.
Another possibility is of rebuilding a new methodology for constructing the **BIOFACS_[domain]** file, which would be more adequate to the new area. Also biogenic emissions can be calculated separately on an hourly basis instead of on a "emission potential X meteo factor" methodology. In this case it is advised to set the numbers in the **BIOFACS_[domain]** file to 0 and to read biogenic emissions and build directly a **BEMISSION** file from an interface program that would be called just after the anthropic emission interface in the top calling script. Sometimes also biogenic emissions are simply included in the anthropic emissions, in which case the **BIOFACS_[domain]** file should be set to 0.

3.6.6 Steps for the definition of a new CHIMERE domain

In the definition of a new CHIMERE domain, a certain number of steps have to be followed:

- Check whether domain is quasi-rectangular. Most projection systems work, even a regular grid in geographic coordinates (longitude-latitude) provided the resolution is not too coarse (say > 2 degree).

[12] http://glcf.umiacs.umd.edu/data/landcover

- Check whether (i) The domain is in Europe as defined above or (ii) not.
- Build a **COORD_[domain]** file containing longitude and latitudes of the grid cell centres.
- In the (i) case, use the **prepdomains** interface or your own interface to build the **LANDUSE_[domain]** and the **BIOFACS_[domain]** files. Using **prepdomains** is achieved by editing the **prep_ng** script and changing the first 3 program lines.
- In the (ii) case, biogenic emission potentials or direct hourly biogenic emissions have to be built in a different way. However the **LANDUSE_[domain]** file can still be built using the **prepdomains** interface, but the input land cover file has to be rebuilt from the GLCF database. See the above Section.
- Once the files are constructed, move them into the **domains** CHIMERE directory.
- In the **domains** CHIMERE directory, edit the **DOMAINS** file and add relevant information for the new domain.

3.7 Anthropic emission interface

This server is not designed to provide a model for constructing emissions. It assumes that the user has his own emissions data bases. This Section explains how to provide data to the interfaces that are in the CHIMERE package. CHIMERE requires a file containing hourly anthropic emissions for the time period to simulate. This file must be placed in the temporary directory and its name is **AEMISSIONS**. The format of this file is such that it can be read using the following loop:

```
do nt=1,ntimes
   read(iunit)idate,
   (((emi(ne,nh,nl),nh=1,nhoriz),nl=1,nlevemis),ne=1,nemisa)
enddo
```

where :

- **idate** is an integer in the form YYYYMMDDHH giving the current date. The emission values refer to those during the hour before this date and are taken constant during this hour
- **ntimes** is the number of hours of the simulation period,
- **nhoriz** is the number of horizontal grid cells ordered as usual from west to east then south to north,
- **nemisa** is the number of anthropic emitted species, which depends exclusively on the chemical mechanism. One assumes exactly the same order in the **nemisa** species as that in the **ANTHROPIC** file.
- **nlevemis** is the number of *emission levels*. Emissions can be given on the first few **nlevemis** model layers, for instance to account for high large source points.

The construction of the **AEMISSIONS** file can be done in two ways: either by using the standard interface provided here, which however requires initial data in a specific format, or by a specific, user-provided interface. The standard procedure proposed here consists in two steps:

- The construction of an emission data base consisting of 12 files containing each hourly emissions for 3 typical days for a given calendar month,
- The construction of the **AEMISSIONS** file from this data base.

The first stage is the most difficult task, and there is no software on this server to perform this stage. We give here only indications on how to proceed in practice in most cases. The output of this stage are 12 ASCII files, named **EMISSIONS-[domain].[MM]** ([MM] is the calendar month) like those provided for European or Ile-de-France in this server in the **dowload** Section. The format is such that it can be read by fortran as:

```
do ne=1,nemisa
do nt=1,ntypedays
do nh=1,nhoriz
   read(iunit,*)((emisprof(ne,nt,nh,ih,nl),ih=1,24),nl=1,nlevemis
enddo
enddo
enddo
```

ntypedays is 3, **nt=1** standing for weekday emissions, **nt=2** for saturday emissions and **nt=3** for sunday emissions. This does not allow, in particular, to have emissions depending on meteorology or other factors than the day of week, but this keeps databases relatively compact.
In general, in order to construct model-species emission files like these from emission inventories, the starting point can be an annual inventory of VOC, NOx, CO, SO2, NH3 and PM per activity sectors. Then the annual totals undergo time disaggregation by applying typical monthly and hourly profiles depending on week day. Such profiles depend very much on the application and no such profiles are given on this server.
Then the VOC and NOx are disaggregated into real compounds using a speciation profile (depending on activity sector). In the data provided on this server, the NOx speciation is 90% of the NO2-equivalent mass of NOx is NO, 9.2% is NO2 and 0.8% is HONO. The VOC speciations come from either the british PORG speciation for the continental model or from the University of Stuttgart (IER) for the Ile-de-France case. The british speciation is public.
The construction of model-species emissions then requires a Middleton (1990) type of aggregation procedure. Each real VOC emissions are aggregated into one or several model VOC(s), with a branching coefficient and a reactivity-dependent coefficient (see the Middleton procedure), for which reactivities and molar masses for both the real and the model VOCs have to be known. In the british speciation, 227 real VOCs are considered, and the following **AGGREGATION** file gives these informations for the MELCHIOR mechanism. Reactivities have been calculated on the basis of the Leeds University MCM Master Mechanism and the Kwok structure/reactivity approach. The **SPECIATION** files gives the VOCs list and the 9-activity speciation, and the way VOCs are preprocessed in CHIMERE for continental scale.
The second stage consists in making the hourly, simulation adapted, **AEMISSIONS** file from the database. This is a very simple procedure, achieved by the **src_[model]/prepemis_std.f** program. Take a look at this very simple program. It calls a calendar routine that identifies day types and possibly extra fixed holidays taken as Sundays (Christmas, 1st January, ...). Beware that these days are for France. One may want to edit this subprogram (idaytype in src_[model]/Bcalendar.f) and changes a few specific dates.
In this second stage there is a possibility to apply a flat reduction (or increase) on all primary pollutants, in order to account for time shift between the simulation year and the inventory year. Unfortunately the possibility is not given to distinguish factors for various pollutants, but this is not a difficult change in **src_[model]/prepemis_std.f**.

3.8 Boundary conditions interface

There are basically three options for giving CHIMERE boundary conditions:

- Either boundary conditions are provided by the user from external data, in which case he should only follow the guidelines on the format of files.
- This is a nest run and a coarse-resolution run is used for boundary conditions (the CHIMERE option).
- This is a "coarse run" and climatological boundary conditions from MOZART II are used. See the model documentation about these climatological values.

Boundary conditions must be provided to CHIMERE in 4 files: Two of them, containing the list of species to be passed at the boundaries, are predefined files and must lie in the **data_[model]_[meca]** directory. The other two files are concentrations files and are constructed dynamically from the top calling script, using two interfaces, the **preplat** interface for lateral boundaries and the **preptop** interface for the top boundary. Here are the files, which must be made available in the temporary directory:

- **LAT_SPEC**: List of species passed along lateral boundaries. Species names must lie in the **ACTIVE_SPECIES** list. The file is copied onto the temporary directory by the top calling script from the **data_[model]_[meca]/LAT_SPEC_[bound]** file depending on chosen options. [bound] can be either CHIMERE or MOZART2.
- **TOP_SPEC**: List of species passed along top boundary. Species names must lie in the **ACTIVE_SPECIES** list. The file is copied onto the temporary directory by the top calling script from the **data_[model]_[meca]/TOP_SPEC_[bound]** file depending on chosen options. In this model version only **[bound]=MOZART2** is available.

- **LAT_CONCS**: A binary file containing lateral mixing ratios in ppbv. A CHIMERE flag (in the **chimere.h** header file (macro _IOPBOUNLAT_) in **chimere.h.sed** controls whether lateral boundaries are fixed or time varying during the simulation. When using MOZART2, the fixed option is automatically chosen. When using CHIMERE the variable option is chosen. In the fixed case, the LAT_CONCS file is written as :

```
      write(51)id0,(((concc(nb,nl,ns),ns=1,nspecs)
$              ,nb=1,nhbound),nl=1,nlayers)
```

  where id0 is a date (starting date), nhbound is the number of lateral boundary cells per layer (2[nx+ny]), nspecs is the number of lateral species, and nlayers is the number of layers. The species order **must exactly correspond to the order of the species listed in the LAT_SPEC file**. In the time varying case, the above binary format remains valid but with one record per simulation hour (+1 for initial condition). This file is dynamically constructed by the interface program **preplat_[bound].f**.
- **TOP_CONCS**: file containing the top boundary mixing ratios in ppb. The file format follows the same logic as for LAT_CONCS, one date per record (actually one record with MOZART2 option), and the record is written as:

```
write(51)id0,((concc(nh,nlayers,ns),ns=1,nspecs),nh=1,nhoriz)
```

  where now **nhoriz** is the number of horizontal grid cells in the same order as in the **COORD_[domain]** file.

When nesting (1-way) a CHIMERE high-resolution simulation into a coarse-resolution one, the simulation must be ready (**out.[label]** file), and the species must be ordered, in this file, exactly as they appear in the **LAT_SPEC** file. The address of the coarse-run simulation file is given in the top calling script, as well as the option for nesting and other necessary parameters of the coarse run.

User-built interfaces should bring to the input of CHIMERE the files described before, with the option of either fixed or time-varying, hourly concentrations.

3.9 Initial Conditions

There are two types of initial conditions that can be used for CHIMERE: Initialized concentrations and non initialized concentrations. In the first case a file named **ini.[label]** and placed in the simulation directory (the same as where the CHIMERE output files **out.[label]**, **par.label** and **bend.[label]** are). In the second case there is no need to provide a file; initial conditions will be taken as interpolated from the boundaries, layer by layer. There is a flag (**iopinit**), to be passed to **chimere.f** by the top calling script saying whether simulation is initialized or not.

It is clear that initialization is important for large-scale simulations (domains larger than about 500 Km X 500 km). Under extremely stagnant conditions, spin-up of one or two days may also be necessary at local (100 Km) scale. At regional scale, the spin-up time is in the order of 4-5 days (for ozone and precursors.

The format of the **ini.[label]** file is exactly the same as that of the **end.[label]** file. It contains the concentrations in the following binary format:

```
do ns=1,nspec
   read(unit)ii,(conc(ns,nb),nb=1,nbox)
enddo
```

where **nspec** is the total number of active species (i.e. the number of lines of the **ACTIVE_SPECIES** file, and **nbox** is the total number of grid cells.

Concatenating short simulations in order to perform a long-term simulation is really easy: At the end of each simulation the **end.[previous_label]** can be copied or linked to the **ini.[new_label]** file.

3.10 Running the model

3.10.1 Running the model directly without the top calling script

3.10.2 Running the model with the top calling script

3.11 Visualization and post-processing

3.11.1 The output files

Three different formats of outputs are delivered: one on screen (just for a run control) and two on files:

- Chemical fields on **out.**<**SIM**>
- Additional parameters on **par.**<**SIM**>

The output on screen

Outputs on screen are predefined in the routine **Boutprint.f** to give some informations about the run. The choices of outputs are:

- The current date in format `yyyymmdhh`
- The first two species defined in the **OUTPUT_SPECIES** file
- The 2m temperature (in Kelvin)
- The surface wind speed (in m/s) (in fact, this value correspond to the first vertical levels wind speed, typically 30-40m AGL).
- The boundary layer height (in meters) diagnosed using the **diagmet.f** pre-processor
- The radiative attenuation: from 0 to 1 with 1 for the complete clear sky.

All these values are given for the prescribed **nsho** value. This value may be changed in the **top calling script** or directly forced in the **chimere.h.sed** common file. An easy way to change these outputs is to modify the **OUTPUT_SPECIES** file. The two first species (in ppb) will appear on screen during the simulation. A more complicated way is, of course, to change directly the fortran source, at the end of the subroutine.

The chemical fields outputs on file

Some three-dimensional chemical fields are written in an output *binary* file called **out.**<**SIM**>. The chosen species correspond to those listed in the **OUTPUT_SPECIES** parameter file. Values are expressed in ppb. The fortran lines to be used to reread this file are as follows:

```
      write(ifnout)idate(ihourrun),((toprint(no,nb),nb=1,nbox)
     $            ,no=1,noutspec),(real(hlayloc(nb)*1d-2),nb=1,nbox)
     $                           ,(real(airmloc(nb)     ),nb=1,nbox)
```

where `idate(ihourrun)` is the current date, `toprint(no,nb)` is the concentration table (only used for output in the model), `hlayloc(nb)` is the corresponding vertical height (for each model mesh and expressed in meters) and `airmloc(nb)` the air density.

The additional fields outputs on file

An additional output file is named **par.**<**SIM**>. Its goal is to provide meteorological outputs, corresponding to the values employed during the run to calculate the concentration fields written in **out.**<**SIM**>. The extracted parameters are hardcoded in the routine `Boutprint.f`. The possible values are displayed in the Table 3.2. To choose the parameters to extract, the user may configure the parameter file named **OUTPUT_PARAMETERS** (with the flag 0 for no outpout and 1 for output).

Model Parameter	Variable	Default flag [0/1]
wind		
iwinhs	Wind speed $\|U\|$ (m/s 1st vertical level)	1
iwinh2	Wind speed $\|U\|$ (m/s 2nd vertical level)	1
iwinvs	Vertical wind speed (m/s 1st vertical level)	0
iwinvt	Vertical wind speed (m/s top domain vertical level)	1
specific humidity, air density and temperature		
isphu	Specific humidity	1
iairm	Air density	0
itemp	Temperature (K)	1
iwinxs	Velocity-equivalent of mixing exchange rate (surface)	0
iwinxt	Velocity-equivalent of mixing exchange rate (top domain)	0
ithlay	Thickness of the model layer (in cm)	0
mixing height, radiation		
ihght	Mixing height (in meters)	1
iatte	Radiation attenuation [0-1]	1
izeni	zenithal angle (degrees)	1
item2	2m temperature (K)	1
diagnostic turbulent parameters		
iusta	Friction velocity u_* (m/s)	1
iaerr	Aerodynamical resistance	1
iobuk	Monin-Oblukov lengtth	1
iwsta	Convective velocity w_*	1
idepo	Dry deposition velocity	1

Table 3.2: *The parameters flags for specific model outputs on file: here are presented the DEFAULT values*

3.11.2 The mapit visualization example program

There is no postprocessing software in this version, only one mapping utility, **mapit**.
In order to be able to use **mapit** the GMT3.x software should be installed. Next, in the **mapit** directory, the **exmap.f** program should be compiled using **f77 exmap.f -o exmap.e**.
Edit the **mapit** shell program and change the few top lines if necessary. Once ready use the following syntax for execution:
mapit *domain_name date pollutant level output_file*
date must be given in a YYYYMMDDHH format, pollutant must be among the **OUTPUT_SPECIES** list, level is the model level and the file is the **out.SIM** file (give absolute or relative path).
Predefined palette files for some pollutants are given in the **palettes** subdirectory. For some other pollutants the user must define a new palette file by using the GMT command **makecpt**.
Once executed the output of **mapit** is the **map.ps** file, which can be viewed by the **gs** command, but which can be redirected in another file. This is a directly printable PostScript file.

The color palettes

The predefined palettes are described in the Table 3.3.
These palettes were defined to be the most adaptated to 'usual' regional surface concentrations. You can create a new palette by using the following command line:

makecpt -C TABLE -T vmin/vmax/pas > OUT_PALETTE_FILE

with:

Filename	Min	Max	Step
CH4.cpt	1700	1900	20
CO.cpt	0	1000	50
HNO3.cpt	0	20	1
NO.cpt	0	50	2.5
NO2.cpt	0	150	10
NOY.cpt	0	150	10
O3.cpt	0	150	10/50
PAN.cpt	0	6	0.3
pHNO3-reg.cpt	0	50	5
pASOA.cpt	0	200	5/10/50
pH2SO4.cpt	0	200	5/10/50
pHNO3.cpt	0	200	5/10/50
PM10.cpt	0	200	5/10/50
PM25.cpt	0	200	5/10/50
pNH3.cpt	0	200	5/10/50
pNH3-reg.cpt	0	20	2
pPPM.cpt	0	200	5/10/50
SO2.cpt	0	100	10

Table 3.3: *The predefined color tables to be used with the GMT software for concentrations maps*

- TABLE : The GMT predefined color tables: type 'makecpt' in a x-terminal to have the complete list. You can use, for example:
  - ☐ gray: Grayramp from black to white [0/1]
  - ☐ rainbow: Rainbow colors: magenta-blue-cyan-green-yellow-red
  - ☐ cool: Linear change from blue to magenta
- vmin / vmax / pas : the minimum/maximum and step values (examples are provided in the Table 3.3)
- OUT_PALETTE_FILE the output file for this new color palette. Please note that if the name is not composed of the species name and the ".cpt" extension, you have to change the corresponding line in the `mapit` script.

Chapter 4

References using CHIMERE

4.1 Peer-Reviewed Articles

1. Beekmann, M.; Derognat, C., 2003: Monte Carlo uncertainty analysis of a regional-scale transport chemistry model constrained by measurements from the Atmospheric Pollution Over the Paris Area (ESQUIF) campaign, J. Geophys. Res., 108, D17, 8559.

2. Blond, N., Bel, L. and Vautard, R., 2003: 3-D ozone data analysis with an air quality model over the Paris area. J. Geophys. Res., in press.

3. Derognat, C.; Beekmann, M.; Baeumle, M.; Martin, D, and Schmidt, H., 2003: Effect of biogenic volatile organic compound emissions on tropospheric chemistry during the Atmospheric Pollution Over the Paris Area (ESQUIF) campaign in the Ile-de-France region, J. Geophys. Res., 108, D17, 8560.

4. Honoré, C., Vautard, R., Beekmann, M., 2000: Low and High Nox chemical regimes in an urban environment, Environmental Modelling and Software, 15, 559-564.

5. Honoré, C., Vautard, R. and Beekmann, M., 2000: Photochemical regimes in urban atmospheres: the influence of dispersion. Geophys. Res. Letters, 27, 1895-1898.

6. Menut, L., Vautard, R., Beekmann, M., Honoré, C., 2000: Sensitivity of photochemical pollution using the adjoint of a simplified chemical-transport model. J. Geophys. Res., 105, D12, 15,379.

7. Menut, L., 2003: Adjoint modeling for atmospheric pollution process sensitivity at regional scale, J. Geophys. Res. 108, No. D17, 8562.

8. Schmidt, H., Derognat, C., Vautard, R., and Beekmann, M., 2001. A comparison of simulated and observed ozone mixing ratios for the summer of 1998 in Western Europe. Atmospheric Environment 35, 2449-2461.

9. Schmidt, H.; Martin, D., Adjoint sensitivity of episodic ozone in the Paris area, to emissions on the continental scale, J. Geophys. Res., 108, D17, 4313.

10. Sillman, S.; Vautard, R., Menut, L., Kley, D., 2003: O3-NOx-VOC sensitivity and NOx-VOC indicators in Paris: Results from models and Atmospheric Pollution Over the Paris Area (ESQUIF) measurements. J. Geophys. Res. 108, D17, 8563.

11. Vautard, R., M. Beekmann, and L. Menut, 2000, Applications of adjoint modelling in atmospheric chemistry: sensitivity and inverse modelling, Environmental Modelling and Software 15, 703-709.

12. Vautard, R., M. Beekmann, J. Roux and D. Gombert, validation of a hybrid forecasting system for the ozone concentrations over the Paris area. Atmos. Environ., 35, 2449-2461.

13. Vautard, R. Martin, D, Beekmann, M. Drobinski, P., Friedrich, R., Jaubertie, A., Kley, D., Lattuati, M., Moral, P., Neininger, B., Theloke, J., 2003, Paris emission inventory diagnostics from ESQUIF airborne measurements and a chemistry transport model, J. Geophys. Res. 108, D17, 8564.

4.2 Submitted manuscripts

1. Bessagnet, B., Hodzic, A., Vautard, R., Beekmann, M., S. Cheinet, Honoré, C., Liousse, C., Rouil, L., 2003, Aerosol modeling with CHIMERE - Preliminary evaluation at the continental scale, Atmospheric Environ., submitted, in revision.

2. Labouze, E., C. Honoré, L. Moulay, B. Couffignal, and M. Beekmann, 2003, Photochemical ozone creation potentials : a new set of characterization factors for different gas species at the scale of the western Europe, Int. Journal of Life Cycle Assessment, submitted.

3. Wackernagel, H., C. Lajaunie, N. Blond, C. Roth and R. Vautard, 2003: Geostatistical risk mapping with chemical transport model output and ozone station data, submitted to Journal of Ecological Modelling.

4.3 Other documents

1. Beekmann, M, R. Vautard, C. Honoré et L. Menut, Climatological uncertainty analysis of the impact of emissions reductions on photo oxidant build-up in the Paris agglomeration, Atmospheric Pollution Modelling and Simulation conference, Champs-sur-Marne, Octobre 1998.

2. Beekmann, M., R. Vautard, C. Honoré et L. Menut, Monte Carlo uncertainty analysis of the effect of precursor emission reductions in an urban agglomeration, EGS Symposium, La Haye, Avril 1999.

3. M. Beekmann, C. Derognat, Monte Carlo uncertainty evaluation of emission reduction scenarios constrained by observations from the ESQUIF campaign, EGS Symposium, Nice, April 2002.

4. M. Beekmann, Regional scale simulation of photooxidants with the CHIMERE model, Chemical weather workshop, DLR, Oberpfaffenhofen, May, 2002

5. Blond, N., 2002 : Assimilation de donnees photochimiques et prevision de la pollution tropospherique. Thèse de doctorat de l'Ecole Polytechnique, Palaiseau, France.

6. Blond ; N. ; 2000 : Elaboration de modèles de prévision des épisodes de pollution photochimique à l'échelle nationale, Validation, Tech. Rap. I, ADEME.

7. Blond ; N., 2000 : Elaboration de modèles de prévision des épisodes de pollution photochimique à l'échelle nationale, Cartographie, Tech. Rap. II, ADEME.

8. Blond, N., Bel, L., and R. Vautard, 2002: Data analysis with a multiscale chemistry-transport forecast model, 82nd American Meteorological Society (AMS) annual meeting, Orlando, Floride, preprint Symposium on Observations, Data assimilation and Probabilistic prediction, Paper 3.4, 2002.

9. Blond, N., and R. Vautard 2001, Data analysis with a multiscale chemistry-transport forecast model, 2nd International Conférence on Air Pollution Modelling and Simulation (APMS), Paris, France, 2001, Sportisse editeur.

10. Blond, N., and R. Vautard 2000, Assimilation de données avec un modèle multi-échelles de photochimie, Atelier de modélisation de l'atmosphère (AMA), Toulouse, France.

11. Derognat C., M. Beekmann, R. Vautard, Elaboration d'un code photochimique applicable à l'étude de la pollution photo oxydante en milieu urbain et rural, Procédures de l'Atelier de Modélisation de Météo-France, Toulouse, Décembre 1999.

12. Derognat, C., H. Schmidt, M. Beekmann, R. Vautard, Biogenic emission impact on photochemical formation of ozone in Western Europe during an extremely hot period, EGS Symposium, Nice, Avril 2000.

13. Honoré, C., 2000 : La photochimie de l'Ozone a l'echelle urbaine, un systeme dynamique non-lineaire. These de doctorat de l'Universite Pierre et Marie Curie, Paris, France

14. Roemer M. , M. Beekmann, R. Bergstrom, G. Boersen, H. Feldmann, F. Flatoy, C. Honore, J. Langner, J. E. Jonson, J. Matthijsen, M. Memmersheimer, D. Simpson, P. Smeets, S. Solberg, R. Stern, D. Stevenson, P. Zandveld, Z. Zlatev, Ozone trends according to ten dispersion models. EUROTRAC special publication, International Scientific Secretariat ISS, GSF National Research Center for Environment and Health, Munich, 2003

15. H. Schmidt, R. Vautard, M. Beekmann, and C. Derognat, Application of a European scale CTM to ozone pollution events in the Paris area, EGS Symposium, Nice, Avril 2000.

16. Vautard, R., M. Beekmann, L., Menut and M. Lattuati, Applications of adjoint modelling in urban air pollution, in Proceedings of EUROTRAC Symposium 98, 1999.