IFS DOCUMENTATION – Cy50r1
Operational implementation 12 May 2026

PART VIII: ATMOSPHERIC COMPOSITION

European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, UK

REVISION HISTORY

Changes from CY49R1 to CY50R1

Chapter 2 - Conversion of aerosols and trace gases

- Update of heterogeneous chemistry parameterization
- Revision of $SO_2 + H_2O_2$ aqueous phase reaction
- Implementation of an equilibration time parameterization for nitrate/ammonium formation
- Revision of hydrophilic growth of OM, SOA and fine mode nitrate
- Correction of hydrophilic growth factor of sea salt for sedimentation

Chapter 3 - Surface fluxes of atmospheric composition

- Activation of online biogenic VOC emissions
- Switch to anthropogenic emissions CAMS-GLOB-ANT M1 (replacing CAMS-GLOB-ANT 6.2)
- Activation of weekly cycle in anthropogenic emissions
- Reduced ground resistances for ozone $R_{gs,O}$
- Sedimentation velocity computed online for all species, using the Stokes formula
- Switch to GFAS-fire emission 1.4.2 (MODIS and VIIRS-based)

Chapter 5 - Diagnostics

- added diagnostics field such as tropospheric columns and dust PM

Chapter 6 - Data assimilation

- Updated background errors for NWP, now taken from the operational EDA
- Increase resolution for minimisation (inner loop)
- Updated background errors for O3 to remove spurious long-range correlations
- Assimilation of Sentinel-3 (A/B) AOD over ocean

Changes from CY48R1 to CY49R1

Chapter 2.2 - Atmospheric Aerosol

- EQSAM4Clim module activated for gas-phase aerosol partitioning
- Computation of aerosols, cloud and precipitation pH
- Aerosol optical properties to account for hydrophilic and aspherical dust
- Size distribution consistent with optics - revised PM2.5 formulae to reflect this
- Implementation of the Gong 03 sea-salt aerosol emission scheme to replace Monahan 1986
- Updates of dust source function

Chapter 2.1 - Tropspheric Chemistry

- Updates to photolysis quantum yields and cross sections for selected trace gases in troposphere

Chapter 2.3 - Stratospheric Chemistry

- Tuning of ozone depletion modelling in the stratosphere

Chapter 3 - Surface fluxes of atmospheric composition

- New version of anthropogenic emissions
- Application of local reduction of boreal NO GFAS emissions and tropical NH_4 GFAS peat emissions
- dry deposition scheme to use IFS stomatal resistance
- Activation of Lopez-lightning emission parameterisation for Lightning Nox emissions
- Wetlands CH_4 emissions
- Revision of Farquhar photosynthesis and ecosystem respiration reference values per PFT
- New reference NEE climatology consistent with 49r1 climate fields

Chapter 5 - Aerosol optical diagnostics

- Aerosol optical properties to account for hydrophilic and aspherical dust

Chapter 6 - data assimilation of atmospheric composition retrievals

- Assimilation of IASI SO_2 retrievals with altitude information
- Updated background errors for CH_4 based on EDA
- Passive monitoring of OMPS limb profiler data
- Passive monitoring of GEMS O_3 and NO_2 data

Chapter 1

Introduction

Table of contents

The simulation of greenhouse gases (GHG) and reactive trace gases and aerosol (COMPO) is a configurable extension of the IFS. These configurations are used in the operational forecast and re-analysis applications of the IFS for the Copernicus Atmosphere Monitoring Service (CAMS). This part of the documentation describes the simulation of the sink and source processes of atmospheric composition with the IFS as well as the specifics of the data assimilation approach for atmospheric composition.

The other parts of IFS documentation also cover tracer related aspects such as transport, radiation and data assimilation. In this part (8), we will refer to the respective parts. The transport by advection is described in Part III Dynamics and Numerical Procedures. Only aspects specific to tracer advection such as the application of global mass fixers are covered in this part (8). The transport of tracers by turbulent diffusion and convection and the use of prognostic aerosols and ozone in the IFS radiation scheme are documented in Part IV Physical processes.

Atmospheric composition is simulated "on-line" in the IFS, i.e. together with the simulation of the dynamical and physical processes using the same data flow, methods and routines as applied for the meteorological tracers such as cloud water and humidity. This means that each forecast or analysis of the IFS with atmospheric composition also provides a meteorological forecast or analysis. The on-line approach enables a close coupling between meteorological and atmospheric composition processes. It also makes it possible to use the data assimilation capability of the IFS to assimilate atmospheric composition observations to correct the simulated concentrations.

1.1 BACKGROUND

Monitoring and forecasting of global atmospheric composition are key objectives of the Copernicus Atmosphere Monitoring Service (CAMS), which is implemented by ECMWF on behalf of the European Commission (Peuch *et al.*, 2022). Starting in the early 2000s, the IFS has been extended for the simulation of reactive trace gases (Flemming *et al.*, 2015; Huijnen *et al.*, 2016), aerosols (Morcrette *et al.*, 2009; Rémy *et al.*, 2019, 2022, 2024) and greenhouse gases (Agustí-Panareda *et al.*, 2014). The on-line approach for the reactive trace gases superseded a two-way coupled system consisting of the IFS and Chemical Transport Models (Flemming *et al.*, 2009). First attempts to assimilate stratospheric ozone at ECMWF started in the late 1990, using a linear parameterization to describe stratospheric ozone chemistry (Part IV Physical processes, Chapter 9). The simulation and assimilation of stratospheric ozone in that way is part of the Numerical Weather Prediction (NWP) applications of the IFS. Up to CY47R3, the COMPO configuration of the IFS applied only the linear parameterization to describe stratospheric ozone. CY48R1 was the first cycle that uses an explicit chemical mechanism to simulate chemistry in the stratosphere.

With the capacity to simulate atmospheric composition in the IFS in place, the 4D-VAR data assimilation method of the IFS could be extended to allow the assimilation of satellite retrieved aerosol optical depth (AOD) (Benedetti *et al.*, 2009), reactive trace gases (Inness *et al.*, 2015) and longer-lived greenhouse gases (Engelen *et al.*, 2009; Massart *et al.*, 2014).

The IFS-COMPO configuration has been used to provide global forecasts of atmospheric composition since 2007 (Flemming *et al.*, 2017b) and in an operational mode since 2014. A further application of IFS-COMPO and IFS-GHG is the production of multi-decadal re-analysis data sets of aerosol, reactive trace gases (Inness *et al.*, 2013; Flemming *et al.*, 2017a; Inness *et al.*, 2019) and gases (Agusti-Panareda *et al.*, 2023). The re-analyses are produced by assimilating a wide range of satellite retrieved products of atmospheric composition starting in 2003. The latest version of the reanalysis (EAC4) is being continued to the present date. MAC4 is produced with IFS-COMPO of CY42R1.

The development of the atmospheric composition aspects in the IFS is a joint effort between ECWMF and other European research institutions, including several meteorological services of ECMWF member states. The joint development approach is pursued by implementing a range of chemistry schemes (Huijnen *et al.*, 2019) and other alternative approaches for the description of aerosol processes. Following the approach of the IFS documentation for other components, only the schemes applied in the operational CAMS forecasting system will be documented here. There are two operational configurations of the CAMS forecasts (o-suites):

(i) reactive gases and aerosols using the COMPO tropospheric and stratospheric chemistry schemes and the AER aerosol scheme (IFS-COMPO o-suite, resolution: TL511).
(ii) greenhouse gases CO_2 and CH_4 (GHG) only at increased resolution (IFS-GHG o-suite, resolution: fc TCO1279, analysis TCO399).

1.2 STRUCTURE OF THE DOCUMENTATION

Part 8 (Atmospheric Composition) documents the scientific basis and implementation of the following processes and diagnostics in the IFS in the following chapters:

- Chapter 1: Basic principles of atmospheric composition simulation with the IFS (this section).
- Chapter 2: Chemical and aerosol conversion in the atmosphere.
- Chapter 3: Surface fluxes of atmospheric composition including emissions and removal of atmospheric trace gases and aerosols by deposition.
- Chapter 4: Transport of trace gases and aerosols.
- Chapter 5: Calculation of aerosol radiative variables and other diagnostics.
- Chapter 6: Data assimilation of atmospheric composition retrievals.

For processes and methods that are commonly applied in the IFS for NWP, COMPO and GHG applications, such as transport of tracers or the data assimilation approach, this document will describe the specifics for the application for COMPO/GHG and will refer for the common aspects to the respective parts of IFS documentation.

1.3 FUNDAMENTALS

1.3.1 Continuity equation for tracers

Running the IFS-COMPO suite adds 123 tracers for reactive gases and 16 tracers for various aerosol species to the simulated 3D fields of any NWP application. The IFS-GHG suite only requires 2 additional tracers. The change of the mass mixing ratio (MMR) $C_A[kg/kg]$ of a tracer A is described by the 3D continuity equation:

$$\frac{\partial C_A}{\partial t} + \vec{v}\nabla C_i - \frac{1}{\rho}\frac{\partial}{\partial z}K_z\frac{\partial C_A}{\partial z} + \frac{1}{\rho}\frac{\partial}{\partial z}M(C_A{}^{cu} - C_A) = \frac{1}{\rho}(R_A + E_A - D_A) \tag{1.1}$$

The change of C_A is caused by transport processes and source and sink processes. The source and sink processes of tracer A are the emissions and surface fluxes (E_A), the loss by wet and dry deposition (D_A) and its conversion by chemical and aerosol microphysical processes (R_A). The emissions are either external data sets or are simulated on-line. The loss by wet or dry deposition D_A depends on the environmental conditions and the surface properties and the MMR of the tracer itself. The chemical or microphysical conversion term can either be a source or sink for A. The simulation of this term is a major contribution to the complexity of the atmospheric composition simulation because the chemical conversion of a gas or aerosol component A depends on the concentration of other tracers $B, C, D, ...$, which introduces a coupling between the simulation of the different tracers.

The transport of C_A is described by the advection with the grid-resolved 3D wind vector $\vec{v}$, by turbulent diffusion and by convection. In the IFS only the vertical component of turbulent diffusion is considered, which is calculated from the vertical turbulent exchange coefficient K_z and the vertical tracer gradient. The vertical motion because of convection is calculated by the IFS convection scheme using the convective mass flux M and the tracer concentration ${C_A}^{cu}$ inside the convection.

1.3.2 Simulation of chemical conversion

This section covers the basic principles of the chemical kinetics as an introduction for readers with a non-chemistry background.

For a chemical (bi-molecular) reaction of species A with species B to produce species C and D:

$$A + B \rightarrow C + D \tag{1.2}$$

the resulting chemical reaction rate, i.e. the change in time of the volume mixing ratios (VMR) of the reacting species (X_A, X_B) and of the product species (X_C, X_D), is given by the product of the volume mixing ratio of the reacting species (X_A, X_B) and the reaction rate constant k_{AB}:

$$-\frac{dX_A}{dt} = -\frac{dX_B}{dt} = \frac{dX_C}{dt} = \frac{dX_D}{dt} = k_{AB} \times X_A \times X_B \tag{1.3}$$

The reaction rate constant k_{AB} depends on the molecular properties of the reactants and on temperature and pressure. Reaction rate constants are determined by laboratory measurements and theoretical considerations and have a varying degree of uncertainty. The VMR of a species A (X_A) is given by the ratio of the number of molecules of A ([A]) and the number of molecules of air per volume unit.

Another case are mono-molecular conversions $A \rightarrow B$, which take place without the interaction with other reactants. The first-order rate of change depends in this case only on X_A :

$$\frac{X_A}{dt} = -k_A \times X_A \tag{1.4}$$

which can be solved analytically, using the initial volume mixing ratio of A:

$$X_A(t) = X_A(t=0)e^{-k_A t} \tag{1.5}$$

Examples of the first-order reaction process are the radioactive decay of radon, aerosol ageing, uni-molecular decay and photolysis. The inverse of the constant k_A can be considered as the e-folding lifetime of species A.

Photolysis is the breakup of molecules by solar radiation into often unstable products. Photolysis is an important initiator for many chemical reaction chains and cycles in the atmosphere. The photolytic rate constant kj depends on the characteristic properties of the species (quantum yield, absorption cross section) and the actinic radiative flux in the atmosphere.

A chemical mechanism, or chemical scheme, is a set of chemical species and set of chemical reactions between the species or their photolytic decay. It is defined by the (i) list of species and (ii) the specific reaction rate constants or photolytic rate constants and their dependency on temperature and radiation.

1.3.3 Units and measures of concentrations

The tracer concentrations are simulated as mass mixing ratios (MMR) in the IFS, which is the same approach as used for specific humidity. The MMR of tracer A (C_A) is given by the mass (m_A) of A with respect to the total mass of air (m_{air}) in each grid cell:

$$C_A = \frac{m_A}{m_{air}} \tag{1.6}$$

Because the mass of air in each grid cell is proportional to the vertical extent of the grid cell in pressure units Δp, the mass of tracer A (m_A) can be calculated using the gravitational acceleration g and the grid box area S in the following way:

$$m_A = C_A \frac{\Delta p}{g} S \tag{1.7}$$

The volume mixing ratio (X, unit= $\frac{mol}{mol}$) of A is the ratio of the number molecules of A ($[A]$) and the number molecules of air ($[air]$). It can be calculated for gases from the MMR (C_A) using the molar mass of A (M_A) and of air (M_{air}) in the following way:

$$X_A = \frac{[A]}{[air]} = C_A \frac{M_{air}}{M_A} \tag{1.8}$$

In most cases taking into account only the molecular weight of dry air ($M_{air-dry} = 0.029 \frac{kg}{mol}$) is sufficient but more accurate conversions need to consider the impact of the specific humidity q on M_{air}.

The mass concentration of A (c_A , unit = $\frac{kg}{m^3}$) is defined as the tracer mass m_A per volume. It can be calculated from C_A by the multiplication with the density of air (ρ_{air})

$$c_A = C_A \rho_{air} \tag{1.9}$$

ρ_{air} depends on the pressure, temperature and humidity (Part IV Physical processes).

The scaling of X_A and c_A to commonly used units such as ppb (parts per billion) and $\frac{\mu g}{m^3}$, respectively, is achieved by the multiplication with 10^9.

1.4 OVERVIEW OF THE CODE

The simulation of atmospheric composition tracers follows the treatment of other grid-point variables such as humidity. The sink and source processes and the transport by vertical turbulent diffusion and convection are included in the physics parameterization package (CALLPAR, Part IV Physical processes, Section 1.2). The simulation of prognostic variables, i.e. also the prognostic tracer fields, in the physics only allows the consideration of dependencies in the vertical dimension.

The input to the physics (CALLPAR) are the MMRs of the prognostic tracers (grid-box mean) and the tendencies from the advection calculated by the dynamics. The tendencies from the different sink and source processes and the turbulent and convective transport are computed in separate routines called from CALLPAR.

The tendencies from the different processes and the advection are added together. The MMRs entering these routines are intermittently updated with the accumulated tendency from the previous process ('method of fractional steps'). This approach to operator splitting introduces a dependency on the order of the different source and sink routines in CALLPAR.

At the end of CALLPAR, the accumulated tendencies of the physics and the tendencies from the advection are added and consequently applied to the initial MMRs entering the physics to calculate the MMRs of the next time step. At that point, check of the total tendencies to avoid MMRs below a "chemical zero" of 1.0×10^{-25} [kg/kg] is carried out.

The sink and source processes for reactive gases and aerosol are simulated in separate routines because they have been developed independently. The further harmonisation of the aerosol, chemistry and GHG routines is an ongoing development effort. For GHG the emissions and the turbulent and convective transport are simulated in the physics. The only sink process for GHG is the application of a climatology of CH_4 loss rates.

In CALLPAR, the following high-level atmospheric composition routines are called in the following order:

COMPO_ APPLY_EMISSIONS_LAYER
: Computes surfaces fluxes from sector specific input data and applies diurnal cycles and application of vertical distributions (see section 3.1.1)

CHEMINI_LAYER
: Initialisation of chemical scheme

AERINI_LAYER
: Calculation of online emissions for sea salt and dust, and aerosol dry deposition and sedimentation velocities

CONVECTION_LAYER or CONVECTION_S_LAYER
: Convection as part of physics package includes convective transport of tracers

TURBULENCE_LAYER or TURBULENCE_S_LAYER
: Turbulence as part of physics package, includes vertical diffusion of tracers with injection of surface fluxes and calculation of dry deposition velocities using input from the land surface scheme

LIGHTNING_LAYER
: Lightning parameterization part of physics package, includes calculation of NO_x lightning emissions

CHEM_MAIN_LAYER
: Calculation of chemical conversions and wet deposition of reactive gases, applies methane loss rates (GHG) and NO_x lightning emissions

AER_PHY3_LAYER
: Calculation of secondary aerosol formation, wet deposition of aerosols and aerosol optical properties and other diagnostics (AOD, PM, etc.)

The MMRs and respective tendencies of chemical species (from CHEM_MAIN_LAYER) required for the calculation of secondary aerosol formation in AER_PHY3_LAYER are stored in the array ZCHEM2AER.

1.5 CONFIGURATION OF ATMOSPHERIC COMPOSITION IN THE IFS

The simulation of atmospheric composition with the IFS allows many different options for the choice of the simulated greenhouse gases, reactive gases and aerosol tracers and their properties. There is the choice of one of several chemistry schemes of varying complexity (MOCAGE, MOZART, RnPb, linear CO) and the number of the considered aerosol tracers.

GHG, aerosols and reactive gases can be simulated together or separately. If aerosols are run without an explicit chemistry scheme, there are options available to account for a simplified treatment of secondary aerosol formation (currently only sulphates) without the need to simulate the gas-phase chemistry.

This document only describes the configuration used by CAMS for (i) the IFS-COMPO operational system, which applies the COMPO tropospheric scheme (based on the CB05 scheme), the COMPO stratospheric chemistry scheme (based on the BASCOE scheme) and the AER aerosol scheme, or (ii) the IFS-GHG configuration (only CO_2 and CH_4).

The IFS is configured by means of namelists. These fortran namelists are compiled by the run time environment (ifs-scripts) based on variables set in prepIFS. Important prepIFS variables are LMACC=true to trigger the treatment of atmospheric composition, LCHEM=true to trigger the calculation of chemistry, LAER=true to trigger the simulation of aerosols and LGHG=true to trigger the simulation of greenhouse gases.

The following IFS namelists are used to configure the atmospheric composition simulation in the IFS:

NAMGFL	Specification of all GFL fields; lists each simulated tracer field and specifies the advection and mass fixing options. Additionally it specifies if chemistry tracers are subject to dry and wet deposition
NAMCOMPO	Configuration options valid for all tracers, and for secondary aerosol formation
NAMCHEM	Configuration of the chemistry simulation
NAEAER	Configuration of the aerosol simulation and properties of every aerosol tracer
NAMCOMPO_EMIS	Configuration of the sectoral emission input

Chapter 2

Conversion of aerosols and trace gases in the atmosphere

Table of contents

The chemical conversion of atmospheric trace gases occurs by the cumulative influence of photolysis, bi-molecular reactions between gaseous components, aqueous-phase reactions in clouds and heterogeneous reactions of gases on the surface of cloud and aerosol particles. Beside the gas-phase chemical conversion, condensable trace gases undergo homogeneous nucleation or heterogeneous condensation on existing aerosol particles to form secondary aerosols, with specific processes that form organic and inorganic aerosol types.

The resulting rate of change of the mass mixing ratios of trace gases and aerosols is simulated by one of several chemical mechanism of the IFS. The various chemical mechanisms differ in the selection of the considered chemical species, photolytic, homogeneous and heterogeneous reactions and in the calculation of the respective rates of each of these processes based on the environmental conditions such as temperature, radiation and presence of aerosols and clouds. The numerical formulation of a chemical

mechanism is a large and often numerically stiff system of ordinary differential equations, which require the application of dedicated mathematical solvers.

Atmospheric aerosol consists of a multitude of components with variable sizes and shapes. Major natural aerosol components are sea salt and dust particles. Further important aerosol types from a variety of natural, biogenic and anthropogenic sources are primary and secondary organic and inorganic aerosols as well as black carbon.

Some of these aerosol particles undergo ageing, while secondary aerosol particles can change due to chemical conversion and volatility.

Trace gases and aerosol can also dissolve in cloud and rain drops depending on their degree of solubility, where aqueous-phase reactions lead to further chemical processing, contributing to the fractional partitioning between the gas and aerosol phase. A dedicated module for aqueous (wet) phase chemistry and gas-aerosol partitioning simulates these processes.

The atmospheric chemistry module used in IFS-COMPO consists of a tropospheric and stratospheric chemistry scheme, supplemented with a few ancillary tracers for specific processes. In total this results in 123 tracers active in the chemistry module. The switch between the application of the tropospheric and stratospheric chemistry schemes is applied at the tropopause, the height of which is diagnosed from the temperature lapse rate in the IFS. The atmospheric aerosol module (AER) in IFS-COMPO consists of 16 species, which are in various ways coupled to the atmospheric chemistry.

2.1 TROPOSPHERIC CHEMISTRY

The tropospheric chemistry used in IFS-COMPO has evolved and expanded over time. It originates from the Carbon Bond mechanism version 5 (CB05) (Yarwood *et al.*, 2005) as formulated for regional air-quality studies, but has been tailored and modified for use in the Chemistry Transport Model TM5 (Williams *et al.*, 2013; Huijnen *et al.*, 2010). The scheme was first implemented in IFS as documented in Flemming *et al.* (2015). The CB05 mechanism designed by Yarwood *et al.* (2005) is characterized by a lumping approach for organic species, i.e. by defining a separate tracer species for specific types of functional groups. The original CB05 scheme implemented in the IFS has gradually been expanded particularly to include more explicit organic chemistry. The complete list of trace gases in the model is given in Tables A.1-A.4. This encompasses trace gases and reactions that describe basic inorganic chemistry and simple organic chemistry responsible for carbon monoxide and methane oxidation (Table A.1), as well as reactions that describe volatile organic compounds (VOCs, Table A.2).

2.1.1 Inorganic chemistry

Basic inorganic chemistry is modelled by the trace gases specified in Table A.1. For application in IFS-COMPO the inorganic chemistry has been extended with NO_y species including $CH_3O_2NO_2$ (Williams *et al.*, 2017; Huijnen *et al.*, 2019). The IFS-COMPO chemistry includes a sulphur cycle, represented by sulphur dioxide (SO_2), di-methyl sulphide (DMS) and methyl sulphonic acid (MSA), as well as ammonia (NH_3) chemistry (Huijnen *et al.*, 2010). This is coupled to inorganic aerosol formation. Hydrogen cyanide (HCN) and acetonitrile (CH_3CN) are included as long-lived tracers indicative of biomass burning activity.

2.1.2 Chemistry of degradation of organic compounds

Basic reactions describing the oxidation of CO and CH_4 are modelled using trace gases as defined in Table A.1. Oxidation up to C3 organic species is treated explicitly, following chemistry mechanism extensions developed in the context of TM5 and IFS-COMPO, as described in Williams *et al.* (2013, 2017); Huijnen *et al.* (2019). The trace gases involved here are given in Table A.2. This includes lumped tracers for specific types of functional groups which are used to model the oxidation of higher volatile organic compounds, such as butane (C4), pentanes (C5), and (higher) aldehydes.

The isoprene oxidation scheme uses explicit approach based on Stavrakou *et al.* (2010), which has been further modified according to Lamarque *et al.* (2012) and Myriokefalitakis *et al.* (2020). Reaction products

from this isoprene oxidation scheme include glyoxal (CHOCHO), glycolaldehyde, isoprene-peroxide, hydroxy-acetone and two hydroxy-aldehydes, as documented in Williams *et al.* (2022).

Basic aromatic chemistry is described through the inclusion of lumped xylene and toluene tracers and their oxidation reaction chain. These reaction products are coupled with secondary organic aerosol (SOA) formation as described below. For this purpose, three gas-phase SOA precursor tracers, referred to as SOGs, are defined.

2.1.3 Mechanism version, reaction rates, and solver

The IFS-COMPO tropospheric chemistry version is referenced as *tc*06*k*. It uses a chemistry table file `tm5_ver19.x`, but run in conjunction with IFS-COMPO stratospheric chemistry the table file is `bascoetm5_ver6.x`. Including SOA precursor tracers, this mechanism consists of 71 trace gases that are active in the tropospheric chemistry, while it has 157 gas-phase reactions, 3 heterogeneous reactions and 2 aqueous phase reactions. The reaction rates follow the recommendations given in either Burkholder *et al.* (2020) or the latest recommendations by IUPAC, as hosted at `http://iupac.pole-ether.fr`.

The tropospheric chemistry is solved based on Kinetic PreProcessor (KPP) routines, using the four stages and third-order Rosenbrock solver (Sandu and Sander, 2006).

2.1.4 Photolysis

For parameterization of the photolysis rates the separate approaches for troposphere and stratosphere are retained, as these have been optimized in the past for these specific applications. The modified band approach (MBA) is adopted for the computation of tropospheric photolysis rates (Williams *et al.*, 2006, 2012). It computes net photolysis rates based on 7 absorption bands across the spectral range 202 - 695 nm. Cross section and quantum yield input data are taken from recent compilations (IUPAC and JPL), as collected in the photolysis input dataset 'tropo_look_up_cbmhybrid' version v1b. For instances of large solar zenith angles (71-85°) a different set of band intervals is used to avoid introducing large biases. The radiative transfer calculation is performed with a two-stream solver using the absorption and scattering components introduced by gases, aerosols and clouds, computed on-line for each of the predefined band intervals.

In the configuration with coupled tropospheric and stratospheric chemistry the photolysis rates for reactions occurring both in the troposphere and stratosphere are merged at the interface, in order to ensure a smooth transition between the tropospheric and stratospheric parameterizations. This is done by an interpolation at four model levels around the interface level between both parameterizations, for solar zenith angles (SZA) $< 85°$. For larger SZA the original value for the photolysis rate is retained in case of stratospheric chemistry, while it is switched off for the troposphere.

The impact of aerosol on the actinic fluxes required for the calculation of the photolysis rates is simulated using the AOD of the aerosol fields of the aerosol scheme AER in the IFS.

2.1.5 Heterogeneous chemistry

Heterogeneous reactions rates are calculated using the Surface Area Density (SAD) which is provided by AER aerosol model to account for scattering and available surface area for heterogeneous conversion. For N_2O_5, HO_2, NO_3, H_2O_2 and HONO the reactions on aerosols are represented as a pseudo-first-order process with reaction rate k_{het} according to:

$$k_{het} = \left(\frac{r}{D_g} + \frac{4}{c\gamma} \right) S \tag{2.1}$$

With D_g being the gas-phase diffusion coefficient of the gaseous species, r the mean aerosol radius, c the mean molecular velocity of the gaseous species, γ the reaction probability of uptake coefficient and S the aerosol surface area density. The γ values used are specified in Table 2.1. Particularly for the uptake of HO_2 the parameterization proposed by Thornton *et al.* (2008) is used, assuming a fixed pH value of

Table 2.1 *γ values used in heterogeneous conversion rates on atmospheric cloud droplets, ice and aerosols for N_2O_5, HO_2 and NO_3. T08 refers to Thornton et al. (2008).*

Particle type	γ (N_2O_5)	γ (HO_2)	γ (NO_3)
cloud droplets	$2.7 \times 10^{-5} exp(1800/T)$	T08	-
ice particles	0.02	0.02	4×10^{-3}
desert dust	0.02 (RH>50%)	0.03 (RH>50%)	0.012
sea salt	0.02 (RH>50%)	0.02 (RH>50%)	0.05
organic matter (hydrophilic)	0.02	4×10^{-4} for RH>50%	10^{-4}
secondary organic aerosol	0.02	4×10^{-4} for RH>50%	10^{-4}
black carbon (hydrophilic)	0.01 for RH>50%	10^{-3} for RH>50%	10^{-4}
sulfate	0.04	0.05	16×10^{-3}
ammonium	0.002	0.05	16×10^{-3}
nitrate	0.002	5×10^{-3}	5×10^{-3}

Table 2.2 *same as table 2.1, but for H_2O_2 and HONO*

Particle type	γ (H_2O_2)	γ (HONO)
ice particles	0.02	10^{-4}
desert dust	6×10^{-4}	1.6×10^{-4}
sea salt	2×10^{-3}	2×10^{-3}
organic matter (hydrophilic)	10^{-4}	10^{-4}
secondary organic aerosol	10^{-4}	10^{-4}
black carbon (hydrophilic)	10^{-4}	10^{-4}
sulfate	10^{-3}	2×10^{-3}
ammonium	10^{-3}	2×10^{-3}
nitrate	10^{-3}	2×10^{-3}

5.5. This results in γ-values ranging between 10^{-4} and 0.1 depending on temperature and ambient HO_2 concentrations.

2.2 ATMOSPHERIC AEROSOL

The aerosol module in the IFS is referred to in literature as AER (Rémy *et al.*, 2022, 2024), and is originally derived from the LOA/LMDZ model (Boucher *et al.*, 2002; Reddy *et al.*, 2005). It is essentially a bulk aerosol scheme, while for sea salt aerosol and desert dust, a sectional approach is employed. As such, it is often denoted as a bulk-bin scheme. The aerosol species use mass mixing ratios as the prognostic variable of the aerosol tracers. The prognostic species are sea salt, desert dust, organic matter (OM), black carbon (BC), sulfate, nitrate, ammonium and secondary organic aerosols (SOA). The SOA species has been introduced in CY48R1 with two tracers (biogenic and anthropogenic); before that, SOA was represented as part of the OM species. This means that up to cycle 47R3, the OM species included both primary and secondary aerosols, while in CY48R1 and beyond, it consists only of the primary aerosols.

AER is run coupled to the COMPO tropospheric chemistry scheme in the IFS-COMPO in the operational configuration. As described in Rémy *et al.* (2019), AER can also be run in a stand-alone mode (IFS-AER), i.e. with a simplified description of tropospheric chemistry describing only the production of sulfate and SOA. In this configuration the nitrate and ammonium species are not included, and a specific tracer representing sulfur dioxide is subsequently added. Furthermore in this configuration two additional tracers are included, to represent secondary organic anthropogenic and biogenic precursors gases.

The assumed number size distribution specifics are shown in Table 2.3. These are used offline in the Mie code to generate the optical properties using an external mixing assumption: wavelength dependent mass extinction, asymmetry parameter, single scattering albedo and lidar ratio for each aerosol species. In the processes that depend on aerosol size, such as dry deposition and for heterogeneous reactions, the mass median diameter (MMD) is used, as computed using the assumed size distribution given in

Table 2.3 *Aerosol species in AER*

Aerosol type	Size bin limits (sphere radius, ¯m)	ρ (kg m^{-3})
Sea Salt (80% RH)	0.03-0.5 0.5-5.0 5.0-20	1183
Dust	0.03-0.55 0.55-0.9 0.9-20	2610
Black carbon hydrophobic	0.005-0.5	1000 0
Black carbon hydrophilic	0.005-0.5	1000 0
Organic matter hydrophobic	0.005-20	1300
Organic matter hydrophilic	0.005-20	1300
Sulfates	0.005-20	1760
Nitrate fine	0.03-0.9	1730
Nitrate coarse	0.9-20	1400
Ammonium	0.005-20	1760
biogenic SOA	0.005-20	1800
anthropogenic SOA	0.005-20	1800

Table 2.3. The assumed number size distribution of dust uses values provided by Ryder *et al.* (2018) from aircraft measurements over the tropical Eastern Atlantic.

Desert dust is represented with three size bins, with radius bin limits at 0.03, 0.55, 0.9 and 20 ¯m. Dust is a slightly hydrophilic species. Sea salt aerosol is also represented with three size bins, with radius bin limits of 0.03, 0.5, 5 and 20 ¯m at 80% relative humidity. All of the sea salt aerosol parameters (concentration, emission, deposition) are expressed at 80% relative humidity; this is in contrast to the other aerosol species in IFS-AER, which are expressed as dry mixing ratio. Users who wish to use dry sea salt aerosol mass mixing ratio, or emissions, burden and sink diagnostics of dry sea-salt, need to be aware that the archived values correspond to sea-salt aerosol at 80% relative humidity, and need to be divided by a factor of 4.3 to convert to dry sea-salt aerosol quantities in order to account for the hygroscopic growth and change in particle density. There is no mass transfer between bins for either dust or sea salt.

The organic matter and black carbon species consist of their hydrophilic and hydrophobic fractions, with the ageing processes transferring mass from the hydrophobic to hydrophilic components. Sulfate aerosols are represented by one prognostic variable. The nitrate species consists of two prognostic variables that represent fine nitrate produced by gas-particle partitioning and coarse nitrate produced by heterogeneous reactions of dust and sea salt particles. Finally, the secondary organics species consists of two tracers, that represent biogenic and anthropogenic SOA. In all, the AER consist of 16 prognostic variables when running coupled with the COMPO tropospheric chemistry scheme (default). When AER is run without coupling to a full chemistry scheme, a simplified chemistry is activated to describe SO_4 and SOA production, while the simulation of nitrate and ammonium aerosol is de-activated.

2.2.1 Hygroscopic growth and ageing

Hygroscopic growth is the process whereby, for some aerosol species, water is mixed in the aerosol particle, increasing its mass and size and decreasing its density. This process is treated implicitly in IFS-AER, since size is not resolved. It plays an important role, however, in the computation of optical properties and also for sinks that are size and/or density dependent, in particular dry deposition. The species subjected to hygroscopic growth in IFS-AER are sea salt, the hydrophilic components of OM and BC, sulfate, nitrate, ammonium and SOA. Desert dust is also subjected to hygroscopic growth following (Chen *et al.*, 2020). The amount of water that is mixed in the aerosol particle depends on the particle size. Table 2.4 details the changes in size for the concerned species. The values are drawn from Tang and Munkelwitz (1994) for sea salt, Tang *et al.* (1997) for sulfate and ammonium, Chin *et al.* (2002) for BC, and Svenningsson *et al.* (2006) for nitrate.

For OM and BC, once emitted, the hydrophobic component is transformed into a hydrophilic one with an exponential lifetime of 2.78 hours.

Table 2.4 *Hygroscopic growth factor depending on ambient relative humidity.*

RH/%	Desert-dust	Sea-salt	OM	BC	Sulfate and Ammo-nium	Nitrate fine	Nitrate coarse	SOA
0–10	1	1	1	1	1	1	1	1
10–20	1.008	1	1	1	1	1	1.1	1
20–30	1.016	1	1	1	1	1	1.2	1
30–40	1.024	1	1	1	1	1	1.3	1
40–50	1.032	1.442	1.	1	1.169	1.1	1.35	1.0
50–60	1.038	1.555	1.05	1	1.220	1.2	1.4	1.2
60–70	1.046	1.666	1.1	1	1.282	1.25	1.5	1.3
70–80	1.054	1.799	1.15	1	1.363	1.3	1.6	1.4
80–85	1.062	1.988	1.2	1.2	1.485	1.35	1.7	1.5
85–90	1.066	2.131	1.25	1.3	1.581	1.5	1.8	1.6
90–95	1.070	2.361	1.3	1.4	1.732	1.7	2.0	1.7
95–100	1.074	2.876	1.4	1.5	2.085	2.1	2.2	1.8

2.2.2 Secondary organic aerosol formation

SOA precursors are included, with the aromatic gas-phase precursors (xylenes, toluene) acting as exclusive anthropogenic and biomass burning sources, using the approach of Karl *et al.* (2009), see also Figure 2.1. In a similar way, the production of a biogenic SOA tracer has been linked to the oxidation of both isoprene and terpenes. For this, we adopted three Secondary Organic aerosol precursor Gas (SOG) classes: two associated to anthropogenic emissions and one associated to biogenic emissions. These precursor gases are linked to the respective anthropogenic and biogenic SOA tracers (two anthropogenic precursor gas tracers contribute to anthropogenic SOA production).

(*a*) *SOA formation from biogenic sources*

In IFS, the modelling of SOA formation from isoprene and terpene oxidation is modeled using a single SOA and SOG family, following the approach proposed by Tsigaridis and Kanakidou (2003). This implies that a single, dominant, low volatility SOG tracer is assumed. Products of VOC oxidation partition between purely gas-phase chemistry and SOG, and its efficiency is described by production yields α and ϵ. The SOG production follows the reactions below:

$$\mathrm{ISOP} + \mathrm{OH} \longrightarrow \alpha_1 \mathrm{SOG}_1 + (1-\alpha_1)(0.65\mathrm{ISOPBO2} + 0.35\mathrm{ISOPDO2}) \tag{2.2}$$

$$\begin{aligned}\mathrm{ISOP} + \mathrm{O_3} \longrightarrow\ & \alpha_2 \mathrm{SOG}_1 \\ & + (1-\alpha_2)(0.35\mathrm{PAR} + 0.65\mathrm{ISPD} + 0.2\mathrm{XO_2} + 0.2\mathrm{C2O3}) \\ & + 0.6\mathrm{HCHO} + 0.066\mathrm{HO_2}\end{aligned} \tag{2.3}$$

$$\begin{aligned}\mathrm{TERP} + \mathrm{OH} \longrightarrow\ & \epsilon_1 \mathrm{SOG}_1 + (1-\epsilon_1)(5\mathrm{PAR}) \\ & + 1.25\mathrm{XO_2} + 0.25\mathrm{XO_2N} + 1.22\mathrm{HCHO} \\ & + 0.47\mathrm{ALD2} + 0.47\mathrm{CO} + 1.22\mathrm{HO_2}\end{aligned} \tag{2.4}$$

$$\begin{aligned}\mathrm{TERP} + \mathrm{O_3} \longrightarrow\ & \epsilon_2 \mathrm{SOG}_1 + (1-\epsilon_2)(6\mathrm{PAR}) \\ & + 0.76\mathrm{XO_2} + 0.18\mathrm{XO_2N} + 1.8\mathrm{HCHO} + 0.21\mathrm{ALD2} \\ & + 0.21\mathrm{CO} + 0.28\mathrm{HO_2} + 0.57\mathrm{OH} \\ & + 0.38\mathrm{C2O3} + 0.39\mathrm{CH_3O_2}\end{aligned} \tag{2.5}$$

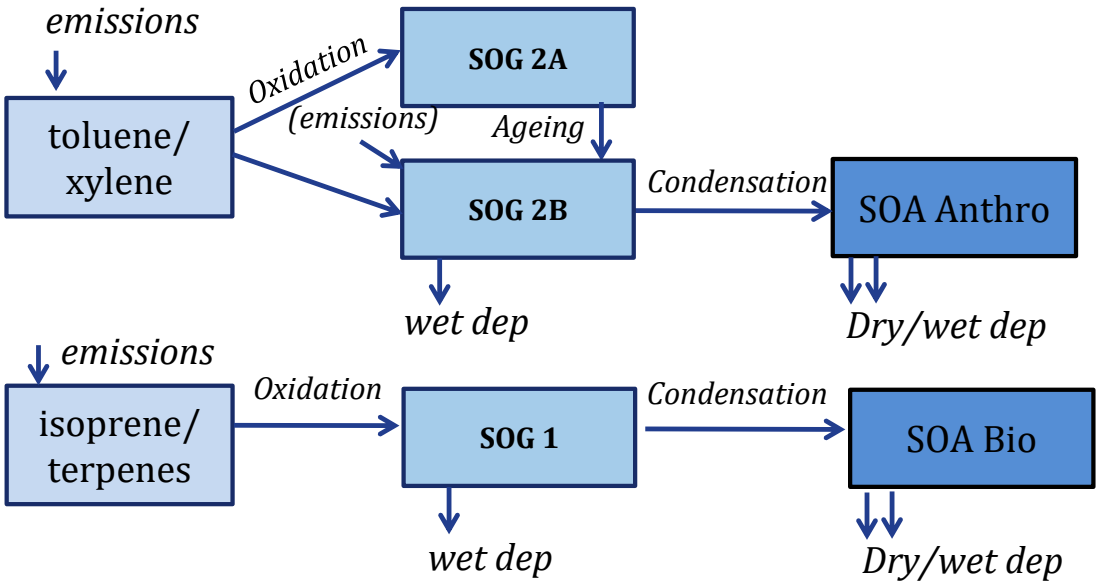


Figure 2.1 *Schematic of the secondary organic aerosol module*

In this formulation we aim at carbon mass conservation across chemistry and aerosol formation. The standard gas-phase reaction rates are used for these individual reactions of isoprene and terpenes with OH and O_3. Here we neglect night-time aerosol precursor formation from reaction with NO_3, considering that this is a minor channel compared to the other parameterizations (Tsigaridis and Kanakidou, 2003), and considering the high level of simplification already adopted by assuming only a single precursor gas (SOG1). Mass yields (α_1, α_2, ϵ_1, and ϵ_2) are specified in Table 2.5.

(*b*) *SOA formation from anthropogenic sources*

To account for SOA formation from anthropogenic sources, a simplified description of aromatic chemistry has been introduced in the COMPO tropospheric chemistry following Karl *et al.* (2009), through transport and oxidation of toluene and lumped xylenes as described by the reactions below, with mass yields specified in Table 2.5.

$$\mathrm{TOL} + \mathrm{OH} \longrightarrow \beta_1 \mathrm{SOG2A} + \beta_2 \mathrm{SOG2B} + (1 - \beta_1 - \beta_2)(5\mathrm{PAR} + \mathrm{AROO2}) \tag{2.6}$$

$$\mathrm{TOL} + \mathrm{O_3} \longrightarrow \phi_1 \mathrm{SOG2A} + \phi_2 \mathrm{SOG2B} + (1 - \phi_1 - \phi_2)(5\mathrm{PAR} + \mathrm{AROO2}) \tag{2.7}$$

$$\mathrm{XYL} + \mathrm{OH} \longrightarrow \gamma_1 \mathrm{SOG2A} + \gamma_2 \mathrm{SOG2B} + (1 - \gamma_1 - \gamma_2)(5\mathrm{PAR} + \mathrm{AROO2}) \tag{2.8}$$

$$\mathrm{XYL} + \mathrm{O_3} \longrightarrow \chi_1 \mathrm{SOG2A} + \chi_2 \mathrm{SOG2B} + (1 - \chi_1 - \chi_2)(5\mathrm{PAR} + \mathrm{AROO2}) \tag{2.9}$$

SOG formation from PAR, OLE and C_3H_6 has not been included here, although combined they may contribute significantly (up to 50%) to the SOA production budget, e.g., Tsimpidi *et al.* (2014). Instead, this is parameterized by adding an extra direct emission source to the SOG2B tracer based on the NMVOC emissions. Note that for anthropogenic SOG two classes have been defined: a high-volatility class (SOG2A) and a low-volatility class (SOG2B). Considering that only a single corresponding aerosol type is assumed, all mass in the high-volatility class is assumed to be in the gas-phase. An ageing reaction, combined with a 15% mass accumulation, is introduced for anthropogenic aerosol, where the high-volatility class is converted to the low-volatility class (Tsimpidi *et al.*, 2014), a according to:

$$\mathrm{SOG2A} + \mathrm{OH} \longrightarrow 1.15 \times \mathrm{SOG2B},\ k = 9 \times 10^{-12} cm^3 molec^{-1} sec^{-1} \tag{2.10}$$

(*c*) *Aerosol mass yields*

SOG production yields from oxidation reactions are given in Table 2.5. In their specification we loosely follow suggestions for net aerosol mass production yields as reported in the literature (Chung and Seinfeld, 2002; Henze and Seinfeld, 2006; Tsimpidi *et al.*, 2016; Hodzic *et al.*, 2016). For the biogenic SOG yields, the low-NOx conditions are assumed best representative on a global scale, while for anthropogenic aromatics the high-NOx conditions are assumed representative. These yields implicitly account for the mass weight conversion from the precursor VOC gases to the SOG. In a future revision, a more explicit formulation including mass weight factors for precursor gases and SOG will be necessary.

Table 2.5 *Secondary aerosol mass yields for various volatility bins in the SOA scheme*

Biogenic/Anthropogenic	Reactions	Mass yields	SOG1/SOG2B C^*-Bins 1-10 [K=1m^3/µg] [log(c^*)=0]	SOG2A C^*-Bins 100-1000 [K=0.01m^3/µg] log(c^*)=2]
Biogenic	ISOP + OH	α_1	0.02	-
Biogenic	ISOP + O_3	α_2	0.02	-
Biogenic	TERP + OH	ϵ_1	0.07	-
Biogenic	TERP + O_3	ϵ_2	0.07	-
Anthropogenic	TOL + OH	β_1, β_2	0.3	0.6
Anthropogenic	TOL + O_3	ϕ_1, ϕ_2	0.3	0.3
Anthropogenic	XYL + OH	γ_1, γ_2	0.3	0.3
Anthropogenic	XYL + O_3	χ_1, χ_2	0.3	0.3

2.2.3 Secondary inorganic aerosol formation

(*a*) *Production of sulfate through gas-phase oxidation*

The oxidation of sulfur dioxide (SO_2) to sulfuric acid (H_2SO_4) in the gas-phase occurs by the reaction with OH as described in Huijnen *et al.* (2010). The reactions involved are:

$$SO_2 + OH + O_2 \longrightarrow SO_3 + HO_2 \tag{2.11}$$

$$SO_3 + H_2O \longrightarrow H_2SO_4 \tag{2.12}$$

In the COMPO tropospheric chemistry scheme, HSO_4 is not a prognostic species. Due to the comparatively high water vapor pressure, SO_3 is assumed to produce sulfate aerosol.

(*b*) *Production of sulfate through aqueous phase chemistry*

The main source term for the production of H_2SO_4 occurs within cloud droplets, predominately by the oxidation by hydrogen peroxide (H_2O_2) and ozone (O_3). Note that each step in this process is dependent on the cloud pH. The pH used is computed using the aerosol pH as provided by EQSAM4Clim (see below). The aqueous chemistry routine thus computes cloud and precipitation pH. The latter is also used in the representation of wet deposition (see Chapter 3). The reactions involved are:

$$SO_2 + H_2O_2 \longrightarrow H_2SO_4 \tag{2.13}$$

$$SO_2 + O_3 + H_2O \longrightarrow H_2SO_4 + O_2 \tag{2.14}$$

The uptake of SO_2 into droplets is expressed as an effective Henry's law coefficient taking into account the dissociation into bisulphite (HSO_3^-) and sulphite ($SO_3^=$), which have pKa values of 1.77 (K_{S1}) and 7.20 (K_{S2}). This increases the fraction dissolved and available for oxidation, according to:

$$H_{eff}(SO_2) = H(SO_2)\left(1 + \left(\frac{K_{S1}}{[H^+]} + \frac{K_{S1}K_{S2}}{[H^+]^2}\right)\right) \tag{2.15}$$

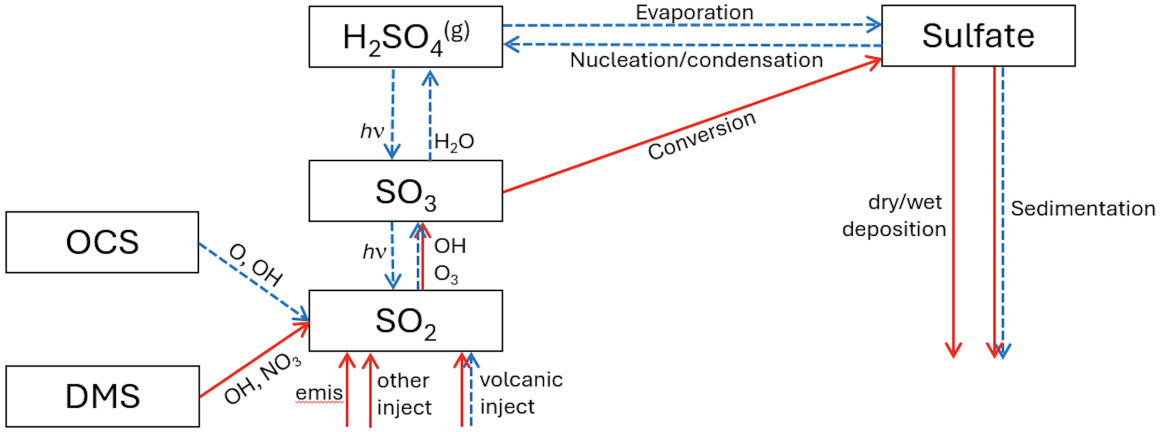


Figure 2.2 *Architecture of the sulfur cycle of IFS-COMPO. Processes may be active in the troposphere (red arrows), in the stratosphere (dashed blue arrows) or in the whole atmosphere (double arrows).*

The rate expressions for the aqueous phase SO_2 oxidation are specified in Table 2.6; they have been updated in cycle 50R1.

Concentrations in the gas phase are updated accordingly once the amount of oxidation has been calculated, which then is used for the aerosol representation.

Table 2.6 *Reaction rate terms used for the description of aqueous phase oxidation*

Aq. reaction	Rate expression
$SO_2 + H_2O_2$	$7.45 \times 10^7 exp\left(-4430(\frac{1}{T} - \frac{1}{298.})\right) \times \frac{H^+}{1+13\times H^+}$
S(IV) + O_3	$4.39 \times 10^{11} exp\left(\frac{-4131}{T}\right) + 2.56 \times 10^3 exp\left(\frac{-966}{T}\right) + \frac{2.56\times 10^3 exp\left(\frac{-966}{T}\right)}{H^+}$

(*c*) *Production of aqueous sulfate in the stratosphere through homogeneous nucleation and condensation*

In the stratosphere, nucleation of gaseous sulfuric acid into aqueous sulfate particles is the main source of particulate sulfate. A simple representation of stratospheric sulfate is implemented, as described in the schematic 2.2

In order to represent stratospheric sulfate with the existing single sulfate tracer there are two processes:

- nucleation/condensation of gaseous sulfuric acid into liquid sulfate particles, and
- evaporation of liquid sulfate particles into gaseous sulfuric acid

The production term, which represents the combined effect of homogoneous nucleation and condensation, is parameterized as a function of temperature and simulated concentration of gaseous sulfuric acid as provided by BASCOE. The production rate of aqueous sulfuric acid in s^{-1} is expressed as:

$$\tau = a[H_2SO_4]_g^b \tag{2.16}$$

Where the gaseous sulfuric acid concentration is in kg/m^3, a and b are two constants with values of 50 and 2 respectively. The production only occurs if the simulated partial pressure of gaseous sulfuric acid is above the saturation pressure. The saturation pressure is computed as a function of temperature, pressure and humidity following (Ayers *et al.*, 1980; Kulmala and Laaksonen, 1990). Where the partial pressure of gaseous sulfuric acid is below the saturation pressure, all liquid particles evaporate. The

production rate τ is then used to compute an updated mass mixing ratio of gaseous H_2SO_4 from nucleation/condensation:

$$[H_2SO_4]_{upd} = [H_2SO_4]exp(-\tau.\delta t) \tag{2.17}$$

Where δt is the model time step. The tendencies of gaseous and liquid sulfuric acid are derived from the updated mass mixing ratio of gaseous sulphuric acid.

The sedimentation of sulfate particle has been adapted to use sedimentation velocity computed from the Stokes formula rather than using fixed sedimentation velocities. Also, the assumed size distribution used to compute the sedimentation velocity has been made variable, depending on the simulated concentration of liquid phase sulfuric acid. This is meant to represent the fact that through coagulation, sulfuric acid particles are usually coarser in volcanic conditions (with higher concentration) and finer in quiescent conditions (with lower concentration). The assumed mass median diameter used to compute the sedimentation velocity is thus computed as:

$$D = D_0(mmr_{H_2SO_{4l}}/5.E8)^{0.25} \tag{2.18}$$

Where D_0 is the sulfate wet diameter, computed from the sulfate dry diameter of 0.6 micron (updated in cycle 50R1 from a value of 0.8 micron in cycle 49R1) and the sulfate hydrophylic growth factor. The assumed size distribution is representative of the volcanic conditions. The same varying size distribution is used to estimate the surface area density (SAD) of sulfate aerosols.

(*d*) *Production of nitrate and ammonium with EQSAM4Clim*

For the production of nitrate and ammonium through gas-particle partitioning we use EQSAM4Clim version 12 – in the following referred to as E4C. E4C includes a revised calculation of aerosol acidity that is described in Metzger *et al.* (2024). The implementation of E4C in IFS-COMPO and its use to estimate aerosol and precipitation pH are described in Rémy *et al.* (2024). Its main features are summarized below, and a detailed description is provided in Appendix B.1. The overall gas/liquid/solid partitioning and parameterization for calculating aerosol water uptake of E4C in IFS-COMPO is identical to that described in Metzger *et al.* (2016). In contrast to the original version of EQSAM (Metzger *et al.*, 2002a), E4C is based on a compound specific single-solute coefficient (ν_i), which was introduced in Metzger *et al.* (2012) to accurately parameterise the single solution hygroscopic growth, considering the Kelvin effect. This ν_i-approach accounts for the water uptake of concentrated nanometer-sized particles up to dilute solutions, i.e. from the compounds relative humidity of deliquescence (RHD) up to supersaturation, using the Köhler theory (Köhler, 1936). E4C extends the ν_i-approach to multicomponent mixtures, including semi-volatile NH_4^+ compounds and major crustal elements such as Ca^{++}, Na^+ and Mg^{2+}. In E4C the entire gas–liquid–solid aerosol phase partitioning and water uptake, including major mineral cations, is solved analytically without iterations and numerical noise, making E4C suited for climate and high resolution NWP applications.

E4C calculates the equilibrium aerosol composition and aerosol associated water mass (AW) through the neutralization of anions by cations, which yields numerous salt compounds. All salt compounds (except calcium sulphate, $CaSO_4$) partition between the liquid and solid aerosol phase, depending on atmospheric temperature (T), relative humidity (RH), AW and the T-dependent RHD of (i) single solute compound solutions, or (ii) of mixed salt solutions (Metzger *et al.*, 2012).

E4C outputs the aerosol pH along with the concentration of aerosol species (lumped cations, anions), as well as the AW.

A schematic of E4C and its integration into IFS-COMPO is shown in Figure 2.3. E4C takes as input, for each model time-step and within a given grid box, particle number and size, (i) T and RH as provided by the meteorological component of IFS-COMPO, (ii) the aerosol precursor gases, i.e., major oxidation products of natural and anthropogenic air pollution represented by NH_3 and HNO_3 from tropospheric chemistry, and (iii) the ionic aerosol concentrations lumped (liquid+solid) anions: sulphate (SO_4^{2-}),

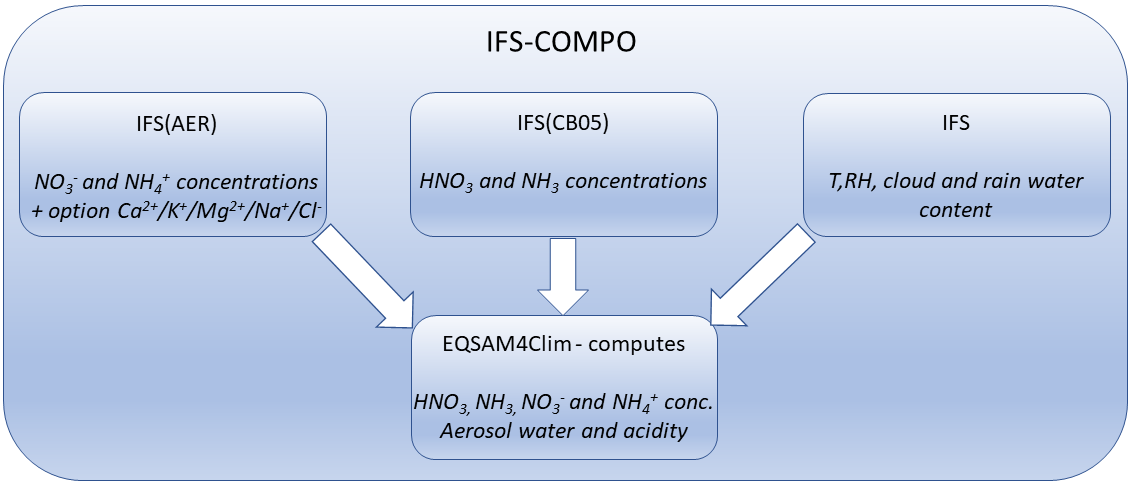


Figure 2.3 *Schematic showing the inputs and outputs of E4C as implemented into IFS-COMPO.*

nitrate (NO_3^-), chloride (Cl^-), and lumped (liquid+solid) cations: ammonium (NH_4^+), sodium (Na^+), potassium (K^+), magnesium (Mg^{2+}), and calcium (Ca^{2+}) as provided by IFS-AER.

It should be noted that input values for the application of E4C, i.e., HNO_3/NO_3^-, NH_3/NH_4^+ and SO_4^{2-} are prognostic species and directly provided by IFS-COMPO, while the mineral anions (Cl^-) and cations (Na^+,K^+, Mg^{2+} and Ca^{2+}), must be derived from the existing tracers. HCl is not yet coupled as the CB05 scheme currently does not contain tropospheric halogen chemistry. For Ca^{2+}, the approach chosen is the same as that described in Rémy *et al.* (2019): an experimental version of IFS-COMPO that decomposes dust into a simplified mineralogical composition has been developed and is used to compute a climatology of airborne calcite. Using an experimental IFS-COMPO version we used as an input the dataset of soil mineralogical composition of Journet *et al.* (2014) to provide an estimate of the calcite content of airborne dust. Geographically dependant scaling factors derived from this climatology are used to estimate the calcium content of dust, which varies between 5-10% for fine/coarse dust and 2-5% for super-coarse dust. For Cl^- and Na^+ input is derived from sea-salt aerosol assuming the mean mass fractions of 55.0% Cl^- and 30.6% Na^+ from dry sea-salt aerosol mass, following Myriokefalitakis *et al.* (2022). Assuming a 80% relative humidity over the ocean, this translates into 12.8% and 7.1% of the mass of sea-salt aerosol at 80% RH. The contribution of Na^+, K^+ Mg^{2+} from desert dust is derived from the desert dust tracers using constant scaling factors of 1.2%, 1.5%, and 0.9%, following the approach in EC-Earth (Myriokefalitakis *et al.*, 2022). This is applied to the sum of all desert dust and sea-salt bins, while in cycle 49R1 this was applied only to the finest bin of each species.

The ammonium nitrate concentration as estimated by E4C is used to compute the tendencies of the ammonium and nitrate from gas-particle partitioning (NO3_1). The nitrate from heterogeneous reactions (NO3_2) tendency is computed from the sum of the E4C concentrations of the other N compounds : $NaNO_3$, $Ca(NO_3)_2$, KNO_3 and $MgNO_3$. Finally, the tendency of HNO_3 is derived from the sum of the tendencies of $NO3_1$ and $NO3_2$, and that of NH_3 from the tendency of NH_4.

(*e*) *Parameterization of aerosol, cloud and precipitation pH*

Using E4C-v12, five different pH values can be computed as detailed in Metzger *et al.* (2024). Version 12 allows the differentiation of the Liquid Water Content (LWC) $[\mathrm{kg(H_2O)/m^3(air)}]$ associated with different type of atmospheric aerosols, haze/fog, or cloud droplets contained in the troposphere as defined in Eqs. (2.19a-2.19e):

$$\mathrm{pH_{equil}} = -\log_{10}\left(\frac{[\mathrm{H^+}]}{\mathrm{LWC_{equil}}}\cdot\frac{1}{\mu_s^o}\right) \tag{2.19a}$$

$$\mathrm{pH_{noneq}} = -\log_{10}\left(\frac{[\mathrm{H^+}]}{\mathrm{LWC_{noneq}}}\cdot\frac{1}{\mu_s^o}\right) \tag{2.19b}$$

$$\mathrm{pH_{cloud}} = -\log_{10}\left(\frac{[\mathrm{H^+}]}{\mathrm{LWC_{cloud}}}\cdot\frac{1}{\mu_s^o}\right) \tag{2.19c}$$

$$\mathrm{pH_{precip}} = -\log_{10}\left(\frac{[\mathrm{H^+}]}{\mathrm{LWC_{precip}}}\cdot\frac{1}{\mu_s^o}\right) \tag{2.19d}$$

$$\mathrm{pH_{total}} = -\log_{10}\left(\frac{[\mathrm{H^+}]}{\mathrm{LWC_{total}}}\cdot\frac{1}{\mu_s^o}\right) \tag{2.19e}$$

Here, (i) $\mathrm{LWC_{equil}}$ $[\mathrm{kg(H_2O)/m^3(air)}]$ denotes the equilibrium water content calculated within E4C (from Eq.(22) in Metzger *et al.* (2016)), (ii) $\mathrm{LWC_{noneq}}$ is the aerosol liquid water content associated with aerosol species not considered in the equilibrium computations of E4C (e.g., from chemical aging of pre-existing organic or black carbon particles as used e.g. in Metzger *et al.* (2018)), (iii) $\mathrm{LWC_{cloud}}$ denotes the cloud liquid water content, (iv) $\mathrm{LWC_{precip}}$ denotes the liquid water content of a given precipitation flux and finally (v) $\mathrm{LWC_{total}}$ denotes the sum of LWC of Eqs. (2.19a-2.19d). The pH values of Eqs. (2.19b-2.19e) are an optional output feature and requires the corresponding input to E4C from IFS. The key equations of the revised E4C pH calculation (Metzger *et al.*, 2024) are given in Appendix Sec. B.9.

Accounting for the different pH values is important for air quality and climate applications, because of the influence of solution pH on aqueous phase chemistry in terms of SO_4^{2-} production and the subsequent deposition processes, as described above. A first evaluation of simulated aerosol and precipitation pH versus climatological and routine observations is shown in (Rémy *et al.*, 2024).

(*f*) *Equilibration time*

Most atmospheric composition models, including IFS-COMPO, assume that a thermodynamic equilibrium is reached instantaneously for the gas-particle partitioning, without considering kinetic limitations on the condensation of HNO_3 or NH_3 onto existing aerosol particles. The ammonia–nitrate system may not reach equilibrium within a time step owing to transport limitations between the gas and aerosol phases (Wexler and Seinfeld, 1990). In cycle 50R1, the equilibration time from (Jones *et al.*, 2021) has been implemented. The time taken to reach equilibrium depends on ambient temperature and relative humidity, as well as the aerosol size and uptake coefficient (γ), where the uptake coefficient is defined as the number of gas molecules condensing on a particle divided by the number impacting onto the particle surface. (Ackermann and Ebel, 1995) find that equilibration time (τ) may range from around 2 min for particles with diameters of 0.1 µm to around 1h for diameters of 0.5 µm, depending on the uptake rate. Here the uptake rate kHNO3 is determined for each nitrate aerosol specie using the first-order uptake theory of (Schwartz, 1986) and by applying the correction factors from (Fuchs and Sutugin, 1970) for molecular effects and for limitations in interfacial mass transport:

$$D_g = \frac{3}{8A_c\rho_a d_a^2}\left[\frac{m_a R_a T}{2\pi}\times\frac{m_a + m_{\mathrm{HNO_3}}}{m_{\mathrm{HNO_3}}}\right]^{\frac{1}{2}}, \tag{2.20a}$$

$$\lambda = \frac{3D_g}{v} = \frac{3D_g}{\sqrt{\frac{8R_aT}{\pi m_{\mathrm{HNO_3}}}}}, Kn = \frac{2\lambda}{D}, \tag{2.20b}$$

$$k_{\mathrm{HNO_3}} = \frac{2\pi D D_g}{1 + \frac{4Kn}{3\gamma}\times\left(1 - \frac{0.47\gamma}{1+Kn}\right)}. \tag{2.20c}$$

The equations above determine the molecular diffusivity coefficient (Dg, $\mathrm{m^2s^{-1}}$), the mean free path (λ, m), the Knudsen number (Kn)., and uptake rate (kHNO3, $\mathrm{m^3s^{-1}}$) respectively. Constants in the algorithm include the Avogadro constant Ac=6.022×1023 $\mathrm{mol^{-1}}$, the gas constant of dry air Ra=8.314 $\mathrm{Jmol^{-1}K^{-1}}$, the molar mass of dry air $\mathrm{m_a}$=0.029 kg $\mathrm{mol^{-1}}$, the molar mass of HNO_3 mHNO3=0.063 kg $\mathrm{mol^{-1}}$, the molecular diameter of dry air molecules da=4.5×10^{-10} m, and the reactive uptake coefficient

(γ) for HNO3. Variables in the above equations also include the air temperature T (K) and air density ρ_a (kg m^{-3}). ν is the mean molecular speed (m s^{-1}). D is the mass median diameter of each particle nitric acid can condense on: nitrate (fine mode), as well as the three sea-salt aerosol and desert dust bins. The uptake rate is thus computed for all of these particles : kHNO3,NO3 , kHNO3,dd1/2/3 and kHNO3,SS1/2/3 . The total equilibration time τ in seconds may then be related to the inverse of product of the uptake rate for one particle kHNO3,NO3/DD/SS and the corresponding aerosol number concentration $N_{NO3/DD/SS}$using the following equation:

$$\tau = \frac{1}{N_{NO31}k_{NO3_1} + N_{SS1}k_{SS_1} + N_{SS2}k_{SS2} + N_{SS3}k_{SS_3} + N_{DD1}k_{DD_1} + N_{DD2}k_{DD_2} + N_{DD3}k_{DD_3}} \tag{2.21}$$

Where k_{SS_1} denotes k_{HNO3,SS_1}, k_{SS_2} denotes k_{HNO3,SS_2} etc. The number concentration for each tracer is computed using the prognostic mass mixing ratio and the assumed size distribution. A maximum of 10000s is set for τ. Rather than assume instantaneous thermodynamic equilibrium in the ammonia–nitrate system, the model assumes an exponential decay of the gas phase towards equilibrium using the equilibration time τ. It should be noted that the equilibration time is thus very dependent on the assumed particle size of nitrate, sea-salt aerosol and desert dust. The nitrate and nitric acid tendencies are computed by a factor F that depends on the model time step Δt :

$$F = 1 - \exp\left(-\frac{\Delta t}{\tau}\right) \tag{2.22}$$

2.3 STRATOSPHERIC CHEMISTRY

The stratospheric chemistry module in IFS-COMPO is a re-implementation of the chemical module originally developed for the Belgian Assimilation System for Chemical ObsErvations (Errera *et al.*, 2019, BASCOE) to assimilate satellite observations of stratospheric composition. BASCOE is based on a Chemistry Transport Model (CTM) of the stratosphere which is used to investigate stratospheric photochemistry (Muncaster *et al.*, 2012; Prignon *et al.*, 2021). From the BASCOE system the chemical scheme and the parameterization for Polar Stratospheric Clouds (PSC) has been implemented in the IFS.

The COMPO stratospheric chemical scheme used here is labelled sb15bs. It includes 64 species interacting through 157 gas-phase, 9 heterogeneous and 53 photolytic reactions. The available tracers are defined in Tables A.1 and A.3 This chemical scheme merges the reaction lists developed by Errera and Fonteyn (2001) to carry out data assimilation of stratospheric trace gases with the list included in the SOCRATES 2-D model for long-term studies of the middle atmosphere (Brasseur *et al.*, 2000; Chabrillat and Fonteyn, 2003). The resulting list of species includes all the ozone-depleting substances and greenhouse gases necessary for multi-decadal simulations of the couplings between dynamics and chemistry in the stratosphere, as well as the reservoir and short-lived species necessary for a complete description of stratospheric ozone photochemistry. Gas-phase and heterogeneous reaction rates have been updated to JPL evaluation 18 (Burkholder *et al.*, 2015).

In addition to the original BASCOE chemical scheme, the COMPO stratospheric chemistry module also includes basic sulphur chemistry to represent the formation of gas-phase sulphuric acid and to allow coupling with the sulfate aerosol module. OCS, SO_2, SO_3 and H_2SO_4 are thus included into the scheme sb15bs, using again the chemical rates compiled by JPL (Burkholder *et al.*, 2015) for the five corresponding gas-phase reactions and the photolysis of OCS and SO_3. The photolysis of H_2SO_4 uses the absorption cross-sections reported by Feierabend *et al.* (2006) in the visible wavelength range, Lane and Kjaergaard (2008) at Lyman-alpha as well as the pressure-dependent quantum yields reported by Miller *et al.* (2007).

2.3.1 Photolysis

Lookup tables of photolysis rates are computed by the TUV package (Madronich and Flocke, 1999) as a function of log-pressure altitude, ozone overhead column and solar zenith angle using a spectral grid of 171 wavelength bins covering the spectral range from 116 nm to 735 nm. The photolysis tables used in chemical scheme sb15bs are based on absorption cross-sections from JPL evaluation 17 (Sander *et al.*, 2011), while the solar spectral irradiance is based on the daily SSI dataset from CMIP6 as published by (Matthes *et al.*, 2017). The official CMIP6 dataset is a compromise (mean spectra) between the empirical model NRLSSI2 and the semiempirical model SATIRE. It covers the years 1850 to 2014 and is extended with forecasts for the years 2015 to 2299.

It is also possible to compute photolysis rates fully online using the same TUV package. While this leads to significant increases in computational costs it was demonstrated that this results in significant improvements for stratospheric ozone especially in the polar regions.

A special case is the photolysis rate of NO_2, for which the tropospheric photolysis module (MBA, see above) is used to calculate the photolysis also for the stratosphere. An advantage is that this module considers the albedo of the surface and of clouds.

2.3.2 Heterogeneous chemistry

The kinetic rates for heterogeneous chemistry are determined by the parameterization of Fonteyn and Larsen (1996), using expressions for the uptake coefficients on sulfate aerosols according to Shi *et al.* (2001) and on PSCs (Sander, 2000).

The SAD of stratospheric aerosols is based on CMIP6-data for the period 1990-2018, while after the year 2018 the 2018 values are used. Ice PSCs are presumed to exist at any grid point in the winter/spring polar regions where water vapour partial pressure exceeds the vapour pressure of water ice (Murphy and Koop, 2005).

Nitric Acid Tri-hydrate (NAT) PSCs are assumed when the nitric acid (HNO_3) partial pressure exceeds the vapour pressure of condensed HNO_3 at the surface of NAT PSC particles (Hanson and Mauersberger, 1988). The surface area density is set to 2×10^{-6} cm^2/cm^3 for ice PSCs and 2×10^{-7} cm^2/cm^3 for NAT PSCs. The sedimentation of PSC particles causes denitrification and dehydration. This process is approximated by an exponential decay of HNO_3 with a characteristic time scale of 20 days for grid points where NAT particles are assumed to exist, and an exponential decay of HNO_3 and H_2O with a characteristic time scale of 9 days for grid points where ice particles are supposed to exist. The actual transport to lower levels by sedimentation is not considered.

2.4 BOUNDARY CONDITIONS FOR LONG-LIVED TRACE GASES

Surface boundary conditions for the tracer MMR are imposed for some long-lived stratospheric gases with surface emissions such as CH_4, N_2O and CFCs. This approach allows to simulate the flux from the troposphere to the stratosphere without the need to simulate the surface emissions of these species.

Meinshausen *et al.* (2017) provide a comprehensive analysis of GHG surface boundary conditions (including the stratospheric source gases) for CMIP6 model simulations as a function of month, year and latitude from pre-industrial times until December 2014. This dataset has been extended from 2015 to 2100 and beyond (Meinshausen *et al.*, 2020) using projections into the future of emissions of greenhouse gases following each of the 9 socioeconomic scenarios provided for the CMIP6 activity (Gidden *et al.*, 2019). We choose the dataset corresponding to the scenario SSP2 4.5, which follows a middle-of-the-road pathway.

For the IFS, a subset of values has been extracted from the combined dataset, at the native 15°latitudinal resolution available between 1995-2030 and encoded in a text datafile (if needed, this can be extended for future climate mode experiments). The surface constraints apply to 16 species, namely: CCl_4, CFC11, CFC113, CFC114, CFC115, CFC12, CH_3Br, CH_3CCl_3, CH_3Cl, CH_4, CO_2, ha1211, ha1301, HCFC22, N_2O and H_2.

The CH_4 surface tropospheric mixing ratios are constrained using a time-dependent latitudinal gradient based on monthly-mean observations derived from remote stations, i.e. no longitudinal variability is accounted for. The option to use CH_4 emissions is also supported, combined with a smaller relaxation at the surface in case CH_4 itself is not assimilated. Optionally also the CH_4 atmospheric concentrations can be forced to a fixed value or taken from the GHG scheme.

2.5 CLIMATOLOGICAL METHANE LOSS (GHG)

In the GHG configuration of the IFS only CO_2 and CH_4 are simulated. The chemical sink of CH_4 mainly by reaction with OH in the troposphere and the stratosphere is simulated by a climatological loss rate derived from the TM5 model (Krol *et al.*, 2005) optimised with methyl chloroform in the troposphere and the 2-D Max Planck Institute photochemical model in the stratosphere, following Bergamaschi *et al.* (2009). The loss rate is provided as climatological 3D monthly mean data set at a resolution of $3° \times 4°$.

2.6 CODE OVERVIEW

The IFS-AER aerosol routines are called from AER_PHY2 and AER_PHY3. The routines in AER_PHY2 deal with online emissions (desert dust, sea salt) and the computation of dry deposition and sedimentation velocities, and will be detailed in Chapter 3. The routines called from AER_PHY3 deal with aerosol conversion processes, wet deposition and optical diagnostics:

AER_CGROWTH
: Conversion of hydrophobic to hydrophilic for OM and BC (called twice)

AER_EQSAM4CLIM_IFS
: Interface to EQSAM4CLIM, which computes nitrate and ammonium from gas-particle conversion and from heterogeneous reactions and related diagnostics

AER_SCAVIN and AER_SCAVBC
: Aerosol in-cloud and below-cloud wet deposition, called twice (for large scale and convective precipitations)

AER_BDGTMSS
: Computes aerosol optical diagnostics: AOD, SSA, extinction profiles, etc.

The COMPO atmospheric chemistry module is called from CHEM_MAIN and is named CHEM_BASCOETM5. From the BASCOETM5 module a large set of routines are called. Here we list them in their approximate order. The most important routines that drive *stratospheric* chemistry aspects are:

BASCOE_J_INTERP
: Interpolation of photolysis rates used for the stratospheric chemistry module

BASCOE_HETCONST
: Compute heterogeneous reaction rates active in the stratosphere

BASCOE_KPP_RATES
: Compute reaction rates active in stratospheric chemistry

BASCOE_KPP_INITIALIZE
: Map array of trace gases towards array ordering used in KPP-solver in the stratosphere

BASCOE_KPP_INTEGRATOR
: Call to KPP solver, this updates trace gases due to chemical reactions in stratosphere

BASCOE_KPP_UPDATE_CIFS_CONC
: Map updated array of trace gases back to standard ordering as used in IFS

BASCOE_PSC_PARAM
: Compute dehydration and denitrification in stratosphere

Apart from the stratospheric photolysis routine these routines are only called for grid boxes which are above the tropopause. The list of the most important routines involved in *tropospheric* chemistry aspects are:

COD_OP_TM5
: Compute cloud absorption and scattering optical depths as input for tropospheric photolysis rate computation

TM5_MACC_AEROSOL
: Compute aerosol absorption and scattering optical depths for tropospheric photolysis rate computation

TM5_PHOTO_FLUX
: Interface routine to compute tropospheric photolysis rates

TM5_WETCHEM
: Compute aqueous phase chemistry

TM5_CALRATES
: Compute reaction rates (including heterogeneous) active in tropospheric chemistry

TM5_KPP_RATES
: Map reaction rates towards array used in KPP-solver in troposphere

TM5_KPP_INITIALIZE
: Map array of trace gases towards array ordering used in KPP-solver in troposphere

TM5_KPP_INTEGRATOR
: Call to KPP solver, updates trace gases due to chemical reactions in troposphere

TM5_KPP_UPDATE_CIFS_CONC
: Map updated array of trace gases back to standard ordering used in IFS

TM5_SOA
: Call to SOA production routine

TM5_O3S
: Call to simple solver to compute tropospheric ozone loss for O3S tracer

Again, apart from photolysis aspects these are only called for grid boxes that are located below the tropopause.

At the end of the routine CHEM_BASCOETM5, after all processes have been finalized, the updated concentrations of all relevant tracers are converted into chemical tendencies (variable PTENC1), which is the most important output of this module.

APPENDIX A. CHEMISTRY SPECIES TABLES

Table A.1 *Trace gases defined in IFS-COMPO atmospheric chemistry module - p1: basic inorganic and organic chemistry.*

Formula (IFS name)	Trace gas name	Characteristics
O	oxygen atom	chemistry modeled in stratosphere only
O1D	excited oxygen atom	chemistry modeled in stratosphere only
O3	ozone	
H	hydrogen atom	chemistry modeled in stratosphere only
H2	hydrogen	chemistry modeled in stratosphere only
OH	hydroxyl radical	
HO2	hydroperoxy radical	
H2O2	hydrogen peroxide	
H2O	water	chemistry modeled in stratosphere only differs from q
N	nitrogen atom	chemistry modeled in stratosphere only
NO	nitrogen monoxide	
NO2	nitrogen dioxide	
NO3	nitrate radical	
HNO3	nitric acid	
HO2NO2	pernitric acid	
N2O5	dinitrogen pentoxide	
N2O	nitrous oxide	chemistry modeled in stratosphere only
PAN	peroxyacetyl nitrate	chemistry modeled in troposphere only
CH3O2NO2	methyl peroxy nitrate	chemistry modeled in troposphere only
HONO	nitrous acid	chemistry modeled in troposphere only
CO	carbonmonoxide	
CO2	carbondioxide	chemistry modeled in stratosphere only differs from GHG config.
CH4	methane	differs from GHG config.
CH3O2	methylperoxy radical	
CH3OOH	methylperoxide	
CH2O	formaldehyde	
CH3	methyl radical	chemistry modeled in stratosphere only
CH3O	methoxy radical	chemistry modeled in stratosphere only
HCO	formyl radical	chemistry modeled in stratosphere only
HCN	hydrogen cyanide	Marker for biomass burning
CH3CN	acetonitrile	Marker for biomass burning
DMS	dimethyl sulfide	chemistry modeled in troposphere only
SO2	sulfur dioxide	
SO3	sulfur trioxide	
MSA	methanesulfonic acid	chemistry modeled in troposphere only
OCS	carbonyl sulfide	chemistry modeled in stratosphere only
H2SO4	sulfuric acid	chemistry modeled in troposphere only
NH3	ammonia	chemistry modeled in troposphere only
NH2	amine	chemistry modeled in troposphere only

Table A.2 *Trace gases defined in the IFS-COMPO atmospheric chemistry module - p2: non-methane hydrocarbons.*

Formula (IFS name)	Trace gas name	Characteristics
C2H4	ethene	chemistry modeled in troposphere only
PAR	paraffins	CB05 tracer with single-bond
OLE	olefins	CB05 tracer with double-bond
ALD2	aldehydes	CB05 tracer representing aldehydes
ROOH	peroxides	CB05 tracer representing peroxides
ONIT	organic nitrates	CB05 tracer representing organic nitrates
C2O3	peroxyacetyl radical	chemistry modeled in troposphere only
ROR	organic ethers	CB05 tracer
RXPAR	PAR budget corrector	CB05 tracer
XO2	NO to NO2 operator	CB05 tracer
XO2N	NO to alkyl nitrate	CB05 tracer
CH3OH	methanol	chemistry modeled in troposphere only
CHOCHO	glyoxal	chemistry modeled in troposphere only
HCOOH	formic acid	chemistry modeled in troposphere only
MCOOH	methacrylic acid	chemistry modeled in troposphere only
C2H6	ethane	chemistry modeled in troposphere only
C2H5OH	ethanol	chemistry modeled in troposphere only
C3H8	propane	chemistry modeled in troposphere only
C3H6	propene	chemistry modeled in troposphere only
CH3COCHO	methyl glyoxal	chemistry modeled in troposphere only
CH3COCH3	acetone	chemistry modeled in troposphere only
ACO2	acetone product	chemistry modeled in troposphere only
HYAC	hydroxyacetone	chemistry modeled in troposphere only
IC3H7O2	IC3H7O2	chemistry modeled in troposphere only
HYPROPO2	HYPROPO2	chemistry modeled in troposphere only
C5H8	isoprene	chemistry modeled in troposphere only
ISPD	methacrolein; MVK	chemistry modeled in troposphere only
ISOPOOH	isoprene hydroperoxides	chemistry modeled in troposphere only
ISOPBO2	Isoprene radical B	chemistry modeled in troposphere only
ISOPDO2	Isoprene radical D	chemistry modeled in troposphere only
HPALD1	hydroperoxy aldehydes type 1	chemistry modeled in troposphere only
HPALD2	hydroperoxy aldehydes type 2	chemistry modeled in troposphere only
GLYALD	glycolaldehyde	chemistry modeled in troposphere only
C10H16	terpenes	chemistry modeled in troposphere only
XYL	xylenes	chemistry modeled in troposphere only
TOL	toluene	chemistry modeled in troposphere only
AROO2	peroxy radical from aromatics	chemistry modeled in troposphere only
SOG1	condensable gas type 1	secondary organic aerosol precursor
SOG2A	condensable gas type 2a	secondary organic aerosol precursor
SOG2B	condensable gas type 2b	secondary organic aerosol precursor

Table A.3 *Trace gases defined in the IFS-COMPO atmospheric chemistry module - p3: Stratosphere.*

Formula (IFS name)	Trace gas name	Characteristics
CL	chlorine atom	reactive in stratosphere only
CL2	chlorine	reactive in stratosphere only
CH3CL	methyl chloride	reactive in stratosphere only
CLO	chlorine monoxide	reactive in stratosphere only
OCLO	chlorine dioxide	reactive in stratosphere only
CLOO	asym. chlorine dioxide radical	reactive in stratosphere only
HCL	hydrogen chloride	reactive in stratosphere only
HOCL	hypochlorous acid	reactive in stratosphere only
CLNO2	nitryl chloride	reactive in stratosphere only
CLONO2	chlorine nitrate	reactive in stratosphere only
CL2O2	dichlorine dioxide	reactive in stratosphere only
BR	bromine atom	reactive in stratosphere only
BR2	bromine atomic ground state	reactive in stratosphere only
CH3BR	methyl bromide	reactive in stratosphere only
CH2BR2	dibromomethane	reactive in stratosphere only
CHBR3	tribromomethane	reactive in stratosphere only
BRO	bromine monoxide	reactive in stratosphere only
HOBR	hypobromous acid	reactive in stratosphere only
BRCL	bromine monochloride	reactive in stratosphere only
HBR	hydrogen bromide	reactive in stratosphere only
BRONO2	bromine nitrate	reactive in stratosphere only
HF	hydrogen fluoride	reactive in stratosphere only
CCL4	tetrachloromethane	reactive in stratosphere only
CH3CCL3	methyl chloroform	reactive in stratosphere only
CFC11	trichlorofluoromethane	reactive in stratosphere only
CFC12	dichlorodifluoromethane	reactive in stratosphere only
CFC113	trichlorotrifluoroethane	reactive in stratosphere only
CFC114	1 2-dichlorotetrafluoroethane	reactive in stratosphere only
CFC115	chloropentafluoroethane	reactive in stratosphere only
HCFC22	chlorodifluoromethane	reactive in stratosphere only
HA1301	bromotrifluoromethane	reactive in stratosphere only
HA1211	bromochlorodifluoromethane	reactive in stratosphere only

Table A.4 *Trace gases defined in the IFS-COMPO atmospheric chemistry module - p4: Marker tracers.*

Formula (IFS name)	Trace gas name	Characteristics
Rn	radon	fixed lifetime; produces lead
Pb	lead	only wet removal
STRATAER	stratospheric aerosol	reactive in stratosphere only
PSC	polar stratosph cloud	reactive in stratosphere only
O3S	stratospheric ozone	Copy of O3 in stratosphere, only loss in troposphere
NOXA	nitrogen oxides Transp	Family tracer for advection
CLXA	Reactive Chlorine	Family tracer for advection
BRXA	Reactive Bromine	Family tracer for advection

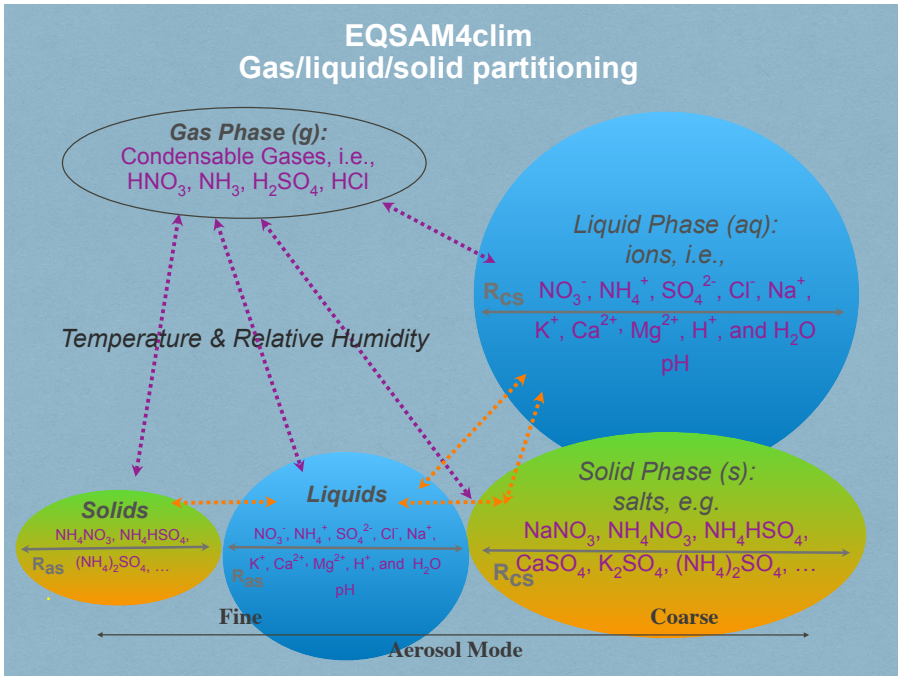


Figure B.1 *A schematic of the components included in EQSAM4Clim (E4C). Details are given in Appendix B.*

B APPENDIX B. EQSAM4CLIM

EQSAM4Clim version 12 – in the following referred to as E4C, enables calculating the gas/liquid/solid partitioning and aerosol water uptake as described in Metzger *et al.* (2012) and Metzger *et al.* (2016), and as schematically depicted in Figure B.1. E4C has been revised by Metzger *et al.* (2024) to improve the calculation of aerosol acidity. A standalone box-model version is available from Metzger (2023).

B.1 Mixed solution parameterisation framework

The mixed solution parameterisation framework was introduced in Metzger *et al.* (2016) – referred to in the following as M2016 – to efficiently calculate the aerosol water uptake for mixtures of semi-volatile and non-volatile compounds with the constraint of using only one parameter, i.e. a solute-specific coefficient ν_i, which was introduced in Metzger *et al.* (2012) – referred to in the following as M2012 – to accurately parameterise the single solution hygroscopic growth, also considering the Kelvin effect. M2012 and M2016 have shown that the ν_i approach is valid for a wide range of atmospheric conditions. Since computational efficiency is a requirement for the E4C parameterisation framework, the overall computational burden is minimised by a set of key-constraints:

(i) Solving the multicomponent and multiphase partitioning analytically, by using a consistent set of equations, based on one compound-specific single solute coefficient, ν_i $[-]$. This set of equations includes the solute molality, μ_s $[\mathrm{mol(solute)\,kg^{-1}(H_2O)}]$, and its equivalent expression in terms of the mass fraction solubility, χ_s $[-]$. Both are the essential thermodynamic properties in the E4C framework and only depend on ν_i; the temperature (T), relative humidity (RH) and the particle dry diameter (D_s) are given;
(ii) Breaking down the complexity of aerosol thermodynamics as much as possible, without the loss of crucial information and critical numerical accuracy, by using chemical domains with a neutralisation order for all salt compounds listed in Table B.5;
(iii) Minimising the dependencies on the required thermodynamic data by using a pre-determined ν_i coefficient for each electrolyte listed in Table B.5;
(iv) Assuming ν_i, constant for the entire range of water activity, a_w.

The relevant single solute equations (of M2012) are summarised in Appendix of Metzger *et al.* (2016). The mixed solution parameterisation framework has been determined by a multi-functional fitting, limited to one parameter per compound, to match the results of ISORROPIA II. The multiphase and multicomponent thermodynamic system has an analytical solution when the set of analytical equations and the ν_i coefficient of M2012 are used. Consequently, the E4C parameterisation framework has a similar applicability as ISORROPIA II.

B.2 Pre-determined ν_i

M2012 have detailed that a (unitless) single solute coefficient, i.e. $\nu_i\,[-]$, can be accurately deduced from one reference data pair of solute molality, μ_s, and the corresponding water activity, $a_w\,[-]$. E4C uses a data pair at solute saturation to pre-determine ν_i, since measurements are available for all major salt compounds that are of interest in atmospheric aerosol modelling. For the salt compounds listed in Table B.5, E4C uses the mass fraction solubility, $w_s\,[-]$, which is an equivalent expression of the saturation molality, $\mu_{s,sat}$. For the corresponding a_w, E4C uses the available relative humidity of deliquescence (RHD) values and obtain ν_i by solving Eq. (5b) of M2012 with a root finding method (bisection). To be consistent with ISORROPIA II, E4C determines ν_i from w_s and RHD data at temperature $T_o = 298\,[\mathrm{K}]$. The water activity data used by ISORROPIA II (and other EQMs) are tabulated only for room temperature. E4C therefore does not consider the T dependency of ν_i.

Table B.5 lists the pre-calculated ν_i values for each salt compound considered, together with the required thermodynamic data: Stoichiometric coefficient $\nu_s\,[-]$, the ion-pair charge $Z_s\,[-]$, the single solute parameter $\nu_i\,[-]$, the mass fraction solubility in percent $W_s\,[\%]$ ($w_s = W_s/100$), the molar masses $M_s\,[\mathrm{kg\,mol^{-1}}]$, the densities $D_s\,[\mathrm{kg\,m^{-3}}]$, RHD($T_o$) $[-]$ at reference temperature T_o = 298.15 [K] and the corresponding temperature coefficients, $T_{\mathrm{coef(RHD)}}\,[-]$. For details, we refer to M2012.

B.3 Chemical domains

To break down the complexity of aerosol thermodynamics as much as possible, E4C minimises the number of chemical compounds and equilibrium reactions that have to be considered. Following the original EQSAM approach (Metzger *et al.*, 2002a), E4C defines chemical domains with a sub-set of neutralisation reactions that are considered for a given T, RH and input concentrations (total of aerosol cations/anions and precursor gases), with all concentration units in $[\mathrm{mol\,m^{-3}(air)}]$. The EQSAM domain definition is listed in Table B.6 and applied in the mixed solution framework with if-else logic and top-down approach. The potential aerosol neutralisation levels depend on the input concentration ratio of total cations, tCAT. The cations are balanced against the total sulfate anions, for which E4C considers, for sulfate-rich cases, the total sulfates as bi-sulfate, $\mathrm{tHSO_4}$, or for sulfate-poor cases as total sulfate, $\mathrm{tSO_4}$. In any case, these totals "t" need to exceed a threshold, $\mathrm{MIN} = 1 \times 10^{-15}\,[\mathrm{mol\,m^{-3}(air)}]$; below the computations are neglected for a given domain. The definition of totals is given by a domain classification (Table B.6), implicitly taking into account the maximum neutralisation level that is theoretically possible for each domain:

- $\mathrm{tSO_4} = \sum (1 \cdot \mathrm{HSO_4^-} + 2 \cdot \mathrm{SO_4^{2-}})$
- $\mathrm{tHSO_4} = \sum (1 \cdot \mathrm{HSO_4^-} + 1 \cdot \mathrm{SO_4^{2-}})$
- $\mathrm{tCAT} = \sum (2 \cdot \mathrm{Ca^{2+}} + 2 \cdot \mathrm{Mg^{2+}} + 1 \cdot \mathrm{Na^+} + 1 \cdot \mathrm{K^+} + 1 \cdot \mathrm{NH_4^+})$.

The domain definition (Table B.6) is subsequently used to define the neutralisation reaction order (see Sect. B.4). The two semi-volatile compounds listed in Table B.5, $\mathrm{NH_4NO_3}$ and $\mathrm{NH_4Cl}$, are considered here only in the sulfate neutral (D1) domain. On the other hand, bi-sulfate is taken into account only for the sulfate-rich (D2) and very-rich (D3) cases, while free sulfuric acid is considered only for the un-neutralised sulfate case (D4).

B.4 Domain-dependent neutralisation reaction order

To avoid the numerical minimisation of the Gibbs free energy, which is required to obtain the equilibrium composition of mixed solutions, E4C defines for each domain (Table B.6) a neutralisation

Table B.5 *Thermodynamic data (Sect. B.2): stoichiometric coefficient ν_s [−], the ion-pair charge Z_s [−], the single solute parameter ν_i [−], the mass fraction solubility in percent W_s [%] ($w_s = W_s/100$), the molar masses M_s [kg mol^{-1}], the densities D_s [kg m^{-3}], RHD(T_o) [−] at reference temperature $T_o = 298.15$ [K], and the corresponding temperature coefficients, $T_{coef(RHD)}$ [−]. The ν_i values have been obtained from the RHD and W_s values (at $T_o = 298$ [K]) following the method described in Metzger et al. (2012). The RHD values are taken from Fountoukis and Nenes (2007); the W_s and all other values are taken from the Handbook of Chemistry and Physics (Lide, 2006).*

Species 01–05	H_2O	H_2SO_4	HNO_3	HCl	NH_3
ν_s — Z_s	2 — 1	3 — 2	2 — 1	2 — 1	1 — 1
ν_i	–	–	–	–	–
W_s	–	–	–	–	–
M_s	0.018020	0.098090	0.063020	0.036460	0.017040
D_s	997	1830	1513	1490	696
RHD	–	–	–	–	–
T_{coef}	–	–	–	–	–
Species 06–10	$(NH_4)_3H(SO_4)_2$	$(NH_4)_2SO_4$	NH_4HSO_4	NH_4NO_3	NH_4Cl
ν_s — Z_s	5 — 3	3 — 2	2 — 1	2 — 1	2 — 1
ν_i	1.616356	1.274822	1.253573	1.051480	1.243054
W_s	53.30	43.31	76.00	68.05	28.34
M_s	0.247300	0.132170	0.115130	0.080060	0.053500
D_s	1775	1770	1780	1720	1519
RHD	0.6900	0.7997	0.4000	0.6183	0.7710
T_{coef}	186.00	80.00	384.00	852.00	239.00
Species 11–15	$Na_3H(SO_4)_2$	Na_2SO_4	$NaHSO_4$	$NaNO_3$	NaCl
ν_s — Z_s	5 — 3	3 — 2	2 — 1	2 — 1	2 — 1
ν_i	–	1.278762	1.293906	1.160345	1.358377
W_s	–	21.94	66.18	47.70	26.47
M_s	0.262120	0.142050	0.120070	0.085000	0.058440
D_s	2565	2700	2430	2260	2170
RHD	–	0.9300	0.5200	0.7379	0.7528
T_{coef}	–	80.00	−45.00	304.00	25.00
Species 16–20	$K_3H(SO_4)_2$	K_2SO_4	$KHSO_4$	KNO_3	KCl
ν_s — Z_s	5 — 3	3 — 2	2 — 1	2 — 1	2 — 1
ν_i	–	1.286445	1.308499	1.014102	1.256989
W_s	–	10.71	33.60	27.69	26.23
M_s	0.310444	0.174266	0.136178	0.101108	0.074548
D_s	2490	2660	2320	2110	1988
RHD	–	0.9750	0.8600	0.9248	0.8426
T_{coef}	–	35.60	0	0	159.00
Species 21–25	n/a	$CaSO_4$	n/a	$Ca(NO_3)_2$	$Ca(Cl)_2$
ν_s — Z_s	– / –	2 — 2	– / –	3 — 2	3 — 2
ν_i	–	1.271828	–	1.586562	2.024869
W_s	–	0.21	–	59.02	44.84
M_s	–	0.136150	–	0.164100	0.110980
D_s	–	2960	–	2500	2150
RHD	–	0.9900	–	0.4906	0.2830
T_{coef}	–	0	–	509.40	551.10
Species 25–30	n/a	$MgSO_4$	n/a	$Mg(NO_3)_2$	$Mg(Cl)_2$
ν_s — Z_s	– / –	2 — 2	– / –	3 — 2	3 — 2
ν_i	–	1.435281	–	1.878693	2.107772
W_s	–	26.31	–	41.59	35.90
M_s	–	0.120375	–	0.148325	0.095205
D_s	–	2660	–	2300	2325
RHD	–	0.8613	–	0.5400	0.3284
T_{coef}	–	−714.45	–	230.20	42.23

Table B.6 *Chemical domains (introduced in Sect. B.3).*

D4	SULFURIC ACID ONLY	tCAT $<$ MIN	AND	$tHSO_4 \geq$ MIN
D3	SULFATE VERY RICH	tCAT $\geq$ MIN	AND	tCAT $< tHSO_4$
D2	SULFATE RICH	tCAT $\geq tHSO_4$	AND	tCAT $< tSO_4$
D1	SULFATE NEUTRAL	tCAT $\geq tSO_4$		

Table B.7 *Neutralisation reaction order for Table B.6 (introduced in Sect. B.4).*

D1	i.e. Sulfate neutral	
1. $CaSO_4$	2. $MgSO_4$	3. K_2SO_4
4. Na_2SO_4	5. $(NH_4)_2SO_4$	6. $Ca(NO_3)_2$
7. $Mg(NO_3)_2$	8. KNO_3	9. $NaNO_3$
10. NH_4NO_3	11. $Ca(Cl)_2$	12. $Mg(Cl)_2$
13. KCl	14. NaCl	15. NH_4Cl
D2	i.e. Sulfate rich	
1. $CaSO_4$	2. $MgSO_4$	3. K_2SO_4
4. $KHSO_4$	5. Na_2SO_4	6. $NaHSO_4$
7. $(NH_4)_2SO_4$	8. NH_4HSO_4	
D3	i.e. Sulfate very rich	
1. $CaSO_4$	2. $MgSO_4$	3. $KHSO_4$
4. $NaHSO_4$	5. NH_4HSO_4	6. $H-HSO_4$
D4	i.e. Sulfuric acid	
1. H_2SO_4		

reaction order (NRO), which can practically be considered as the salting-out effect of salt solutes (Metzger and Lelieveld, 2007). E4C ranks the cations and anions according to their preferred neutralisation reaction by

- Anions: $SO_4^{2-} - HSO_4^- - NO_3^- - Cl^-$
- Cations: $Ca^{2+} - Mg^{2+} - K^+ - Na^+ - NH_4^+ - H^+$,

which yields the NRO listed in Table B.7. The ordering is based on numerous modelling studies, both extensive box-modelling comparisons (Metzger *et al.*, 2002a, 2006, 2012) and global applications (Metzger *et al.*, 2002b; Metzger and Lelieveld, 2007). Note that E4C has constrained the ordering to achieve the closet agreement with ISORROPIA II for two reasons: (1) ISORROPIA II was that time the only EQM that was widely applied in global modelling; (2) while it also considered the mineral cations Ca^{2+}, Mg^{2+} and K^+.

To solve the mixed solution framework, E4C applies the NRO to balance cation–anion pairs that have a non-zero ion–ion product. Within a chemical domain (Table B.6), the electrolyte compounds listed in Table B.7 are subsequently formed for non-zero ion–ion product, until all cation–anion pairs are paired, or either all cations or anions are fully neutralised. To analytically solve the entire gas–liquid–solid partitioning, E4C considers at this stage all electrolytes in solution (computing totals of gas and ions). The gas–solid and gas–liquid partitioning of semi-volatile compounds, the liquid–solid partitioning and the water uptake are determined in that order in subsequent and independent computational steps (a schematic is given in Sect. S2 of the Supplement of M2016).

B.5 Treatment of semi-volatile compounds

Table B.5 includes two semi-volatile compounds that exhibit the gas–liquid–solid partitioning, i.e. ammonium nitrate, NH_4NO_3, and ammonium chloride, NH_4Cl. Both are allowed in the E4C framework

Table B.8 *Equilibrium dissociation constants* [ppbv^2] *and T coefficients* [−]*; from Fountoukis and Nenes (2007).*

NH_4NO_3	$K_p^o = 57.46$	$a = -74.38$	$b = 6.120$
NH_4Cl	$K_p^o = 1.086$	$a = -71.00$	$b = 2.400$

only in D1, provided that a surplus ammonium, NH_4^+, is available. The implicit assumption is that all sulfates are neutralised first through the reactions with cations (Sect. B.4). Only excess ammonium may further neutralise anions, nitrate, NO_3^- and/or chloride, Cl^-. Thus, semi-volatile compounds can only partition into the particles, if the concentration product exceeds a threshold that is given by the temperature- and humidity-dependent equilibrium dissociation constant, $K_p(T, \mathrm{RH})$. The equilibrium partitioning is detailed in Seinfeld and Pandis (2006) – hereafter referred to as SP2006 (Sect. 10.4.3 ff); 1st edition Seinfeld and Pandis (1998) – hereafter referred to as SP1998.

B.5.1 RH < RHD – pure and mixed compounds

When the RH is below the RHD and the partial pressure product of gaseous (g) ammonia, $NH_3(g)$, and nitric acid, $HNO_3(g)$, with units either in [ppbv] or [mol m^{-3}(air)], equals or exceeds the temperature-dependent equilibrium dissociation constant, $K_p(T)$, solid (s) ammonium nitrate (AN), $NH_4NO_3(s)$, is assumed to be formed instantaneously:

$$\begin{aligned} &NH_3(g) + HNO_3(g) \Longleftrightarrow NH_4NO_3(s) \\ &\text{EQ: } [NH_3(g)] \cdot [HNO_3(g)] = K_{p,AN}(T). \end{aligned} \tag{B.1}$$

For Reaction (B.1) the equilibrium concentration of solid ammonium nitrate can be analytically computed from the gaseous concentrations by solving a quadratic equation – for the examples in the Supplement of M2016 [ppbv] is used. E4C computes the temperature dependency of the equilibrium dissociation constants, $K_p(T)$, following Fountoukis and Nenes (2007):

$$K_p(T) = K_p^o \times \exp[a \times (T_o/T - 1) + b \times (1 + \ln(T_o/T) - T_o/T)], \tag{B.2}$$

where T and $T_o = 298.15$ K, the ambient temperature and reference temperature, respectively. The $K_p^o(T_o)$ values are given in Table B.8 in [ppbv^2] for T_o and reference pressure, $P_o = 1$ [atm] = 101 325 [Pa], together with the dimensionless temperature coefficients, a and b [−]. For applications on a mole basis (e.g. for E4C), $K_p(T)$ [ppbv^2] can be converted to [(mol m^{-3}(air))^2], using $K_{p,\mathrm{mol}}(T) = K_p(T)/(R/P \times T)^2$, with $R = 8.314409$ [Pa m^3 mol^{-1} K^{-1}] the gas constant and $P = 101\,325$ [Pa] ($R/P_o \times T_o = 24.465$ [L mol^{-1}]).

The equilibrium dissociation constant of NH_4NO_3 is sensitive to temperature changes and varies over more than 2 orders of magnitude for typical ambient conditions. This is illustrated in Fig. 10.19 of SP2006, which is recommend to the reader for a detailed discussion. For comparison, Fig. B.2 shows the same $K_{p,AN}$ values as a function of T at RH < RHD for the E4C and ISORROPIA II applications. Although the results are similar, those of SP2006 are slightly lower since their values are obtained from a slightly different equation, i.e. $K_{p,AN}(T) = \exp(84.6 - 24\,220/T - 6.1 \times \ln(T/T_o))$ – see Eq. (10.91) of SP2006 (respectively Eq. 9.91 and Fig. 9.19 of SP1998).

Note that Reaction (B.1) applies to the gas–aerosol partitioning over dry aerosols – pure $NH_4NO_3(s)$, or any mixture of $NH_4NO_3(s)$ with other dry salt compounds. An example is given in Sect. S1.1 in the Supplement of M2016.

B.5.2 RH ≥ RHD – pure compound

For the wet case, with RH above the compound RHD or mixed solution RHD (see below), the situation is more complicated. In contrast to the gas–solid partitioning described above, the gas–liquid equilibrium partitioning of, e.g. gaseous ammonia, $NH_3(g)$ and nitric acid, $HNO_3(g)$, is in equilibrium with aqueous ammonium nitrate, $NH_4NO_3(aq)$, when the vapour pressure product of the gases exceeds its temperature and humidity-dependent equilibrium dissociation constant, $K_{p,AN}(T, \mathrm{RH})$. The salt

compound formed is – when equilibrium is reached – additionally dissociated into a cation $NH_4^+(aq)$ and anion $NO_3^-(aq)$ pair.

(i) Following SP2006 (their Sect. 10.4.3), Reaction (B.1) expands to

$$\begin{aligned} &NH_3(g) + HNO_3(g) \Longleftrightarrow NH_4^+(aq) + NO_3^-(aq) \\ &\text{EQ: } [NH_3(g)] \times [HNO_3(g)] = K_{p,AN}(T, RH) = \frac{\gamma_{AN}^2 \times \mu_{NH_4^+(aq)} \times \mu_{NO_3^-(aq)}}{K_{AN}}. \end{aligned} \tag{B.3}$$

For Reaction (B.3) the equilibrium partitioning dissociation constant is now a function of T and RH. In the notation of SP2006 (see their Eq. 10.99), $K_{p,AN}(T, RH)$ is related to the temperature-dependent equilibrium constant K_{AN}. K_{AN} $[mol^2\,kg^{-2}\,atm^{-2}]$ depends on the ion molalities $[mol\,kg^{-1}(H_2O)]$ of ammonium and nitrate, i.e. $\mu_{NH_4^+(aq)}$ and $\mu_{NO_3^-(aq)}$, and on the corresponding mean molal binary activity coefficient of aqueous ammonium nitrate, i.e. γ_{AN}^2, squared because of the cation–anion product. Solving Reaction (B.3) requires iterations. To determine the aqueous phase concentration of all compounds that can exist in solution at given T and RH requires knowledge of the total aerosol water mass (see below), which in turn depends on the solute concentrations and according to Reaction (B.3) on activity coefficients. Thus, since γ_{AN} is a function of the aqueous phase concentration, $K_{p,AN}(T, RH)$ has no analytical solution. According to the thermodynamic literature, the standard treatment is therefore quite comprehensive and requires complex thermodynamic codes.

(ii) For E4C, the product $(\gamma_{AN}^2 \times \mu_{NH_4^+(aq)} \times \mu_{NO_3^-(aq)})$ of Reaction (B.3) is expressed to be only a function of ν_i and RH, which is motivated by M2012, since their μ_s is only a function of ν_i and RH. To be able to solve Reaction (B.3) analytically, E4C parameterises $K_{p,AN}(T, RH)$ by introducing a solute-specific correction term for Reaction (B.1), which only depends on RH:

$$K_p(T, RH) = K_p(T) \times COEF(RH). \tag{B.4}$$

At given T and RH, $K_p(T, RH)$ is then a priori known, if $COEF(RH)$ is independent of the solute concentration and associated water mass. This can be achieved either by fitting data (Metzger *et al.*, 1999, 2002a; Hauglustaine *et al.*, 2014), or expressing $COEF(RH)$ in terms of the ν_i and the RH-dependent solute molality parameterisation of M2012, using their Eq. (5a). With a further comparison to Eq. (A11) of M2012, E4C uses the solute molality in terms of the solute mass fraction, χ_s, and defines $COEF(RH)$ for pure $NH_4NO_3(aq)$ solutions in terms of $\chi_s(RH)$, i.e. for RH $\geq$ RHD:

$$COEF(RH) := 2 \times \chi_s^2(RH), \tag{B.5}$$

where $COEF(RH) := 1$ for RH $<$ RHD. Equation (B.5) has been empirically derived to approximate the results obtained by ISORROPIA II (shown in M2016). $\chi_s(RH)$ denotes the solute mass fraction, which requires in the EQSAM mixed solution parameterisation framework only knowledge of RH and the solute-specific coefficient, ν_i. In accord with the dry case (Reaction B.1), aqueous solutions (Reaction B.3) can now be analogously solved. Using Eqs. (B.4) and (B.5) to obtain $K_{p,AN}(T, RH)$ at a given T and RH, the quadratic equation, which has an analytical solution for the dry case, can now also applied to pure $NH_4NO_3(aq)$ solutions. The direct solution of Reaction (B.3) by using Eqs. (B.4) and (B.5) is exemplified in the Supplement of M2016.

The T-dependent equilibrium dissociation constant of NH_4NO_3, shown in Fig. B.2, is also sensitive to changes in relative humidity and varies over orders of magnitude for typical ambient conditions. This is illustrated in Fig. 10.21 of SP2006, which is recommend to the reader for a detailed discussion. For comparison, Fig. B.3 extends Fig. B.2 showing the $K_{p,AN}(T, RH)$ values as a function of RH at $T =$ 298.15 [K] for the E4C and ISORROPIA II applications. The line points, which refer to pure $NH_4NO_{3(aq)}$ solutions, are relatively close for the E4C and ISORROPIA II results, but both are (for RH $\leq$ RHD) roughly a factor 2 higher than the corresponding values of SP2006 (see Fig. B.2); the constant $K_{p,AN}(T)$ of SP2006 is included for reference (at $T = 298.15$ [K]). Note that with Eq. (B.4) and the quadratic form of Eq. (B.5), E4C analytically approximates the solution of $K_{p,AN}(T, RH)$ for $Y = 1.0$ (Fig. B.3).

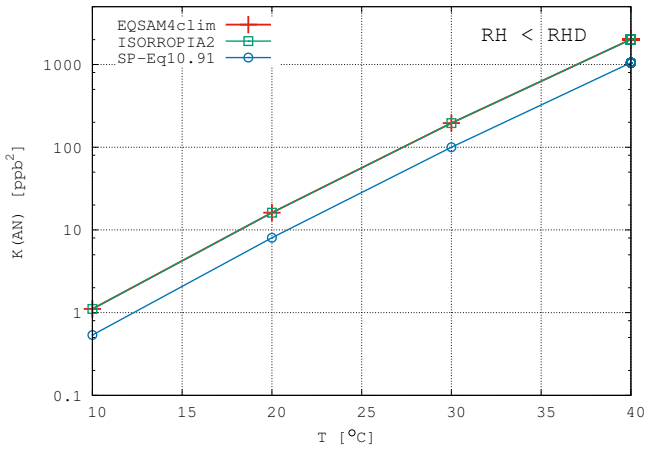


Figure B.2 NH_4NO_3 *equilibrium dissociation constant as function of T at RH < RH(AN). Red crosses show the values of E4C, green squares refer to ISORROPIA II and the blue circles show Eq. (10.91) of Seinfeld and Pandis (2006); see their Fig. 10.19.*

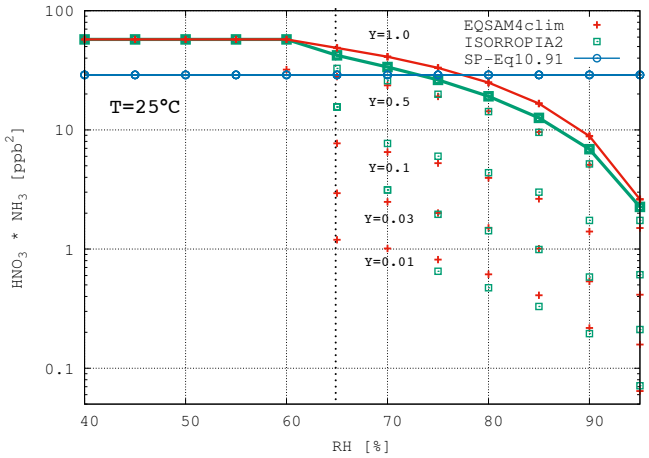


Figure B.3 NH_4NO_3 *equilibrium dissociation constant as function of RH (at* $T = 298.15$ [K]*) for various ionic strength factors* (Y) *defined in Seinfeld and Pandis (2006); see their Fig. 10.21. Red crosses show the results of E4C, green squares those of ISORROPIA II. The line points refer to pure ammonium nitrate* $(Y = 1)$*. The vertical line at* $RH = 65$ [%] *indicates the solid–liquid phase partitioning threshold for the mixed solution, shown in Seinfeld and Pandis (2006). The ordinate values refer to the product of* $[NH_3]_{(g,AN)} \times [HNO_3]_{(g,AN)}$ *that are obtained at end of the gas–liquid–solid* NH_4NO_3 *partitioning of ISORROPIA II, and by Eq. (B.11) for E4C.*

B.5.3 RH ≥ RHD – mixed compound

According to SP2006 (and references therein), Reaction (B.3) needs to be extended for mixed aqueous solutions to include an ionic strength factor.

(i) Following the notation of SP2006 (see their Eq. 10.100), the equilibrium concentration (either in [ppbv] or [mol m^{-3}(air)]) of $[NH_4NO_3]$ in mixed aqueous solutions is controlled by the presence of ammonium sulfate, $[(NH_4)_2SO_4]$, and depends on a dimensionless ionic strength factor Y, which is defined by the ratio:

$$Y := \frac{[NH_4NO_3]}{[NH_4NO_3] + 3 \times [(NH_4)_2SO_4]} \tag{B.6}$$

To extend the calculation of the T and RH-dependent equilibrium dissociation constant to the case of multicomponent aqueous solutions of NH_4NO_3, shown in Fig. B.3, Eq. (B.6) needs to be considered such that $K_{p,AN}(T, RH)$ becomes $K_{p,AN}(T, RH, Y)$.

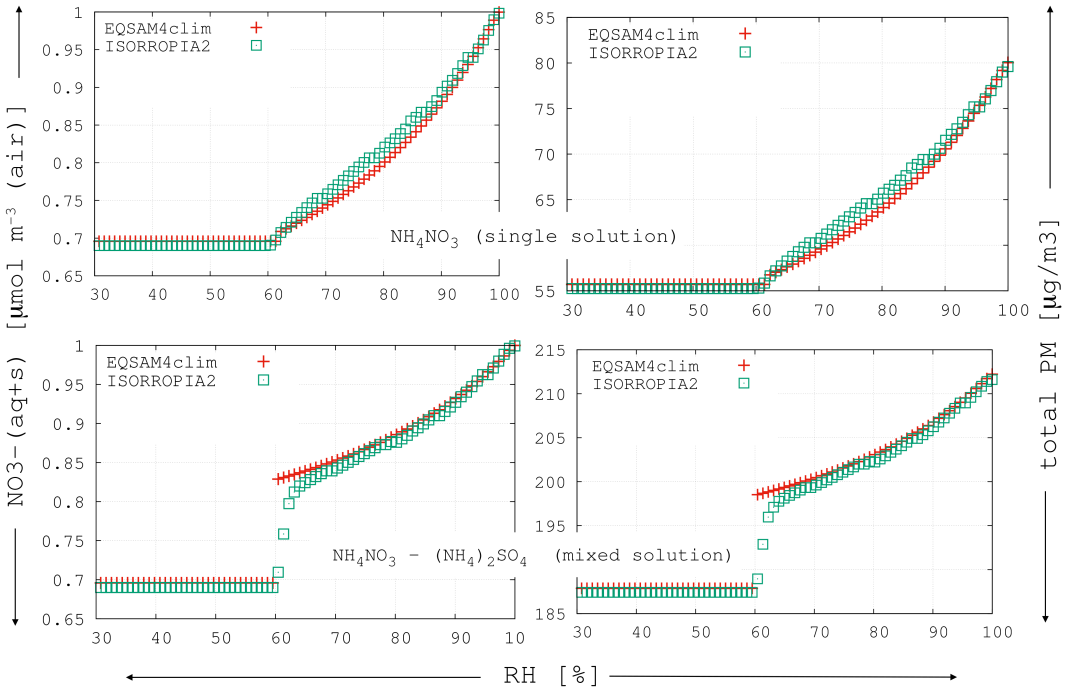


Figure B.4 *Results of E4C (red crosses) and ISORROPIA II (green squares) for two idealised gas–liquid–solid partitioning examples: single solute (binary) solution of pure* NH_4NO_3 *(upper panels) and mixed solution of* NH_4NO_3 *and* $(NH_4)_2SO_4$ *with the concentration of each compound fixed to 1* $[\mu\text{mol}\,\text{m}^{-3}(\text{air})]$ *at* $T = 298.15$ K. *The left panels show the* NH_4NO_3 *concentration in* $[\mu\text{mol}\,\text{m}^{-3}(\text{air})]$, *the right panels show the corresponding particulate mass* $[\mu\text{g}\,\text{m}^{-3}(\text{air})]$. *The mixed solution RHD described in Sect. B.7 has been neglected for E4C for this case, since this figure presents only an example for Sects. B.5 and B.6 (Sect. B.7 is considered below).*

(ii) To satisfy the key constraint (see Sect. B.1), E4C avoids iterations in its mixed solution parameterisation. E4C therefore parameterises $K_{\text{p,AN}}(T, \text{RH}, Y)$ by expanding Eq. (B.4) to be additionally a function of Y (Eq. B.6):

$$K_{\text{p}}(T, \text{RH}, Y) = K_{\text{p}}(T) \times \text{COEF}(\text{RH}, Y). \tag{B.7}$$

For Eq. (B.7), $\text{COEF}(\text{RH}, Y) := \text{COEF}(\text{RH}) \times Y^{0.8}$, where $\text{COEF}(\text{RH})$ is given by Eq. (B.5) and Y by Eq. (B.6), for which E4C uses the concentration given by the NRO (Sect. B.4), i.e. $Y := \frac{[\text{NH}_4\text{NO}_3]_{(\text{nro,max})}}{[\text{NH}_4\text{NO}_3]_{(\text{nro,max})} + 3 \times [(\text{NH}_4)_2\text{SO}_4]_{(\text{nro})}}$. The $Y^{0.8}$ term has been empirically determined to approximate the results of ISORROPIA II by keeping the initial NH_4NO_3 and $(NH_4)_2SO_4$ concentrations constant. Since E4C does not use at this computation step any dissociation constant, it obtains the maximum concentrations (NRO, max) directly from the NRO (see Sect. B.4). Equation (B.7) and the quadratic equation can be solved non-iteratively. The solution is detailed below (Sect. B.6); examples are given in Sect. S1 in the Supplement of M2016.

Figure B.3 shows that the results of E4C and ISORROPIA II exhibit a similar dependency on Y for $K_{\text{p,AN}}(T, \text{RH}, Y)$, where the values decrease with decreasing Y according to the results and the discussion of SP2006 (see their Fig. 10.21). $K_{\text{p,AN}}(T, \text{RH}, Y)$ is given here by the product of the gaseous concentrations of ammonia, $[\text{NH}_3]_{(\text{g,AN})}$, and nitric acid, $[\text{HNO}_3]_{(\text{g,AN})}$, which are in equilibrium with either the solid $[\text{NH}_4\text{NO}_{3(\text{s})}]$ concentration, if $\text{RH} < \text{RHD}$, or in equilibrium with the aqueous $[\text{NH}_4\text{NO}_{3(\text{aq})}]$ concentration when $\text{RH} \geq \text{RHD}$ in case of pure $[\text{NH}_4\text{NO}_3]$ (zero $[(\text{NH}_4)_2\text{SO}_4]$, where $Y = 1$). Below the RHD, $K_{\text{p}}(T, \text{RH}, Y)$ reduces to $K_{\text{p}}(T)$ as given by Eq. (B.2) in Reaction (B.1). Differences, which occur mainly in the mixed deliquescence humidity range, are discussed below (Sect. B.7).

B.6 Solving NH_4NO_3 / NH_4Cl-thermodynamic equilibrium

To analytically compute the equilibrium concentrations of the two semi-volatile compounds, NH_4NO_3 and NH_4Cl, for a given RH and T with the mixed solution parameterisation, E4C first solves all neutralisation reactions at once for the domain by using the NRO (Table B.7, Sect. B.4) and the totals (gas + aerosol) of the cation and anion input concentrations. Thus, E4C obtains the free ammonium $\mathrm{TA} = [\mathrm{NH_4^+}]_{(\mathrm{nro,free})}$ and nitrate $\mathrm{TN} = [\mathrm{NO_3^-}]_{(\mathrm{nro,free})}$, after all higher ranked cation–anion paris are paired. To enable a non-iterative solution, E4C does not use at this computation step any dissociation constant, so that it directly obtains from [TA] and [TN] the corresponding maximum ammonium nitrate concentration $[\mathrm{NH_4NO_{3(nro,max)}}] = \mathrm{MIN}([\mathrm{TA}], [\mathrm{TN}])$, which is possible for $K_\mathrm{p}(T, \mathrm{RH}, Y) = 1$ and the given input concentration, T and RH. Analogously, E4C computes the maximum ammonium chloride concentration from the final free ammonium and free chloride, $[\mathrm{TC}] = [\mathrm{Cl}]^-_{(\mathrm{nro,free})}$, with $[\mathrm{NH_4Cl(nro,max)}] = \mathrm{MIN}(\mathrm{TA}, \mathrm{TC})$.

With the initial (maximum) values of $[\mathrm{NH_4NO_{3(nro,max)}}]$ and $[(\mathrm{NH_4})_2\mathrm{SO_{4(nro,max)}}]$, E4C allows to solve Eq. (B.6). To obtain the final equilibrium concentrations, E4C computes the evaporative loss. For $[\mathrm{NH_4NO_{3(nro)}}]$, it computes the gaseous ammonia, $[\mathrm{NH_3}]_{(\mathrm{g,AN})}$, and nitric acid, $[\mathrm{HNO_3}]_{(\mathrm{g,AN})}$ from $[\mathrm{NH_4NO_{3,(nro,max)}}]$:

$$[\mathrm{TN}] = [\mathrm{NO_3^-}]_{(\mathrm{nro,free})}, \tag{B.8}$$

$$[\mathrm{TA}] = [\mathrm{NH_4^+}]_{(\mathrm{nro,free})}, \tag{B.9}$$

$$[X] = \frac{1}{2} \times \left(-([\mathrm{TA}] + [\mathrm{TN}]) + \sqrt{([\mathrm{TA}] + [\mathrm{TN}])^2 + 4 \times K_\mathrm{p}(T, \mathrm{RH}, Y)/(\mathrm{RT})^2} \right), \tag{B.10}$$

where the variable $[X]$ is used to obtain

$$[\mathrm{NH_3}]_{(\mathrm{g,AN})} = [\mathrm{HNO_3}]_{(\mathrm{g,AN})} = \mathrm{MIN}([\mathrm{NH_4NO_3}]_{(\mathrm{nro,max})}, [X]). \tag{B.11}$$

With Eq. (B.11), E4C computes the final gaseous concentrations of $\mathrm{HNO_{3(g)}}$ and $\mathrm{NH_{3(g)}}$ from

$$[\mathrm{NH_3}]_{(\mathrm{g})} = [\mathrm{TA}] + [\mathrm{NH_3}]_{(\mathrm{g,AN})}, \tag{B.12}$$

$$[\mathrm{HNO_3}]_{(\mathrm{g})} = [\mathrm{TN}] + [\mathrm{HNO_3}]_{(\mathrm{g,AN})}, \tag{B.13}$$

and the final ammonium nitrate equilibrium concentration from

$$[\mathrm{NH_4NO_3}]_{(\mathrm{nro})} = [\mathrm{NH_4NO_3}]_{(\mathrm{nro,max})} - [\mathrm{HNO_3}]_{(\mathrm{g,AN})}. \tag{B.14}$$

One can now solve with Eqs. (B.8)–(B.14) the quadratic equation for the dry, pure or mixed solution cases. But, in contrast to SP1998 (see their Eq. 9.103), E4C computes with Eq. (B.10) the evaporative losses of gaseous concentrations at equilibrium.

Figures B.4 and B.5 show a comparison of idealised box model calculations of E4C versus ISORROPIA II and E-AIM (web version); see M2016 for details. Despite fundamental differences in the underlying modelling approaches, the comparison of these results is satisfactory for single and mixed solute concentration from which the aerosol water mass is derived in a subsequent calculation step.

B.7 Mixed solution RHD

To calculate the liquid–solid partitioning, E4C follows (Fountoukis and Nenes, 2007) and consider a mutual deliquescence RH range. In the EQSAM framework, it depends on a minimum and maximum threshold: RHDMIN and RHDMAX, which are defined below. When the RH is below RHDMIN the aerosol is considered to be dry, while for RH above RHDMAX the aerosol is considered wet with all ionic compounds dissolved. In between a mixture can exist, with some compounds dissolved while other compounds are precipitated from the solution.

For mixed solutions (two or more compounds and water), only the amount that exists for RH > RHDMIN is considered in solution and allowed to contribute to the mixed solution water uptake.

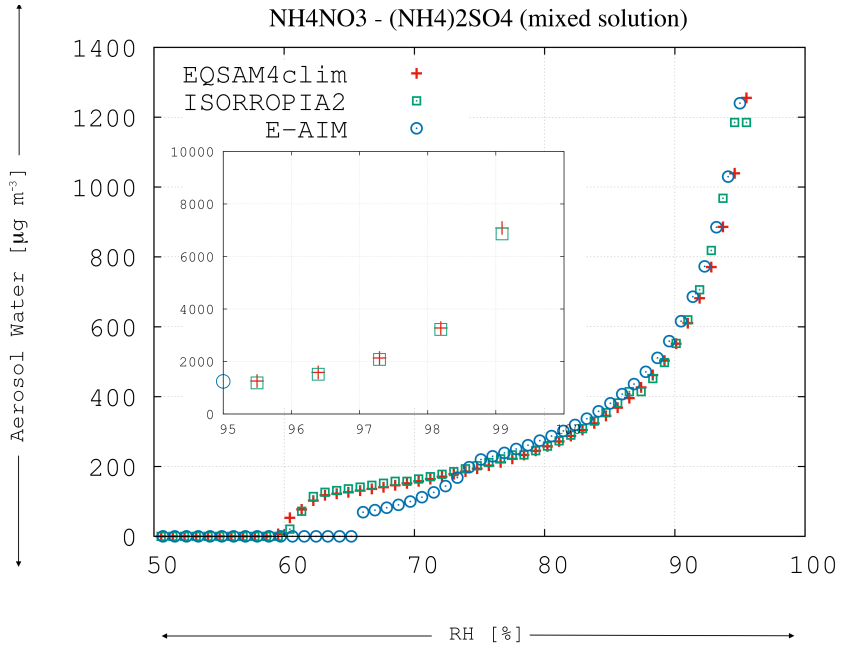


Figure B.5 *Results of E4C (red crosses) and ISORROPIA II (green squares) for the total aerosol water mass* $m_{w,mix}$ $[\mathrm{kg\,m^{-3}(air)}]$ *obtained by Eq. (B.24) for the mixed solution of* NH_4NO_3 *and* $(NH_4)_2SO_4$ *shown in Fig. B.4 (lower panels). The results are shown for* RH = 50–97 [%] *(large panels) and for the* RH = 95–99.5 [%] *(small panel). The results of E-AIM (web version) (blue circles) are included for comparison. The mixed solution RHD has been obtained for E4C from Eqs. (B.15)–(B.24) and are based on measured MDRH values for ISORROPIA II. The mutual deliquescence range of E4C and ISORROPIA II (described in Sect. B.7) differs from those of E-AIM (http://www.aim.env.uea.ac.uk/aim/aim.php). This figure is extended by Fig. S3 of M2016.*

Otherwise, the compounds are considered to be instantaneously solid and precipitated from the solution. For all non-precipitated compounds, a weighted solute concentration is computed from which subsequently all partial water masses are obtained. The sum of all partial water masses yields the total water for the given aerosol composition, size, T and RH (see Sect. B.8).

However, comparing the water uptake calculation of E4C with reference calculations of, e.g. ISORROPIA II and E-AIM is somewhat precarious. The reason is that for mixed solutions the calculated water mass mainly depends on the threshold at which the mixture is considered to take up water. The assumptions made to define the mixed solution RHD, or the mutual deliquescence RH range, are generally a major source of uncertainty in modelling the aerosol associated water uptake. First we discuss the procedure of Fountoukis and Nenes (2007), and then we describe the E4C single parameter approach.

(i) For ISORROPIA II, if the RH is within a mutual deliquescence RH range, the so-called MDRH region, the solution is assumed to be the sum of two weighted solutions; a "dry solution" (considering a pure dry case) and a "saturated liquid" solution (considering a pure liquid case). Then a numerical solution needs to be found based on a weighting factor (WF) for the dry and liquid solution that could be present in the given sub-domain (solute composition). The weighting factors are obtained from $(\mathrm{RHD} - \mathrm{RH})/(\mathrm{RHD} - \mathrm{MDRH})$ using prescribed MDRH values, which have been measured and tabulated for certain mixtures of salt solutes (see Fountoukis and Nenes, 2007). When the RH is below the MDRH, only a solid phase is possible. Otherwise a liquid and solid phase may coexist (with $\mathrm{MDRH} < \mathrm{RH} < \mathrm{RHD}$). For the latter case the aqueous phase concentrations are determined by WF. The summation over all partial water masses then yields the total aerosol water mass. But the gas–liquid–solid partitioning is reiterated until the solution converges and the concentrations do not change further (equilibrate). For details see Fountoukis and Nenes (2007).

(ii) For E4C, we follow the idea of a weighted mixed solution approach of ISORROPIA II, but approximately solve the liquid–solid partitioning by computing the weighting factor non-iteratively. E4C computes the liquid–solid partitioning after solving the NRO (Sect. B.4) and the gas–liquid partitioning (Sect. B.5). For each salt compound (j), E4C analytically obtains the solid concentration $n_{j(\mathrm{s})}$ from its aqueous $n_{j(\mathrm{nro})}$ concentration (determined in the previous computation steps). Analogously to Fountoukis and Nenes (2007), E4C uses a mixed solution weighting factor, $\mathrm{WF}_{j,\mathrm{mix}}$:

$$n_{j(\mathrm{s})} = n_{j(\mathrm{nro})} \times \mathrm{WF}_{j,\mathrm{mix}}, \tag{B.15}$$

with

$$n_{j(\mathrm{aq})} = n_{j(\mathrm{nro})} - n_{j(\mathrm{s})}, \tag{B.16}$$

where $n_{j(\mathrm{nro})}$ denotes an aqueous concentration of, e.g. $[\mathrm{NH_4NO_3}]_{(\mathrm{nro})}$ from Eq. (B.14).

$\mathrm{WF}_{j,\mathrm{mix}}$ is defined for each compound (the jth salt solute in Table B.7) by

$$\mathrm{WF}_{j,\mathrm{mix}} := (\mathrm{RHDMAX}_j - \mathrm{RH}) / (\mathrm{RHDMAX}_j - \mathrm{RHDMIN}) \tag{B.17}$$

always with a positive sign: $\mathrm{RHDMIN} < \mathrm{RHDMAX}_j$ and $\mathrm{RH} < \mathrm{RHDMAX}_j$. RHDMIN and RHDMAX_j are defined below. For $\mathrm{RH} \geq \mathrm{RHDMAX}_j$, $\mathrm{WF}_{j,\mathrm{mix}} = 0$; $\mathrm{RH} \leq \mathrm{RHDMIN}$, $\mathrm{WF}_{j,\mathrm{mix}} = 1$. Note that E4C uses a different notation of RHDMIN and RHDMAX_j (instead of the MDRH and RHD used by Fountoukis and Nenes, 2007) to indicate that E4C is using different values and underlying mixed solution calculations, which do not necessarily have to yield the same results despite the constraint that the overall liquid–solid partitioning aims to be comparable.
To solve the liquid–solid partitioning analytically, i.e. without iteration, E4C modifies the approach of Fountoukis and Nenes (2007). Each binary concentration $n_{j(\mathrm{nro})}$ is weighted by the total solute concentration, $n_{\mathrm{s,sum(nro)}} = \sum_{j=1,N_{\max}} n_{j(\mathrm{nro})}$ with all concentration units in $[\mathrm{mol\,m^{-3}(air)}]$. But in the EQSAM framework, $n_{\mathrm{s,sum(nro)}}$ is directly obtained from the sum of all single solute concentrations that are formed by solving the neutralisation reaction order (Sect. B.4). In case a semi-volatile compound has been initially neutralised, e.g. $[\mathrm{NH_4NO_3}]_{(\mathrm{nro,max})}$, E4C additionally solves Eqs. (B.8)–(B.14) to obtain $n_{j(\mathrm{nro})} = [\mathrm{NH_4NO_3}]_{(\mathrm{nro})}$ (Sect. B.5), before E4C obtains a solute-specific weighting factor, WF_j, from:

$$\mathrm{WF}_j := \frac{n_{j(\mathrm{nro})}}{n_{\mathrm{s,sum(nro)}}} = \frac{n_{j(\mathrm{nro})}}{\sum_{j=1,N_{\max}} n_{j(\mathrm{nro})}}. \tag{B.18}$$

The maximum value of $N_{\max}$ is limited by the domain-dependent NRO (see Table B.7). It refers to aqueous solutes at this stage. The liquid–solid partitioning is computed below.
The concentration-weighted maximum RHD_j (upper threshold), which normally needs to be computed iteratively for each compound, is directly obtained for E4C using WF_j:

$$\mathrm{RHDMAX}_j := \mathrm{RHDMIN} \times \mathrm{WF}_j^{0.25} + \mathrm{RHD}_j \times (1 - \mathrm{WF}_j^{0.25}). \tag{B.19}$$

In case of mixed solutions, Eq. (B.19) is used to obtain the upper RH threshold, otherwise the compound's RHD_j given in Table B.5 are used, while RHDMIN is computed directly for $n_{\mathrm{s,sum(nro)}}$ only from ν_i (see below). WF_j was introduced for E4C with an exponent (empirically derived) to parameterise the results of ISORROPIA II, which uses iterations to solve the liquid–solid partitioning using MDRH measurements as the lower RH-threshold in Eq. (B.19).
To adhere to the EQSAM key-constraints (Sect. B.1, i.e. to minimise the dependency on the required thermodynamic data), E4C computes the RHDMIN from Eq. (5b) of M2012 (with a summary given in the Appendix of M2016), by using the mixed solution values for μ_{s} and ν_i that correspond to $n_{\mathrm{s,sum(nro)}}$ (Eq. B.18). Assuming $K_{\mathrm{e}} = 1$, $A = 1$ and $B = 0$, the single RHDMIN value can be obtained from

$$\mathrm{RHDMIN} := \left(1 + \mu_{\mathrm{s}}^{\mathrm{o}} \times M_{\mathrm{w}} \times \nu_{i,\mathrm{mix}} \times \left(\frac{1}{\mu_{\mathrm{s}}^{\mathrm{o}}} \times \mu_{\mathrm{s,sat,mix}}\right)^{\nu_{i,\mathrm{mix}}}\right)^{-1}. \tag{B.20}$$

$\mu_{s,sat,mix}$ is the saturation solute molality and $\nu_{i,mix}$ the solute-specific constant of the mixed solution. $\mu_{s,sat,mix}$ and $\nu_{i,mix}$ are introduced with E4C. $\mu_s^o = 1\,[\mathrm{mol\,kg^{-1}}]$ is the reference molality to match units. $M_w\,[\mathrm{kg\,mol^{-1}}]$ is the molar mass of water.
Due to a lack of experimental data, E4C approximates $\mu_{s,sat,mix}$ from the summation over all single solute molalities, $\mu_{s,sat,single}\,[\mathrm{mol\,kg^{-1}}]$, using the relation to the mass fraction solubility (see, e.g., Eq. A11 of M2012). E4C therefore obtains $\mu_{s,sat,mix}$ from

$$\mu_{s,sat,mix} := \sum_{j=1,N_{max}} \left[\frac{1}{M_{j(single)} \times \left(100/W_{j(single)} - 1\right)} \right]. \tag{B.21}$$

$M_{j(single)}\,[\mathrm{kg\,mol^{-1}}]$ is the single solute molar mass, $W_{j(single)}$ [%] its mass fraction solubility. The data are given in Table B.5 for all compounds considered in E4C.
With the mixed solution molality, $\mu_{s,sat,mix}$, E4C can directly compute the mixed solution solubility, $w_{s,mix}$. It uses, as the corresponding total molar mass, the sum of the molar masses over all (N_{max}) compounds that can dissolve in the mixed solution, i.e. $M_{s,mix} = \sum_{j=1,N_{max}} M_{j(single)}$ (same compounds and N_{max} as in Eq. B.18):

$$w_{s,mix} := \frac{1}{(\mu_{s,sat,mix} \times M_{s,mix})^{-1} + 1}, \tag{B.22}$$

where $0.1 < w_{s,mix} < 1$. Finally, E4C obtains with Eq. (B.22) the solute-specific constant that corresponds to the mixed solution, $\nu_{i,mix}$, using an empirical equation that approximates $\nu_{i,mix}$ from the corresponding mixed solution solubility $w_{s,mix}$:

$$\nu_{i,mix} := (0.25 \times \ln(w_{s,mix}) + 1)^{-1}. \tag{B.23}$$

Thus, with Eq. (B.21), E4C solves Eq. (B.22) and with Eq. (B.22) it solves Eq. (B.23). With Eqs. (B.23) and (B.21), E4C solves Eq. (B.20) to obtain RHDMIN. Furthermore, with Eq. (B.18) it solves Eq. (B.19) to obtain RHDMAX_j. Together with RHDMIN, E4C solves Eq. (B.17) to obtain $\mathrm{WF}_{j,mix}$. $\mathrm{WF}_{j,mix}$ is then used to compute the liquid–solid partitioning from Eqs. (B.15)–(B.16) after solving the NRO (Sect. B.4) and the gas–liquid partitioning (Sect. B.5), in case of semi-volatile compounds. Finally, the aerosol water uptake is computed for each salt compound that exists in the aqueous phase at the given T and RH from $n_{j(aq)}$ (Eq. B.16). See Figure B.6 (Fig. S2.2 in the Supplement of M2016) for a flow chart of the E4C calculations (labels refers to the corresponding text of M2016).

B.8 Aerosol water uptake

To calculate the mixed solution aerosol water uptake, the standard procedure employs the widely used ZSR-mixing rule (see, e.g., SP2006, Eq. 10.98). Assuming that solute concentrations are in equilibrium with the ambient air, the total aerosol water mass, $m_{w(mix)}\,[\mathrm{kg\,m^{-3}(air)}]$, can be directly obtained from the sum of all pure compound partial aerosol water masses in the case of a mixed solution (N compounds dissolved):

$$m_{w,mix} = \sum_{j=1,N} m_{w,j} = \sum_{j=1,N} \frac{n_{j(aq)}}{\mu_{j(aq)}}. \tag{B.24}$$

Here, E4C follows the standard procedure, while the liquid–solid partitioning and the N compounds in the aqueous phase are solved non-iteratively with Sect. B.7. N can differ from N_{max} considered in Eq. (B.18), because certain salt solutes may precipitate from the mixed solution during the liquid–solid partitioning so that $N \leq N_{max}$. With increasing RH (from RHDMIN up to RHDMAX_j), an increasing number of compounds are considered for the water uptake calculations by Eq. (B.24). The partial aerosol water masses, $m_{w,j}$, which are associated with each binary solution (one compound and water), $n_{j(aq)}\,[\mathrm{mol\,m^{-3}(air)}]$, can be directly obtained from tabulated single solute molalities, $\mu_{j(aq)}$

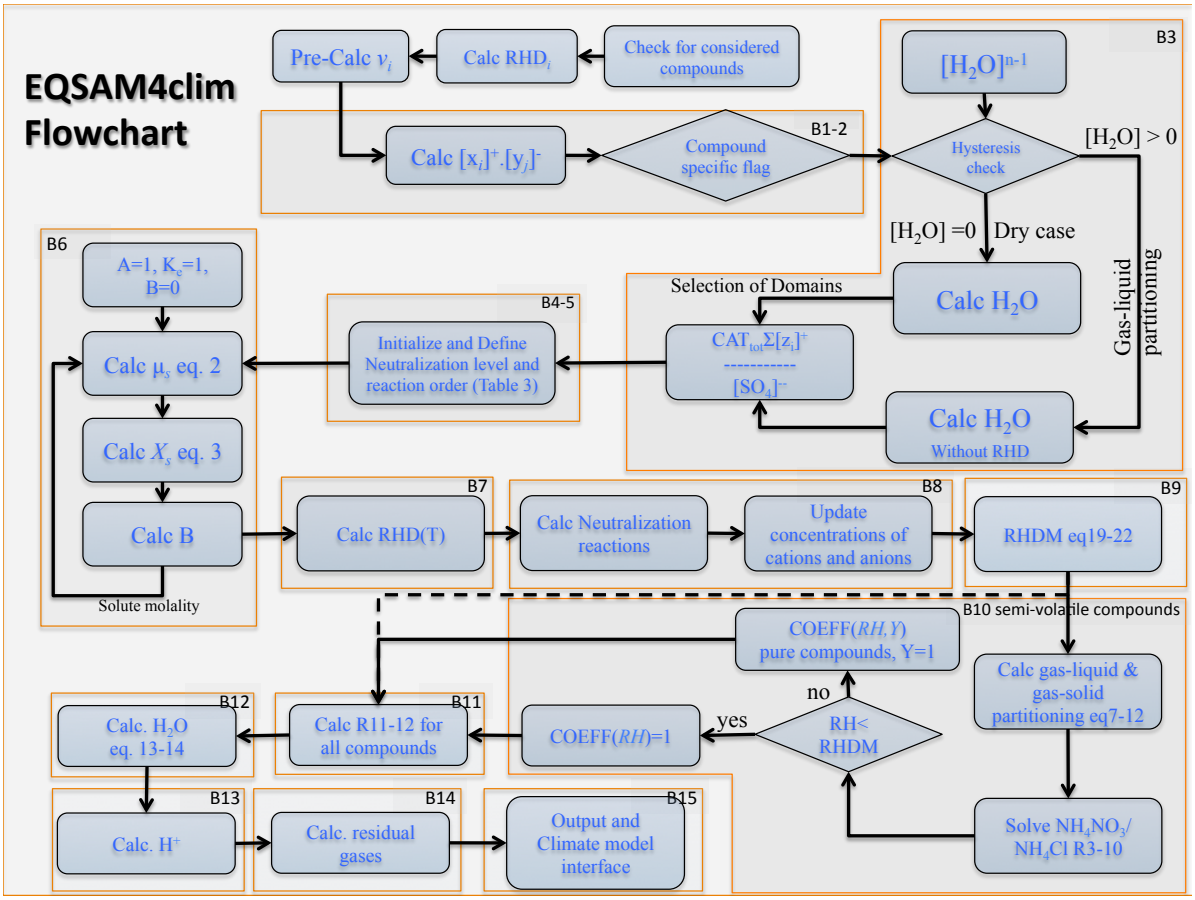


Figure B.6 *EQSAM4clim flowchart, see description of Fig. S2 of M2016.*

$[\mathrm{mol(solute)\,kg^{-1}(H_2O)}]$ (see Appendix A1 of M2016), or parameterised based on Eq. (5a) of M2012 (Appendix A2, Eq. A3 of M2016).

The parameterisation of solute molality μ_s $[\mathrm{mol(solute)\,kg^{-1}(H_2O)}]$ used by E4C is given by inverting Eq. (5a) of M2012, which allows to express μ_s as a function of RH and ν_i, with $a_w := \frac{RH}{K_e}$:

$$\mu_s = \mu_s^o \times \left(\left[\frac{1}{\mu_s^o \times M_w \times \nu_i} \times \left(\frac{K_e}{RH} - A \right) \right]^{\frac{1}{\nu_i}} - B \right) \tag{B.25}$$

$\mu_s^o = 1$ $[\mathrm{mol\,kg^{-1}}]$ denotes a reference to match units. M_w $[\mathrm{kg\,mol^{-1}}]$ is the molar mass of water and ν_i a single solute-specific constant. K_e denotes the Kelvin term (see Sect. A5) and depends on the mass equivalent hygroscopic growth factor, GF (see Sect. A6 of M2016).

The terms A and B are defined by M2012; see their Eqs. (2) and (3) and are further simplified for E4C. E4C uses a B term that has been empirically determined to be a function of ν_i with the constraint that $A := 1$. Here, B is expressed in terms of the solute mass fraction χ_s and defined as

$$B := \chi_s^{\left[\frac{1}{1+\nu_i+\chi_s}\right]}. \tag{B.26}$$

The relation between solute mass fraction χ_s and solute molality μ_s is central in the E4C mixed solution parameterisation framework. Both can be expressed through each other (see, e.g., Eq. A11 of M2012).

The solute mass fraction, χ_s [−], is defined as the mass [kg] of solute, $m_s = n_s \times M_s$, relative to the total mass [kg] of the solution composed out of the mass of solute m_s and water, $m_w = n_w \times M_w$:

$$\chi_s := \frac{m_s}{(m_s+m_w)} = \left(\frac{m_w}{m_s}+1\right)^{-1} = \left(\frac{n_w \times M_w}{n_s \times M_s}+1\right)^{-1} = \left(\frac{1}{M_s \times \mu_s}+1\right)^{-1} \quad \text{(B.27)}$$

n_s and n_w [mol] are the number of moles of solute and solvent (water), M_s and M_w [kg mol^{-1}] are the corresponding molar masses of the solute and water, respectively.

To express χ_s in Eq. (B.26), E4C uses Eq. (B.25) for μ_s (right term of Eq. B.27).

Note that the representation of water activity (M2012) relates a_w to the solute molality μ_s through a single solute-specific constant, ν_i. This is a major advantage compared to other parameterisations, because the number of unknowns is reduced to one. For more details we refer to M2012 and M2016.

In case the RH is below the T-dependent RHD or the RHDMIN, E4C assumes the compound to be dry and the partial aerosol water mass to be zero. Using the RH-dependent $\mu_{j(aq)}$ parameterisation of M2012 (their Eq. 5a), E4C can solve Eq. (B.24) without iterations.

The E4C mixed solution framework is independent of the total aerosol water mass because

(i) $n_{j(aq)}$ is independent of $m_{w,mix}$, since it is directly given by the NRO (Sect. B.4);
(ii) $K_{p,AN}(T, RH, Y)$ is independent of $m_{w,mix}$, because of the EQSAM χ_s(RH)-based parameterisations of semi-volatile compounds, i.e. Eqs. (B.4)–(B.7) (Sect. B.5)
(iii) χ_s(RH) is independent of $m_{w,mix}$, since it is directly given by $\mu_{j(aq)}$(RH) (depending also only on ν_i, and RH due to the relation Eq. A11 of M2012);
(iv) $\mu_{j(aq)}$, Eq. (B.25), is independent of $m_{w,mix}$, since it only depends on ν_i, and RH ($\mu_{j(aq)}$ is based on Eq. 5a of M2012 and included with χ_s in Appendix A2, Eq. A3 of M2016);
(v) Mixed solution RHDMIN is independent of $m_{w,mix}$, because of the mixed solution weighting factor, WF_{mix}, parameterisations, which also only depends on ν_i, and RH;
(vi) $m_{w,j}$ is independent of $m_{w,mix}$, because of the independence of points 1–5.

Finally, Eq. (B.24) is solved for diagnostic output only, since $m_{w,mix}$ does not need to be recalculated. Different from ISORROPIA II, $m_{w,j}$ and $m_{w,mix}$ are not central in the EQSAM gas–liquid–solid partitioning computations because of the ν_i concept, which requires that the thermodynamic key-properties, $\mu_{j(aq)}$ and χ_s, only depend on RH and ν_i. Note that Eq. (B.24) can be equally used for so-called metastable aerosols, for which the formation of solid salts is generally not considered.

B.9 Updates to the acidity component

E4C takes as input (i) the meteorological parameters air temperature (T) and relative humidity (RH), (ii) the aerosol precursor gases, i.e., major oxidation products of emissions from natural sources and anthropogenic air pollution represented by ammonia (NH_3), hydrochloric acid (HCl), nitric acid (HNO_3), sulphuric acid (H_2SO_4), and (iii) the ionic aerosol concentrations, i.e., lumped (both liquid and solid) anions, sulphate (SO_4^{2-}), bi-sulphate (HSO_4^-), nitrate (NO_3^-), chloride (Cl^-), and lumped (liquid+solid) cations, i.e., NH_4^+, sodium (Na^+), potassium (K^+), magnesium (Mg^{2+}) and calcium (Ca^{2+}).

The equilibrium aerosol composition and aerosol Associated Water mass (AW) is calculated by E4C through the neutralization of anions by cations, which yields numerous salt compounds, i.e., the sodium salts Na_2SO_4, $NaHSO_4$, $NaNO_3$, NaCl, the potassium salts K_2SO_4, $KHSO_4$, KNO_3, KCl, the ammonium salts $(NH_4)_2SO_4$, NH_4HSO_4, NH_4NO_3, NH_4Cl, the magnesium salts $MgSO_4$, $Mg(NO_3)_2$, $MgCl_2$, and the calcium salts $CaSO_4$, $Ca(NO_3)_2$, $CaCl_2$. All salt compounds (except $CaSO_4$) can partition between the liquid and solid aerosol phase, depending on T, RH, AW and the temperature-dependent Relative Humidities of Deliquescence of (a) single solute compound solutions (RHD) and (b) of mixed salt solutions (Metzger *et al.*, 2016).

Based on the RHD of the single solutes, the (mixed) solution liquid/solid partitioning is calculated, whereby all compounds for which the RH is below the RHD are assumed to be precipitated, such that a

solid and liquid phase can coexist. The liquid-solid partitioning is strongly influenced by mineral cations and in turn largely determines the aerosol pH (Metzger *et al.*, 2024).

E4C estimates the concentration of the hydronium ion (H^+) $[mol/m^3(air)]$ and, subsequently, the pH of the solution from electroneutrality (Z^0 $[mol/m^3(air)]$) after neutralization of all anions by all cations in the system (following the neutralization reaction order given by Table 3 of Metzger *et al.* (2016)), by using the effective hydrogen concentrations $H^{+,*}$ and Z^* that are derived from Eqs. (B.28–B.35).

Note that the auto dissociation of H_2O is taken into account, but currently no dissolution and dissociation of aerosol precursor gases such as sulphur dioxide (SO_2), nitric acid (HNO_3), hydrogen chloride (HCl), or ammonia (NH_3) is taken into account, as this is typically considered in the aqueous phase chemistry module of any global chemistry forecast model. The initial $H^{+,0}$ concentration $[mol/m^3(air)]$ after cation-anion neutralization is obtained from:

$$Z^0 = \text{tAnions} - \text{tCations} = \sum_i [Z^-]_i - \sum_j [Z^+]_j \tag{B.28}$$

$$[H^{+,0}] = Z^0 = 2[SO_4^{2-}] + [HSO_4^-] + [NO_3^-] + [Cl^-] - [K^+] - 2[Ca^{2+}] - 2[Mg^{2+}] - [Na^+] - [NH_4^+] \tag{B.29}$$

with tAnions and tCations (hereafter referred to as tCAT) representing the total absolute charge number density of all anions and cations $[mol/m^3(air)]$, respectively, that are present in the given aerosol composition. The concentration is denoted by square brackets, while Z^- and Z^+ denote the charge of tAnions and tCations, respectively, and $[H^{+,0}]$ denotes the initial hydronium ion concentration per volume air and which also depends on the auto-dissociation of water K_w $[mol^2/kg^2(H_2O)]$. This is derived from Eq. (B.30) considering the temperature dependency as widely assumed in equilibrium models.

$$K_w = 1.010 \times 10^{-14} \cdot \exp\left(-22.52 \cdot (\frac{T_0}{T} - 1) + 26.920 \cdot A_T\right) \qquad \text{where} \qquad A_T = \left(1 + \log(\frac{T_0}{T}) - \frac{T_0}{T}\right), \tag{B.30}$$

with $T_0 = 298K$.

B.9.1 Dependency of H^+ on the Chemical Domain

The neutralization equation does not correct for non-ideal solutions, such as described in Pye *et al.* (2020) and the references therein. For that purpose, with version 12, a new factor F_N was introduced with E4C, which depends on the degree of neutralization of the given aerosol composition and is used to correct the initial $[H^{+,0}]$ (Eq. B.29). F_N is obtained from:

$$F_N = [X]/[Y] \tag{B.31}$$

with X denoting the sum of all anions noted above, while Y = tNH_4, i.e., the sum of NH_3 and NH_4^+. F_N is applied without further scaling factors for ranges of $F_N < 0.9$ with ambient temperatures below 293K.

For cases outside this range ($F_N \geq 0.9$ or $T \geq 293K$), F_N needs to be scaled by 10 and multiplied by the factor K_D given in Table B.9, in order to account for chemical processes which are not resolved by the parameterizations (particularly concerning HSO_4^- and free H_2SO_4). Following Table 2 of Metzger *et al.* (2016), four chemical domains are considered to correct $[H^{+,0}]$ obtained with Eq. (B.29). No additional correction ($K_{1,2}$=1) is needed for the neutral cases (D1-D2), i.e. where cations are in excess of total SO_4^{2-}, thus preventing the formation of all HSO_4^- salts (see Table 1 of Metzger *et al.* (2016)). For the SO_4^{2-} rich case (D3), F_N and K_D from Table B.9 are multiplied, while for the SO_4^{2-} very rich case (D4), only

Table B.9 *H^+ correction factors introduced with E4C-v12 for the chemical domains introduced in Metzger et al. (2016).*

Domain	Characterization	Regime				Correction factor K_D	Relation
D1	CATION RICH	tCAT-tNH4 $\geq$ tSO_4				K_1=1	F_N
D2	SO_4^{2-} NEUTRAL	tCAT $\geq$ tSO_4				K_2=1	F_N
D3	SO_4^{2-} RICH	tCAT $\geq$ $tHSO_4$	AND	tCAT	$<$ tSO_4	K_3=1e1	F_N
D4	SO_4^{2-} VERY RICH	tCAT $\geq$ MIN	AND	tCAT	$<$ $tHSO_4$	K_4=1e3	–

a constant correction factor (K_D) is applied to correct Eq. (B.29). In Table B.9, tCAT denotes the sum of cations given in Eq. (B.29), tSO_4 is the sum of all SO_4^{2-} including HSO_4^- and H_2SO_4, while $tHSO_4$ denotes the sum of HSO_4^- and H_2SO_4.

Additionally, E4C considers three cases for estimating the H^+ concentration, according to the possible solutions of Eq. (B.28), i.e.:

$$Z^* < 0 \quad | \quad [H^{+,*}] = \frac{\mathrm{LWC}_{tot}}{10^{(7.0+\log(-Z^* \cdot \frac{10^4}{\mathrm{LWC}_o \cdot \mu_s^o}))}} \cdot \mu_s^o \tag{B.32a}$$

$$Z^* = 0 \quad | \quad [H^{+,*}] = [H^{+,neutral}] \times 10^{-6} \tag{B.32b}$$

$$Z^* > 0 \quad | \quad [H^{+,*}] = Z^* \times 10^{-6} \tag{B.32c}$$

with LWC_{tot} being the total Liquid Water Content $[\mathrm{kg}(H_2O)/\mathrm{m}^3(\mathrm{air})]$ as defined below in Eqs. (2.19a-2.19d). $\mathrm{LWC_o}$ = 1 $[\mathrm{kg}/\mathrm{m}^3(\mathrm{air})]$ and $\bar{}_s^o = 1\,[\mathrm{mol}/\mathrm{kg}(H_2O)$, a reference solution and reference molality, respectively, to match units (Metzger *et al.*, 2012, 2016; Pye *et al.*, 2020). Z^* is given by Eq. (B.33) and denotes the sum of our initial hydrogen concentration $[H^{+,0}]$ and $[H^{+,neutral}]$, an effective hydrogen concentration in a neutral solution (pH=7), which is given by Eq. (B.34), but empirically derived for the E4C parameterizations:

$$Z^* = [H^{+,neutral}] + [H^{+,0}] \tag{B.33}$$

$$[H^{+,neutral}] = \frac{B \cdot \mathrm{LWC}_o \cdot \mathrm{K}_w{}^{0.5}}{(1.0 - \mathrm{RH}^2)} \tag{B.34}$$

with K_w from Eq. (B.30), a constant $B = 1/(\bar{}_s^o \cdot \mathrm{m_w}) = 55.51\ [-]$, and the molar mass of water, $\mathrm{m_w}$ $[\mathrm{kg}/\mathrm{mol}]$; RH denotes the fractional relative humidity [0-1].

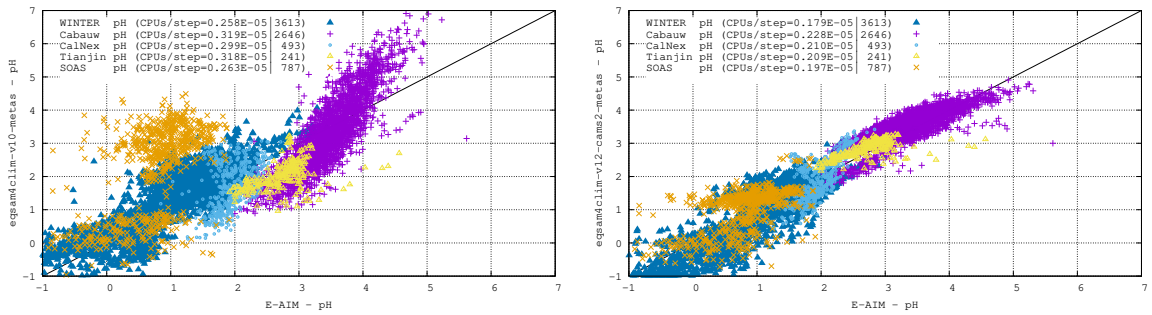
Finally, the H^+ concentration of a given solution is obtained from:

$$[H^+] = [H^{+,*}] \cdot \mathrm{F_N} \tag{B.35}$$

The cloud droplet pH is calculated from Eq. (B.35) using the cloud water content as described in Sect. B.9.2. Additionally the dependency on dissolved CO_2, methane sulphonic acid (MSA), HSO_3^- and SO_2, and the solubility and dissociation of the two main inorganic acids in the chemical scheme, HCOOH and CH_3COOH is accounted for, which, in IFS-COMPO, is coded in the scope of the aqueous phase chemistry module. These inorganic acids have pKa values of 3.75 and 4.76 respectively. Depending on the acidity of the droplet, only a fraction of the acids are dissociated.

B.9.2 Dependency of pH on the Liquid Water Content

For E4C, five different pH values can be computed from the revised H^+ $[mol/m^3(air)]$ computation, Eqs. (B.31-B.35), for diagnostic output. Therefore, E4C allows the differentiation of the various Liquid Water Content (LWC) $[kg(H_2O)/m^3(air)]$ values associated with different type of atmospheric aerosols, haze/fog, or cloud droplets as defined in Metzger *et al.* (2024) and summarized in Sec. (e).



(a) (b)

Figure B.7 *Comparison of the EQSAM4Clim pH results of v10 (panel a) and v12 (panel b) versus the pH results of E-AIM for all five cases. The CPU consumption per step is included for each case. Chip: Apple M1 Ultra; Memory: 128 GB; llvm-11/flang compiler with O3.*

Figure B.7 shows a comparison of the E4C pH results, version 10 (left), version 12 (right) versus the pH results of E-AIM for five cases described in Metzger *et al.* (2024). Clearly, the pH results of EQSAM4Clim-v12 pH are closer to E-AIM compared to the v10, now more closely following the one-by-one line for a wide range of atmospheric conditions, although some scatter still remains. Note that this scatter is acceptable for the EQSAM4Clim parameterization concept. A more explicit treatment of the phase partitioning will be subject of a follow-up study. Also note that both versions only differ by Eqs. (B.28-2.19e) with the results shown being sensitive to the Eq. (B.35) and the correction factors given in Table B.9. Finally, note that what is most important for 3D applications is the fact that version 12 introduces a refined parameterization that separates the pH of aerosol, cloud and precipitation and addresses a limitation of previous versions through Eqs. (2.19a-2.19e). For a in-depth analysis we refer to the global modelling studies of Rémy *et al.* (2024) and Williams *et al.* (2024).

Chapter 3

Emission and deposition of trace gases and aerosols

Table of contents

Emissions are the fluxes of tracer mass from the surface into the atmosphere. Emissions are caused by human activity (anthropogenic emissions), release and uptake processes from the vegetation (biogenic emissions/fluxes) and from the solid earth surface or the oceans (natural emissions/fluxes).

The production of nitrogen monoxide (NO) from lighting activity (3.1.3) that occurs throughout the atmosphere is strictly speaking a chemical conversion (oxidation) of nitrogen. These lightning NO_x emissions are still covered in this section because they are an important source for nitrogen oxides independent of other trace gases.

The removal of the trace gases at the surface is called deposition. Dry deposition occurs by contact with the surface or vegetation at the ground level. Wet deposition is caused by precipitation that removes dissolved trace gases and aerosols. The aerosol and trace gases are dissolved in cloud water or ice (rain out) or captured by rain and snow while precipitating through air (wash out). Sedimentation is the gravity driven downward motion of heavier aerosol particles.

3.1 EMISSIONS AND SURFACE FLUXES

Emission data are provided as mapped gridded data sets, often called **inventories**, which are processed as input to the IFS simulation. Emission inventories are typically averages over longer time periods such as annual means or monthly means over several years to represent seasonality. Some emissions, which strongly depend on meteorological conditions, are **simulated on-line** at every time-step in the IFS to capture their temporal variability. Currently the IFS simulates on-line the following emissions:

- Saltation of dust aerosol from dust source regions (3.1.2, (a))
- Release of sea salt aerosol over the oceans (3.1.2, (b))
- Biogenic fluxes of volatile organic carbons (3.1.4)
- Biogenic fluxes of CO_2 (3.1.5)

As the simulation of time varying emissions is often too complex to be simulated on-line, at least some aspects of the temporal variability are on-line simulated in IFS such as the application of diurnal profiles (3.1.1, (c)) to the processed inventory emissions.

3.1.1 Emissions from inventories (anthropogenic, natural, soil and ocean)

The emissions inventories used in the IFS come from a wide range of data sets and are listed in Table 3.1.

Table 3.1 *Emission inventory data sets used.*

Type	Inventory	Species	Reference
Biomass burning	GFAS 1.4.2	CO, SO_2, NO_x, CH_4, CO_2, NH_3, VOCs	Kaiser *et al.* (2012)
Anthropogenic	CAMS-GLOB-ANT M1	CO, SO_2, NO_x, CH_4, CO_2, NH_3, VOCs	Soulie *et al.* (2023), Li *et al.* (2024)
Aviation	CAMS-GLOB-AIR v1.1	CO_2, NO_x	Granier *et al.* (2022)
Natural/Ocean	CAMS-GLOB-OCE v4.1	DMS	Denier van der Gon *et al.* (2021)
Natural/Volcanic	Climatology	SO_2	Carn *et al.* (2017)
Natural/Ocean	CAMS-GLOB-SOIL v2.4	C_2H_4, C_2H_6, C_3H_6,C_3H_8, CO, NH_3	Granier *et al.* (2005)
Natural/Soil	CAMS-GLOB-SOIL v2.4	NO	Lathière *et al.* (2006)
Natural/Soil	WMO	Rn	Schery (2004)
Biogenic		DMS	Spiro *et al.* (1992)
Natural/Ocean	Jena-Carboscope	CO_2	Rödenbeck *et al.* (2013)
Natural/Termites	Climatology	CH_4	Sanderson (1996)
Natural/Wild animals	Climatology	CH_4	Houweling *et al.* (1999)
Natural/Ocean	Climatology	CH_4	Lambert and Schmidt (1993)
Natural/soil	Climatology	CH_4	Ridgwell *et al.* (1999)
Natural/wetland	Climatology	CH_4	Spahni *et al.* (2011)

Most emissions inventory data are provided on a monthly time resolution to capture the seasonal cycle. Only the GFAS-based fire emissions are provided as daily mean values. GFAS emissions are estimated by assimilating fire radiative power retrievals from satellite-based sensors (MODIS and VIIRS) and by means of land cover-dependent conversion and emission factors. The anthropogenic emissions include trends based on projections into the future, which allows the use of year-specific monthly-mean emission data. The CAMS-GLOB-ANT M1 emissions are based on the CAMS-GLOB-ANT 6.2 (Soulie *et al.*, 2023). but the emissions for North America, Europe and China are replaced by regional inventories (EPA, CAMS-REGIONAL and MEIC). For natural emissions monthly-mean climatologies are constructed for the use in most application because timely data is not available.

Out-gassing volcanic SO_2 emissions are known to be varying considerably over the years, Therefore a climatology has been constructed based on a recent satellite-based inventory of large emitters (Carn *et al.*, 2017). In this climatology the minimum of the mean value over time, or the trend over the full time series is used, to avoid over-estimating present-day emissions.

The anthropogenic emissions are given for a set of source categories (sectors), that are listed in Table 3.2. The emissions for each anthropogenic sector and the biomass burning emissions are assigned to sector-specific diurnal profile and injection methods, which are also listed in Table 3.2.

A weekly cycle in anthropogenic emissions has been introduced for a selection of sectors, based on a simplified version of country-dependent CAMS-GLOB-TEMPO weekly profiles (Guevara *et al.*, 2021). For this, the original country-dependent profiles are averaged at the world region level. For the averaging of the country-specific cycles a weighing based on the CO_2 national emissions reported by EDGARv8 was adopted. A total of 10 world regions have been defined based on a modified version of the world region definition reported by the United Nations. One of the regions includes countries adopting Sunday-Thursday as their regular workweek. Weekly profiles have been constructed for the following anthropogenic sectors:

- Road transport (tro): see figure 3.1
- Energy (ene): for the "Sunday-Thursday working week countries" region the "Europe" profile was shifted by 1 day to show the drop during Friday and Saturday.

Table 3.2 *Emission sectors and application of diurnal cycle and injection height range specification.*

Sector	Diurnal Cycle	Injection
Ships (shp)	none	30 - 100 m
Fugitives (fef)	none	surface
Power generation (ene)	sector	200 - 800 m
Off Road transportation (tnr)	none	surface
Road transportation (tro)	sector	surface
Residential and other sectors (res)	sector	surface
Industrial process (ind)	sector	20 - 300 m
Solvents (slv)	sector	surface
Agriculture livestock (agl)	sector	surface
Agriculture soils (ags)	sector	surface
Solid waste and waste water (swd)	none	20 - 100 m
Aviation	none	3D model levels
Biomass burning	sector	daily injection heights maps
Soil and ocean emissions	none	surface
Out-gassing volcanoes	none	2D Map -fixed

- Residential (res) and Domestic (dom) combustion: The same profile considered in all world regions. For the "Sunday-Thursday working week countries" region the profile is shifted to show the drop during Friday and Saturday
- Industry (ind): Same profile considered in all world regions. For the "Sunday-Thursday working week countries" region the profile is shifted to show the drop during Friday and Saturday

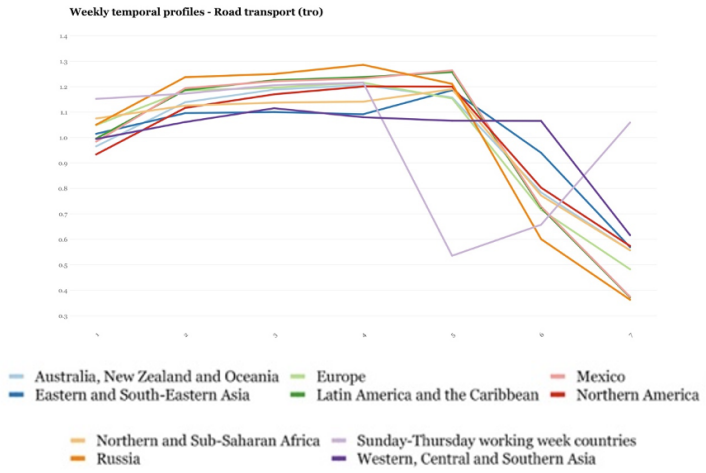


Figure 3.1 *Representation of weekly cycle for the road transport sector in different world regions*

Land based biogenic emissions for DMS with a total of approx. 0.9 Tg S yr^{-1} are applied according to Spiro *et al.* (1992).

(*a*) *Emission processing*

The emission fields read by the IFS are sector specific (3.2) daily mean fields for each considered species at the horizontal resolution of the IFS valid for the first day of the forecast. Reading the emissions fields separately for each sector and species makes it possible to apply sector-specific diurnal cycle profiles and injection procedures in the IFS. The external pre-processing of emission inventory data to compile the required IFS input field is carried out before each IFS forecast or analysis cycle. This processing consists of the following steps:

- Grib conversion of netcdf inventory files.
- Mars retrieval of daily biomass burning emissions and injection heights produced by the Global Fire Assimilation System (GFAS).
- Mapping of species provided in the inventories to the species simulated in the IFS using sector-specific pre-scaling factors (derived from tuning).
- Spatial interpolation to IFS grid.
- Temporal interpolation to forecast start day from the monthly mean values provide in the inventories (not for biomass burning emissions).
- Capping of grid-point emission values per sector below safety threshold (if applicable).
- Compilation of an emission processing name-list for the IFS.

The processing of the emissions by the IFS is configured by an emission-processing name-list. After the emission files have been read in as part of the initial conditions, the following processing steps are carried out in the IFS at every time step.

- Application of a sector-specific diurnal cycle profile with respect to local solar time.
- Application of specified scaling factors.
- Application of the injection above the surface according to a sector-specific injection height profile or
- Aggregation of sector emissions to a single surface flux field per species, that is injected as part of the turbulent tracer diffusion scheme (4.2).

For medium-range forecasts such as the operational 5-day forecasts of reactive gases and aerosols or the 10-day forecast of of greenhouse gases, the daily mean emissions for the first day are also used for all consecutive forecast days. For long-range simulation, there is the option to read in daily emission files at the start of every new simulation day (LMCC_COMPO=true).

(*b*) *Species mapping, scaling factors and emission capping*

The external emission processing infrastructure is documented here: https://confluence.ecmwf.int/display/CA/Flexible+IFS+emissions+input+for+atmospheric+composition#FlexibleIFSemissionsinputforatmosphericcomposition-Emissionspecfilesyntax

The specification of the emission processing, i.e. the applied species and sector mappings, the scaling factors and capping values, is contained in a configuration file. The operational configuration for the reactive gases and aerosol for CY50r1 are given in this file: /home/eccams/data/cifs_input/emis_spec/compo_emissions_nrt_50r1_pre.txt.

In this section, we only document some aspects of the external emission processing that leads to a modification of the emissions from the inventory data.

The emission inventories may not contain data for every single species simulated in the IFS. For example, the emissions for volatile organic compounds represented in the COMPO chemistry scheme have to be derived from the groups of volatile organic compounds (VOC) provided in the inventories for anthropogenic, biogenic and biomass burning emissions. An other example of species mapping are the emissions for hydrophilic and hydrophobic organic matter, which are derived from organic carbon emissions and an assumption about the split between the hydrophobic and hydrophilic fractions at the time of the emission release.

The emissions for species in the IFS are constructed from the species or species-groups represented in the emission inventories by using scaling factors during the pre-processing stage as well as inside the IFS. As part of the pre-processing, the emissions for certain IFS species are derived from fractions or linear combination of emissions for the species represented in the inventories. This concerns trace gases and aerosol for which no direct inventory for a given emission type is available, such as certain VOCs, HCN, CH_3CN and OM.

The scaling inside the IFS is applied to split BC and OM emissions in their hydrophilic and hydrophobic components and to approximate emissions for some species based on the CO emissions for certain sectors.

For HCN and CH_3CN we adopt fire emission fluxes based on previous global modelling studies (Singh *et al.*, 2003), in absence of direct GFAS-based estimates. In addition, a small fraction of anthropogenic emissions is included to add up to approx. 0.18 Tg N yr^{-1} for CH_3CN , and 0.15 Tg N yr^{-1} for HCN.

For acetone (CH_3COCH_3) a scaling of CO oceanic emissions (20 Tg CO yr^{-1}) to arrive at 28 Tg yr^{-1} is applied.

A further application of the scaling factors inside the IFS is to change the sign of the emission values according to the convention in the IFS. All scaling factors are applied globally in a uniform way for the respective sectors and are not suited to represent any spatial variability of the species-to-species mapping. The IFS scaling factor are configured as part of the emissions specification name-list.

The species to species mapping procedure is often uncertain and can be also be used to modify emissions for a better model performance. For example the conversion factor between OM and OC from the inventory has been optimised based on the model performance and will have to be further reviewed.

Besides the scaling of emissions, the capping emissions flux values above a predefined threshold value is applied for the emissions of OM, BC and SOG2B as part of the pre-processing to avoid excessive mass mixing ratio values. These emission caps are a simple and ad-hoc safety measure and should be reviewed, especially considering that their impact is dependent on the horizontal resolution. The scaling factors are not suited to represent the spatial variability of the species to species relation.

Table 3.3 provides a summary of scaling factors and capping values for selected species.

Table 3.3 *Scaling factors and capping values (Cap) applied to derive emissions of IFS species from the species represented in the emission inventories, selection.*

IFS species	Splitting	Conversion / Tuning	Inventory species	Sectors	Cap [kg/m/s]
SOG2B		0.1	NMVOC	anthropogenic	
OM_A	0.5	1.8	OC	anthropogenic	
OM_B	0.5	1.8	OC	anthropogenic	
BC_A	0.2	1	BC	anthropogenic	
BC_B	0.6	1	BC	anthropogenic	
OM_A	0.5	1.5	OC	Biomass burning	5×10^{-11}
OM_B	0.5	1.5	OC	Biomass burning	5×10^{-11}
BC_A	0.2	1.5	BC	Biomass burning	
BC_B	0.6	1.5	BC	Biomass burning	
CH_3COCH_3		28/20	CO	natural	
HCN		0.001	CO	ene / ind / ref / res	
CH_3CN		0.0015	CO	ene / ind / ref / res	
HCN		0.006	CO	Biomass burning	
CH_3CN		0.004	CO	Biomass burning	
NO		0.3	NO	Biomass burning - boreal region	
NH4		0.3	NH4	Biomass burning - Tropical peat burning region	

(*c*) *Application of a diurnal cycle*

The diurnal cycle profiles are formulated with respect to solar time, which is 12:00 noon at the time of the highest solar elevation. The solar time definition can deviate up to three hours form the local time because of the time-zones boundaries and daylight-saving time shifts. For the anthropogenic emission

sectors, sector specific diurnal profiles are applied. The anthropogenic diurnal profiles (CAMS-GLOB-TEMPO, (Guevara *et al.*, 2021)) were derived from European data sets but are applied globally in the IFS.

The diurnal profile for biomass burning emissions and biogenic VOC emissions (when not computed online) is a function $f(h)$ of the local time h defined by 3 parameters: a constant night time value a, the time of local daily maximum b and a measure c for the spread around the time of the local maximum.

$$f(h) = A - \frac{24(1-a)}{C * \sqrt{2\pi}} \exp\left[-0.5\left(\frac{h-b}{c}\right)^2\right] \tag{3.1}$$

For biomass burning, the parameters are the same as used as in the GFAS code for the calculation of hourly emissions values. The parameters a, b and c for different source groups are listed in 3.4.

Table 3.4 *Parameters of the diurnal profile function: a night time value, b time of local daily maximum and c spread around the time of the local maximum.*

Group	A	B	C	Species
Biomass burning	0.2	13.5	2.0	all

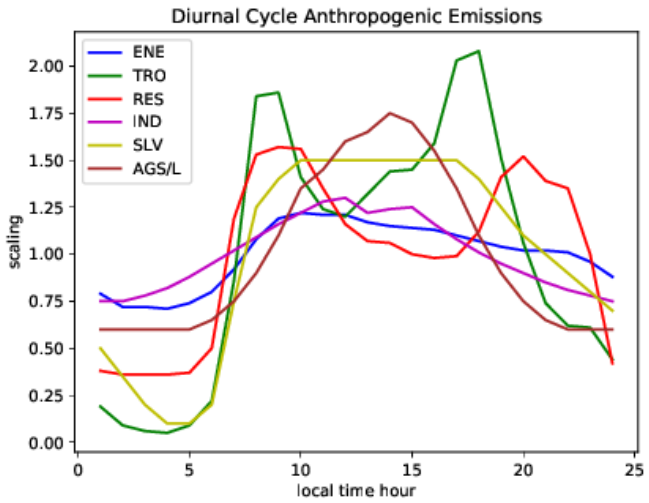


Figure 3.2 *Diurnal profile of anthropogenic emission sectors with respect to local solar time (CAMS-GLOB-TEMPO)*

(*d*) *Injection of elevated emissions*

Elevated emissions (Table 3.2) are directly converted into tendencies of mass mixing ratios (C) at each model time step. The injection is carried out uniformly w.r.t to pressure in a vertical injection range defined by a lower and upper model level $k_{\min}$ and $k_{\max}$. The resulting tendency of C at each model level k is calculated from the 2D emission flux E in the following way:

$$\frac{\partial C_k}{\partial t} = ff_k \times E \times \frac{G}{\Delta p_k} \text{ with } ff_k = \frac{\Delta p_k}{\sum_{k=k_{\min}}^{k_{\max}} \Delta p_k} \text{ for } k_{\min} \leq k \leq k_{\max} \tag{3.2}$$

The injection height profile ranges ($k_{\min} \leq k \leq k_{\max}$) are configured based on:

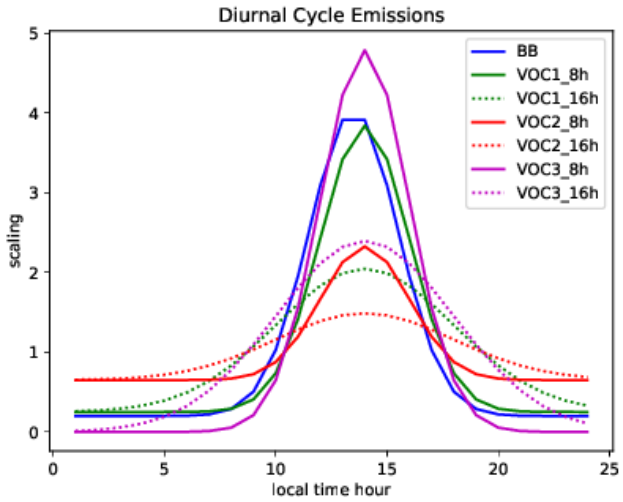


Figure 3.3 *Diurnal profile of biomass burning and biogenic VOC emissions (when not computed online) for a 8 h and 16 h daylight period.*

- Upper and lower injection heights (anthropogenic sectors, see Table 3.2) or
- Gridded map of injections heights with a specified number of model levels below and above (volcanic emission and biomass burning emissions).

Alternatively to the injection range specification, a 3D set of model level fields containing the emissions fluxes (units kg/m/s) for each model level can be read in by the IFS. This approach is used for aviation emissions of CO_2 and NO_x

Surface emissions are injected in the atmosphere by means of the turbulent diffusion scheme. The emissions are combined with the dry deposition fluxes to surface flux that is used as lower boundary condition for the vertical turbulent tracer diffusion scheme (Part IV Physical processes, Chapter 3.7).

3.1.2 On-line emissions of desert dust and sea salt

(*a*) *Emissions of desert dust*

The dust emission scheme combines the approaches of Marticorena and Bergametti (1995) for the representation of the saltation process and of Kok (2011) for the size distribution of desert dust emissions. In cycle 49R1, the dust emission scheme has been moved into the vertical diffusion scheme, so as to use the latest available wind gusts from the surface scheme. The dust emissions scheme was originally adapted from the scheme implemented in TACTIC (Michou *et al.*, 2015; Nabat *et al.*, 2012).

The emissions of dust particles of a given size D_p through sandblasting occurs if the friction velocity u^* is above a threshold value $u_t^*(D_p)$, written as

$$u_t^*(D_p) = \frac{u_{ts}^*(D_p) f_w}{f_{eff}}, \tag{3.3}$$

where $u_t^* s(D_p)$ represents an minimum threshold friction velocity and is determined according to the parameterization of Marticorena and Bergametti (1995) as a function of the Reynolds number R_e as

$$u_{ts}^*(D_p) = 0.129 \times K \times \begin{cases} [1 - 0.858 \times exp(-0.0617(R_e - 10))] & R_e > 10 \\ \left(1.928 \times R_e^{0.092} - 1\right)^{-0.5} & R_e \leq 10 \end{cases} \tag{3.4}$$

Here, the Reynolds number R_e is parameterized following (Marticorena and Bergametti, 1995) as

$$R_e = 1331.647 \times D_p^{1.561228} + 0.38194 \tag{3.5}$$

and

$$K = \sqrt{\frac{2 \times g \times \rho_p \times D_p}{\rho_a} \times \left[1 + \frac{0.006}{\rho_g \times g \times (2 \times D_p)^{2.5}}\right]} \tag{3.6}$$

Where D_p is the particle diameter in cm, ρ_p is the dust aggregate density taken as 2.6 kg/m^3, ρ_a is the surface air density and g the gravitational constant. It should be noted that the 1331.647 factor has a unit of $cm^{-1.561228}$ to make sure that R_e is dimensionless. Similarly, the 0.006 factor has a unit of $g/cm^{0.5}s^{-2}$. All other factors are unitless. The term f_{eff} is a correction factor accounting for the effect of surface roughness, expressed as:

$$f_{eff} = 1 - \left[\frac{ln(\frac{z_m}{z_{0s}})}{ln(0.35(\frac{10}{z_{0s}})^{0.8})}\right] \tag{3.7}$$

Where z_m is the aerodynamic roughness length and z_{0s} is the roughness length of smooth erodible surfaces, both in m. Finally, f_w accounts for the effect of soil moisture content on the threshold friction velocity. Following Fecan *et al.* (1999), it is parameterized as:

$$f_w = \begin{cases} [1 + 1.21 \times (w - w')^{0.68}]^{0.5} & \text{for } w > w' \\ 1 & \text{for} w < w' \end{cases} \tag{3.8}$$

Where w is the surface soil moisture, provided by the IFS surface scheme, and w' is a threshold gravimetric water content of the top soil layer above which w increases the threshold friction velocity u_t^*, expressed as:

$$w' = 0.0014 \times (\%clay)^2 + 0.17 \times (\%clay) \tag{3.9}$$

Where %clay is the fraction of soil that is composed of clay. Both w and w' are in percent. The information on the clay, silt and sand fraction is provided externally by the Global Soil Data set for use in Earth system models (GSDE, (Shangguan *et al.*, 2014). The horizontal flux of dust from saltation is expressed as:

$$G(D_p) = E_{soil} \times \frac{\rho_a}{g} \times S_{rel}(D_p) \times (u^*)^3 * \left(1 - \left(\frac{u_t^*(D_p)}{u^*}\right)^2\right) \times \left(1 + \frac{u_t^*(D_p)}{u^*}\right) \tag{3.10}$$

Where E_{soil} is the soil "erodibility" and S_{rel} is the ratio of the surface of the dust aggregate of diameter D_p over the sum of the surface of aggregates of all diameters. Both are unitless. The soil erodibility can be defined as the soil erosion efficiency of a surface under a given meteorological forcing (Zender *et al.*, 2003). It is also often denoted as "dust source function". Because soil erodibility is hard to estimate, several methods have been tested in dust emission schemes, one of the most commonly used is the topographic approach from Ginoux *et al.* (2001), which assumes that the topographic depressions are the largest source of dust. In cycle 47R1, the soil erodibility/dust source function (DSF) is provided empirically by a climatological data set of the frequency of occurrence of dust AOD > 0.4, as provided by Paul Ginoux and introduced in Ginoux *et al.* (2012). In cycle 46R1, the climatological frequency of dust AOD > 0.2 was used for the DSF, which led to an overestimation of simulated dust AOD. In cycle 48R1, the DSF has been recomputed by comparing daily dust AOD from a simulation and provided by the MIDAS product (Gkikas *et al.*, 2020) over a three years period. Using a longer period of time to recompute the DFS allowed for the provision of a monthly DSF, instead of a fixed (yearly) one used from cycles 46R1 to 47R3. The DSF for January and May is shown in Figure 3.4. There are significant differences between the two, such as higher values over the Sahel in January, possibly partly over cultivated areas (Ginoux *et al.*, 2012), and over Australia, and higher values over the Taklamakan in May.

The friction velocity u^* used in equation 3.10 is computed using the sum of the wind gusts are provided by the surface scheme and from deep convection.

$$u^* = (WG + WG_{conv}) * \frac{k}{log(\frac{10}{z_0})} \tag{3.11}$$

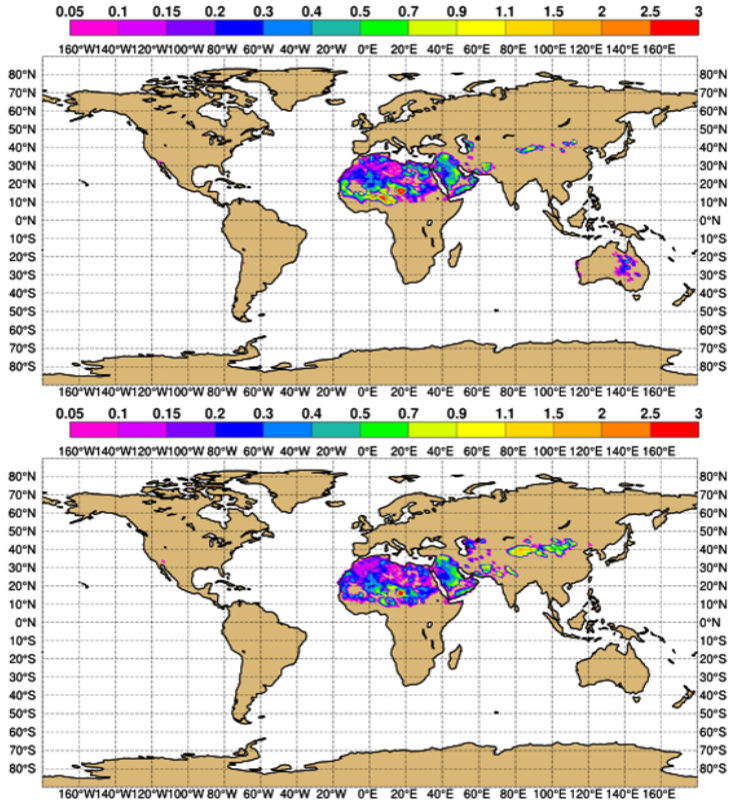


Figure 3.4 *Dust source function (DSF) used for January (top) and May (bottom)*

Where WG are the wind gusts provided by the surface scheme, WG_{conv} the wind gusts from deep convection, k is the Von Karman constant and z_0 is the roughness length.

Finally, the flux of vertically emitted dust is computed from the horizontal flux using Gilette (1979):

$$H(D_p) = G(D_p) \times F_{bare} \times C \times \begin{cases} 10^{0.134\times(\%clay)-6.0)} & \text{for } \%clay \leq 17\% \\ 10^{-0.09\times(\%clay)-2.19)} & \text{for } \%clay > 17\% \end{cases} \tag{3.12}$$

where F_{bare} is the fraction of the soil that is bare; C is a normalization constant set to 0.034, nearly similar to the value used in Nabat *et al.* (2012) who used 0.035. This formula is integrated for all particle diameters D_p and provides the total flux of emitted dust. In order to distribute this flux into the three bins, the size distribution at emissions of Kok (2011) is used up to cycle 49R2 included, which means a much larger share of emissions being distributed to the super coarse bin as compared to the Ginoux *et al.* (2001) scheme used operationally before cycle 46R1. In cycle 50R1, the assumed size distribution of Meng *et al.* (2022) (an updated version of the Kok (2011) brittle fragmentation theory) is used, with slightly more finer particles as compared to Kok (2011), but still much less than the original dust emission scheme in place before cycle 46R1. This is illustrated by Table 3.5, and as a consequence the simulated lifetime of total dust is significantly lower with the new dust emission scheme implemented in cycle 46R1 as compared to the old scheme, because the super coarse dust bin has a much shorter lifetime from increased dry deposition and sedimentation. The reduced emissions in cycle 49R1 are a consequence of using wind gusts directly from the surface scheme, while before that they were computed specifically for the purpose of dust emissions and were on average higher. The higher emissions of bins1 and 2 in cycle 50R1 are a consequence of the change in assumed size distribution at emissions.

Activated if the switch LAERDUSTSIZEVAR the distribution of the total dust emissions into the three bins is modulated regionally using an external file, which provides the fraction of total emissions that goes into bin1 and bin2. This fraction has been computed using long simulations of dust mineralogy (the

dust mineralogical species being emitted preferentially to bin1 and 2 or bin3). This results on average on slightly more emissions being distributed to bin1 and 2 as compared to cycle 47R3.

Table 3.5 shows the impact of the new optical properties and of the new DSF implemented in cycle 48R1: less extincting dust leads to higher dust emissions (for a global dust AOD roughly unchanged) and burden in cycle 48R1 as compared to cycle 47R3. Also, the distribution of the total dust emissions between the three bins is slightly changed in cycle 48R1, with relatively slightly more emissions to bins 1 and 2 as compared to bin 3, coming from the regional modulation of the dust size distribution at emissions implemented in cycle 48R1.

Table 3.5 *Desert dust emissions, burden and lifetime for 2017 simulated by IFS-AER cycles 45R1, 47R1, 47R3, 48R1, 49R1/R2 and 50R1 (forecast only). The emissions are in Tg yr^{-1}, the burdens are in Tg and the lifetimes are in days.*

process	bin1 (0.05 - 0.55 µm)	bin2 (0.55 - 0.9 µm)	bin3 (0.9 - 20 µm)	total
Emissions (45R1)	87.9	292	2054.9	2434.8
Burden (45R1)	1.7	5.9	8.5	16.1
Lifetime (45R1)	7.0	7.2	1.5	2.4
Emissions (47R1)	4.9	45.2	3248.5	3298
Burden (47R1)	0.12	1.0	13.5	14.6
Lifetime (47R1)	8.9	8.1	1.5	1.6
Emissions (47R3)	5.1	47.8	3456.5	3509.4
Burden (47R3)	0.12	1.1	16.4	17.62
Lifetime (47R3)	8.6	8.4	1.7	1.8
Emissions (48R1)	11.1	103.3	5537.1	5651.4
Burden (48R1)	0.26	2.4	27.2	29.86
Lifetime (48R1)	8.5	8.5	1.8	1.9
Emissions (49R1/R2)	8.35	76.7	4186.5	4271.5
Burden (49R1/R2)	0.21	1.82	21.3	23.33
Lifetime (49R1/R2)	9.2	8.7	1.9	2.0
Emissions (50R1)	12.4	127	4179.2	4318.6
Burden (50R1)	0.29	2.76	21.3	24.35
Lifetime (50R1)	8.5	7.9	1.9	2.1

(*b*) *Emissions of sea salt*

In addition to the M86 (Monahan *et al.*, 1986) and the G14 (Grythe *et al.*, 2014) sea salt aerosol emission schemes used in previous cycles, a new sea salt emission scheme "A16" based on Albert *et al.* (2016) has been developed and is used operationally since cycle 47R1. It is similar to the M86 scheme in the sense that as a prerequisite, the oceanic whitecap fraction is first estimated; in the M86 scheme this is done following the work of Monahan and Muircheartaigh (1980). In the A16 scheme, this is done by a statistical fit between a dataset of one year of whitecap fraction estimated from remote sensing observations of ocean surface brightness by radiometers onboard the WindSat satellite, at two frequencies: 10 and 37 GHz (Anguelova and Webster, 2006), and 10m wind speed provided by Quickscat as well as sea-surface temperature provided by ERA interim. The whitecap fraction W is expressed as a function of 10m windspeed U_{10} and SST by :

$$W = a(SST)\,[U_{10} + b(SST)]^2 \tag{3.13}$$

where

$$a(SST) = a_0 + a_1 SST + a_2 SST^2 \tag{3.14}$$

$$b(SST) = b_0 + b_1 SST \tag{3.15}$$

The $a_{0,1,2}$ and $b_{0,1}$ parameters are given in Albert *et al.* (2016) for the whitecap fraction estimated with WindSat 10 and 37 GHz brightness temperature. As the coverage of the retrieved whitecap fraction data set is very good, the sample size is very large, which makes the fit quite robust. In the IFS-AER

implementation of this scheme, using the fit to whitecap from 37 GHz brightness temperature gave better results, and the $a_{0,1,2}$ and $b_{0,1}$ parameters for this wavelength were chosen.

Using the oceanic whitecap fraction as an input, the production flux of sea salt aerosol is then computed by the following formula from Monahan *et al.* (1986) as extended by Gong (2003):

The modified form is:

$$\frac{dF}{dD_p} = W \times 3.5755 \times 10^5 D_p^{-A}(1 + 0.057 D_p^{3.45}) \times 10^{1.607\exp(-B^2)} \tag{3.16}$$

where

$$\begin{aligned} A &= 4.7(1 + \theta D_p)^{-0.017 D_p^{-1.44}} \\ B &= \frac{0.433 - \log(D_p)}{0.433} \end{aligned} \tag{3.17}$$

Where D_p is the particle diameter. Here, θ is an adjustable parameter, which (Gong, 2003) introduced to fit the sea-spray shape function values to field observations of O'Dowd *et al.* (1997). It provides additional control of the shape of the sub-micron size distribution of the generated sea spray. Following the original parameterization of Gong (2003), we set θ = 30.

Up to cycle 48R1 included, the A16 scheme was used together with the Monahan *et al.* (1986) formulation to derive sea-salt aerosol emissions from whitecap fraction, while now, the formulation of Gong (2003) is used instead, which is valid for the desired size range.

Table 3.6 shows the simulated emissions, burden and lifetime of the three sea salt bins for the three available emission schemes. The lifetime of sea-salt aerosol decreases for larger particles, because sedimentation is a more effective sink for larger particles, and because the simulated dry deposition velocity increases with particle size for particles above 1 micron diameter. The emissions of super coarse sea salt aerosol are much higher with the G14 scheme as compared to the two others. Similar to the M86 scheme, the A16 scheme shows a relatively smaller increase in emissions with bin size. The lifetime of coarse and super coarse sea salt bins is the lowest with the A16 scheme. The M86 scheme has been used operationally until cycle 43R3. The G14 scheme has been used operationally in cycles 45R1 and 46R1, while the new A16 scheme has been implemented in operational CY47R1 IFS-AER. More detail on the A16 scheme and the use of the Gong (2003) sea-spray shape function can be found in (Remy and Anguelova, 2021). In cycle 50R1, several bugs were found impacting the sedimentation of sea-salt aerosols, which explain the significantly higher burden and lifetime with this cycle, particularly for bins 2 (lifetime doubled) and 3 (lifetime multiplied by 5.5).

3.1.3 Lightning as source of Nitrogen Oxides

NO production from lightning is a considerable contribution to the global atmospheric NOx budget. Estimates of the global annual source vary between 2 and 8 Tg N yr^{-1} (Schumann and Huntrieser, 2007). 5 Tg N yr^{-1} is the most commonly assumed value for global CTMs, which is about 6–7 times the value of NO emissions from aircraft (Gauss *et al.*, 2006), or about one fifth of the total anthropogenic NO emissions. NO emissions from lightning play an important role in the chemistry of the atmosphere because they are released in the rather clean air of the free troposphere. The parameterisation of the lightning NO production in IFS consists of estimates of

- the flash rate density,
- the flash energy release and
- the vertical NO emission profile for each model grid column.

The IFS includes several paramaterisations to diagnose the flash rate density using input parameters from the convective scheme. For the simulation of the lightning NO emissions, the flash-rate density is estimated to be proportional to the convective rain-flux at the surface (Meijer *et al.*, 2001) using a conversion factor of 1728 over land and 172.8 over ocean. Further, lightning occurrences are only

Table 3.6 *Dry sea salt aerosol emissions, burden and lifetime simulated by IFS-AER without data assimilation with the M86, G14, A16 (with the M86 sea-spray shape function) and A16E (with the G03 sea-spray shape function) schemes. Cycle 50R1 is also included. The emissions are in Tg yr^{-1}, the burdens are in Tg and the lifetimes are in days.*

process	bin1 (0.05 - 0.5 µm)	bin2 (0.5 - 5 µm)	bin3 (5 - 20 µm)	total
Emissions (up to 43R3,M86)	32.2	2767.2	3363.8	6163.2
Burden (up to 43R3,M86)	0.09	3.53	1.43	5.05
Lifetime (up to 43R3,M86)	1.0	0.46	0.16	0.29
Emissions (45R1-46R1,G14)	41.6	1799.5	45531.6	47372.7
Burden (45R1-46R1,G14)	0.14	2.86	22.5	25.5
Lifetime (45R1-46R1,G14)	1.3	0.58	0.18	0.2
Emissions (47R1-48R1,A16)	110.3	6595.5	13657.8	20363.6
Burden (47R1-48R1,A16)	0.39	4.46	1.41	6.2
Lifetime (47R1-48R1,A16)	1.3	0.25	0.04	0.11
Emissions (from 49R1,A16E)	197.8	2444.6	9190.3	11832.7
Burden (from 49R1,A16E)	0.7	3.45	1.87	6.02
Lifetime (from 49R1,A16E)	1.29	0.52	0.07	0.19
Emissions (from 50R1,A16E)	84.5	2238.4	8254.9	10577.8
Burden (from 50R1,A16E)	0.34	6.83	9.2	16.4
Lifetime (from 50R1,A16E)	1.48	1.11	0.4	0.57

computed if the cloud base height is below 4 km, the cloud top height above 5 km and the temperature profile reaches values below -25°C. The energy produced by each flash and the associated NO release is calculated using a global conversion factor for the energy and the released number of NO molecules per energy unit. The combined factor is further multiplied with an empirical tuning factor to ensure that the total annual lightning emissions are equal to 5 Tg N. The tuning factor is resolution dependent and compensates modifications to the convective activity introduced by upgrades of the model physics. It is determined by a 1-year test simulation. The vertical distribution of the NO lightning emissions adopts a "backward C-shape" profile (Ott *et al.*, 2010), which locates most of the emission in the middle of the troposphere (Fig. 3.5) As lightning NO emissions occur mostly in situations with strong convective transport, tests showed that differences in the injection profile had little impact.

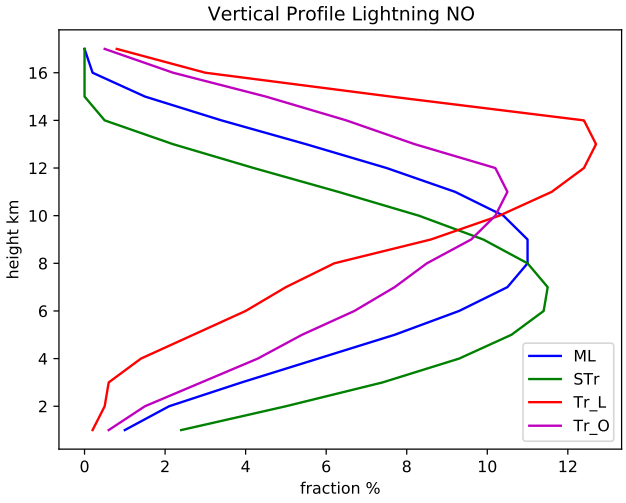


Figure 3.5 *Vertical Profile of the Lightning NO production (fraction in % for Mid-latitudes (ML), Subtropics (STr), Tropics over land (Tr_L) and Tropics over Oceans (Tr_O).*

3.1.4 Biogenic fluxes of volatile organic carbons

The modelling of biogenic volatile organic carbons (BVOC) emissions in IFS-COMPO is done through a version of the Model of Emission of Gases and Aerosols from Nature (MEGAN, (Guenther *et al.*, 2006, 2012), here referred to as G06 and G12). In the atmospheric composition modelling community the MEGAN module is amongst the most popular approaches to simulate BVOC emissions, although different research groups apply variations of the basic parameters depending on model specifics. Here we provide here further details on the actual implementation of MEGAN in IFS-COMPO. In MEGAN the effective emission of individual trace gases depends on a reference emission factor (EF) for a specific plant functional type (PFT) under standard conditions which is modulated depending on environmental conditions:

$$E_j = \gamma_j \sum_i^{PFT} f_{PFTi} \epsilon_{ij} \tag{3.18}$$

Here γ is a dimensionless activity factor, f is a fraction of a grid box covered by PFT_i and ϵ_{ij} is the emission factor which represents the total emissions under standard conditions in [μg m^{-2} h^{-1}]. The values for ϵ_{ij} used here are based on MEGAN PFTs as specified in G12, except for α-pinene, which use a factor 2 lower values than reported there (Messina *et al.*, 2016). This implementation is referred to as IFS-MEGAN, and is based on the G06 and G12 parameterizations. The net activity factor is computed based on an emission response to light, temperature, leaf area index, leaf age and CO_2 inhibition, essentially following G06 with a Parameterized Canopy Environment Emission Activity (PCEEA) algorithm. Specifically, the net activity factor is computed based on an emission response to light (γ_P), temperature (γ_T), leaf area index (γ_{LAI}), leaf age (γ_A) and CO_2 inhibition (γ_C) following:

$$\gamma = \gamma_P \gamma_T \gamma_{LAI} \gamma_A \gamma_C \tag{3.19}$$

The BVOC emissions parameterization is implemented as a separate module as part of ECLand.

(*a*) *Vegetation description and emission factors*

The MEGAN PFT categories are given in Table 3.7, for which a set of ϵ_{ij} is defined. These PFT categories are different from the ones that are available in ECLand, which implies that a set of conversions is required. A set of Biosphere-Atmosphere Transfer Scheme (BATS) classes are defined in ECLand, which are associated with two vegetation tiles: "high vegetation" and "low vegetation", as well as a wet skin vegetation tile. The vegetation types in the BATS classes are converted to MEGAN-type PFTs following Table 3.8. Here a very basic set of rules was also defined. Basically, the Köppen-Geiger classification maps, which govern the climate zone, have been represented by very basic latitudinal bands. This implies particularly that both longitudinal variations in vegetation types, and vegetation in high-altitude (mountainous) regions are currently not represented.

This approach is followed for all BVOC emissions except for isoprene (C_5H_8). The corresponding emission factors for the MEGAN PFT categories are given in Table 3.9. For C_5H_8 instead of using an emission factor per land use category we make use of a gridded map of emission potential data. This means that the equation 3.18 reduces to:

$$E_{C_5H_8} = \gamma_{C_5H_8} EP_{C_5H_8} f \tag{3.20}$$

Here $EP_{C_5H_8}$ refers to the emission potential for isoprene and f the gridbox fraction covered by vegetation. Here we use a gridded dataset which is the same as the one used in CAMS-GLOB-BIO v3.1 (Sindelarova *et al.*, 2022), except that also for Europe we adopt a constant value throughout the year. An additional scaling factor of 1.5 is needed to match the regional total isoprene emissions as simulated in IFS-COMPO to those from CAMS-GLOB-BIO v3.1.

Table 3.7 *MEGAN-style plant functional types*

Index	MEGAN PFT Category
1	Needleleaf evergreen temperate tree
2	Needleleaf deciduous boreal tree
3	Needleleaf evergreen boreal tree
4	Broadleaf evergreen tropical tree
5	Broadleaf evergreen temperate tree
6	Broadleaf deciduous tropical tree
7	Broadleaf deciduous temperate tree
8	Broadleaf deciduous boreal tree
9	Broadleaf evergreen temperate shrub
10	Broadleaf deciduous temperate shrub
11	Broadleaf deciduous boreal shrub
12	Cold C3 grass
13	Cool C3 grass
14	Warm C3 grass
15	Other crops
16	Corn

Table 3.8 *Matching between ECLand vegetation types and MEGAN plant functional type (PFT) categories.*

ECLand Vegetation Category	**MEGAN PFT Category**
1 Crops, Mixed Farming (L)	15 (Other crops)
2 Short Grass (L)	LAT $>$ 60N: 12 (cold grass) 30N $<$ LAT $<$ 60N or LAT $>$ 30S: 13 (cool grass) -30 $<$ LAT $<$ 30: 14 (warm grass)
3 Evergreen Needleleaf Trees (H)	LAT $<$ 50N: 1 (needleleaf evergreen temperate) LAT $>$ 50N: 3 (evergreen boreal tree)
4 Deciduous Needleleaf Tree (H)	2 (Needleleaf deciduous boreal tree)
5 Deciduous Broadleaf Trees (H)	30S $<$ LAT $<$ 30N: 6 (Broadleaf evergreen / deciduous tropical) 30N $<$ LAT $<$ 50N or LAT $>$ 30S: 7 (broadleaf deciduous temperate) LAT $>$ 50N: 8 (broadleaf deciduous boreal)
6 Evergreen Broadleaf Trees (H)	30S $<$ LAT $<$ 30N: 4 (Broadleaf evergreen tropical) Else: 5 (broadleaf temperate tree)
7 Mixed Crops (L)	15 (other crops)
9 Tundra (L)	12 (cold grass)
10 Irrigated Crops (L)	15 (other crops)
11 Semidesert (L)	13 (cool C3 grass)
13 Bogs and Marshes (L)	13 (cool C3 grass)
16 Evergreen Shrubs (L)	9 (broadleaf evergreen temperate shrub)
17 Deciduous Shrubs (L)	LAT $<$ 50N: 10 (broadleaf deciduous temperate) LAT $>$ 50N: 11 (boreal shrub)
18 Broad-Savanna	7 (broadleaf deciduous temperate tree)
19 Interrupted Forest (H)	7 (broadleaf deciduous temperate tree)
20 Water and Land Mixtures (L)	N/A

(*b*) *Parameterisation of activation factors*

The net activity factor is computed based on equation 3.19, essentially following G06 with a Parameterized Canopy Environment Emission Activity (PCEEA) algorithm. The parameterization of the light response to emissions (γ_P) is simulated as:

Table 3.9 *List of modelled compound classes and emission factors (µg m-2 s-1) for each of the plant functional type defined in Table 3.7*

Compound Class	EF1	EF2	EF3	EF4	EF5	EF6	EF7	EF8	EF9	EF10	EF11	EF12	EF13	EF14	EF15	EF16
Isoprene	600	1	3000	7000	10000	7000	10000	11000	2000	4000	4000	1600	800	200	1	1
Myrcene	70	60	70	80	30	80	30	30	30	50	30	0.3	0.3	0.3	0.3	0.3
Sabinene	70	40	70	80	50	80	50	50	50	70	50	0.7	0.7	0.7	0.7	0.7
Limonene	100	130	100	80	80	80	80	80	60	100	60	0.7	0.7	0.7	0.7	0.7
3-D-Carene	160	80	160	40	30	40	30	30	30	100	30	0.3	0.3	0.3	0.3	0.3
trans-b-ocimene	70	60	70	150	120	150	120	120	90	150	90	2	2	2	2	2
b-pinene	300	200	300	120	130	120	130	130	100	150	100	1.5	1.5	1.5	1.5	1.5
a-pinene	373	698	373	386	380	386	204	259	200	300	200	2	2	2	2	2
Other Monoterpenes	180	170	180	150	150	150	150	150	110	200	110	5	5	5	5	5
a-Farnesene	40	40	40	60	40	60	40	40	40	40	40	3	3	3	4	4
b-Caryophyllene	80	80	80	60	40	60	40	40	50	50	50	1	1	1	2	4
Other Sesqiterpenes	120	120	120	120	100	120	100	100	100	100	100	2	2	2	2	2
232-MBO	700	0.01	60	0.01	0.01	0.01	0.01	2	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Methanol	900	900	900	500	900	500	900	900	900	900	900	500	500	500	900	900
Acetone	240	240	240	240	240	240	240	240	240	240	240	80	80	80	80	80
CO	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600
Bidirectional VOC	500	500	500	500	500	500	500	500	500	500	500	80	80	80	80	80
Stress VOC	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300
Other VOC	140	140	140	140	140	140	140	140	140	140	140	140	140	140	140	140

$$\gamma_P = (1 - \mathrm{LDF}_j) + \mathrm{LDF}_j \cdot \gamma_{P\_LDF} \tag{3.21}$$

with LDF_j the light-dependent fraction specific for each emission compound, as specified in Table 3.10

Table 3.10 *Light dependent fractions of each of the modelled compound species / classes*

Compound Class	LDF
Isoprene	1.0
Myrcene	0.6
Sabinene	0.6
Limonene	0.4
3-D-Carene	0.4
trans-b-ocimene	0.4
b-pinene	0.4
a-pinene	0.6
Other Monoterpenes	0.4
a-Farnesene	0.5
b-Caryophyllene	0.5
Other Sesqiterpenes	0.5
232-MBO	1.0
Methanol	0.8
Acetone	0.2
CO	1.0
Bidirectional VOC	0.8
Stress VOC	0.8
Other VOC	0.2

γ_{P_LDF} is defined as (following the corrigendum to G06):

$$\gamma_{P\_LDF} = \sin(a)\left[2.46\Phi\left(1 + 0.0005(P_{\mathrm{daily}} - 400)\right) - 0.9\Phi^2\right] \quad \text{for } \sin(a) > 0 \tag{3.22}$$

Here a is a solar angle and P_{daily} represents the daily average above canopy photosynthetic photon flux density (PPFD) representative of the simulation period. In IFS-MEGAN P_{daily} is approximated through a climatological input value of the PPDF, which has a seasonal cycle only. Specifically, it is taken from the monthly mean PPFD based on ERA5 valid for the year 2019. Furthermore, Φ is the above canopy PPFD transmission, which is estimated as

$$\Phi = \frac{P}{\sin(a) \cdot P_{\text{toa}}} \tag{3.23}$$

Here P is above-canopy PPFD, and Ptoa is the PPFD at the top of the atmosphere. For this we adopt the expression given in G06, which depends on the day of the year (DOY) only:

$$P_{\text{toa}} = 3000 + 99 \cos\left(\frac{6.28(DOY - 10)}{365}\right) \tag{3.24}$$

For the temperature activity factor we follow G06:

$$\gamma_T = E_{\text{opt}}\left(\frac{C_{T2}\exp(C_{T1}x)}{C_{T2} - C_{T1}\left[1 - \exp(C_{T2}x)\right]}\right) \tag{3.25}$$

Here

$$x = \left(\frac{1}{T_{\text{opt}}} - \frac{1}{T_{\text{hr}}}\right) \cdot \frac{1}{0.00831} \tag{3.26}$$

and

$$E_{\text{opt}} = 1.75 \cdot \exp\left[0.08 \cdot (T_{\text{daily}} - 297)\right] \tag{3.27}$$

$$T_{\text{opt}} = 313 + \left[0.6 \cdot (T_{\text{daily}} - 297)\right] \tag{3.28}$$

Furthermore, This parameterization requires as input T_{hr} as hourly average air temperature (K), T_{daily} as daily average air temperature (K) representative of the simulation period (week or month), C_{T1} (=80) and C_{T2} (=200) empirical coefficients. For T_{hr} the instantaneous 2-meter temperature is taken, while for T_{daily} the soil temperature at level 3 is selected, which is assumed to match best with the past 30-day mean air temperature, following McNorton *et al.* (2023).

The emission responses to leaf area index γ_{LAI} is estimated as

$$\gamma_{LAI} = 0.49LAI/\sqrt{1 + 0.2LAI^2} \tag{3.29}$$

The leaf age activity factor γ_A follows exactly the specification as detailed in G06. This parameterization depends strongly on short-term changes in Leaf Area Index (LAI), by default quantified by changes in LAI over a time range of 10 days; hence the LAI from a specified set of days (by default 10 days) in the past is made available to the BVOC emissions module. However, to date the LAI is specified as a climatological monthly dataset, which lags relevant variability, which causes γ_A to have negligible impact.

Finally, the CO_2 inhibition factor γ_C is only applied to isoprene, and is computed as:

$$\gamma_C = I_{s_{\text{max}}} - \left[\frac{I_{s_{\text{max}}} \cdot (0.7\, C_a)^h}{(C^*)^h + (0.7\, C_a)^h}\right] \tag{3.30}$$

Here $I_{s_{\text{max}}} = 1.334$, h (=1.4614) and C* (=585) are empirical coefficients and C_a is ambient concentration of CO_2 (ppmv). The parametrization assumes that the internal CO_2 concentration inside the plant tissue is equal to 70% of the ambient concentration. The γ_C decreases non-linearly with increasing CO_2 concentration and is normalized to 1 for CO_2 = 400 ppmv.

(*c*) *conversion to emissions used in IFS-COMPO*

As specified in Table 3.9 the BVOC emissions module provides instantaneous emissions from biogenic sources for a selection of gases defined in MEGAN. These still have to be matched with the trace gases used in IFS-COMPO. For several key trace gases (e.g. isoprene, methanol, acetone, and CO) this is a one-to-one relationship, but for several other BVOCs a linear combination of multiple emission contributions

is defined. This conversion is given in Table 3.11. Note that for the CB05-bulk carbon bond tracers PAR, OLE and ALD2 we sum up the contributions from different BVOC emission types, and furthermore account for the different molar masses assumed in the emissions output module and assumed in IFS-COMPO.

3.1.5 CO_2 biogenic fluxes

The total CO_2 biogenic flux over land ecosystems is represented by the Net Ecosystem Exchange (NEE). NEE is composed of the sum of two fluxes with opposite sign: the Gross Primary Production (GPP) flux associated with the CO_2 sink from plant photosynthesis and the ecosystem respiration (R_{eco}) corresponding to the CO_2 emissions from autotrophic and heterotrophic ecosystem respiration. Both GPP and R_{eco} are modelled online in the IFS Land Surface Scheme (Part IV Physical processes, Chapter 8). The Farquhar *et al.* (1980) photosynthesis parametrization is used to model GPP. A detailed description of the photosynthesis model is available in Part IV Physical processes, Chapter 8, section 8.7.2 for CY48R1 and Part IV Physical processes, Chapter 8, section 8.7.1 for previous cycles. The ecosystem respiration R_{eco} is the sum of two fluxes: autotrophic dark respiration R_d, which is modelled within the photosynthesis module, and $R_{soilstr}$ encompassing both heterotrophic respiration from the soil and autotrophic respiration from the above and below ground structural biomass. $R_{soilstr}$ is parametrized using an empirical equation with a plant-functional-type (PFT) dependent reference ecosystem respiration constant together with three functions which represent the temperature, soil moisture and snow cover dependencies, as documented in Part IV Physical processes, Chapter 8, section 8.7.3. The sign of both GPP and R_{eco} follows the IFS Land Surface Scheme convention, which means that GPP is always positive and R_{eco} is negative.

The coupling of the NEE flux to the atmospheric CO_2 forecast is enabled by the flag LNEEONLINE which is set to True by default. Modelling the NEE online has the benefit of having consistent temporal and spatial resolution with the transport model, but it can also result in large-scale biases (Agustí-Panareda *et al.*, 2014). Thus, the NEE needs to be bias corrected in order to avoid any large-scale biases in NEE leading to drifts in the atmospheric CO_2 forecast. The Biogenic Adjustment Flux Scheme (BFAS) provides an online PFT-dependent bias correction to NEE by correcting either R_{eco}, GPP or both. Currently, only the modelled R_{eco} component is corrected, as it is deemed to have the largest uncertainty. BFAS has been updated to be consistent with the 49R1 IFS climate fields and vegetation maps (climate.v021) and the CAMS inversion products v23r2. A detailed description of BFAS can be found in Agustí-Panareda *et al.* (2016).

Table 3.11 *Matching of IFS-COMPO tracer emissions with MEGAN BVOC output, and corresponding scaling factors.*

IFS-COMPO tracer	MEGAN Emission Output
C_5H_8	isoprene
$C_{10}H_{16}$ (Terpenes)	myrcene + sabenine + limonene + 3-D-carenene + trans-beta-ocimine
HCHO	0.08 × bi-directional VOC
CH_3OH (methanol)	methanol
CH_3COOH	0.05 × bi-directional VOC
HCOOH	0.05 × bi-directional VOC
C_2H_6	0.01 × other-bvoc
C_3H_6	0.48 × other-bvoc
lumped parafins (alkanes; PAR)	0.003 × other-bvoc
lumped Olefins (alkenes; OLE)	0.24 × other-bvoc
lumped aldehydes (ALD2)	0.007 × other-bvoc + 0.17 × stress-bvoc
C_2H_4	0.58 × stress-bvoc
C_2H_5OH	0.4 × bi-directional VOC
C_3H_8	0.001 × other-bvoc
CH_3COCH_3	acetone

A new CH_4 wetland model has been introduced online, while previously these biogenic fluxes were prescribed. It is based on climatological monthly wetland fraction maps based on GIEMS (Global Inundation Extent from Multi-Satellites: https://lerma.obspm.fr/spip.php?article91&lang=fr), an available organic matter, a Q10 factor relevant to the sensitivity of CH_4 flux to temperature, and a global scaling factor to constrain the global budget.

3.1.6 Code overview

The processing of the emission inventory input is done in the routine COMPO_APPLY_EMISSIONS. It calls the routine APPLY_2D_EMISSIONS, which loops over the emissions specification structure YEMIS2D_DESC. The structure (part of namelist NAMCOMPO_EMIS) defines the target species, sectors, scaling factors, choice of diurnal cycle profile and injection height method. The emission fields are multiplied with the local diurnal cycle factor, calculated in COMPO_DIURNAL, and further scaling factors before they are added to the tracer-specific surface flux if the emissions are configured as surface emissions in YEMIS2D_DESC. The surface flux is input to the vertical tracer diffusion scheme (VDFDIFC). When the application of a injection profile method is configured in YEMIS2D_DESC, the emission field is converted to a vertical profile of MMR tendencies, which are added to the tracer tendencies in APPLY_2D_EMISSIONS. Aviation emissions are read in as 3D model level fields and are added to the tracer tendencies in APPLY_3D_EMISSIONS.

The sea salt emissions are calculated in AER_SSALT_ALBERT, which is called in the following sequence: AERINI_LAYER, AER_PHY2 and AER_SRC. The desert dust emissions are calculated in AER_DUSTEMIS, called from VDFMAIN.

CULINOX is the routine, which computes the NO emissions from lightning based on the flash densities provided by the lightning code of the IFS.

The CO_2 fluxes are calculated by the COTWORESTRESS and SRFCOTWO modules in the land surface scheme (Part IV Physical processes, Chapter 8). The photosynthesis model is in routine FARQUHAR which is called by COTWORESTRESS, and COTWORESTRESS is called by VSURF. The ecosystem respiration is calculated by SRFCOTWO. Both VSURF and SRFCOTWO are called by the interface routine of the land surface model (SURFEXCDRIVER).

3.2 REMOVAL BY DEPOSITION

3.2.1 Dry deposition

Dry deposition is a major removal pathway for many reactive gases and aerosols, which removes the tracers from the atmosphere by either sticking to or reacting with the surface or vegetation. The spatial and temporal variability of dry deposition is controlled by turbulent mixing and by surface properties.

Solubility and reactivity are properties that determine the effectiveness of dry deposition for the chemical species. Ozone is a prototype for a very reactive and insoluble species and SO_2 is the prototype for a soluble species. The effectiveness of the dry deposition of aerosols is governed by their size and shape, i.e their aerodynamic properties and whether they are heavy enough to be subject to gravitational sedimentation. Dry deposition is a highly uncertain process and deposition fluxes are difficult to measure. Modelling of dry deposition is based on empirically determined parameters (reference resistances) that are provided for different land use and vegetation classes and prototypes for reactive and soluble species.

(*a*) *Dry deposition for reactive gases*

The deposition dry flux (Φ_D) is approximated using a dry deposition velocity (V_D) that describes the effectiveness of the dry deposition process and the MMR at the lowest model level C.

$$\Phi = -\rho V_D C \tag{3.31}$$

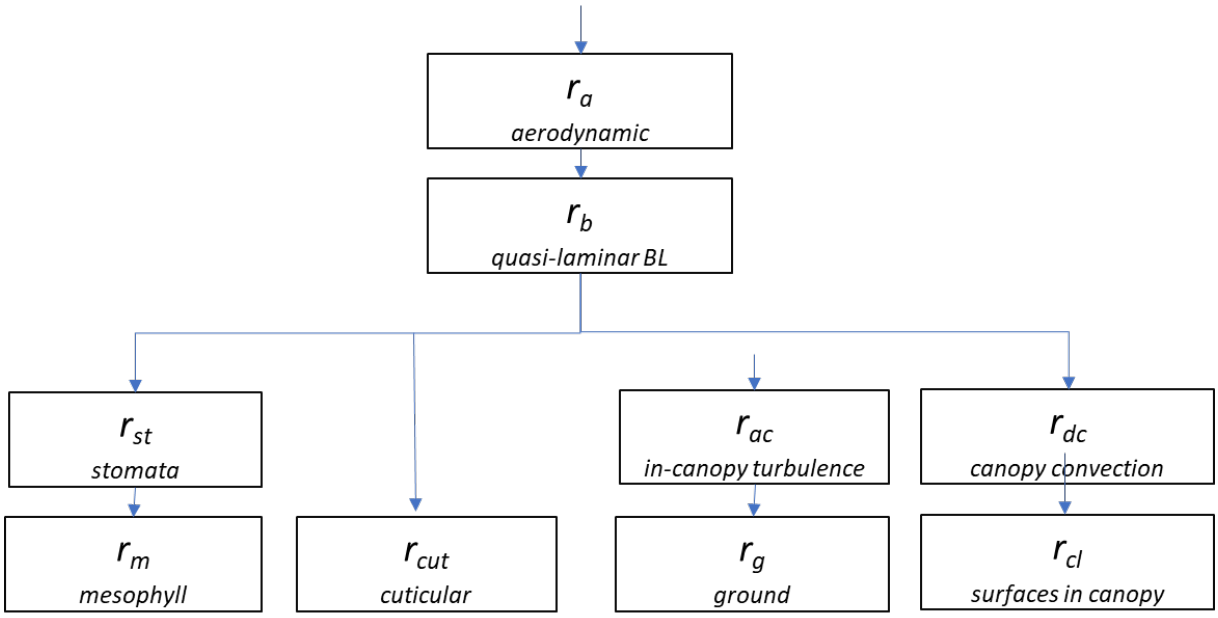


Figure 3.6 *Schematic of the components of the "big-leaf" resistance approach following Wesely (1989) applied in the IFS. The resistance terms are: aerodynamic resistance (r_a) , resistance to the quasi-laminar boundary layer around the bulk surface (r_b), stomatal resistance (r_{st}), mesophyll resistance (r_m), cuticular resistance (r_{cut}), resistance associated with within-canopy convection (r_{dc}), resistances to surfaces in the lower canopy (r_{cl}), resistance to in-canopy turbulence (r_{ac}) and resistance at the ground (r_g)*

.

V_D is modelled following a "big-leaf" approach introduced by Wesely (1989). Fig. 3.6 shows the parallel and serial pathways assumed for the dry deposition modelling in the IFS. The "big-leaf" approach treats the canopy as one entity without vertical structure. The dry deposition velocity is simulated based on a combination of resistances, namely the aerodynamic resistance r_a, the quasi-laminar sub-layer resistance r_b and the canopy/surface resistance r_c in the following way:

$$V_D = \frac{1}{r_a + r_b + r_c} \tag{3.32}$$

r_a is simulated under the surface layer assumptions outlined in Part IV Physical processes, Chapter 3, Section2 r_a [s m^{-1}] is derived from the bulk transfer coefficient for tracer mass C_c (dimensionless) and the surface wind speed U_n (Part IV Physical processes, Eq. 3.19)

$$r_a = \frac{1}{C_c U_n} \tag{3.33}$$

C_c is formulated in an equivalent as the transfer coefficients for moment, heat and moisture (C_M, C_H, C_Q, (Part IV Physical processes, Eq. 3.16-18)) based on the stability function for heat and moisture considering stable and unstable situations using Monin-Obuchov similarity theory. While the transfer coefficients are calculated over all tile fraction in the grid box, only the value from the dominate tile is used for the dry deposition calculations.

The quasi-laminar boundary layer resistance around the bulk surface (r_b, [s m^{-1}]) follows Wesely and Hicks (1977):

$$r_b = \frac{2}{k u_*} \left(\frac{\kappa}{D_A} \right)^{\frac{2}{3}} \tag{3.34}$$

The parameter $\kappa = 2.0 \times 10^{-5}$ [m^2 s^{-1}] is the thermal diffusivity of air; D_X [m^2 s^{-1}] is the diffusivity of gas X in air. k is the von Kármán constant and u^* the friction velocity calculated from the IFS land surface scheme (Part IV Physical processes, Chapter 8).

The components of the bulk surface resistance (r_c) depend on the underlying surface and vegetation properties. Hence, the deposition velocities are calculated for each surface tile fraction separately and a resulting dry deposition velocity is constructed from the weighted mean of tile fractions. r_c is composed of different pathways of interaction of the deposited gases with the surface, the canopy, the cuticle of plant and its uptake via the stomata. For vegetated surfaces, r_c is calculated from a combination of series and parallel resistances:

$$r_c = \left(\frac{1}{r_{st} + r_m} + \frac{1}{r_{cut}} + \frac{1}{r_{dc} + r_{cl}} + \frac{1}{r_{ac} + r_g} \right)^{-1} \tag{3.35}$$

r_st is the stomatal resistance, r_m is the mesophylic resistance, r_{cut} is the cuticular resistance, r_{dc} is the resistance associated with within-canopy convection, r_{cl} is the resistance to surface removal in the lower canopy, r_{ac} is the resistance to in-canopy turbulence, and r_g is the ground resistance.

For non-vegetated surface tiles such bare ground, ice, snow-covered vegetation, ocean and lake the canopy related terms are neglected resulting in $r_c = r_g$. The bulk surface resistance for the highly soluble gases HNO_3 and H_2O_2 is set to a fixed value of $r_c = 10$ s m^{-1} independent of the land cover and vegetation type of the surface, which means an efficient dry deposition if that is not limited by the aerodynamic or quasi-laminar boundary layer resistance.

In particular, the variability of stomatal uptake is simulated depending on the meteorological conditions.

The numerical values of the reference resistances parameters are taken from the GEOS-Chem model and are listed in table 3.12. The mapping of the IFS land use tiles and vegetation types to the GEOS-Chem land use type is listed in 3.13

$$r_{st} = \frac{D_w}{D_X} \frac{r_i}{LAI_{eff} f(T_a)} \tag{3.36}$$

The parameter r_i is the initial stomatal resistance that is tabulated for the GEOS-Chem 3.12 ; D_w [m^2 s^{-1}] is the diffusivity of water vapour or tracer X in air. It should be noted that the diffusivities are uncertain values. Usually, the diffusivity ratio for O_3 is assumed to has the value of 1.6 whereas the choice of the parameter values in GEOS-Chem results in a diffusivity ratio value of 1.3 leading to smaller stomatal resistance. LAI_{eff} [m^2 m^{-2}] is effective leaf area of actively transpiring leaves. The variable LAI_{eff} is calculated with a function (called 'biofit' in the source code) from leaf area index (LAI), solar zenith angle and cloud fraction. When incident shortwave radiation is zero (i.e. at night), LAI_{eff} is set to 0.01 m^2 m^{-2}. The function was fitted to match the results of a multi-layer canopy radiative transfer model. The function describing the impact of air temperature (T_a, [°C]) is given by:

$$f(T_a) = T_a \frac{40 - T_a}{400} \tag{3.37}$$

The mesophyll resistance (r_m, [s m^{-1}]) is given by:

$$r_m = \left(\frac{H}{3000} + 100 f_0 \right)^{-1} \tag{3.38}$$

The variable H is the Henry's law constant of the species which varies with air temperature [0.01 M atm^{-1} at 298 K for ozone]. f_0 is the reactivity factor, which expresses the tendency to chemically react with respect to the reactivity to ozone. $f_0 = 1$ is the value for ozone and H_2O_2, $f_0 = 0.3$ for CH_2OOH and $f_0 = 0.1$ for selected nitrogen species and hydrocarbons. $f_0 = 0$ is the value for all other species, which means that only their solubility, expressed by H, determines the effectiveness of their dry deposition.

The resistance associated with in-canopy convection (r_{dc}) follows:

$$r_{dc} = 100\left(1 + \frac{1}{G}\right) \tag{3.39}$$

The variable G [W m^{-2}] is the incoming shortwave radiation.

Cuticular resistance (r_{cut}) is given by:

$$r_{cut} = \frac{r_{lu}}{LAI}\left(\frac{H}{10^5} + f_0\right)^{-1} \tag{3.40}$$

The parameter r_{lu} is the initial resistance for cuticular uptake.

The resistances to surfaces in the lower canopy (r_{cl}) and the ground (r_g) are calculated using a similar structure consisting of processes related to solubility and reactivity using SO_2 (subscript S) and ozone (subscript O) as reference species. Specifically, the ground resistances is given by:

$$r_g = \left(\frac{H}{10^5}\frac{1}{r_{S,g}} + \frac{f_0}{r_{O,g}}\right)^{-1} \tag{3.41}$$

and the resistances to surface in the canopy:

$$r_{cl} = \left(\frac{H}{10^5}\frac{1}{r_{S,cl}} + \frac{f_0}{r_{O,cl}}\right)^{-1} \tag{3.42}$$

The resistance because of in-canopy turbulence r_{ac} is a land-use specific value (3.12).

To consider the effects of air temperature (T_a) a resistance term $r + T = 1000\exp\left(-T_a - 4 - 273\right)$ is added to to the tabulated resistance parameters r_{lu}, $r_{S,g}$, $r_{O,g}$, $r_{S,cl}$ and $r_{O,cl}$.

Table 3.12 *Reference resistances used for the dry deposition of gases based on GEOS-Chem land types. R_i: the minimal stomatal resistances (differs from values used in IFS Land Surface Scheme), R_{lu} : the uptake pathway for leaf cuticles, R_{ac} : resistances within the canopy, $R_{gs,O}$: uptake pathways for the ground including soil, leaf litter for O_3. $R_{cl,o}$: the uptake pathways by in the lower canopy by leaves, twigs, bark etc. for O_3.*

GC	Snow Ice	Deciduous Forest	Conifer Forest	Agricultural Land	Shrub Grassland	Amazon Forest	Tundra	Desert	Wetland	Urban	Water
R_i		200	400	200	200	200	200		200		
R_{lu}		9000	9000	9000	9000	1000	4000		9000		
R_{ac}	0	2000	2000	200	100	2000	0	0	300	100	0
$R_{gs,O}$		150	150	110	150	150	260	300	750	225	2000
$R_{cl,O}$	1000	1000	1000	1000	1000				1000		

(*b*) *Dry deposition for aerosols*

A new parameterization of aerosol dry deposition following Zhang and He (2014) has been implemented in cycle 47R1 IFS-AER and is used operationally since CY47R3. In cycle 49R1, the aerosol dry deposition routine has been moved to be called from the vertical diffusion routine, similarly to what is done for the dry deposition of chemical species. This also allows to use the latest output from the surface scheme, in particular for the friction velocity u^* which is a key input of aerosol dry deposition.

The Zhang and He (2014) dry deposition parameterization has been implemented because it gave good results in a recent inter-comparison of dry deposition schemes (Khan and Perlinger, 2017), and also because instead of use the particle size as an input, it divides particles in size ranges: fine, coarse and giant (super-coarse). Only the surface resistance differs as compared to the Zhang *et al.* (2001) scheme. The inverse of the surface resistance is also referred to as surface deposition velocity and denoted as V_D. It is computed as as a function of the particle diameter D_p and friction velocity u^* as :

Table 3.13 *Mapping of IFS vegetation and land use classes to GEOS-Chem Land use types*

IFS Vegetation Type	IFS tile	GEOS-Chem Type
Crops, Mixed Farming	Low vegetation	Agricultural
Short Grass	Low vegetation	Shrub/Grassland
Evergreen Needleleaf Trees	High vegetation	Coniferous Forest
Deciduous Needleleaf Trees	High vegetation	Coniferous Forest
Deciduous Broadleaf Trees	High vegetation	Deciduous Forest
Evergreen Broadleaf Trees	High vegetation	Amazon
Tall Grass	Low vegetation	Shrub/Grassland
Desert	Low vegetation	Desert
Tundra	Low vegetation	Tundra
Irrigated Crops	Low vegetation	Agricultural
Semidesert	Barren ground	Desert
Ice Caps and Glaciers	Ice	Snow/Ice
Bogs and Marshes	Low vegetation	Wetland
Inland Water	Water	Water
Ocean	Ocean	Water
Evergreen Shrubs	Low vegetation	Shrub/Grassland
Deciduous Shrubs	Low vegetation	Shrub/Grassland
Mixed Forest/Woodland	High vegetation	Deciduous Forest
Interrupted Forest	High vegetation	Deciduous Forest
Water and Land Mixtures	Water	Wetland
Lake	Lake	Water
Urban		Urban, if implemented
	Snow on vegetation	Snow/Ice

$$V_D = \begin{cases} a_1 \times u_* & \text{for } D_p \leq 2.5\mu m \\ \left(b_1 \times u_* + b_2 \times u_*^2 + b_3 \times u_*^3\right) \times \exp\left(K_1 \times \left(\frac{LAI}{LAI_{MAX}} - 1\right)\right) & \text{for } 2.5\mu m < D_p \leq 10\mu m \\ \left(d_1 \times u_* + d_2 \times u_*^2 + d_3 \times u_*^3\right) \times \exp\left(K_2 \times \left(\frac{LAI}{LAI_{MAX}} - 1\right)\right) & \text{for } D_p > 10\mu m \end{cases} \quad (3.43)$$

$$K_1 = c_1 \times u_* + c_2 \times u_*^2 + c_3 \times u_*^3$$
$$K_2 = e_1 \times u_* + e_2 \times u_*^2 + e_3 \times u_*^3$$

Where a_i, b_i, c_i, d_i, e_i are land-surface dependent coefficients are provided in Tables 2a and 2b of Zhang and He (2014). LAI_{MAX} is the maximum leaf area index for a given land surface category, which has been estimated from the LAI climatology used in the IFS. u^* in equation 3.43 is computed from u_0^* as provided by the surface scheme by adding a gustiness component:

$$u^* = u_0^* if F_{bu} <= 0$$
$$u^* = u_0^* (\frac{W_s^2 + F_{bu}}{W_s^2})^{0.25} if F_{bu} > 0$$

Where F_bu is the surface buoyancy flux as provided by the surface scheme and W_s is the surface wind speed.

Figure 3.7 shows a comparison of the simulated dry deposition velocity by the Zhang *et al.* (2001) and Zhang and He (2014) schemes over a particular land surface category (desert).

In cycle 48R1, a parameterization of the rebound effect of super-coarse dust particles over continental surfaces has been. implemented following Zhang *et al.* (2001). The surface resistance for super coarse

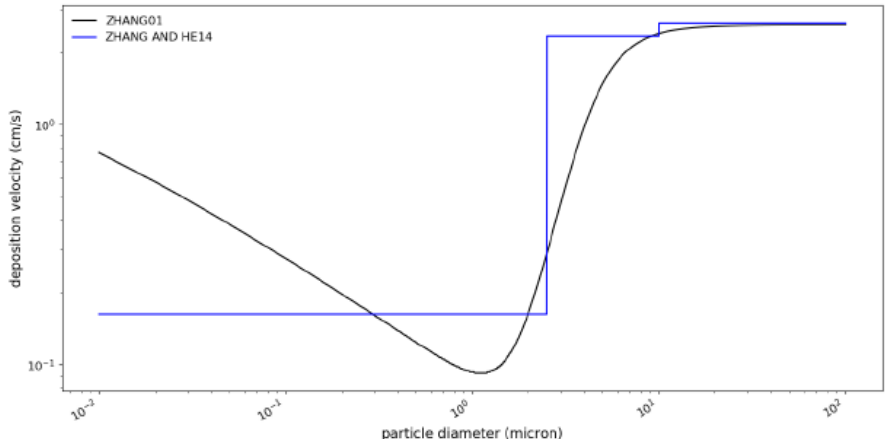


Figure 3.7 *Dry deposition velocity with a friction velocity of 1 m/s, over a desert surface as a function of particle size, parameterized by the Zhang* et al. *(2001) and the Zhang and He (2014) schemes.*

dust particles, as parameterized in the Zhang and He (2014) scheme, is multiplied by a factor R over non-water surfaces, if the friction velocity is above 0.335 m s^{-1}, following results from field campaigns as detailed in Bergametti *et al.* (2018):

$$R = \exp\left(R_0 S_t^{1/2}\right) \tag{3.44}$$

Where R_0 is a constant set to 0.1 and S_t is the Stokes number. Following Zhang *et al.* (2001), S_t is computed differently between vegetated surfaces and smooth surfaces (or surfaces with bluff roughness elements):

$$S_t = V_D u^* / gA \text{ over vegetated surfaces} \tag{3.45}$$

$$S_t = V_D u^{*2} / \mu \text{ over smooth surfaces} \tag{3.46}$$

Where A is the characteristic radius of collectors (provided in Table 3 of Zhang *et al.* (2001), g is the gravity constant,μ is the kinematic velocity of air and V_s is the gravitational settling velocity computed as a function of particle radius r and density ρ_p. C_F is the Cunningham correction factor.

$$V_s = \frac{2\rho_p g}{9\mu} r^2 C_F \tag{3.47}$$

3.2.2 Wet deposition

Wet deposition is the transport and removal of soluble or scavenged trace gases and aerosol by precipitation. It includes the following processes:

- In-cloud scavenging and removal by rain and snow (rain-out).
- Release by evaporation of rain and snow.
- Below-cloud scavenging by precipitation falling through without formation of precipitation (wash-out).

It is important to take the sub-grid scale of cloud and precipitation formation into account for the simulation of wet deposition. The IFS cloud scheme provides information on the cloud and the precipitation fraction for each grid box. It uses a random overlap assumption (Jakob and Klein, 2000) to derive cloud and precipitation area fraction. The precipitation fluxes for the simulation of wet removal in IFS are divided by the precipitation fraction of the grid cell in order to be valid over the precipitation fraction of the respective grid box.

The loss of trace gas and aerosol by rain-out and wash-out is limited to the area of the grid box covered by precipitation. Likewise, the cloud water and ice content is scaled to the respective cloud area fraction.

If the sub-grid-scale distribution was not considered in this way, wet deposition would be lower for highly soluble species such as HNO_3 and aerosol because the species is only removed from the cloudy or rainy grid box fraction.

Even if wet deposition removes tracer mass only in the precipitation area, the mass mixing ratio representing the entire grid box is currently changed accordingly after each model time step. This is equivalent to the assumption that there is instantaneous mixing within the grid box on the timescale of the model time step.

The module for wet deposition in IFS is based on the Harvard wet deposition scheme (Jacob *et al.*, 2000; Liu *et al.*, 2001) with additions derived from Luo *et al.* (2019). In cycle 49R1, the implementations of the wet deposition scheme for aerosols and chemistry have been grouped into a common routine, using similar meteorological inputs. For aerosols only, an aerosol activation scheme based on (Verheggen *et al.*, 2007) has been implemented, which provides the fraction of aerosols that are subjected to in-cloud wet deposition in mixed clouds (temperature between 0°C and -40°c). Also, instead of using a fixed precipitation pH value over oceans and continents as done in cycle 48R1, cycle 49R1 implements the use of diagnostic precipitation pH as computed in EQSAM4Clim and updated in the aqueous chemistry routine. The input fields to the wet deposition routine are the following prognostic variables, calculated by the IFS cloud scheme (Forbes *et al.*, 2011): total cloud and ice water content, grid-scale rain and snow water content and cloud and grid-scale precipitation fraction as well as the derived fluxes for convective and grid-scale precipitation fluxes at the grid cell interfaces. For convective precipitation, a precipitation fraction of 0.05 is assumed and the convective rain and snow water content is calculated using the convective liquid precipitation rate and assuming a droplet fall speed of 5 m s^{-1}.

All aerosols and the soluble chemical species as listed in Table 3.14 are subject to wet deposition.

Rain-out, evaporation and wash-out are calculated after each other for large-scale and convective precipitation. Following Jacob *et al.* (2000), the fractions ($f_{i,L}$, $f_{i,I}$) of a trace gas i dissolved in cloud / rain droplets (L) or cloud ice (I) is calculated using Henry's law equilibrium:

$$\frac{C_{i,L}}{C_{i,G}} = H_i \times LRT \tag{3.48}$$

in the following way:

$$f_{i,L} = \frac{C_{i,L}}{C_{i,T}} = \frac{\frac{C_{i,L}}{C_{i,G}}}{1 + \frac{C_{i,L}}{C_{i,G}} + \frac{C_{i,I}}{C_{i,G}}} \tag{3.49}$$

$$f_{i,I} = \frac{C_{i,I}}{C_{i,T}} = \frac{\frac{C_{i,I}}{C_{i,G}}}{1 + \frac{C_{i,L}}{C_{i,G}} + \frac{C_{i,I}}{C_{i,G}}} \tag{3.50}$$

$C_{i,L}$, $C_{i,I}$, $C_{i,G}$ and $C_{i,T}$ are the concentrations of i in the droplet (L), cloud ice (I) and air (G) and the total concentration (T), respectively. H_i is the effective Henry coefficient for species i, L the (scaled) cloud liquid water content (or precipitation, in the case of below-cloud scavenging), R the ideal gas constant and T temperature. Only H_2O_2 and HNO_3 are assumed to dissolve in ice clouds (Lawrence and Crutzen, 1998), and hence subject to in-cloud ice precipitation.

The effective Henry coefficient for SO_2 and NH_3 accounts for its dissociation in water, and is calculated following (Seinfeld and Pandis, 1998), using the rain water acidity diagnostic as computed in EQSAM4Clim and the aqueous chemistry routine. The other Henry's law coefficients are taken from the compilation by (Sander, 2015).

(*a*) *In-cloud scavenging*

The loss by in-cloud scavenging (rain-out) is governed by the precipitation formation rate P_r.

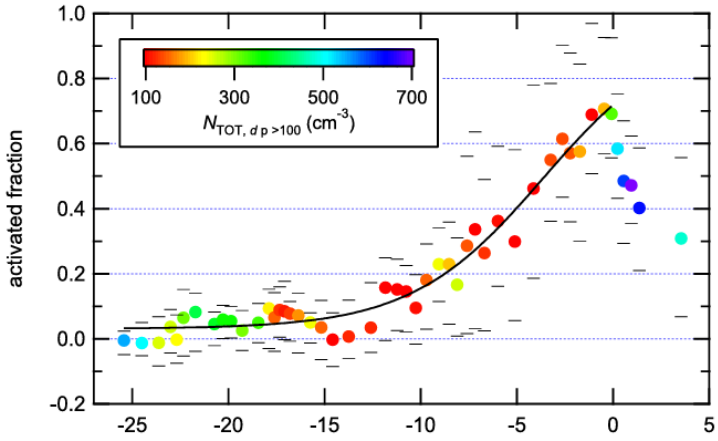


Figure 3.8 *Figure from (Verheggen et al., 2007): Dependence of the aerosol activated fraction on the temperature based on 910 hours of in-cloud data. Each circle denotes the average, while horizontal stripes denote the 25 and 75 percentile values of 18 hours of data. Circle colours indicate the average concentration of particles with diameter above 100nm; the solid line is the sigmoid fit through the average values with temperature below 0°c, used in equation 3.53.*

The in-cloud scavenging rate $W^{\mathrm{I}}_{i,k}$ [s^{-1}] for species i at level k is adapted from the approach of Luo *et al.* (2019). For liquid precipitation of trace gases,

$$W^{\mathrm{I}}_{i,k} = \frac{cf \times \beta_{r,k}}{k \times q_{k,r,tot}} \exp\left(-kR_c f_{i,L} \Delta_t\right) \tag{3.51}$$

Rc is the retention coefficient which accounts for the retention of dissolved gas in the liquid cloud condensate as it is converted to precipitation. $R = 1.0$ for all species in warm clouds (T $>$ 268 K). For mixed clouds (T $<$ 268 K), R is 0.02 for all species but 1.0 for HNO_3 and 0.6 for H_2O_2 (von Blohn *et al.*, 2011).

A similar expression is used for in-cloud precipitation due to ice cloud, but replacing $f_{i,L}$ with $f_{i,I}$ and computing k, the rain-out loss rate, based on ice precipitation formation.

For aerosol expression 3.51 reads:

$$W^{\mathrm{I}}_{i,k} = \frac{cf \times \beta_{r,k}}{k \times q_{k,r,tot}} \exp\left(-kD_{w,i} \Delta_t\right) \tag{3.52}$$

Here $D_{w,i}$ is the fraction of aerosol that is embedded in the cloud liquid/solid water, provided by table 3.15 for liquid clouds (temperature above 0°c). In mixed clouds (temperature below 0°c), $D_{w,i}$ corresponds to the fraction of activated aerosols as estimated as a function of temperature by (Verheggen *et al.*, 2007):

$$D_{w,i} = 0.031 + \frac{0.93}{(1 + exp(\frac{-t+269.51}{3.42}))} \tag{3.53}$$

Figure 3.8 shows the underlying observational data as well as the dependency of the activated aerosol fraction as a function of temperature as estimated by equation 3.53.

In eqs 3.51 and 3.52, $\beta_{r,k}$ is the rate of new precipitation formation (rain only) and cf is the cloud fraction. $q_{k,r,tot}$ represents the condensed water content (liquid) within the grid cell and is derived from the liquid

water mass mixing ratio q_k by

$$q_{k,r,tot} = q_k + \Delta_t \times \beta_{r,k} \tag{3.54}$$

where Δ_t is the time step and $\beta_{r,k}$ is defined as in Giorgi and Chameides (1986) using the rain flux $P_{r,k}$ at level k:

$$\beta_{r,k} = (P_{r,k+1} - P_{r,k}) \times (\rho_k \Delta z_k) \tag{3.55}$$

The first-order rain-out loss rate k_{loss}, which represents the conversion of cloud water (or ice water) to precipitation water, is then computed as:

$$k_{loss} = \left[k_{min} + \frac{\beta_{k,r}}{q_{k,r,tot}} \right] \tag{3.56}$$

k_{min} is the minimum value of rain-out loss rate, set to $0.0001 s^{-1}$ following Luo *et al.* (2019).

The formulation of Luo *et al.* (2019) applies only to liquid precipitation. It has been extended for solid precipitation, but taking into account the smaller fraction of aerosols included in solid precipitation, the value the D_i parameter is divided by two for solid precipitation. The scavenging rates for solid and liquid precipitation are then added.

(*b*) *Re-evaporation*

The release of trace gases and aerosols contained in rain drops at level k occurs if evaporation of precipitation is diagnosed, i.e. if the precipitation flux at level k is higher than at level $k+1$, where level $k+1$ is below level k. If there is no precipitation at level $k+1$, then all aerosols that have been subjected to in-cloud scavenging at or above level k are released. If the precipitation flux at level $k+1$ is not null, then the re-evaporation is partial. Before cycle 46R1, it was assumed arbitrarily that half of the scavenged trace gas and aerosols at or above level k are then released. Since cycle 46R1, a more complex parameterization has been implemented, following de Bruine *et al.* (2018). The mass of an aerosol or trace gas species i that is re-evaporated at level k is computed as a function of the fraction of evaporated precipitation defined with the precipitation flux at level k (P_k), and the fraction of evaporated precipitation $\epsilon_k = \frac{P_{k+1} - P_k}{P_k}$:

$$\delta M_{i,k}^{evap} = \epsilon_k \times \left(\left[1 - exp^{-2\sqrt{\epsilon_k}} \left(1 + 2 \times \epsilon_k^{\frac{1}{2}} + 2 \times \epsilon_k + \frac{4}{3} \epsilon_k^{\frac{3}{2}} \right) \right] \times (1 - \epsilon_k) + \epsilon_k^2 \right) \times Loss_{scav,i,k} \tag{3.57}$$

Where $Loss_{scav,i,k}$ is the sum of the mass of aerosol or trace gas that is subjected to in-cloud wet deposition from level k to the model top.

(*c*) *Below-cloud scavenging*

The fraction W_i^B of a highly soluble tracer i that is scavenged by below-cloud scavenging (wash-out), is computed as

$$W_i^B = f \left(1 - exp \left(-k_i' \frac{P_{r,k}}{f} \Delta_t \right) \right) \tag{3.58}$$

with $P_{r,k}$ the grid-scale precipitation rate and f the precipitation area faction of the grid cell. For HNO_3, and for other trace gases for which the scavenging is limited by mass transfer ($f_{i,L} > W_i^B / f$), equation 3.58 is adopted with a first order wash-out loss rate constant of $k_i' = 1$ cm^{-1}. If scavenging is limited by Henry solubility ($f_{i,L} < W_i^B / f$) then the total mass of scavenged tracer transported out through the bottom of the gridbox is:

$$\Delta m_{i,bottom} = f_{i,L} (f m_i + \Delta m_{im,top}) \tag{3.59}$$

For aerosols, the below cloud scavenging rate is computed following (Croft *et al.*, 2009):

$$W^{B}_{i,k} = fp_k \times \left[(1 - \exp\left(P^2_{r,k} \times \alpha_r\right)) + (1 - \exp\left(P_{i,k} \times \alpha_i\right)) \right] \tag{3.60}$$

Where $P_{r,k}$ and $P_{i,k}$ are the fluxes of liquid and solid precipitation respectively, fp_k is the fraction of grid cell at level k in which precipitation occurs, a is a constant scaling factor with a value of 0.00273 to account for a unit conversion between the units used in Croft *et al.* (2009) (mm/hr) and in the IFS (mm/s), and α_r and α_i are the efficiency with which aerosol variables are washed out by rain and snow, respectively. The values used have been derived from Croft *et al.* (2009) and are summarized in table 3.16.

3.2.3 Sedimentation of aerosols

Sedimentation has been left broadly unchanged since cycle 32R2 (Morcrette *et al.* (2009)). In cycle 47R3 and before, It was applied only for super-coarse dust and sea-salt, for which it is an important sink. In cycle 48R1, it is applied to all aerosol tracers. In cycle 49R1 and 49R2, the Stokes formula is applied to compute the sedimentation velocity of sulfate aerosols, with a varying size distribution (see details in Chapter 2). In cycle 50R1, the Stokes formula is applied to compute the sedimentation velocity of all aerosol species. The change in mass mixing ratio from sedimentation follows the approach of Tompkins (2005) for ice sedimentation. The change in mass concentration caused by a transport in flux form at velocity V_s is given by:

$$\frac{dC}{dt} = \frac{1}{\rho}\frac{d(\rho V_{\mathrm{s}} C)}{dz} \tag{3.61}$$

where ρ is the air density. The integration of this gives for each level k and time step j:

$$C^{j}_{k+1} = \frac{\dfrac{\rho^{j-1} V_{\mathrm{s}} C^{j-1}_{k+1}}{\rho^{j}\Delta Z}\Delta t + C^{j}_{k}}{1 + \dfrac{\rho^{j} V_{\mathrm{s}}}{\rho^{j}\Delta Z}\Delta t} \tag{3.62}$$

which is solved from top to bottom. Before cycle 50R1, the gravitational velocity V_{s} was provided from an external source, the aerosol table file, and is horizontally and vertically invariant for all species except for sulfate and for bin2 and 3 of sea-salt aerosol, for which it is computed online using Stokes' law if the namelist entry LAERSEDIMSS is set to true (which is the case in cycle 48R1). Since cycle 50R1, the sedimentation velocity is computed online for all species, and the LAERSEDIMSS namelist entry is deprecated. The effect of asphericity on desert dust sedimentation is accounted for by decreasing the estimated sedimentation velocity by 20%. This reduction is inspired by the work of Drakaki *et al.* (2022), who obtained the best results for the transatlantic transport of coarse and super coarse particles with sedimentation velocities reduced by 40 to 80%. The settling velocity is computed with the Stokes formula as:

$$V_{\mathrm{s}} = \frac{1\rho_{\mathrm{p}} g}{18\mu} D^2_p \, C_{\mathrm{F}} \tag{3.63}$$

where D_p and ρ_{p} are the particle diameter and density respectively, g the gravitational constant, μ the air viscosity and C_{F} the Cunningham correction factor. For sulfate, a varying size distribution is assumed in order to represent correctly both volcanic and quiescent conditions in the stratosphere; D_p is computed as:

$$D_p = D_0 (mmr_{SO_4}/5.E8)^{0.25} \tag{3.64}$$

Where D_0 is the sulfate wet diameter, computed from the sulfate dry diameter of 0.8 micron and the sulfate hydrophylic growth factor. For the other species, table 3.17 provides the mass median diameters used in the Stokes formula.

For cycles 49R2 and before, table 3.18 provides the fixed sedimentation velocities used for all species except sulfate and sea-salt aerosols bins2 and 3, for which the Stokes formula is used.

3.2.4 Code overview

The dry deposition velocities for reactive gases are calculated in DEPVEL_GC (and subroutines within), which is called in VDFMAIN before the simulation of the vertical diffusion of tracers in VDFDIFC. DEPVEL_GC requires input from the land surface scheme (Part IV Physical processes, Chapter 6), which is provided via the interface routine SURFEXCDRIVER.

DEPVEL_GC is also the routine used as an interface for the calculation of the aerosol dry deposition velocities. The computation itself is carried out in AER_DRYDEPVELZH14 The update of the 3D aerosol tendencies because of sedimentation is calculated in AER_SEDIMNT, which is called from AER_PHY2.

Table 3.14 *Trace gases subject to dry and wet deposition, along with the assumed Henry constants*

Trace gas	H_s^{cp} $[10^2 \frac{mol}{m^3 Pa}]$	$\frac{dlnH_s^{cp}}{d(1/T)}$ [K]
O3	1×10^{-2}	2800
SO2	1×10^{5}	3000
NH3	59	4200
H2O2	8.3×10^{4}	7600
HNO3	3.2×10^{11}	8700
NO	1.9×10^{-3}	1600
HO2	680	0
NO2	1.2×10^{-2}	2400
NO3	3.8×10^{-2}	0
N2O5	21	3400
HONO	48	4800
HO2NO2	1.2×10^{4}	6900
CH3OOH	2.9×10^{2}	5200
CH2O	3.2×10^{3}	6800
CH3OH	200	5600
CH3O2	15	3700
HCOOH	8.8×10^{3}	6100
MCOOH	4.1×10^{3}	6300
C2H5OH	190	6400
CHOCHO	4.1×10^{5}	7500
CH3COCHO	3.4×10^{3}	7500
CH3COCH3	27	5500
ALD2	17	5000
PAN	2.9	5700
ROOH	340	6000
ONIT	1e3	6485
ISOPOOH	3.2×10^{11}	8700
ISPD	4.3	5300
HPALD1	17	5000
HPALD2	17	5000
GLYALD	4.1×10^{4}	4600
HYAC	6.3×10^{3}	6460
SOG1	1×10^{5}	6000
SOG2A	1×10^{3}	6000
SOG2B	1×10^{5}	6000
MSA	3.2×10^{11}	8700
HCN	12	5000
CH3CN	49	4000
HCL	3.2×10^{11}	8700
CLONO2	3.2×10^{11}	8700
HOCL	6.5×10^{2}	5600
HBR	3.2×10^{11}	8700
BRONO2	3.2×10^{11}	8700
HOBR	6.5×10^{2}	5600
Pb	3.2×10^{11}	8700

Table 3.15 *Value of the parameter D_w, representing the the fraction of the aerosol that is embedded in the cloud liquid water.*

Species i	$D_{w,i}$ value
Sea salt	0.9
Dust	0.7
OM hydrophilic	0.7
BC hydrophilic	0.7
Sulfate	0.7
Nitrate	0.8
Ammonium	0.9

Table 3.16 *Value of the parameters α_r and α_i.*

Species	α_r	α_i
Seasalt fine	0.15	0.01
Seasalt coarse	0.36	0.018
Seasalt super coarse	0.36	0.05
Dust fine	0.15	0.01
Dust coarse	0.36	0.018
Dust super coarse	0.36	0.05
OM hydrophilic	0.001	0.001
OM hydrophobic	0.001	0.001
BC hydrophilic	0.001	0.001
BC hydrophobic	0.001	0.001
Sulphate	0.001	0.001
Nitrate fine	0.001	0.001
Nitrate coarse	0.36	0.05
Ammonium	0.001	0.001
SOA Biogenic	0.001	0.001
SOA Anthropogenic	0.001	0.001

Table 3.17 *Mass median diameters used in the Stokes formula in cycle 50R1.*

Species	D_p (micron)
Sea salt 1	0.68
Sea salt 2	2.58
Sea salt 3	12.92
Dust 1	0.51
Dust 2	1.41
Dust 3	5.06
OM/SOA	0.312
BC	0.288
NI1/AM	0.352
Nitrate 2	2.58

Table 3.18 *Sedimentation velocity for aerosol species other than sulfate and sea-salt aerosol bin2 and 3*

Species	V_s (m/s)
Sea salt 1	2.4E-5
Dust 1	6.9E-5
Dust 2	1.982E-4
Dust 3	1.962E-3
OM/BC/NI1/AM/SOA	2E-4
Nitrate 2	1.33E-3

Chapter 4

Transport processes

Table of contents

The transport of tracers by advection, vertical turbulent diffusion and convective mass fluxes is conceptually the same for aerosol and traces gases and other tracer such as humidity, cloud water and ice. This section documents the specifics for the aerosol and trace gases transport .

4.1 ADVECTION

4.1.1 Overview

The simulation of tracer advection by the three-dimensional wind fields applies the semi-Lagrangian (SL) method as outlined in Part III Dynamics and Numerical Procedures, Chapter 3, which is also used to simulate the advection of humidity and temperature. In contrast to mass-conserving flux-form advection schemes, which calculate the tracer mass flux through all grid-box boundaries, the SL scheme approximates the change of the mass mixing ratio by advection by assigning the interpolated mass mixing ratio of the departure point to the grid box value. The departure point is the point where the trajectory of an air parcel originated at the beginning of the time step. The departure point, which is the same for all advected variables, is calculated from the wind fields in an iterative way. The mass mixing ratio at the departure point is interpolated from the surrounding grid points. Different interpolation methods and limiters can be configured for each tracer field.

The SL method is computationally very efficient because the time step length is not limited by the Courant–Friedrichs–Lewy (CFL) condition. The CFL condition constrains the time step depending on the wind speed and the grid box size (resolution) for flux form advection schemes. A specific computational advantage of the SL advection scheme for atmospheric composition simulation is that the identification of the departure point, which is computationally expensive, is only required to be carried out one time independently of the number of advected tracers.

The semi-Lagrangian advection method does not formally conserve the tracer mass in contrast to flux-form based advection schemes. The local amount of mass non-conservation of the SL advection can not be correctly diagnosed and both erroneous creating and destruction of tracer mass occurs. It is only possible to calculate the global amount of tracer mass conservation by calculating the global mass of a tracer before and after the advection time step.

To enforce global mass conservation, the IFS has different options for global mass fixers (MF) (Diamantakis and Flemming, 2014). Global mass fixers (i) diagnose the global amount of mass conservation after each advection time and (ii) modify the resulting MMR to match the global tracer mass at the start of the time step. As the global mass integration depends on the surface pressure, conservation of the global surface pressure integral has to be ensured to accurately diagnose the mass non-conservation.

The local corrections of any MF approach are always positive or always negative depending on the sign of the global mass non-conservation. The global mass fixers implemented in the IFS uses different

methods to locally modify the MMR. The proportional MF in the IFS scales all grid point MMR values with the same correction factor. There are also MF that apply locally varying amount of mass depending on conditions such as the smoothness of the tracer field or the approximated accuracy of the interpolation of the departure point value (Part III Dynamics and Numerical Procedures, Section 3.2.5).

4.1.2 Specification of SL advection and Mass fixer for tracers

The specification of the options for SL-advection and MF are specified in a name-list (NAMGFL). Only one type of global MF can be selected for all tracers, which can be specifically configured for each tracer. The configuration is summarised in table 4.1 and explained below.

All Greenhouse gases, aerosols and most reactive gas species are advected. Only very short-lived chemical species such as OH, HO2 and H radical are not advected because the chemical lifetime is smaller than the advection time step. The MMR tendencies of these very-short-lived species are entirely controlled by the chemical conversion.

Table 4.1 *configuration of SL interpolation and global mass fixers*

Item	GHG	Aerosols	Chemistry
Advected species	all	all	excluding very-short-lived
Family advection	none	none	NO_y, Br_y, Cl_y
Quasi Monotonic limiter	3D	3D	3D
COMAD interpolation	horizontal	horizontal	horizontal
Mass fixer	3D Bermejo-Conde	proportional	proportional

(*a*) *Departure point interpolation*

The COntinuous Mapping About Departure points (Malardel and Ricard, 2015, COMAD) scheme is used for the interpolation to the departure point. The COMAD scheme improves the mass conserving property of the SL advection scheme by modifying the interpolation weights. This is achieved by using the size of the grid cell at the arrival point for computations at the departure point. The COMAD scheme is only applied for the horizontal interpolation.

For the vertical interpolation quasi-monotonic cubic interpolation method is used.

Further, a quasi-monotone limiter of the interpolated departure point value is applied prevent overshoots/undershoots by limiting the interpolated values to the range of surrounding grid points in horizontal and vertical direction (LQM3D).

(*b*) *Family advection*

Groups of chemical species (families) such as oxidised nitrogen components are characterised by quick chemical conversions leading to large spatial gradients for example near the day-night terminator. As the SL advection has the tendency to suffer from larger mass conservation errors in areas of large gradient, the individual advection of each family member can lead to drift and issues of non-conservation. Therefore the chemical species of the family are added to a single family tracer, which is advected and subject to MF application. The individual MMR for each family member is calculated after the advection from the proportional contribution to the family at the start of the time step. The family advection approach is applied to the following chemical families:

- Bromine family (BrCl, HOBr, $BrONO_2$, Br, HBr, BrO and Br_2)
- Chlorine family (Cl_2O_2, OClO, BrCl, HOCl, ClONO2, Cl, HCl, ClO, $ClNO_2$, Cl_2 and ClOO)
- NOy family (N, NO, NO_2, NO_3, HO_2NO_2, N_2O_5, HNO_3, $ClNO_2$, $ClONO_2$ and $BrONO_2$)

(*c*) *Global mass fixing*

The IFS tracer MF extends the method of Bermejo and Conde (2002) (BC, LTRCMFBC=true) as described in Agusti-Panareda *et al.* (2017) and Diamantakis and Agusti-Panareda (2017) for Greenhouse Gases.

For the aerosol and reactive gases the proportional MF is applied. The BC MF computes a correction to the transported field with its magnitude depending on the local smoothness of the field with larger corrections in locations of large gradients. A detailed description of the BC MF scheme is provided in Part III Dynamics and Numerical Procedures, Section 3.2.5.

The BC MF type is the multiplicative standard configuration (NOPTMFBC=3). The species specific-smoothness parameter (BETAMFBC=2) is set for GHG in such a way that the MF applies smaller corrections in areas, where the tracer gradient is close to zero. As only one mass fixer type can be selected in the IFS, setting BETAMFC=-999 for all advected aerosol and chemistry species makes the BC configuration act as the proportional MF, which applies the mass conservation correction by a globally constant scaling factor.

The choices for SL scheme interpolation and the MF have been determined empirically for GHGs, aerosol and reactive gases and will be further reviewed. The more conservative choice for the proportional MF for aerosol and reactive gases, is motivated by the lack of stronger localised changes of the MMR compared to the BC MF. The main motivation for the choice of the BC MF for GHG is its tendency to modified less the very smooth background field throughout the atmosphere.

4.2 TURBULENT DIFFUSION, INJECTION OF SURFACE FLUXES AND CONVECTION

The vertical transport of tracer by diffusion follows the treatment of heat and moisture and is explained in more detail in Part IV Physical processes, Section 3.7.

The flux boundary conditions at surface are the added fluxes of emissions 3.1 from inventories and on-line simulation and from the dry deposition flux $-\rho V_d C$. It should be noted that the direct coupling of the surface emissions to the tracer diffusion scheme differs from the injection of emissions from elevated sources (see (d)). The elevated emissions lead to a positive mass mixing ration tendency in the model layers and the process of turbulent diffusion is applied on the updated MMR profile.

The tracer transport by cumulus convection is outlined in Part IV Physical processes, Section 6.2 based on the parameterised updraught and downdraught mass fluxes.

The routine for the tracer vertical diffusion is VDFDIFC called from VDFMAIN. The routine for the tracer convective transport is CUTRACER called from CUMASTRN.

Chapter 5

Diagnostic variables and radiative impact of aerosol

Table of contents

5.1 COMPUTATION OF AEROSOL OPTICAL DIAGNOSTICS

Aerosol optical depth (AOD) observations from the AERONET network (Holben *et al.*, 1998) or from VIIRS, MODIS, Sentinel 3 and other remote sensing instruments are the most common observation type to evaluate aerosol products, and AOD values retrieved from satellite observations are used for the aerosol data assimilation. Hence, the correctness of the calculation of the radiative impact of aerosol on atmospheric radiation is important for the evaluation and the successful assimilation of AOD satellite retrievals.

Further frequently observed in-situ parameter are particulate matter concentration below a diameter of 10 or 2.5 μm

5.1.1 Introduction

The aerosol optical diagnostics are computed in the routine AER_BDGTMSS, called from AER_PHY3. As shown in Figure 5.1, the aerosol optical diagnostics are computed using aerosol optical properties, which are computed offline with a Mie code for each aerosol species and bin, and for a selection of wavelengths. These aerosol optical properties are saved into a netcdf file, which is read as part of the radiation scheme. These aerosol optical properties are then stored into IFS arrays, which are used together with the simulated mass mixing ratio of each aerosol tracer and air density to compute the extinction by each aerosol species: it thus depends on its abundance (mass mixing ratio) and its mass extinction, which is dependent on the wavelength and differs significantly from one species to the other. In this section, we describe the inputs of the offline Mie code, and how they are used online to compute aerosol optical diagnostics.

5.1.2 Offline computations of aerosol optical properties

The Mie theory (Wiscombe, 1980) is a mathematical-physical theory of scattering of scattering of electromagnetic waves by homogeneous spherical particles. It is an adequate method to estimate the scattering, the absorbing and the extinction efficiency of spherical homogeneous particles, for particles

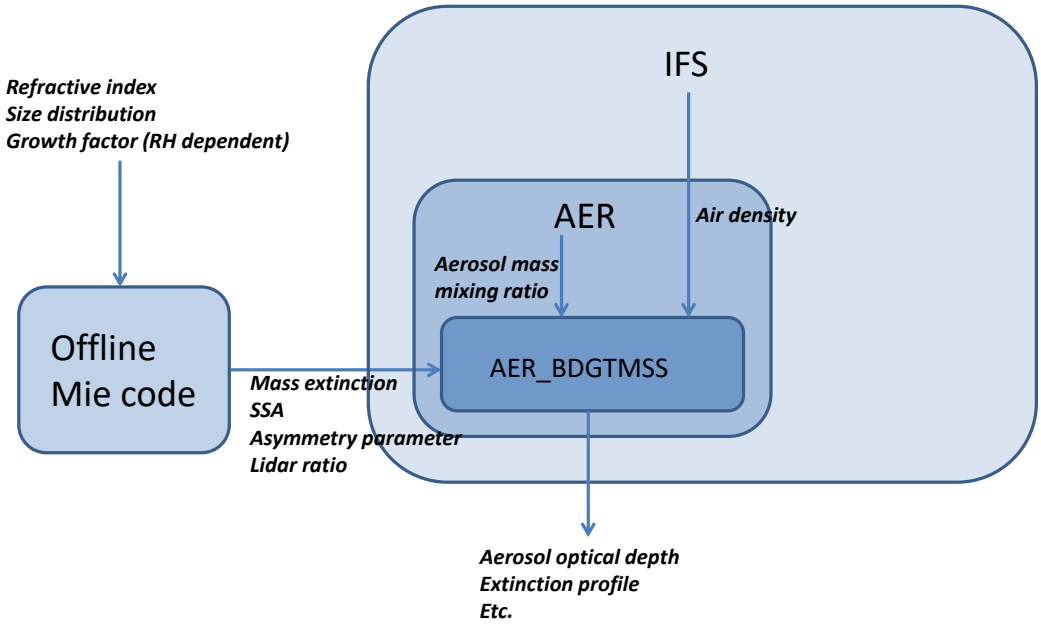


Figure 5.1 *Schematic showing how the aerosol optical diagnostics are computed, using aerosol optical properties that are computed offline with a Mie code, loaded and stored as arrays in the IFS, and used together with the mass mixing ratio of each aerosol species, and air density.*

that have a size in the same order of magnitude as the wavelength considered. From these efficiencies the following four bulk optical aerosol properties are derived which are used in the IFS to compute aerosol optical diagnostics: (i) mass extinction, (ii) single scattering albedo (SSA), (iii) asymmetry parameter and (iv) lidar ratio. The Mie theory uses the complex refractive indices as input, which depend on the chemical composition of the particles and information about the assumed size distribution of the particles and their density.

The sofware (fortran) that does the Mie theory calculation is used offline to compute the bulk aerosol optical properties for each species for each of the 14 short-wave (SW) and 16 long-wave (LW) bands of the RRTM radiation scheme on which the IFS radiation scheme is based (ecRad, (Hogan and Bozzo, 2018)) as well as for each of 20 wavelengths used for IFS-AER output. Spherical shape is assumed for all species. The number size distribution $n(r)$ is described by a log-normal function in general similar to the original version of the aerosol scheme (Reddy *et al.*, 2005):

$$n(r) = \frac{dN(r)}{dr} = \frac{N}{\sqrt{2\pi}r\ln(\sigma)} exp\left(-\frac{\ln^2(r/r_{mod})}{2\ln^2(\sigma)}\right) \tag{5.1}$$

with N total particle number concentration, σ geometric standard deviation and r_{mod} mode radius.

Table 5.1 lists the relevant parameters of the assumed number size distribution for each species. The assumed size distribution of sulfate has been updated in cycle 49R1. **It should be noted that this assumed size distribution is used for the offline computation of aerosol optical properties with the Mie code only**. For dust, the number size distribution specific are from (Ryder *et al.*, 2018). The bulk optical properties (mass extinction coefficient, single scattering albedo (ω) and asymmetry parameter (g)) are computed with a standard code for Mie scattering based on (Wiscombe, 1980). The size bin limits and density of the aerosol species is given in 2.3. For the hydrophilic types the optical properties change with the relative humidity due to the swelling of the water soluble component in wetter environments. The refractive index (m) and density (ρ) of the aerosol particle change according to the relations:

Table 5.1 *Aerosol species and parameters of the number size distribution associated to each aerosol type of AER as used in the off-line Mie calculations. (r_{mod} =mode radius, ρ=particle density, σ=geometric standard deviation), Number (where relevant).* ***Values are for the dry aerosol apart from sea salt which is given at 80% RH.****. The number size distribution is assumed to be monomodal for all species except for dust (four modes), sea salt aerosols and coarse mode nitrate for which a bimodal size distribution is assumed.*

Aerosol type	r_{mod} (μm)	σ	Number cm^{-3}
Sea Salt (80% RH)	0.1992,1.992	1.9,2.0	70,3
Dust	0.05,0.42,0.79,16.2	2.2,1.18,1.93,1.53	391,8.39,11.6,0.000138
Black carbon	0.0118	2.0	
Organic matter	0.09	1.6	
Sulfates	0.11	1.6	
Nitrate fine	0.0355	2.0	
Nitrate coarse	0.199,1.992	1.9,2.0	70,3
Ammonium	0.0355	2.0	
SOA	0.09	1.6	

$$\rho = \rho_{dry} * r_{dry}^3/r^3 + \rho_{water} * (r^3 - r_{dry}^3)/r^3 \tag{5.2a}$$

$$m = m_{water} + (m_{dry} - m_{water}) * r_{dry}^3/r^3 \tag{5.2b}$$

with r_{dry} and r the mode radius respectively of the dry particle and at a relative humidity value. The size distribution is modified applying growth factors to the mode radius and to the limits of integration, maintaining the same geometric standard deviation.

Information about the size distribution, particle density and refractive index is used in the Mie code to compute mass extinction, single scattering albedo (SSA) and asymmetry parameter. Figure 5.2 gives an example of mass extinction and absorption coefficients for desert dust.

5.1.3 Refractive index and growth factors for off-line Mie calculations

(*a*) *Sea-salt aerosol*

The sea-salt refractive index are from the Optical Properties of Aerosols and Clouds (OPAC) database (Hess *et al.*, 1998). They are shown in Table A.1.

The assumed hygroscopic growth is shown in Table 5.2

(*b*) *Desert dust*

The refractive index for dust for the 200-1000nm spectral range have been derived from a simulation using an experimental version of IFS-COMPO that represents twelve dust mineralogical species. For the 1000 to 3000nm range, we used values provided by (Balkanski *et al.*, 2007), who propose different values with different assumptions for the relative abundance of hematite (0.9, 1.5 and 2.7% hematite fraction of total volume). The refractive index chosen was with 0.9% hematite , which is their low value for hematite but which is much higher than the simulated fraction of dust that consists of hematite as provided by the climatology of dust mineralogy (0.22% average for 2017-2020). Finally, for the 3000-10000nm, the refractive indexes proposed by (Di Biagio *et al.*, 2017) have been used. Regional values are proposed that reflect the changing composition of dust; the values for the Bodele depression have been chosen.

The refractive index of dry dust is shown in table A.3.

In cycle 49R1, hydrophilic growth for dust has been introduced, following Chen *et al.* (2020), using a growth factor that increases linearly from 0.8% at 10% RH to 7.4% at 95% RH (see table 5.2. In cycle 50R1, the growth factors of organic matter, SOA, ammonium and fine mode nitrate have been updated.

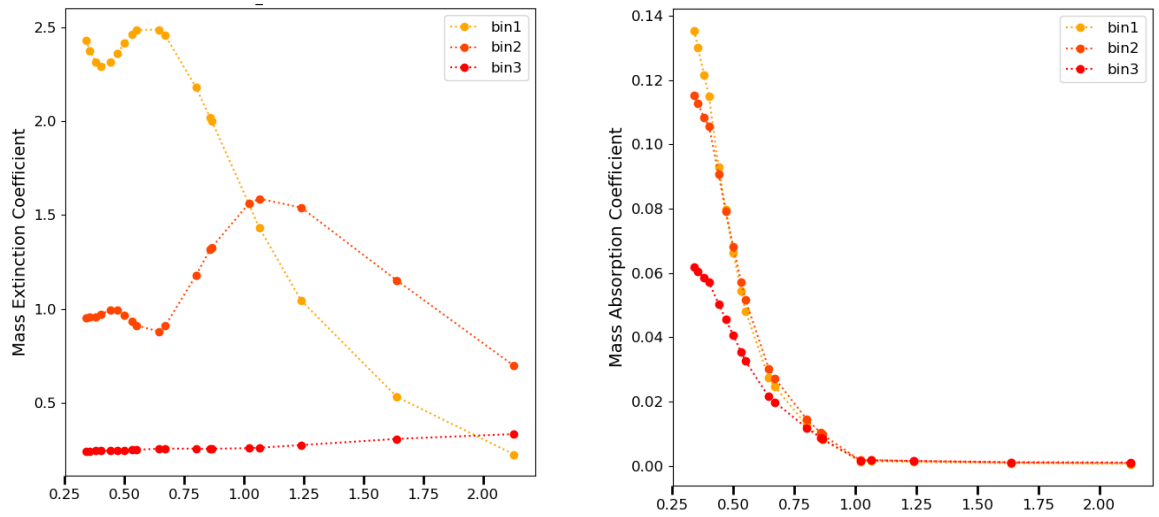


Figure 5.2 *Desert dust: mass extinction (left) and absorption (right) coefficients of spherical dust (as computed with the offline Mie code) as a function of wavelength in µm. Values are shown for the three bins, from fine to super-coarse. The mass absorption coefficient is computed as mass extinction x (1 - SSA). Values are shown for three dust bins.*

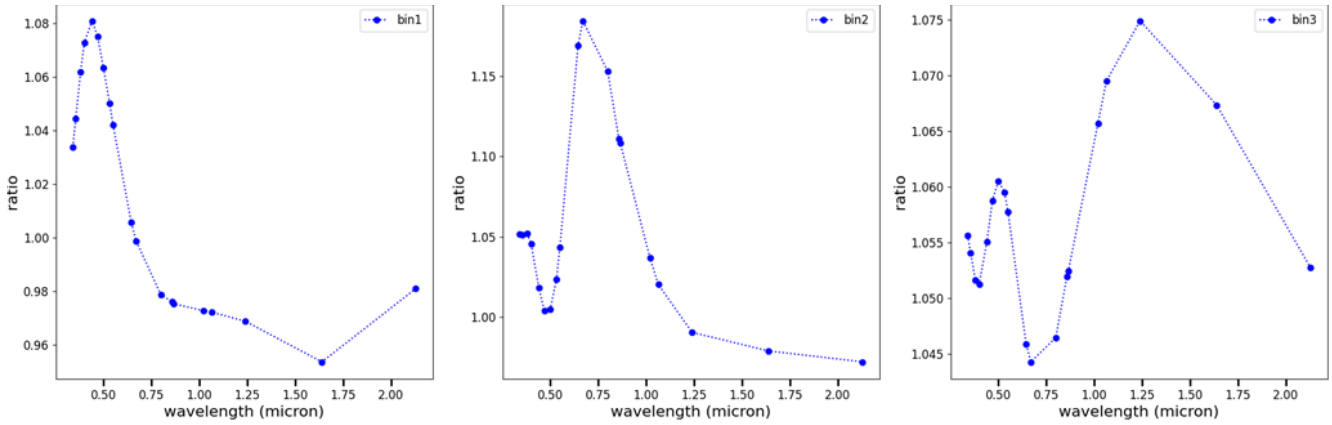


Figure 5.3 *Ratio of non-spherical vs spherical dust mass extinction as a function of wavelength, for each of the IFS(AER) three dust bins. The scaling coefficients shown here are applied to the mass extinction shown in Figure 5.2 to compute the aspherical mass extinction in IFS-COMPO.*

It is well know that desert dust are composed of mineralogical components and aggregates with very different shapes, which means that the assumed spherical shape of dust used in the Mie code (Wiscombe, 1980) that computes offline the aerosol optical properties (mass extinction, asymmetry parameter, single scattering albedo and lidar ratio) is clearly invalid. Using the online tool MOP-SMAP (Gasteiger and Wiegner, 2018), a scaling factor that accounts for the effect of asphericity has been computed, which is applied on the spherical desert dust mass extinction, SSA and asymmetry parameter computed with the Mie code. The assumed shape of the desert dust particles is derived from Kandler *et al.* (2009): a prolate spheroïd with an aspect ratio of 1.6. The ratio of non-spherical vs spherical dust mass extinction as a function of wavelength is shown in Figure 5.3 for each of the 3 dust bins. The implementation of non-spherical dust in cycle 49R1 leads to an increase in the dust mass extinction of 5-20% in the ultraviolet and visible parts of the spectrum for the fine and coarse bins, and of 5-10% for the super coarse bin.

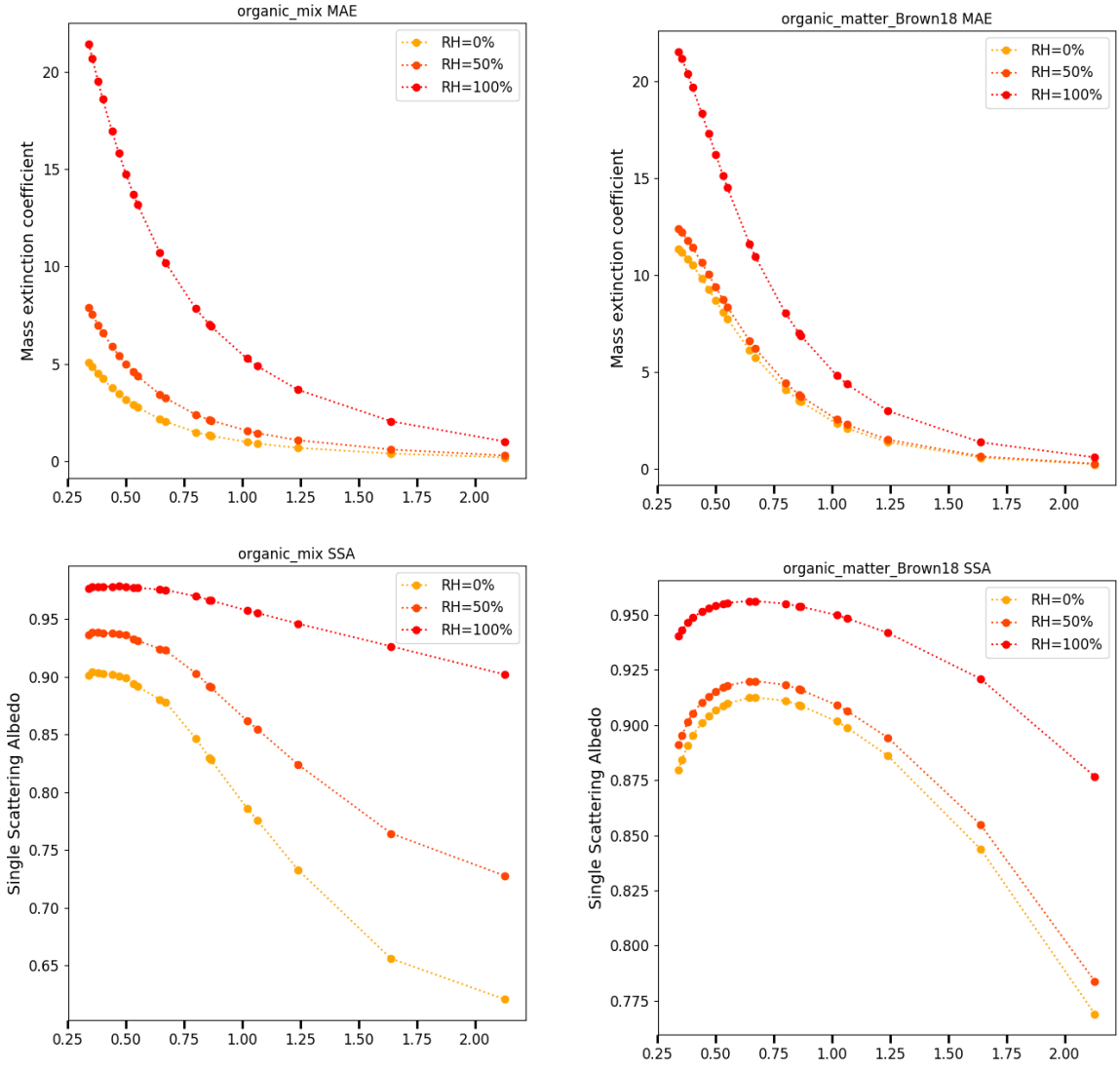


Figure 5.4 *Organic matter: mass extinction coefficient or efficiency in m^2/g (top) and single scattering albedo (bottom) as a function of wavelength in µm for cycle 47R3 and before (left) and 48R1 and later cycles (right). Values are shown for 0, 50 and 100 % relative humidity.*

(*c*) *Organic matter*

With the introduction of a specific Secondary Organic Aerosol species in Cy48r1, OM included only primary OM, while in previous cycles, OM included both primary and secondary organics. The optical properties from (Brown *et al.*, 2018) have been implemented for the OM species. Figure 5.4 shows how mass extinction and single scattering albedo are impacted by the new set of optical properties. Extinction is significantly higher with the new optical properties, especially for low relative humidities. The SSA, which determines how absorbing is OM, differs also significantly. The current set of optical properties give SSA that are continuously decreasing with increasing wavelength. The (Brown *et al.*, 2018) refractive index give a SSA that is lower and thus more absorbing OM for smaller wavelengths, in the UV and near-UV, which is a signature of brown carbon. This in turn translates into much higher mass absorption coefficient (MAC) for UV and to a lesser extent visible wavelengths. MAC in the IR is on the other hand lower with the (Brown *et al.*, 2018) refractive index as compared to the OM optics used in cycle 47R3 and before.

Table 5.2 *Hygroscopic growth of IFS-COMPO aerosol species*

RH/%	Desert-dust	Sea-salt	OM	BC	Sulfate	Ammonium	Nit. fine	Nit. coarse	SOA bio	SOA anthro
0–10	1	1	1	1	1	1	1	1	1	1
10–20	1.008	1	1	1	1	1	1	1	1	1
20–30	1.016	1	1.05	1	1	1	1	1.2	1.05	1.05
30–40	1.024	1	1.1	1	1	1	1	1.3	1.1	1.1
40–50	1.032	1.442	1.12	1	1.169	1.	1.1	1.35	1.15	1.115
50–60	1.038	1.555	1.13	1	1.220	1.	1.2	1.4	1.2	1.12
60–70	1.046	1.666	1.14	1	1.282	1.	1.25	1.5	1.21	1.125
70–80	1.054	1.799	1.15	1	1.363	1.31	1.3	1.6	1.24	1.13
80–85	1.062	1.988	1.18	1.2	1.485	1.43	1.35	1.7	1.29	1.145
85–90	1.066	2.131	1.2	1.3	1.581	1.57	1.5	1.8	1.35	1.155
90–95	1.070	2.361	1.25	1.4	1.732	1.8	1.7	2.0	1.46	1.17
95–100	1.074	2.876	1.4	1.5	2.085	1.9	2.1	2.2	1.47	1.18

Tables A.4 and 5.2 show the refractive index and hygroscopic growth factors used for organic matter. The hydrophobic tracer uses values at 20% relative humidity.

(*d*) *Black carbon*

The black carbon refractive index are from (Bond and Bergstrom, 2006), with a value at 500nm of 1.95 for the real part and 0.79i for the imaginary part.

(*e*) *Sulfate*

The refractive index is taken from the Global Aerosol Climatology Project (GACP, `http://gacp.giss.nasa.gov/data_sets/`) and it is representative of dry ammonium sulfate. The growth factors shown in Table 5.2 are from (Tang and Munkelwitz, 1994)

(*f*) *Nitrate*

The refractive index for nitrate from gas/particle partitioning (”nitrate 1” or ”fine mode nitrate”) use a spectrally constant value of 1.611 for the real part, and 0i for the imaginary part. The refractive index of nitrate from heterogeneous reactions (”nitrate 2” or ”coarse mode nitrate”) uses a spectrally constant value of 1.51 for the real part and 0.001i for the imaginary part of the complex refractive index. Its values are shown in Table A.5. The hygroscopic growth is shown in Table 5.2.

(*g*) *Ammonium*

Ammonium uses similar refractive index as for sulfate. No scaling factor is applied for extinction. Similarly, the growth factor used is the same as that of sulfate.

(*h*) *Secondary organic aerosol*

Following (Moise *et al.*, 2015), the anthropogenic and biogenic SOA tracers use slightly different real parts of the refractive index: a spectrally constant value of 1.5 for anthropogenic SOA and 1.4 for biogenic SOA. For the two tracers, a spectrally constant value of 0.01i is used for the imaginary part. The hygroscopic growth of SOA is taken from (Simon L. Clegg, 2019) and is shown in table 5.2

5.1.4 Aerosol optical depth calculation

AOD is computed online in the IFS in the AER_BDGTMSS routine. The mass extinction for each species (computed offline and stored as look-up tables in arrays) is multiplied to air density as well as the mass mixing ratio to obtain model level extinction for each aerosol species, which is then summed over all aerosol species to obtain total aerosol extinction at each model level. Vertical integration is then carried out to compute total AOD at each of the 20 following wavelengths: 340, 355, 380, 400, 440, 469, 500, 532, 555, 645, 670, 800, 858, 865, 1020, 1064, 1240, 1640, 2130 and 10000 nm.

The aerosol absorption is computed for each species and each model level by multiplying the simulated aerosol extinction by (1 - SSA) where SSA is the single scattering albedo computed offline by the Mie code. Similarly to AOD, this is then summed over species and integrated over the vertical to provide absorption AOD (AAOD) at each of the 20 wavelengths.

5.1.5 Other vertically integrated diagnostics in the IFS: single scattering albedo and asymmetry parameter

In addition to AOD, vertically integrated SSA and asymmetry parameter are also provided. They are computed as the sum of the values for each aerosol species. For each aerosol species, the vertically integrated SSA and asymmetry parameter is computed as the integral over the vertical of the species's SSA and asymmetry parameter (which is independent of height) weighted by the model level extinction coefficient for the concerned aerosol species. The vertically integrated SSA and asymmetry parameter are proposed for the 20 wavelengths mentioned above.

5.1.6 3D diagnostics: extinction coefficient, SSA and backscatter

If the namelist swith LAERLISI is set to true, a selection of 3D diagnostic of aerosol optical properties are also available over all of the model levels, only at the 355, 532 and 1064 nm wavelengths. These consists of:

- Extinction coefficient (sum of all species),
- Single Scattering Albedo, defined as 1 - AAOD/Extinction coefficient,
- Aerosol backscatter from top of atmosphere,
- Aerosol backscatter from surface,
- Unattenuated molecular backscatter coefficient,
- Unattenuated aerosol backscatter coefficient,

The 3D Extinction coefficient and SSA are computed in AER_BDGTMSS. The other 3D diagnostics are called in the routine AER_LIDSIM, called from AER_BDGTMSS if the LAERLISI namelist switch is true. The last four fields are computed using the lidar ratio, itself computed offline for each aerosol species by the Mie code.

5.2 PARTICULATE MATTER DIAGNOSTICS

Particulate Matter smaller than 1, 2.5 and 10 μm (PM1, PM2.5, PM10) are important outputs. They are computed with the following formulae that uses the mass mixing ratio from each aerosol tracer as an input, denoted

$[SS_{1,2,3}]$ for seasalt aerosol, $[DD_{1,2,3}]$ for desert dust, $[NI_{1,2}]$ for nitrate, $[OM]$, $[BC]$, $[SU]$, $[NI]$, $[AM]$, $[SOA]$ for Organic Matter, Black Carbon, Sulfate, Nitrate, Ammonium and SOA respectively :

$$PM_1 = \rho\Big(\frac{[SS_1]}{4.3} + 0.5[DD_1] + 0.96[OM] + 0.96[BC] + 0.91[SU] + 0.91[NI_1] + 0.91[AM] + 0.96[SOA]\Big)$$

$$PM_{2.5} = \rho\Big(\frac{[SS_1]}{4.3} + 0.6\frac{[SS_2]}{4.3} + [DD_1] + 0.15[DD_2] + [OM] + [BC] + [SU] + [NI_1] + 0.5[NI_2] + [AM] + [SOA]\Big)$$

$$PM_{10} = \rho\Big(\frac{[SS_1]}{4.3} + \frac{[SS_2]}{4.3} + 0.05\frac{[SS_3]}{4.3} + [DD_1] + [DD_2] + 0.4[DD_3] + [OM] + [BC] + [SU] + [NI_1] + [NI_2] + [AM] + [SOA]\Big)$$

where ρ is the air density. The sea-salt aerosol tracers are divided by 4.3 so as to transform the mass mixing ratio at 80% ambient relative humidity to dry mass mixing ratio. The PM formulae have been

updated in cycle 49R1 so that the assumed size distribution is consistent between the PM formula and the input of the offline Mie code used to compute the aerosol optical properties.

PM values only considering dust are a further diagnostic output of IFS-COMPO.

5.3 AEROSOL AND CHEMISTRY MASS DIAGNOSTICS

The following diagnostics are calculated for both aerosol and chemistry species:

- **Total column (all species) :** Vertical integration of the species mass .
- **Tropospheric column (chemical species only:** Vertical integration of the concentration of a given chemical species across all the tropospheric model levels. The tropopause is calculated based on the definition of the thermal tropopause height from the lapse rate. According to this definition, the tropopause height is the first level where the lapse rate is below 2 K/km, between 500 hPa and 70 hPa. Otherwise, it is given by the first model level with a full-level pressure below 500 hPa.
- **Accumulated dry deposition (chemical species)**
- **Accumulated wet deposition (chemical species)**
- **Instantaneous dry deposition (all species)**
- **Instantaneous wet deposition (chemical species)**
- **Instantaneous wet deposition by convective precipitation (aerosols)**
- **Instantaneous wet deposition by grid-scale precipitation (aerosols)**
- **Dry deposition velocity (chemical species)**
- **Surface mass concentration (chemical species):** Mass concentration [$\frac{kg}{m^3}$] at the lowest model level.

5.4 GLOBAL MASS DIAGNOSTICS

Besides diagnostic output as fields at the model resolution, the IFS can calculate global diagnostic of the (i) atmospheric and (ii) tropospheric burden and the global values of sink and source terms such as (iii) emissions, (iv) dry deposition, (v) wet deposition and (iv) chemical conversion. These values are stored in text files and are useful to compare global sink and source terms in a convenient way, without the need to retrieve grid-point fields. The global mass diagnostics work for all tracers and are switched on with LCHEM_DIA=true.

APPENDIX A. COMPLE REFRACTIVE INDEXES

Table A.1 *Refractive index for sea-salt aerosol at 0% relative humidity, as a function of wavelength*

Wavelength (m)	Real part	Imaginary part
0.250E-06	1.510000	5.000000e-06
0.300E-06	1.510000	2.000000e-06
0.350E-06	1.510000	3.240000e-07
0.400E-06	1.500000	3.000000e-08
0.450E-06	1.500000	2.430000e-08
0.500E-06	1.500000	1.550000e-08
0.550E-06	1.500000	1.000000e-08
0.600E-06	1.490000	1.600000e-08
0.650E-06	1.490000	4.240000e-08
0.700E-06	1.490000	2.000000e-07
0.750E-06	1.490000	1.080000e-06
0.800E-06	1.480000	1.950000e-06
0.900E-06	1.480000	4.240000e-05
1.000E-06	1.470000	1.410000e-04
1.250E-06	1.470000	3.580000e-04
1.500E-06	1.460000	5.700000e-04
1.750E-06	1.450000	7.620000e-04
2.000E-06	1.450000	1.000000e-03
2.500E-06	1.430000	4.000000e-03
3.000E-06	1.610000	1.000000e-02
3.200E-06	1.490000	3.000000e-03
3.390E-06	1.480000	2.050000e-03
3.500E-06	1.480000	1.600000e-03
3.750E-06	1.470000	1.400000e-03
4.000E-06	1.480000	1.400000e-03
4.500E-06	1.490000	1.400000e-03
5.000E-06	1.470000	2.500000e-03
5.500E-06	1.420000	3.600000e-03
6.000E-06	1.410000	1.100000e-02
6.200E-06	1.600000	2.200000e-02
6.500E-06	1.460000	5.000000e-03
7.200E-06	1.420000	7.000000e-03
7.900E-06	1.400000	1.300000e-02
8.200E-06	1.420000	2.000000e-02
8.500E-06	1.480000	2.600000e-02
8.700E-06	1.600000	3.000000e-02
9.000E-06	1.650000	2.800000e-02
9.200E-06	1.610000	2.620000e-02
9.500E-06	1.580000	1.800000e-02
9.800E-06	1.560000	1.600000e-02
10.000E-06	1.540000	1.500000e-02

Table A.2 *Refractive index for sea-salt aerosol at 0% relative humidity, as a function of wavelength*

Wavelength (m)	Real part	Imaginary part
10.600E-06	1.500000	1.400000e-02
11.000E-06	1.480000	1.400000e-02
11.500E-06	1.480000	1.400000e-02
12.500E-06	1.420000	1.600000e-02
13.000E-06	1.410000	1.800000e-02
14.000E-06	1.410000	2.300000e-02
14.800E-06	1.430000	3.000000e-02
15.000E-06	1.450000	3.500000e-02
16.400E-06	1.560000	9.000000e-02
17.200E-06	1.740000	1.200000e-01
18.000E-06	1.780000	1.300000e-01
18.500E-06	1.770000	1.350000e-01
20.000E-06	1.760000	1.520000e-01
21.300E-06	1.760000	1.650000e-01
22.500E-06	1.760000	1.800000e-01
25.000E-06	1.760000	2.050000e-01
27.900E-06	1.770000	2.750000e-01
30.000E-06	1.770000	3.000000e-01
35.000E-06	1.760000	5.000000e-01
40.000E-06	1.740000	1.000000e+00

Table A.3 *Refractive index for desert dust at 0% relative humidity, as a function of wavelength*

Wavelength (m)	Real part	Imaginary part
0.2000E-06	0.1530E+01	0.8000E-01
0.3000E-06	0.1530E+01	0.1800E-01
0.4000E-06	0.1530E+01	0.9000E-02
0.5000E-06	0.1530E+01	0.5700E-02
0.6000E-06	0.1530E+01	0.4700E-02
0.7000E-06	0.1530E+01	0.3600E-02
0.8000E-06	0.1520E+01	0.4300E-02
0.9000E-06	0.1520E+01	0.4300E-02
0.1000E-05	0.1520E+01	0.4500E-02
0.1536E-05	0.1400E+01	0.4500E-02
0.2000E-05	0.1260E+01	0.4500E-02
0.2250E-05	0.1220E+01	0.4500E-02
0.3000E-05	0.1160E+01	0.3500E-01
0.4000E-05	0.1260E+01	0.2100E-01
0.5000E-05	0.1250E+01	0.1500E-01
0.6000E-05	0.1150E+01	0.6000E-01
0.6500E-05	0.1130E+01	0.1000E+00
0.7200E-05	0.1400E+01	0.1700E+00
0.8000E-05	0.1150E+01	0.7800E-01
0.8200E-05	0.1130E+01	0.1100E+00
0.9000E-05	0.1700E+01	0.2650E+00
0.1000E-04	0.1750E+01	0.5360E+00
0.1150E-04	0.1590E+01	0.1200E+00
0.1200E-04	0.1550E+01	0.1240E+00
0.1300E-04	0.1470E+01	0.2380E+00
0.1720E-04	0.1630E+01	0.1920E+00
0.2000E-04	0.1680E+01	0.5360E+00
0.2500E-04	0.1970E+01	0.6000E+00
0.3000E-04	0.1800E+01	0.6300E+00
0.3500E-04	0.1900E+01	0.6300E+00
0.4000E-04	0.2100E+01	0.6300E+00

Table A.4 *Refractive index for organic matter, as a function of wavelength*

Wavelength (m)	Real part	Imaginary part
0.200E-06	1.7105	0.0294772786667372
0.400E-06	1.7045	0.0244591212418513
0.600E-06	1.6985	0.0219296481995005
0.800E-06	1.6925	0.0202952456597922
1.000E-06	1.6865	0.0191118687903374
1.200E-06	1.6805	0.0181963854318748
1.400E-06	1.6745	0.0174566535040615
1.600E-06	1.6685	0.0168402205589679
1.800E-06	1.6625	0.0163145816678297
2.000E-06	1.6565	0.0158582995800324
2.200E-06	1.6505	0.0154565453837116
2.400E-06	1.6445	0.0150986664160388
2.600E-06	1.6385	0.0147767716294706
2.800E-06	1.6325	0.0144848650840567
3.000E-06	1.6265	0.0142182921207056
3.200E-06	1.6205	0.0139733725439216
3.400E-06	1.6145	0.0137471501921122
3.600E-06	1.6085	0.0135372174458498
3.800E-06	1.6025	0.0133415894234761
4.000E-06	1.5965	0.0131586119771397
4.200E-06	1.5905	0.0129868932072291
4.400E-06	1.5845	0.0128252516724681
4.600E-06	1.5785	0.0126726766666984
4.800E-06	1.5725	0.0125282973586326
5.000E-06	1.5665	0.0123913585371517
5.200E-06	1.5605	0.0122612013454353
5.400E-06	1.5545	0.012137247828825
5.490E-06	1.5518	0.0120833546953075
5.990E-06	1.5368	0.0118030958451163
6.490E-06	1.5218	0.0115510613542545
6.990E-06	1.5068	0.0113225413536594
7.490E-06	1.4918	0.0111138810095052
7.990E-06	1.4768	0.0109221916183854
8.490E-06	1.4618	0.0107451534799538
8.990E-06	1.4468	0.0105808778246553
9.490E-06	1.4318	0.0104278078762683
9.990E-06	1.4168	0.0102846465235916
10.490E-06	1.4018	0.0101503024989424
11.990E-06	1.3568	0.00979155880834447
13.490E-06	1.3118	0.0094856952696089
14.990E-06	1.2668	0.00922021773661187
16.490E-06	1.2218	0.00898648599305493
17.990E-06	1.1768	0.00877829455865847
19.490E-06	1.1318	0.00859104752568938
20.990E-06	1.0868	0.00842125327771081
22.490E-06	1.0418	0.00826620162355337
23.990E-06	0.9968	0.00812374999946372

Table A.5 *Refractive index for sulfate at 0% RH, as a function of wavelength*

Wavelength (m)	Real part	Imaginary part
0.200E-06	1.565000	1.000000e-07
0.400E-06	1.540300	1.000000e-07
0.600E-06	1.527300	1.000000e-07
0.800E-06	1.519700	1.000000e-07
1.000E-06	1.512300	3.500260e-07
1.200E-06	1.504800	3.400170e-06
1.400E-06	1.497300	1.100020e-05
1.600E-06	1.489800	1.900200e-04
1.800E-06	1.482000	7.599760e-05
2.000E-06	1.473000	1.000000e-03
2.200E-06	1.463000	1.699810e-03
2.400E-06	1.449900	4.499870e-04
2.600E-06	1.428000	6.807700e-04
2.800E-06	1.391600	6.552390e-03
3.000E-06	1.289700	1.024240e-01
3.200E-06	1.490700	2.508420e-01
3.400E-06	1.607100	1.639830e-01
3.600E-06	1.627000	8.410070e-02
3.800E-06	1.570000	2.734010e-02
4.000E-06	1.543000	1.528970e-02
4.200E-06	1.523000	1.142350e-02
4.400E-06	1.505400	8.402340e-03
4.600E-06	1.490000	6.686520e-03
4.800E-06	1.476600	5.934720e-03
5.000E-06	1.462900	6.148940e-03
5.200E-06	1.445700	6.807330e-03
5.400E-06	1.426582	7.490120e-03
5.490E-06	1.418048	7.810950e-03
5.990E-06	1.353414	1.174680e-02
6.490E-06	1.224115	4.102270e-02
6.990E-06	1.004190	7.888260e-01
7.490E-06	1.610615	1.047550e-01
7.990E-06	1.315695	7.914190e-02
8.490E-06	0.859320	2.750770e-01
8.990E-06	0.923574	1.626800e+00
9.490E-06	2.725528	6.138340e-01
9.990E-06	2.197668	1.329510e-01
10.490E-06	1.993193	6.076220e-02
11.990E-06	1.792637	2.020310e-02
13.490E-06	1.688341	2.003190e-02
14.990E-06	1.543961	2.223400e-02
16.490E-06	2.757855	2.956680e-01
17.990E-06	1.783815	2.191460e-02
19.490E-06	1.708317	1.877450e-02
20.990E-06	1.65598	2.357060e-02
22.490E-06	1.612242	3.164590e-02
23.990E-06	1.577303	4.125900e-02

Chapter 6

Data assimilation of atmospheric composition retrievals

Table of contents

6.1 DATA ASSIMILATION METHOD

6.1.1 4DVAR method

The IFS uses an incremental four-dimensional variational (4D-Var) data assimilation system going back to Courtier *et al.* (1994). It is described in detail in Part II Data assimilation of the IFS documentation and we only repeat the basics here before describing composition-specific aspects, with particular focus on the background errors and the observation operators used for the atmospheric composition control variables. Several atmospheric composition fields, namely O3, NO2, CO, SO2, Volcanic SO2, HCHO, CO2, CH4 and aerosol total mixing ratio are included in the control vector and minimised together with the meteorological control variables. The IFS-COMPO o-suite uses 12-hour assimilation windows from 03 UTC to 15 UTC and 15 UTC to 03 UTC, and two minimisations at spectral truncations T159 ($\sim$ 110 km) and T255 ($\sim$ 78 km). The IFS-GHG o-suite uses the same assimilation window used in the NWP setup (09 UTC to 21 UTC and 21 UTC to 09 UTC) and performs three minimisations at spectral truncations T159 ($\sim$ 110 km), T191 ($\sim$ 100 km) and T255 ($\sim$ 78 km). The assimilation window of the IFS-COMPO setup differs from the ECMWF NWP setup to allow it to run timely enough to provide boundary conditions for the CAMS regional ensemble. The IFS-GHG o-suite (assimilating CH4 and CO2 retrievals) runs independently from the IFS-COMPO o-suite.

In the current atmospheric composition 4D-Var setup, a cost function that measures the differences between the model fields and the corresponding observations is minimised to obtain the best possible forecast through the length of the assimilation window by adjusting the initial conditions. In its incremental formulation (Courtier *et al.*, 1994) this cost function can be written as:

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^{\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{d})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}) \tag{6.1}$$

Here $\delta\mathbf{x}$ is the increment and at the minimum the resulting analysis increment $\delta\mathbf{x}^{\mathrm{a}}$ is added to the background $\mathbf{x}^{\mathrm{b}}$ in order to provide the analysis $\mathbf{x}^{\mathrm{a}}$ given by

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \delta\mathbf{x}^{\mathrm{a}} \tag{6.2}$$

B is the covariance matrix of background error while **d** is the innovation vector

$$\mathbf{d} = \mathbf{y}^{\mathrm{o}} - H\mathbf{x}^{\mathrm{b}} \tag{6.3}$$

where $\mathbf{y}^{\mathrm{o}}$ is the observation vector. **H** is a suitable low-resolution linear approximation of the observation operator H in the vicinity of $\mathbf{x}^{\mathrm{b}}$, and **R** is the covariance matrix of observation errors. The incremental formulation of 3D/4D-Var consists therefore of solving for $\delta\mathbf{x}$ the inverse problem defined by the (direct) observation operator **H**, given the innovation vector **d** and the background constraint.

6.1.2 COMPO/GHG control variables

The majority of the COMPO and GHG variables in the control vector are straightforward representations of the fields in question, e.g. O3, CO, CO2, CH4 and HCHO. However, aerosols, NO2, SO2 and Volcanic SO2 (VSO2) are more complex. The aerosol model used in the IFS-COMPO configuration of the IFS consists of 16 bins (Chapter 2.2) and yet only total AOD observations are used in the system. Total AOD does not contain sufficient information to constrain each of the individual bins and so instead a total aerosol mass mixing ratio, defined as the sum of the aerosol species, is used as the control variable in the assimilation process.

For NO2, a logarithmic control variable is used because if the analysis were based on a linear mixing ratio scale it would be prone to large extrapolation errors due to the high variability of NO2 in space and time.

SO2 and Volcanic SO2 both currently refer only to the SO2 coming from volcanic eruptions and NOT anthropogenic SO2. The difference between the two fields is that the SO2 is part of the full coupled chemistry-aerosol system and the observations are total column SO2 observations with a flag identifying them as volcanic. In contrast the Volcanic SO2 tracer simply has a prescribed lifetime of 7 days and is used with volcanic SO2 observations that contain information about the altitude of the volcanic plume.

6.1.3 Background error representation and estimation

In the IFS 4D-var data assimilation system, the background error covariance matrix is represented in operator form (Bonavita *et al.*, 2012), decomposing it into standard deviations and correlations,

$$\mathbf{B} = \mathbf{T}^{-1}\mathbf{\Sigma}_b^{1/2}\mathbf{C}\mathbf{\Sigma}_b^{1/2}\mathbf{T}^{-T} \tag{6.4}$$

where **T** is a matrix representation of the balance operator, $\mathbf{\Sigma}_b^{1/2}$ is the diagonal matrix of background error standard deviations and **C** is the background error correlation operator. The correlations, **C**, are given in a wavelet formulation (Fisher, 2004, 2006), which allows for both spatial and spectral variations of the background error covariances. The wavelet formulation is covered in detail in Part II Data assimilation, Section 4.2.1. The total background error covariance matrix is assumed to be block diagonal, so there are no correlations between variables.

In both the IFS-COMPO and IFS-GHG configurations, the correlations contained in the wavelet file are separated into those relevant for NWP control variables and those relevant for the COMPO or GHG control variables. The correlations used for the NWP fields are taken from the hybrid wavelet file produced operationally for NWP, in which climatological information is combined with the errors of the day coming from the EDA (see Part II Data assimilation, Section 6.7). The formulation of the part of the correlation operator related to both the GHG and the COMPO fields is discussed in more detail below.

The background error standard deviations, $\mathbf{\Sigma}_b^{1/2}$, are currently climatological time-constant values for both COMPO and GHG fields. They vary vertically through the atmosphere but are a globally constant value at each model level. The NWP standard deviations are derived from the errors of the day produced by the operational EDA and have values which vary both horizontally and vertically.

(*a*) *CO, HCHO, Aerosol*

The background error correlations and standard deviations for CO, HCHO and aerosols were calculated using the National Meteorological Center (NMC) method (Parrish and Derber, 1992). For this, 150 days

of 2-day IFS forecasts were run, and differences between pairs of 24- and 48-hour forecast fields were calculated whose statistical characteristics serve as proxy for the background errors.

For the chemical species, CO and HCHO, the background errors are directly calculated for each of the fields. For the total aerosol mixing ratio control variable, the background error standard deviations and correlations are for the total aerosol mass mixing ratio. The increment produced by the assimilation process for this total mass mixing ratio is repartitioned into the individual aerosol components according to their fractional contribution to the total aerosol mass (Benedetti *et al.*, 2009). This can lead to issues, which are documented in Flemming *et al.* (2017a).

(*b*) *O3, NO2, SO2, Volcanic SO2*

The background error correlations stored in the wavelet file for O3, log(NO2), SO2 and Volcanic SO2 are prescribed analytical horizontal-only correlations. There are no vertical correlations between neighbouring levels. Using an ensemble or NMC approach for O3 leads to long-range correlations that can produce unphysical increments and degrade the O3 profiles. Similarly, using an NMC ensemble approach for NO2 produces a lot of spurious long-range correlations that lead to a single observation of NO2 having an impact globally. This is not appropriate for a field that has very localised behaviour. SO2 and Volcanic SO2 observations are currently only assimilated in the IFS-COMPO o-suite in the event of volcanic eruptions. An NMC or ensemble approach would not give useful SO2 background error statistics in these cases as the forecast model does not have information about individual volcanic eruptions, even though it does include emissions from outgassing volcanoes.

For all four fields the wavelet file is formed of diagonal vertical wavenumber correlation matrices, with the value on the diagonal controlled by a horizontal Gaussian correlation function. The values of the elements on the diagonal of each of these vertical correlation matrices are the same at every level but vary for each wavenumber as prescribed by the Gaussian correlation function. For O3, the Gaussian correlation function has a lengthscale of 250km, for the log(NO2) variable it is 200km and for SO2 and Volcanic SO2 it is 100km. These lengthscales were chosen as they are a compromise between the wavelength that can be represented by the T159 and T255 inner loop grid resolutions of the IFS-COMPO configuration and the spread of the information from an observation.

The globally constant background error standard deviations are defined separately for the analytical wavelet fields and imposed on the correlations. For O3 this is an EDA derived vertical profile taken from the original MACC system (Inness and the MACC team, 2013). In line with the control variable, the background error standard deviations for the log(NO2) variable are in log space. They are artificially curtailed to be practically zero in the stratosphere. This is because only tropospheric NO2 columns are assimilated and reducing the background standard deviation to close to zero constrains the influence of the observations to the troposphere in the assimilation process.

The background error standard deviations for SO2 and Volcanic SO2 are also artificially prescribed profiles. Calculating these using the NMC or ensemble methods will lead to peaks near the surface where anthropogenic SO2 concentrations are largest and will hence lead to the largest increments near the surface. This is not appropriate for volcanic eruptions which are generally elevated in the atmosphere. For the SO2 control variable, the background error standard deviation profile is a delta function that peaks in the mid troposphere around model level 98 (about 550hPA) in the 137 level model version. This corresponds to an SO2 plume height of about 5 km (Inness *et al.*, 2022). This delta function is used since the total column, volcanic-flagged observations of SO2 contain no information about the height of the SO2 plume. Constraining the background errors to be zero everywhere apart from the immediate vicinity of model level 98 forces the increment to be applied at this height. For the 'Volcanic SO2' control variable, which makes use of the altitude of the volcanic SO2 layer, the background error standard deviation profile is a constant in height and the SO2 observation operator defines a model equivalent to the observation by calculating a total column between the pressure values that correspond to the bottom and the top of the retrieved volcanic SO2 layer.

(*c*) *CO2, CH4*

The background error correlations and standard deviations for CO2 and CH4 have been re-computed offline using a multi-seasonal sample based on 4 month-long EDA experiments of 10 members covering each a different season (January, April, July and October). The sample consisted of a total of 1200 forecast fields (300 per month), initialised at 00 and 12 every two days, with a lead time of 3 hours. The EDA ensemble was preferred to a NMC ensemble method which was strongly impacted by model biases. The vertical profile of standard deviation of CO2 has been tapered below 800 hPa to reduce the impact of undesired vertical correlations on the surface increments. A similar approach has been taken for CH4 in the stratosphere (e.g. above 100 hPa).

6.1.4 COMPO/GHG Observation operators

The observation operators provide the link between the analysis variables and the observations (Lorenc, 1986; Pailleux, 1990). The observation operator is applied to components of the model state to obtain the model equivalent of the observation, so that the model and observation can be compared like for like. The operator **H** in Eq. 6.1 above signifies the ensemble of all operators transforming the control variable **x** into the equivalent of each observed quantity, $\mathbf{y}^{o}$, at observation locations. More information about the observation related processing in the IFS can be found in Part II Data assimilation, Chapter 5. Here, only observation operators dealing with atmospheric composition data are described further. Currently, only satellite retrievals of atmospheric composition are assimilated in the IFS-COMPO and IFS-GHG o-suites, and in-situ atmospheric composition data are only used for validation.

(*a*) *Observation operators for chemical species*

The satellite retrievals for the chemical species that are part of the control vector (O3, CO, NO2, SO2, HCHO) are total or partial column data, i.e. integrated layers bounded by a top and a bottom pressure. The model's background column value is either calculated as a simple vertical integral between the top and the bottom pressure of the partial or total column, or it is determined by applying the averaging kernels of the retrievals, at the time and location of the observation. The column retrievals can be written using the averaging kernel **A**, that relates the true vertical profile $\mathbf{x}_t$ to the retrieved columns $\mathbf{y}^{o}$ as

$$\mathbf{y}^{o} = \mathbf{x}_{ap} + \mathbf{A}(\mathbf{x}_t - \mathbf{x}_{ap}) + \epsilon \tag{6.5}$$

where $\mathbf{x}_{ap}$ is an a-priori profile used in the retrieval of $\mathbf{y}^{o}$ and **ffl** an error term for measurement errors and errors in the forward model. In the observation operator we apply the averaging kernels **A** to the model profiles, $\mathbf{x}_m$, to smooth the model profiles according to the sensitivity of the retrievals. This means Eq. 6.3 can be written as

$$\mathbf{d} = \mathbf{x}_{ap} + \mathbf{A}(\mathbf{x}_t - \mathbf{x}_{ap}) + \epsilon = \mathbf{A}(\mathbf{x}_t - \hat{\mathbf{H}}(\mathbf{x}_m) \tag{6.6}$$

where $\hat{\mathbf{H}}$ is an operator to calculate layer values of the chemical species from the model profiles on the vertical grid of the a-priori profile. Using this observation operator, we remove the explicit influence of the a-priori profile in the calculation of the departures, but knowledge of the a-priori profile is still needed in the observation operator calculations. Also, the impact of the a-priori remains implicitly through the dependence of the retrieval $\mathbf{y}^{o}$ and the retrieval error on the a-priori, since a badly chosen a-priori will generally lead to larger retrieval errors and larger departures.

The master routine controlling the calls to the individual observation operators is called HOP and it calls the routine OBSOP_COMPOSITION which in turn calls the various observation operators for atmospheric composition data. Most chemical species are treated in the routine GRG_AK_OP (and the corresponding tangent-linear and adjoint routines GRG_AK_TL and GRG_AK_AD) where averaging kernels are applied to calculate the model equivalent of the observations. MOPITT CO data are treated separately in the routine MOPITT_AK_OP (and the corresponding tangent-linear and adjoint routines MOPITT_AK_TL and MOPITT_AK_AD) because the MOPITT averaging kernels work in log(VMR) space and a modified observation operator is required. There is a further routine MOPITT_PROFILE_AK (and the corresponding tangent-linear and adjoint routines

MOPITT_PROFILE_AK_TL and MOPITT_PROFILE_AK_AD) that deals with MOPITT profile retrievals. However, in the operational IFS-COMPO o-suite only MOPITT total column CO data (from the thermal infrared (TIR) retrieval) are assimilated. For some chemical species (currently O3 and SO2) no averaging kernels are applied and the model equivalent of the observation is calculated as a simple vertical integral between the top and bottom pressure values of the layer. This uses the routine PPNEW. There is also an observation operator to facilitate the assimilation of in-situ atmospheric composition data, for example IAGOS aircraft data. These data are currently not assimilated in the IFS-COMPO o-suite, but the relevant routines ISAC_GRG (and the corresponding tangent-linear and adjoint routines ISAC_GRGTL and ISAC_GRGAD) are in place and are also called from OBSOB_COMPOSITION.

(*b*) *Observation operators for aerosols*

The satellite retrievals used for aerosols are of Aerosol Optical Depth (AOD). The observation operator for AOD is based on precomputed optical properties (mass extinction coefficient, α_e, single scattering albedo, ω, and asymmetry parameter, g) for each of the aerosol species at each of the MODIS wavelengths λ (see Section 5.1.2 for full details of how the optical properties are calculated). The aerosols are assumed to be externally mixed. That is, the individual species are assumed to coexist in the volume of air considered and to retain their individual optical and chemical characteristics.

For the calculation of the model equivalent optical depth, the relative humidity (RH) is first computed from the model temperature, pressure and specific humidity. The appropriate mass extinction coefficients are then retrieved from the look-up table for the wavelength of interest (550nm as standard), multiplied by the aerosol mass mixing ratio which has been previously interpolated at the observation locations, and then integrated vertically. The total optical depth is the sum of the single-species optical depths as given by

$$\tau_\lambda = \sum_{i=1}^{N} \int_{p_{surf}}^{0} \alpha_{ei}(\lambda, RH(p)) r_i(p) \frac{dp}{g}, \tag{6.7}$$

where r is the mass mixing ratio, dp is the pressure of the model layer and g is the constant of gravity; p_{surf} represents the surface pressure. The number of aerosol species included in the calculation is given by N. The SO2 and Volcanic SO2 precursors are excluded from this calculation as they are not visible in the AOD observations.

The aerosol observation operator routine AOD_OP and the corresponding tangent-linear and adjoint routines AOD_TL and AOD_AD are all called from OBSOP_COMPOSITION.

6.2 COMPO/GHG ASSIMILATED OBSERVATION

More background information regarding the general treatment of observations in the IFS can be found in Part I Observations.

6.2.1 Observation data sets assimilated in the IFS-COMPO and IFS-GHG o-suites

Table 6.1 lists the satellite retrievals that have been assimilated in the IFS-COMPO o-suite between September 2014 and October 2024 and those assimilated in the (pre-operational) IFS-GHG o-suite since September 2021.

Table 6.1 *Satellite retrievals of reactive, greenhouse gases and aerosol optical depth that were actively assimilated in the IFS-COMPO and IFS-GHG o-suites up to Cy50R1.*

Instrument	Satellite	Provider	Version	Type	Status
MLS	AURA	NASA	V4 V5 V5-NN	O_3 profiles	20130107 - 20220207 20220207 - 20230216 20230216 -
OMI	AURA	NASA	V883	O_3 total column	20090901 -
GOME-2	Metop-A	EUMETSAT	GDP 4.8	O_3 total column	20131007 - 20181231
GOME-2	Metop-B	EUMETSAT	GDP 4.8	O_3 total column	20140512 - 20240703
GOME-2	Metop-C	EUMETSAT	GDP 4.9	O_3 total column	20200505 -
SBUV-2	NOAA-19	NOAA	V8	O_3 21-layer profiles	20121007 - 20201005
OMPS-NP	Suomi-NPP	NOAA / EUMETSAT		O_3 13-layer profiles	20170124 - 20190409 20201006 - 20250225
OMPS-NP	NOAA-20	NOAA / EUMETSAT		O_3 13-layer profiles	20201006 - 20201215 20230201 -
OMPS-NP	NOAA-21	NOAA / EUMETSAT		O_3 13-layer profiles	20250226 -
OMPS-LP	Suomi-NPP	NOAA / EUMETSAT	v2.5	O_3 limb profiles	20260201 -
TROPOMI	Sentinel-5P	ESA		O_3 column	20181204 -
IASI	MetOp-A	LATMOS/ULB EUMETSAT		CO total column	20090901 - 20180621 20180622 - 20191118
IASI	MetOp-B	LATMOS/ULB EUMETSAT		CO total column	20140918 - 20180621 20180622 -
IASI	MetOp-C	EUMETSAT		CO total column	20191119 -
MOPITT	TERRA	NCAR	V5-TIR V7-TIR V7-TIR Lance V8-TIR V9-TIR	CO total column	20130129 - 20160124 - 20180626 20180626 - 20190702 - 20211009 20211010 - 20250131
TROPOMI	Sentinel-5P	ESA		CO total column	20230627 -
OMI	AURA	KNMI	DOMINO V2.0	NO_2 tropospheric column	20120705 - 20210331
GOME-2	MetOp-A	EUMETSAT	GDP 4.8	NO_2 tropospheric column	20180626 - 20200504
GOME-2	MetOp-B	EUMETSAT	GDP 4.8	NO_2 tropospheric column	20180626 -
GOME-2	MetOp-C	EUMETSAT	GDP 4.9	NO_2 tropospheric column	20200505 -
TROPOMI	Sentinel-5P	ESA	v2.2.0/v2.3.1	NO_2 tropospheric column	20211013 -
GOME-2	MetOp-A	EUMETSAT	GDP 4.8	SO_2 total column	20150902 - 20191210
GOME-2	MetOp-B	EUMETSAT	GDP 4.8	SO_2 total column	20150902 - 20240703
GOME-2	MetOp-C	EUMETSAT	GDP 4.9	SO_2 total column	20200505 -
IASI	Metop-B	AC-SAF/EUMETSAT		Volcanic SO_2 total column	20241112 -
IASI	Metop-C	AC-SAF/EUMETSAT		Volcanic SO_2 total column	20241112 -
TROPOMI	Sentinel-5P	ESA		Volcanic SO_2 total column	20201006 -
MODIS	Aqua	NASA	Col 5 Col 6 Dark Target and Deep Blue Col 6.1 Dark Target and Deep Blue	Aerosol optical depth	20090901 - 20161114 - 20171015 -
MODIS	Terra	NASA	Col 5 Col 6 Dark Target and Deep Blue Col 6.1 Dark Target and Deep Blue	Aerosol optical depth	20090901 - 20161114 - 20171015 -
VIIRS	SNPP	NASA	V3r0 V3r2	Aerosol optical depth	20230201 - 20240516 20240516 -
VIIRS	NOAA-20	NASA	V3r0 V3r2	Aerosol optical depth	20230201 - 20240516 20240516 -
SLSTR	Sentinel-3A Sentinel-3B	EUMETSAT	Col 3.1	Aerosol optical depth	20260201 -
PMAp	METOP-A	EUMETSAT	v2.1	Aerosol Optical Depth	20170124 - 20210505
PMAp	METOP-B	EUMETSAT	v2.2.4 v2.2.5 v2.2.6	Aerosol Optical Depth	20210719 - 20230331 - 20240715 -
PMAp	METOP-C	EUMETSAT	v2.2.4 v2.2.5 v2.2.6	Aerosol Optical Depth	20210719 - 20230331 - 20240715 -
IASI	METOP-B	LMD	v4.0 v10.1	CO2 mid-tropospheric column	20210901 - 20220712 20220712 -
IASI	METOP-B	LMD	v8.1 v10.1	CH4 mid-tropospheric column	20210901 - 20220715 20220715 -
TANSO	GOSAT-1	IUP-UB	BESD FOCAL v3.0	CO2 total column	20210901 - 20220621 20220621-
TANSO	GOSAT-1	SRON	Proxy/SRPR Full Physics v2.3.8	CH4 total column	20210901 - 20220711 20220711 - 20241112
TANSO	GOSAT-2	SRON	Proxy/SRPR Full Physics v2.3.8	CH4 total column	20241112 -
TROPOMI	Sentinel-5p-1	ESA	v3.2.8	CH_4 total column	20260201 -

Instrument	Satellite	Provider	Version	Type
GEMS	KOMPSAT-2B	KMA	V3	O_3 total column
GEMS	KOMPSAT-2B	KMA	V3	NO_2 tropospheric column
GEMS	KOMPSAT-2B	KMA	V3	HCHO total column

Figure 6.1 *Timeseries of satellite retrievals assimilated in the IFS-COMPO o-suite between 2014 and September 2025. Green lines show observations for which averaging kernels are used. The numbers to the right of the plot list the report types that are used to identify the data in the observational data base (ODB), see https://apps.ecmwf.int/odbgov/all/ for more information.*

6.2.2 Quality control (variational quality control, first-guess check, blacklisting) and observation errors

The observation error and background error covariance matrices determine the relative weight given to the observation and the background in the analysis (see Eq. 6.1). For the chemical species, observation errors as given by the data providers are used. If the error values for species other than CO2 and CH4 are below 5 %, a minimum value of 5% is taken. The observation errors in both the IFS-COMPO and IFS-GHG o-suites are assumed to include any observation operator errors and representativeness errors that could arise because of differences in resolution of observation and the model, and that accounts for scales unresolved by the model. The observation errors for both COMPO and GHG species are given in the observation BUFR files and are further modified for some species in the routine REO3SIN. The screening processes to determine if an observation will be used in the analysis happen in the first trajectory run of each analysis cycle. They include the data selection criteria defined in the blocklist (formerly blacklist) as well as first-guess checks, variational quality control and data thinning. Most data selection criteria are coded in so called blocklist files, written in a convenient, readable blocklist language (see the Blocklist Documentation; Järvinen *et al.* (1996)). The blocklist mechanism is very flexible and allows nearly complete control of which data to use/not use in the assimilation. The observations are scanned through for blocklisting in the subroutine BLACK. At the set-up stage the blocklist interface is initialised (BLINIT) to the external blocklist library. The blocklist files consist formally of two parts. Firstly, the selection of variables for assimilation is specified in the 'data selection' part of the blocklist. This controls which observation types, variables, vertical ranges etc. will be selected for the assimilation. Some more complicated decisions are also performed through the data selection file; for instance, an orographic rejection limit is applied in the case of the observation being too deep inside the model orography;

data might be temporarily blocklisted because of algorithm upgrades. This part of the blocklist also provides a handy tool for experimentation with the observing system, as well as with the assimilation system itself. Secondly, a 'monthly monitoring' blocklist is provided for discarding the stations that have recently been reporting in an excessively noisy or biased manner compared with the ECMWF background field. This 'monthly blocklist' is maintained by the Forecast Department. Depending on the blocklisting criteria, flags are communicated to the routine BLACK, and those are written to the ECMA ODB data base. Blocklist-rejected data are subsequently excluded from the CCMA ODB and will not be present in the 4DVAR minimisation job steps. The operational blocklist history is kept in an archive.

The IFS-COMPO o-suite maintains its own copy of the operational blocklist, with a dedicated part containing blocklist criteria for the chemical species and aerosols, as well as the deactivation of IR ozone channels from AIRS, IASI, HIRS and CriS that are used in the ECMWF NWP system but not used by in the IFS-COMPO configuration. The IFS-GHG o-suite uses the operational NWP blocklist and adds a section for CO2 and CH4 observations. This happens in the script ADD_CAMS_BLACK.

Variational quality control (VarQC) and background error checks are carried out for the atmospheric composition data. General information about background quality control can be found in https://www.ecmwf.int/en/elibrary/19745-ifs-documentation-cy47r1-part-i-observations and about VarQC in https://www.ecmwf.int/en/elibrary/19746-ifs-documentation-cy47r1-part-ii-data-assimilation. Here we only highlight some of the points that apply to the atmospheric composition data, with more details given for specific species when necessary. The variational quality control, VarQC, has been described by Andersson and Järvinen (1999). It is a quality control mechanism which is incorporated within the variational analysis itself. A modification of the observation cost function to take into account the non-Gaussian nature of gross errors has the effect of reducing the analysis weight given to data with large departures from the current iterant (or preliminary analysis). Data are not irrevocably rejected but can regain influence on the analysis during later iterations if supported by surrounding data.

The switch LVARQC in DEFRUN (can be modified via the namelist NAMJO) determines if VarQC is carried out for a variable. VarQC is currently carried out for all atmospheric composition control variables except for SO2 in the IFS-COMPO configuration. Only volcanic SO2 observations are assimilated in the IFS-COMPO o-suite and these usually deviate a lot from the background SO2 values which do not include information about the volcanic eruption. (In a way the volcanic SO2 data are used as a source to bring the information about a volcanic eruption into the model.) Therefore, these observations will have large first-guess departures but should not be rejected or given reduced weight by the VarQC.

All observations are assigned an estimate of the background error in observation space for later use in the background quality control, and this estimate is stored in the ODB under fg_error. This estimate is only used to determine the expected variance of the background departures in the quality control against the background, and it is technically separate from the background error used during the assimilation for the control variables to determine the weighting of observations. In the routine FGCHK first-guess errors are set to 20% of the first-guess values for AOD and the chemical species, to 1% of the first-guess value for CO2 and to 2% of the first-guess value for CH4. The background quality control (FIRST) is performed for the variables that are intended to be used in the assimilation. The procedure is as follows. The variance of the background departure $\mathbf{y}^{o} - H\mathbf{x}^{b}$ can be estimated as a sum of observation and background-error variances $\boldsymbol{\sigma}_{o}^{2} + \boldsymbol{\sigma}_{b}^{2}$, assuming that the observation and the background errors are uncorrelated. After normalising with $\boldsymbol{\sigma}_{b}$, the estimate of variance for the normalised departure is given by $\mathbf{1} + (\boldsymbol{\sigma}_{o}^{2})/\boldsymbol{\sigma}_{b}^{2}$. In the background quality control, the square of the normalised background departure is considered as suspect when it exceeds its expected variance by more than a predefined multiple (FGCHK, SUFGLIM). These predefined limits (RBGQC) are set in DEFRUN and can be changed in namelist NAMJO. For the chemical species these factors currently are {9, 16, 25}; for aerosols {4, 9, 16}; while for GHG they depend on the species: they are {9, 16, 25} for CO2 and {2.25, 4, 9} for CH4. These values are also the default values used in the variational quality control to determine which observations are either rejected or given less weight. The RBGQC values mean that, e.g. for the chemical species, an observation is rejected if the square of the normalised background departure is greater than 5 and given

reduced weight if it is greater than 3. No first-guess check is carried out for SO2, again because the volcanic SO2 data will per se be showing large differences from the model background.

6.2.3 Bias correction and anchoring

Retrievals of the same parameter from different satellite instruments can have biases with respect to each other or to the model. Assimilating biased data violates one of the underlying assumptions of data assimilation, namely that the data should be unbiased, and therefore a bias correction scheme has to be applied to the data. Variational bias correction (VarBC, Dee (2004)) of observations was first introduced into the IFS in Cy31r1. VarBC works by including additional degrees of freedom (bias parameters) in the observational term of the 4D-Var cost function to account for possible systematic errors in selected observations and/or observation operators. The systematic errors (or biases) are represented by linear predictor models, which can be formulated separately for different groups of observations. More information about the general working of VarBC can be found in Part II Data assimilation, Chapter 5. VarBC is applied to several of the atmospheric composition data assimilated in the IFS-COMPO o-suite (see Table 6.3). Table 6.3 also lists the predictors used for the various instruments. Some data sets are used to 'anchor' the bias correction for a chemical species, i.e. are assimilated without bias correction. Past experience had shown that it is important to have such anchors for the bias correction to avoid drifts in the fields (Inness *et al.*, 2013). The VarBC settings for the atmospheric composition data are defined in the routine VARBC_TO3 and can be modified via the namelist NAMVARBC_TO3 in the scripts IFSTRAJ, IFSMIN or OOPSVAR. The predictors available for use in VarBC are defined in the routine VARBC_PRED.

6.2.4 Data thinning and super-obbing

Thinning of the atmospheric composition data is carried out in several places. First, there is the so-called pre-screening (carried out in the prepare_obs tasks under the obs family). Any data thinned here are discarded, i.e. not included in the ODB, and will not be available for use (either actively or passively) in the analysis. For atmospheric composition data this pre-screening happens in the task PREROE3 and uses the routine satrad/programs/REO3_PRESCREEN.

Alternatively, observation numbers can be reduced by averaging data to so-called super-observations. This happens in the task PREREO3_SUPEROB and is currently applied to TROPOMI, VIIRS, SLSTR and GEMS data. The super-obbing uses the routines satrad/programs/BUFR_GRID_SCREEN and satrad/programs/BUFR_SCREEN_REO3_SUPEROB.

In addition to the thinning in the pre-screening, data can also be thinned in the IFS (routine NEW_THINN_REO3). Data thinned here are included in the ODB as passive observations and therefore available for comparison against the model background. This thinning is controlled by settings in a file on the super computer $XDATA_CAMS/$IFS_CYCLE/sat/thin_reo3 (linked in the script MKLINKS). The chemical species and aerosol data that are not super-obbed are thinned here to 0.5°. This is currently applied to O3 data from GOME-2 and OMI, CO data from IASI and MOPITT, NO2 data from GOME-2, AOD data from MODIS and PMAp. GHG observations are currently thinned to 1° x 1° applied to CH4 and CO2 from IASI and TANSO. This screening is done by randomly selecting an observation in the grid box (i.e. the first observation the thinning comes across).

6.3 TANGENT LINEAR AND ADJOINT REPRESENTATION OF ATMOSPHERIC COMPOSITION PROCESSES

The tangent linear (TL) and adjoint (AD) models for the chemical processes in the IFS are based on a simplified version of the forward chemical model. The simplified chemical model consists of two types of reactions, namely, the NO_x photochemical equilibrium and the production of nitric acid (HNO_3) from NO_2 oxidation (NO_x sink):

$$NO + O_3 \longrightarrow NO_2 \tag{6.8}$$

$$NO_2 + h\nu \longrightarrow NO + O_3 \tag{6.9}$$

Table 6.3 *Variational Bias correction information for atmospheric composition data used in the IFS-COMPO o-suite (CY50R1). SOE denotes solar elevation. Thermal contrast is the temperature difference between the surface and the lowest model level. No bias correction is currently applied to GHG observations.*

Species	Satellite	Instrument	VarBC applied	Predictors
O3	Metop-C	GOME-2	yes	Constant, SOE
O3	Aura	OMI	yes	Constant, SOE
O3	Sentinel-5p	TROPOMI	yes	Constant, SOE
O3	NOAA-20	OMPS-NP	yes	Constant, SOE
O3	NOAA-21	OMPS-NP	yes	Constant, SOE
O3	SNPP	OMPS-LP	no	
O3	Aura	MLS	no	
CO	Metop-B	IASI	yes	Constant, 1000-300 hPa thickness, thermal contrast
CO	Metop-C	IASI	no	
CO	Sentinel-5p	TROPOMI	no	
NO2	Metop-B	GOME-2	yes	Constant, SOE
NO2	Metop-C	GOME-2	no	
NO2	Sentinel-5p	TROPOMI	no	
SO2 (volcanic)	Metop-C	GOME-2	no	
SO2 (volcanic)	Sentinel-5p	TROPOMI	no	
SO2 (volcanic)	Metop-B	IASI	no	
SO2 (volcanic)	Metop-C	IASI	no	
AOD	Aqua	MODIS	yes	Constant, surface wind over sea
AOD	Terra	MODIS	yes	Constant, surface wind over sea
AOD	SNPP	VIIRS	yes	Constant, surface wind over sea
AOD	NOAA-20	VIIRS	no	
AOD	Sentinel-3A	SLSTR	yes	Constant, surface wind over sea
AOD	Sentinel-3B	SLSTR	yes	Constant, surface wind over sea
AOD	METOP-B	PMAp	yes	Constant, surface wind over sea
AOD	METOP-C	PMAp	yes	Constant, surface wind over sea

$$NO_2 + OH \longrightarrow HNO_3 \tag{6.10}$$

Furthermore, in the reactions above, at each chemical time step, the concentrations of O_3 and OH are considered constant and are taken from the outer-loop trajectory computed in task IFSTRAJ. Therefore, at each time step, only increments in NO_2 and NO are propagated by the TL model during the minimisation (task IFSMIN), and, reciprocally, only adjoint forcings associated with NO_2 observations are propagated backward in time to the initial NO_2 and NO concentrations.

Starting from the main physics routines, the following routine calls implement the simplified chemistry TL/AD models:

- Tangent-linear integration:
  CALLPARTL>CHEM_MAIN_LAYER_TL>
  CHEM_MAIN_TL>CHEM_TM5_TL>TM5_O3CHEM_TL
- Adjoint integration:
  CALLPARAD>CHEM_MAIN_LAYER_AD>
  CHEM_MAIN_AD>CHEM_TM5_AD>TM5_O3CHEM_AD

The actual TL and AD codes associated with reactions 6.8 to 6.10 are implemented in TM5_O3CHEM_TL and TM5_O3CHEM_AD, respectively.

The chemistry TL/AD can be activated by setting LCHEM_TL=TRUE in prepIFS in the Atmospheric Composition section. It is activated by default in operation.

References

Ackermann, H. H. M. M. Z. C., I.J and Ebel, A. (1995). The parameterization of the sulfate-nitrate-ammonia system in the long-range transport model eurad. *Meteorol. Atmos. Phys.*, **57**, 101–114.

Agusti-Panareda, A., Barré, J., Massart, S., Inness, A., Aben, I., Ades, M., Baier, B. C., Balsamo, G., Borsdorff, T., Bousserez, N., Boussetta, S., Buchwitz, M., Cantarello, L., Crevoisier, C., Engelen, R., Eskes, H., Flemming, J., Garrigues, S., Hasekamp, O., Huijnen, V., Jones, L., Kipling, Z., Langerock, B., McNorton, J., Meilhac, N., Noel, S., Parrington, M., Peuch, V.-H., Ramonet, M., Ratzinger, M., Reuter, M., Ribas, R., Suttie, M., Sweeney, C., Tarniewicz, J. and Wu, L. (2023). Technical note: The cams greenhouse gas reanalysis from 2003 to 2020. *Atmos. Chem. Phys.*, **23**, 3829–3859, doi:https://doi.org/10.5194/acp-23-3829-2023.

Agusti-Panareda, A., Diamantakis, M., Bayona, V., Klappenbach, F. and Butz, A. (2017). Improving the inter-hemispheric gradient of total column atmospheric co2 and ch4 in simulations with the ecmwf semi-lagrangian atmospheric global model. *Geoscientific Model Development*, **10**(1), 1–18, doi:10.5194/gmd-10-1-2017, URL https://gmd.copernicus.org/articles/10/1/2017/.

Agustí-Panareda, A., Massart, S., Chevallier, F., Balsamo, G., Boussetta, S., Dutra, E. and Beljaars, A. (2016). A biogenic co_2 flux adjustment scheme for the mitigation of large-scale biases in global atmospheric co2 analyses and forecasts. *Atmospheric Chemistry and Physics*, **16**(16), 10399–10418, doi:10.5194/acp-16-10399-2016, URL https://acp.copernicus.org/articles/16/10399/2016/.

Agustí-Panareda, A., Massart, S., Chevallier, F., Boussetta, S., Balsamo, G., Beljaars, A., Ciais, P., Deutscher, N. M., Engelen, R., Jones, L., Kivi, R., Paris, J.-D., Peuch, V.-H., Sherlock, V., Vermeulen, A. T., Wennberg, P. O. and Wunch, D. (2014). Forecasting global atmospheric co_2. *Atmos. Chem. Phys.*, **14**, 11959–11983.

Albert, M. F. M. A., Anguelova, M. D., Manders, A. M. M., Schaap, M. and de Leeuw, G. (2016). Parameterization of oceanic whitecap fraction based on satellite observations. *Atmos. Chem. Phys.*, **16**, 13725–13751.

Andersson, E. and Järvinen, H. (1999). Variational quality control. *Q. J. R. Meteorol. Soc.*, **125**, 697–722.

Anguelova, M. D. and Webster, F. (2006). Whitecap coverage from satellite measurements: A first step toward modeling the variability of oceanic whitecaps. *J. Geophys. Res*, **111**, C03017.

Ayers, G., Gillett, R. and Gras, J. (1980). On the vapor pressure of sulfuric acid. *Geophysical Research Letters*, **7(6)**, 433–436.

Balkanski, Y., Schulz, M., Claquin, T. and Guibert, S. (2007). Reevaluation of mineral aerosol radiative forcings suggests a better agreement with satellite and aeronet data. *Atmos. Chem. Phys.*, **7**, 81–95.

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J. *et al.* (2009). Aerosol analysis and forecast in the european centre for medium-range weather forecasts integrated forecast system: 2. data assimilation. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **114**(D13).

Bergamaschi, P., Frankenberg, C., Meirink, J. F., Krol, M., Villani, M. G., Houweling, S., Dentener, F., Dlugokencky, E. J., Miller, J. B., Gatti, L. V., Engel, A. and Levin, I. (2009). Inverse modeling of global and regional ch4 emissions using sciamachy satellite retrievals. *Journal of Geophysical*

Research: Atmospheres, **114**(D22), doi:https://doi.org/10.1029/2009JD012287, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD012287.

Bergametti, G., Marticorena, B., Rajot, J. L., Foret, G., Alfaro, S. C. and Laurent, B. (2018). Size-resolved dry deposition velocities of dust particles: In situ measurements and parameterizations testing. *J. Geophys. Res.*, **123**, 11080–11099.

Bermejo, R. and Conde, J. (2002). A conservative quasi-monotone semi-lagrangian scheme. *Monthly Weather Review*, **130**(2), 423 – 430, doi:https://doi.org/10.1175/1520-0493(2002)130⟨0423:ACQMSL⟩2.0.CO;2, URL https://journals.ametsoc.org/view/journals/mwre/130/2/1520-0493_2002_130_0423_acqmsl_2.0.co_2.xml.

Bonavita, M., Isaksen, L. and Holm, E. (2012). On the use of eda background error variances in the ecmwf 4d-var. *QJRMS*, **138**, 1540–1559.

Bond, T. and Bergstrom, R. (2006). Light absorption by carbonaceous particles: An investigative review. *Aerosol Science and Technology*, **40(1)**, 27–67.

Boucher, O., Pham, M. and Venkataraman, C. (2002). Simulation of the atmospheric sulfur cycle in the laboratoire de meteorologie dynamique general circulation model: Model description, model evaluation, and global and european budgets. *Note Scientifique de l'IPSL*, **23**, 27 pp.

Brasseur, G. P., Smith, A. K., Khosravi, R., Huang, T., Walters, S., Chabrillat, S. and Kockarts, G. (2000). Natural and human-induced perturbations in the middle atmosphere: A short tutorial. *Washington DC American Geophysical Union Geophysical Monograph Series*, **123**, 7–20.

Brown, H., Liu, X., Feng, Y., Jiang, Y., Wu, M., Lu, Z., Wu, C., Murphy, S. and Pokhrel, R. (2018). Radiative effect and climate impacts of brown carbon with the community atmosphere model (cam5). *Atmos. Chem. Phys.*, **18**, 17745–17768.

Burkholder, J., Sander, S., Abbatt, J., Barker, J., Cappa, C., Crounse, J., Dibble, T., Huie, R., Kolb, C., Kurylo, M. *et al.* (2020). Chemical kinetics and photochemical data for use in atmospheric studies; evaluation number 19. *Publication 19–5*, Jet Propulsion Laboratory, NASA, URL http://jpldataeval.jpl.nasa.gov/.

Burkholder, J., Sander, S., Abbatt, J., Barker, J., Huie, R., Kolb, C., Kurylo, M., Orkin, V., Wilmouth, D. and Wine, P. (2015). Chemical kinetics and photochemical data for use in atmospheric studies: evaluation number 18. *Technical report*, Pasadena, CA: Jet Propulsion Laboratory, National Aeronautics and Space Administration, URL https://jpldataeval.jpl.nasa.gov/previous_evaluations.html.

Carn, S., Fioletov, V., McLinden, C., Li, C. and Krotkov, N. (2017). A decade of global volcanic so2 emissions measured from space. *Scientific reports*, **7**(1), 44095.

Chabrillat, S. and Fonteyn, D. (2003). Modelling long-term changes of mesospheric temperature and chemistry. *Advances in Space Research*, **32**(9), 1689–1700, doi:https://doi.org/10.1016/S0273-1177(03)90464-9, URL https://www.sciencedirect.com/science/article/pii/S0273117703904649.

Chen, L., Peng, C., Gu, W., Fu, H., Jian, X., Zhang, H., Zhang, G., Zhu, J., Wang, X. and Tang, M. (2020). On mineral dust aerosol hygroscopicity. *Atmos. Chem. Phys.*, **20**, 13611–13626, doi:10.5194/acp-20-13611-2020.

Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B., Duncan, B. and Martin, R. (2002). Tropospheric aerosol optical thickness from the gocart model and comparisons with satellite and sun photometer measurements. *Journal Atmospheric Sciences*, **59**, 461–483.

Chung, S. H. and Seinfeld, J. H. (2002). Global distribution and climate forcing of carbonaceous aerosols. *Journal of Geophysical Research: Atmospheres*, **107**(D19), AAC 14–1–AAC 14–33, doi:https://doi.org/10.1029/2001JD001397, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JD001397.

Courtier, P., Thépaut, J.-N. and Hollingsworth, A. (1994). A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120**, 1367–1388.

Croft, B., Lohmann, U., Martin, R. V., Stier, P., Wurzler, S., Feichter, J., Posselt, R. and Ferrachat, S. (2009). Aerosol size-dependent below-cloud scavenging by rain and snow in the echam5-ham. *Atmos. Chem. Phys.*, **9**, 4653–4675.

de Bruine, M., Krol, M., van Noije, T., Le Sager, P. and Röckmann, T. (2018). The impact of precipitation evaporation on the atmospheric aerosol distribution in ec-earth v3.2.0. *Geosci. Model Dev.*, **11**, 1443–1465.

Dee, D. (2004). Variational bias correction of radiance data in the ECMWF system. In *Proc. of the ECMWF Workshop on Assimilation of High Spectral Resolution Sounders in NWP*, pp. 97–112, Reading, 28 June–1 July, 2004.

Denier van der Gon, H., Arellano, S., Darras, S., Dellaert, S., Gauss, M., Granier, C., Guevara, M., Jalkanen, J.-P.-., Kuenen, J., Liousse, C., Markova, J., Quack, B., Simpson, D., Sindelarova, K. and Soulie, A. (2021). Global and regional emissions documentation of products. *CAMS reports*.

Di Biagio, C., Formenti, P., Balkanski, Y., Caponi, L., Cazaunau, M., Pangui, E., Journet, E., Nowak, S., Caquineau, S., Andreae, M. O., Kandler, K., Saeed, T., Piketh, S., Seibert, D., Williams, E. and Doussin, J.-F. (2017). Global scale variability of the mineral dust long-wave refractive index: a new dataset of in situ measurements for climate modeling and remote sensing. *Atmos. Chem. Phys.*, **17**, 1901–1929.

Diamantakis, M. and Agusti-Panareda, A. (2017). A positive definite tracer mass fixer for high resolution weather and atmospheric composition forecasts. doi:10.21957/qpogzoy, URL https://www.ecmwf.int/node/17914.

Diamantakis, M. and Flemming, J. (2014). Global mass fixer algorithms for conservative tracer transport in the ecmwf model. *Geoscientific Model Development*, **7**, 965–979.

Drakaki, E., Amiridis, V., Tsekeri, A., Gkikas, A., Proestakis, E., Mallios, S., Solomos, S., Spyrou, C., Marinou, E., Ryder, C., Bouris, D. and Katsafados, P. (2022). Modeling ocarse and giant desert dust particles. *Atmos. Chem. Phys.*, **22**, 12727–12748.

ECMWF (2021a). *IFS Documentation CY47R3 - Part I: Observations*. 1, ECMWF, doi:10.21957/ycow5yjr1, URL https://www.ecmwf.int/node/20195.

ECMWF (2021b). *IFS Documentation CY47R3 - Part II: Data assimilation*. 2, ECMWF, doi:10.21957/t445u8kna, URL https://www.ecmwf.int/node/20196.

ECMWF (2021c). *IFS Documentation CY47R3 - Part III Dynamics and numerical procedures*. 3, ECMWF, doi:10.21957/b18qsx663, URL https://www.ecmwf.int/node/20197.

ECMWF (2021d). *IFS Documentation CY47R3 - Part IV Physical processes*. 4, ECMWF, doi:10.21957/eyrpir4vj, URL https://www.ecmwf.int/node/20198.

Engelen, R. J., Serrar, S. and Chevallier, F. (2009). Four-dimensional data assimilation of atmospheric co_2 using airs observations. *Journal of Geophysical Research: Atmospheres*, **114**(D3), doi:10.1029/2008JD010739, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD010739.

Errera, Q., Chabrillat, S., Christophe, Y., Debosscher, J., Hubert, D., Lahoz, W., Santee, M. L., Shiotani, M., Skachko, S., von Clarmann, T. and Walker, K. (2019). Technical note: Reanalysis of aura mls chemical observations. *Atmospheric Chemistry and Physics*, **19**(21), 13647–13679, doi:10.5194/acp-19-13647-2019, URL https://acp.copernicus.org/articles/19/13647/2019/.

Errera, Q. and Fonteyn, D. (2001). Four-dimensional variational chemical assimilation of crista stratospheric measurements. *Journal of Geophysical Research: Atmospheres*, **106**(D11), 12253–12265, doi:https://doi.org/10.1029/2001JD900010, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JD900010.

Farquhar, G. D., von Caemmerer, S. and Berry, J. A. (1980). A biochemical model of photosynthetic co2 assimilation in leaves of c3 species. *Planta*, **149**(1), 78–90, doi:10.1007/BF00386231.

Fecan, F., Marticorena, B. and Bergametti, G. (1999). Parameterization of the increase of aeolian erosion threshold wind friction velocity due to soil moisture for arid and semi-arid areas. *Ann. Geophys.*, **17**, 149–157.

Feierabend, K. J., Havey, D. K., Brown, S. S. and Vaida, V. (2006). Experimental absolute intensities of the 4nu9 and 5nu9 o–h stretching overtones of h2so4. *Chemical Physics Letters*, **420**(4), 438 – 442, doi:10.1016/j.cplett.2006.01.013, URL http://www.sciencedirect.com/science/article/pii/S0009261406000546.

Fisher, M. (2004). Generalized frames on the sphere, with application to the background error covariance modelling. In *Proc. ECMWF Seminar on Recent Developments in Numerical Methods for Atmosphereic and Ocean Modelling*, pp. 87–102, Reading, 6–10 September 2004.

Fisher, M. (2006). Wavelet jb - a new way to model the statistics of background errors. *ECMWF Newsletter No. 106*, pp. 23–28.

Flemming, J., Benedetti, A., Inness, A., Engelen, R. J., Jones, L., Huijnen, V., Remy, S., Parrington, M., Suttie, M., Bozzo, A. *et al.* (2017a). The cams interim reanalysis of carbon monoxide, ozone and aerosol for 2003–2015. *Atmospheric Chemistry and Physics*, **17**(3), 1945–1983.

Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Josse, B., Katragkou, E., Marecal, V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O. and Tsikerdekis, A. (2015). Tropospheric chemistry in the integrated forecasting system of ecmwf. *Geosci. Model Dev.*, **8**, 975–1003.

Flemming, J., Inness, A., Flentje, H., Huijnen, V., Moinat, P., Schultz, M. G. and Stein, O. (2009). Coupling global chemistry transport models to ecmwf's integrated forecast system. *Geoscientific Model Development*, **2**, 253–265.

Flemming, J., Peuch, V. and Jones, L. (2017b). Ten years of forecasting atmospheric composition at ecmwf. *ECMWF Newsletter*, **152**, URL https://www.ecmwf.int/en/newsletter/152/news/ten-years-forecasting-atmospheric-composition-ecmwf.

Fonteyn, D. and Larsen, N. (1996). Detailed PSC formation in a two-dimensional chemical transport model of the stratosphere. **14**, 315–328.

Forbes, R., Tompkins, A. and Untch, A. (2011). A new prognostic bulk-microphysics scheme for the ifs. *ECMWF Tech. Memo.*, **649**.

Fountoukis, C. and Nenes, A. (2007). Isorropia ii: a computationally efficient thermodynamic equilibrium model for k^+–ca^{2+}–mg^{2+}–nh_4^+–na^+–so_4^{2-}–no_3^-–cl^-–h_2o aerosols. *Atmos. Chem. Phys.*, **7**, 4639–4659.

Fuchs, N. and Sutugin, A. (1970). *Highly Dispersed Aerosols*. Butterworth-Heinemann, Newton, Mass., USA.

Gasteiger, J. and Wiegner, M. (2018). Mopsmap v1.0: a versatile tool for the modeling of aerosol optical properties. *Geosci. Model Dev.*, **11**, 2739–2762, doi:10.5194/gmd-11-2739-2018.

Gauss, M., Isaksen, I. S. A., Lee, D. S. and Søvde, O. A. (2006). Impact of aircraft nox emissions on the atmosphere ; tradeoffs to reduce the impact. *Atmospheric Chemistry and Physics*, **6**(6), 1529–1548, doi:10.5194/acp-6-1529-2006, URL https://acp.copernicus.org/articles/6/1529/2006/.

Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., van Vuuren, D. P., van den Berg, M., Feng, L., Klein, D., Calvin, K., Doelman, J. C., Frank, S., Fricko, O., Harmsen, M., Hasegawa, T., Havlik, P., Hilaire, J., Hoesly, R., Horing, J., Popp, A., Stehfest, E. and Takahashi, K. (2019). Global emissions pathways under different socioeconomic scenarios for use in cmip6: a dataset of harmonized emissions trajectories through the end of the century. *Geosci. Model Dev.*, **12**, 1443–1475.

Gilette, D. A. (1979). Environmental factors affecting dust emission by wind erosion.

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O. and Lin, S.-J. (2001). Sources and distributions of dust aerosols simulated with the gocart model. *Journal of Geophysical Research: Atmospheres*, **106**(D17), 20255–20273.

Ginoux, P., Prospero, J. M., Gill, T. E., Hsu, N. C. and Zhao, M. (2012). Global-scale attribution of anthropogenic and natural dust sources and their emission rates based on modis deep blue aerosol products. *Reviews of Geophysics*, **50**(3), doi:10.1029/2012RG000388, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012RG000388.

Giorgi, F. and Chameides, W. L. (1986). Rainout lifetimes of highly soluble aerosols and gases as inferred from simulations with a general circulation model. *J. Geophys. Res.*, **91**, 367–376.

Gkikas, A., Proestakis, E., Amiridis, V., Kazadzis, S., Di Tomaso, E., Tsekeri, A., Marinou, E., Hatzianastassiou, N. and Pérez García-Pando, C. (2020). Modis dust aerosol (midas): a global fine-resolution dust optical depth data set. *Atmos. Meas. Tech.*, **14**, 309–334.

Gong, S. (2003). A parameterization of sea-salt aerosol source function for sub- and super-micron particles. *Global Biogeochem. Cycles*, **17**, 1097.

Granier, C., Lamarque, J. F., Mieville, A., Muller, J. F., Olivier, J., Orlando, J., Peters, J., Petron, G., Tyndall, G. and Wallens, S. (2005). Poet, a database of surface emissions of ozone precursors. *available at: http://eccad.aeris-data.fr/ (last access: 25 April 2019)*.

Granier, C., Soulie, A. and Darras, S. (2022). Global emissions dataset for latest year available (2023). *CAMS reports*.

Grythe, H., Ström, J., Krejci, R., Quinn, P. and Stohl, A. (2014). A review of sea-spray aerosol source functions using a large global set of sea salt aerosol concentration measurements. *Atmos. Chem. Phys.*, **14**, 1277–1297.

Guenther, A., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T. a., Emmons, L. and Wang, X. (2012). The model of emissions of gases and aerosols from nature version 2.1 (megan2. 1): an extended and updated framework for modeling biogenic emissions. *Geoscientific Model Development*, **5**(6), 1471–1492.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I. and Geron, C. (2006). Estimates of global terrestrial isoprene emissions using megan (model of emissions of gases and aerosols from nature). *Atmospheric Chemistry and Physics*, **6**(11), 3181–3210.

Guevara, M., Jorba, O., Tena, C., Denier van der Gon, H., Kuenen, J., Elguindi, N., Darras, S., Granier, C. and Pérez García-Pando, C. (2021). Copernicus atmosphere monitoring service temporal profiles (cams-tempo): global and european emission temporal profile maps for atmospheric chemistry modelling. *Earth System Science Data*, **13**(2), 367–404, doi:10.5194/essd-13-367-2021, URL https://essd.copernicus.org/articles/13/367/2021/.

Hanson, D. and Mauersberger, K. (1988). Laboratory studies of the nitric acid trihydrate: Implications for the south polar stratosphere. *Geophysical Research Letters*, **15**(8), 855–858, doi:https://doi.org/10.1029/GL015i008p00855, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/GL015i008p00855.

Hauglustaine, D. A., Balkanski, Y. and Schulz, M. (2014). A global model simulation of present and future nitrate aerosols and their direct radiative forcing of climate. *Atmos. Chem. Phys.*, **14**, 11031–11064.

Henze, D. K. and Seinfeld, J. H. (2006). Global secondary organic aerosol from isoprene oxidation. *Geophysical Research Letters*, **33**(9), doi:https://doi.org/10.1029/2006GL025976, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006GL025976.

Hess, M., Koepke, P. and Schult, I. (1998). Optical properties of aerosols and clouds: The software package opac. *Bull. Amer. Met. Soc.*, **79**, 831–844.

Hodzic, A., Kasibhatla, P. S., Jo, D. S., Cappa, C. D., Jimenez, J. L., Madronich, S. and Park, R. J. (2016). Rethinking the global secondary organic aerosol (soa) budget: stronger production, faster removal, shorter lifetime. *Atmos. Chem. Phys.*, **16**, 7917–7941.

Hogan, R. J. and Bozzo, A. (2018). A flexible and efficient radiation scheme for the ecmwf model. *Journal of Advances in Modeling Earth Systems*, **10**(8), 1990–2008, doi:10.1029/2018MS001364, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001364.

Holben, B. N., Eck, T., Slutsker, I., Tanre, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., Nakajima, T. *et al.* (1998). Aeronet—a federated instrument network and data archive for aerosol characterization. *Remote sensing of environment*, **66**(1), 1–16.

Houweling, S., Kaminski, T., Dentener, F., Lelieveld, J. and Heimann, M. (1999). Inverse modeling of methane sources and sinks using the adjoint of a global transport model. *Journal of Geophysical Research: Atmospheres*, **104**(D21), 26137–26160, doi:https://doi.org/10.1029/1999JD900428, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1999JD900428.

Huijnen, V., Flemming, J., Chabrillat, S., Errera, Q., Christophe, Y., Blechschmidt, A.-M., Richter, A. and Eskes, H. (2016). C-ifs-cb05-bascoe: stratospheric chemistry in the integrated forecasting system of ecmwf. *Geoscientific Model Development*, **9**, 3071–3091.

Huijnen, V., Pozzer, A., Arteta, J., Brasseur, G., Bouarar, I., Chabrillat, S., Christophe, Y., Doumbia, T., Flemming, J., Guth, J., Josse, B., Karydis, V. A., Marécal, V. and Pelletier, S. (2019). Quantifying uncertainties due to chemistry modelling – evaluation of tropospheric composition simulations in the cams model (cycle 43r1). *Geosci. Model Dev.*, **12**, 1725–1752.

Huijnen, V., Williams, J., van Weele, M., van Noije, T., Krol, M., Dentener, F., Segers, A., Houweling, S., Peters, W., de Laat, J., Boersma, F., Bergamaschi, P., van Velthoven, P., Le Sager, P., Eskes, H., Alkemade, F., Scheele, R., Nedelec, P. and Patz, H.-W. (2010). The global chemistry transport model tm5: description and evaluation of the tropospheric chemistry version 3.0. *Geoscientific Model Development*, **3**, 445–473.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M. (2019). The cams reanalysis of atmospheric composition. *Atmospheric Chemistry and Physics*, **19**, 3515–3556.

Inness, A., Ades, M., Balis, D., Efremenko, D., Flemming, J., Hedelt, P., Koukouli, M.-E., Loyola, D. and Ribas, R. (2022). Evaluating the assimilation of s5p/tropomi near real-time so2 columns and layer height data into the cams integrated forecasting system (cy47r1), based on a case study of the 2019 raikoke eruption. *Geosci. Model Dev.*, **15**, 971–994.

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., P., Coheur, Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, D., V. andHurtmans, Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N., Thouret, V., Vrekoussis, M. and the MACC team (2013). The macc reanalysis: an 8 yr data set of atmospheric composition. *Atmospheric Chemistry and Physics*, **13**, 4073–4109.

Inness, A., Blechschmidt, A.-M., Bouarar, I., Chabrillat, S., Crepulja, M., Engelen, R. J., Eskes, H., Flemming, J., Gaudel, A., Hendrick, F., Huijnen, V., Jones, L., Kapsomenakis, J., Katragkou, E., Keppens, A., Langerock, B., de Mazière, M., Melas, D., Parrington, M., Peuch, V. H., Razinger, M., Richter, A., Schultz, M. G., Suttie, M., Thouret, V., Vrekoussis, M., Wagner, A. and Zerefos, C. (2015). Data assimilation of satellite-retrieved ozone, carbon monoxide and nitrogen dioxide with ecmwf's composition-ifs. *Atmospheric Chemistry and Physics*, **15**(9), 5275–5303, doi:10.5194/acp-15-5275-2015, URL https://acp.copernicus.org/articles/15/5275/2015/.

Inness, B. F. B. A. B. I. C. S. C. H. C. C. C. P. E. R. J. E. Q. F. J. G. M. G. C. H.-L. J. H. V. H. D. J. L. K. J. W. K. J. L. K. L. J. R. M. R. A. S. M. G. S. A. J. S. M. S. O. T. J.-N. T. V. V. M. Z. C., A. and the MACC team (2013). The macc reanalysis: an 8 yr data set of atmospheric composition. *Atmospheric Chemistry and Physics*, **13**, 4073–4109.

Jacob, D., Liu, H., Mari, C. and Yantosca, R. (2000). Harvard wet deposition scheme for gmi.

Jakob, C. and Klein, S. (2000). A parametrization of the effects of cloud and precipitation overlap for use in general-circulation models. *Q.J.R. Meteorol. Soc.*, **126**, 2525–2544.

Järvinen, H., Saarinen, S. and Undén, P. (1996). *User's Guide for Blacklisting*. Available on request from ECMWF, Shinfield Park, RG2 9AX, Reading, Berkshire, UK.

Jones, A. C., Hill, A., Remy, S., Abraham, N. L., Dalvi, M., Hardacre, C., Hewitt, A. J., Johnson, B., Mulcahy, J. P. and Turnock, S. T. (2021). Exploring the sensitivity of atmospheric nitrate concentrations to nitric acid uptake rate using the met office's unified model. *Atmos. Chem. Phys.*, **21**, 15901–15927.

Journet, E., Balkanski, Y. and Harrison, S. P. (2014). A new data set of soil mineralogy for dust-cycle modeling. *Atmos. Chem. Phys.*, **14**, 3801–3816.

Kaiser, J., Heil, A., Andreae, M., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M., Suttie, M. *et al.* (2012). Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*, **9**(1), 527.

Kandler, K., Schutz, L., Deutscher, C., Ebert, M., Hofmann, H., Jackel, S., Jaenicke, R., Knippertz, P., Lieke, K., Massling, A., Petzold, A., Schladitz, A., Weinzierl, B., Wiedenshohler, A., Zorn, S. and Weinbruch, S. (2009). Size distribution, mass concentration, chemical and mineralogical composition and derived optical parameters of the boundary layer aerosol at tinfou, morocco, during samum 2006. *Tellus B*, **61**, 32–50, doi:10.1111/j.1600-0889.2008.00385.x.

Karl, M., Tsigaridis, K., Vignati, E. and Dentener, F. (2009). Formation of secondary organic aerosol from isoprene oxidation over europe. *Atmospheric Chemistry and Physics*, **9**(18), 7003–7030, doi:10.5194/acp-9-7003-2009, URL https://acp.copernicus.org/articles/9/7003/2009/.

Khan, T. and Perlinger, J. A. (2017). Evaluation of five dry particle deposition parameterizations for incorporation into atmospheric transport models. *Geoscientific Model Development*, **10**, 3861–3888.

Köhler, H. (1936). The nucleus in and the growth of hygroscopic droplets. *Transactions of the Faraday Society*, **32**, 1152–1161.

Kok, J. F. (2011). A scaling theory for the size distribution of emitted dust aerosols suggests climate models underestimate the size of the global dust cycle. *Proceedings of the National Academy of Sciences*, **108**, 1016–1021.

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F. and Bergamaschi, P. (2005). The two-way nested global chemistry-transport zoom model tm5: algorithm and applications. *Atmospheric Chemistry and Physics*, **5**(2), 417–432, doi:10.5194/acp-5-417-2005, URL https://acp.copernicus.org/articles/5/417/2005/.

Kulmala, M. and Laaksonen, A. (1990). Binary nucleation of water-sulfuric acid system: Comparison of classical theories with different h2so4 saturation vapor pressures. *J. Chem. Phys.*, **93(1)**, 696–701, doi:10.1063/1.459519.

Lamarque, J.-F., Emmons, L. K., Hess, P. G., Kinnison, D. E., Tilmes, S., Vitt, F., Heald, C. L., Holland, E. A., Lauritzen, P. H., Neu, J., Orlando, J. J., Rasch, P. J. and Tyndall, G. K. (2012). Cam-chem: description and evaluation of interactive atmospheric chemistry in the community earth system model. *Geoscientific Model Development*, **5**(2), 369–411, doi:10.5194/gmd-5-369-2012, URL https://gmd.copernicus.org/articles/5/369/2012/.

Lambert, G. and Schmidt, S. (1993). Reevaluation of the oceanic flux of methane: Uncertainties and long term variations. *Chemosphere*, **26**(1), 579–589, doi:https://doi.org/10.1016/0045-6535(93)90443-9, URL https://www.sciencedirect.com/science/article/pii/0045653593904439, proceedings of the NATO advanced research workshop.

Lane, J. R. and Kjaergaard, H. G. (2008). Calculated electronic transitions in sulfuric acid and implications for its photodissociation in the atmosphere. *J. Phys. Chem. A*, **112**, 4958–4964, doi:10.1021/jp710863r.

Lathière, J., Hauglustaine, D. A., Friend, A. D., De Noblet-Ducoudré, N., Viovy, N. and Folberth, G. A. (2006). Impact of climate variability and land use changes on global biogenic volatile organic compound emissions. *Atmospheric Chemistry and Physics*, **6**(8), 2129–2146, doi:10.5194/acp-6-2129-2006, URL https://acp.copernicus.org/articles/6/2129/2006/.

Lawrence, M. G. and Crutzen, P. J. (1998). The impact of cloud gravitational settling on soluble trace gas distributions. *Tellus*, **50B**, 263–289.

Li, C., Granier, C. and Zilberman, N. (2024). 2000-2023 global emissions based on another methodology, using available officially reported emissions in several countries. *Cams 61 deliverable d2.5.1*.

Lide, D. R. (2006). *Chemical Rubber Company (CRC): Handbook of Chemistry and Physics, 86th Edn*. Taylor and Francis Group LLC, CD-ROM version.

Liu, H., Jacob, D., Bey, I. and Yantosca, R. (2001). Constraints from 210pb and 7be on wet deposition and transport in a global three-dimensional chemical tracer model driven by assimilated meteorological fields. *J. Geophys. Res.*, **106**, 109–128.

Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **112**, 1177–1194.

Luo, G., Yu, F. and Schwab, J. (2019). Revised treatment of wet scavenging processes dramatically improves geos-chem 12.0.0 simulations of surface nitric acid, nitrate, and ammonium over the united states. *Geosci. Model Dev.*, **12**, 3439–3447.

Madronich, S. and Flocke, S. (1999). The role of solar radiation in atmospheric chemistry. In *Environmental photochemistry*, pp. 1–26, Springer.

Malardel, S. and Ricard, D. (2015). An alternative cell-averaged departure point reconstruction for pointwise semi-lagrangian transport schemes. *Quarterly Journal of the Royal Meteorological Society*, **141**(691), 2114–2126, doi:https://doi.org/10.1002/qj.2509, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2509.

Marticorena, B. and Bergametti, G. (1995). Modeling the atmosphere dust cycle: 1. design of a soil-derived dust emission scheme. *J. Geophys. Res*, **100**, 16415–16430.

Massart, S., Agusti-Panareda, A., Aben, I., Butz, A., Chevallier, F., Crevoisier, C., Engelen, R., Frankenberg, C. and Hasekamp, O. (2014). Assimilation of atmospheric methane products into the macc-ii system: from sciamachy to tanso and iasi. *Atmospheric Chemistry and Physics*, **14**(12), 6139–6158, doi:10.5194/acp-14-6139-2014, URL https://acp.copernicus.org/articles/14/6139/2014/.

Matthes, K., Funke, B., Andersson, M. E., Barnard, L., Beer, J., Charbonneau, P., Clilverd, M. A., Dudok de Wit, T., Haberreiter, M., Hendry, A., Jackman, C. H., Kretzschmar, M., Kruschke, T., Kunze, M., Langematz, U., Marsh, D. R., Maycock, A. C., Misios, S., Rodger, C. J., Scaife, A. A., Seppälä, A., Shangguan, M., Sinnhuber, M., Tourpali, K., Usoskin, I., van de Kamp, M., Verronen, P. T. and Versick, S. (2017). Solar forcing for cmip6 (v3.2). *Geoscientific Model Development*, **10**(6), 2247–2302, doi:10.5194/gmd-10-2247-2017, URL https://www.geosci-model-dev.net/10/2247/2017/.

McNorton, J., Agustí-Panareda, A., Arduini, G., Balsamo, G., Bousserez, N., Boussetta, S., Chericoni, M., Choulga, M., Engelen, R. and Guevara, M. (2023). An urban scheme for the ecmwf integrated forecasting system: Global forecasts and residential co2 emissions. *Journal of Advances in Modeling Earth Systems*, **15**(3), e2022MS003286.

Meijer, E., van Velthoven, P., Brunner, D., Huntrieser, H. and Kelder, H. (2001). Improvement and evaluation of the parameterisation of nitrogen oxide production by lightning. *Physics and Chemistry of the Earth, Part C: Solar, Terrestrial and Planetary Science*, **26**(8), 577–583, doi:https://doi.org/10.1016/S1464-1917(01)00050-2, URL https://www.sciencedirect.com/science/article/pii/S1464191701000502.

Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Krummel, P. B., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P. J., Reimann, S., Smith, S. J., van den Berg, M., Velders, G. J. M., Vollmer, M. K. and Wang, R. H. J. (2020). The shared socio-economic pathway (ssp) greenhouse gas concentrations and their extensions to 2500. *Geoscientific Model Development*, **13**(8), 3571–3605, doi: 10.5194/gmd-13-3571-2020, URL https://gmd.copernicus.org/articles/13/3571/2020/.

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J. and Weiss, R. (2017). Historical greenhouse gas concentrations for climate modelling (cmip6). *Geoscientific Model Development*, **10**(5), 2057–2116, doi: 10.5194/gmd-10-2057-2017, URL https://www.geosci-model-dev.net/10/2057/2017/.

Meng, J., Huang, Y., Leund, D., Li, L., Adebiyi, A., Ryder, C., Mahowald, N. and Kok, J. (2022). Improved parameterization for the size distibution of emitted dust aerosols reduces model underestimation of super coarse dust. *Geophysical Research Letters*, **49**.

Messina, P., Lathière, J., Sindelarova, K., Vuichard, N., Granier, C., Ghattas, J., Cozic, A. and Hauglustaine, D. A. (2016). Global biogenic volatile organic compound emissions in the orchidee and megan models and sensitivity to key parameters. *Atmospheric Chemistry and Physics*, **16**(22), 14169–14202.

Metzger, S. (2023). The eqsam box model (for eqsam4clim-v12). doi:10.5281/zenodo.10276178, URL https://doi.org/10.5281/zenodo.10276178.

Metzger, S., Abdelkader, M., Steil, B. and Klingmüller, K. (2018). Aerosol water parameterization: long-term evaluation and importance for climate studies. *Atmos. Chem. Phys.*, **18**, 16747–16774, URL https://doi.org/10.5194/acp-18-16747-2018.

Metzger, S., Dentener, F. and Lelieveld, J. (1999). Aerosol multiphase equilibrium composition: results of a parameterization applied to a global chemistry/tracer transport model. *Aerosol Sci.*, **30**, S877, URL https://doi.org/10.1016/S0021-8502(99)80449-2.

Metzger, S., Dentener, F., Pandis, S. and Lelieveld, J. (2002a). Gas/aerosol partitioning: 1. a computationally efficient model. *J. Geophys. Res.*, **107**, ACH 16–1–ACH 16–24, URL https://doi.org/10.1029/2001JD001102.

Metzger, S., Dentener, K. M. J. A., F. and Lelieveld, J. (2002b). Gas/aerosol partitioning: 2. global modeling results. *J. Geophys. Res.*, **107**, ACH 17–1–ACH 17–23, URL https://doi.org/10.1029/2001JD001103.

Metzger, S. and Lelieveld, J. (2007). Reformulating atmospheric aerosol thermodynamics and hygroscopic growth into fog, haze and clouds. *Atmos. Chem. Phys.*, **7**, 3163–3193, URL https://acp.copernicus.org/articles/7/3163/2007/.

Metzger, S., Mihalopoulos, N. and Lelieveld, J. (2006). Importance of mineral cations and organics in gas-aerosol partitioning of reactive nitrogen compounds: case study based on MINOS results. *Atmospheric Chemistry and Physics*, **6**(9), 2549–2567, doi:10.5194/acp-6-2549-2006, URL https://acp.copernicus.org/articles/6/2549/2006/.

Metzger, S., Rémy, S., Williams, J. E., Huijnen, V. and Flemming, J. (2024). A computationally efficient parameterization of aerosol, cloud and precipitation ph for application at global and regional scale (eqsam4clim-v12). *Geoscientific Model Development*, **17**(12), 5009–5021, doi:10.5194/gmd-17-5009-2024, URL https://gmd.copernicus.org/articles/17/5009/2024/.

Metzger, S., Steil, B., Abdelkader, M., Klingmüller, K., Xu, L., Penner, J. E., Fountoukis, C., Nenes, A. and Lelieveld, J. (2016). Aerosol water parameterisation: a single parameter framework. *Atmos. Chem. Phys.*, **16**, 7213–7237, URL https://doi.org/10.5194/acp-16-7213-2016.

Metzger, S., Steil, B., Xu, L., Penner, J. E. and Lelieveld, J. (2012). New representation of water activity based on a single solute specific constant to parameterize the hygroscopic growth of aerosols in atmospheric models. *Atmospheric Chemistry and Physics*, **12**(12), 5429–5446, doi:10.5194/acp-12-5429-2012, URL https://acp.copernicus.org/articles/12/5429/2012/.

Michou, M., Nabat, P. and Saint-Martin, D. (2015). Development and basic evaluation of a prognostic aerosol scheme (v1) in the cnrm climate model cnrm-cm6. *Geoscientific Model Development*, **8**, 501–531.

Miller, Y., Gerber, R. B. and Vaida, V. (2007). Photodissociation yields for vibrationally excited states of sulfuric acid under atmospheric conditions. *Geophysical Research Letters*, **34**(16), doi:10.1029/2007GL030529, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL030529.

Moise, T., Flores, J. and Rudich, Y. (2015). Optical properties of secondary organic aerosols and their changes by chemical processes. *Chem. Rev.*, **115**, 4400–4439.

Monahan, E. C. and Muircheartaigh, I. O. (1980). Optimal power-law description of oceanic whitecap coverage dependence on wind speed. *J. Phys. Oceanogr.*, **10**, 2094–2099.

Monahan, E. C., Spiel, D. E. and Davidson, K. L. (1986). A model of marine aerosol generation via whitecaps and wave disruption. In E. C. Monahan and G. MacNiocaill (Eds), *Oceanic Whitecaps and Their Role in Air–Sea Exchange Processes*, pp. 167–174, D. Reidel.

Morcrette, J.-J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J., Razinger, M. *et al.* (2009). Aerosol analysis and forecast in the european centre for medium-range weather forecasts integrated forecast system: Forward modeling. *Journal of Geophysical Research: Atmospheres*, **114**(D6).

Muncaster, R., Bourqui, M. S., Chabrillat, S., Viscardy, S., Melo, S. M. L. and Charbonneau, P. (2012). A simple framework for modelling the photochemical response to solar spectral irradiance variability in the stratosphere. *Atmospheric Chemistry and Physics*, **12**(16), 7707–7724, doi:10.5194/acp-12-7707-2012, URL https://acp.copernicus.org/articles/12/7707/2012/.

Murphy, D. M. and Koop, T. (2005). Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Quart. J. Roy. Meteor. Soc.*, **131**(608), 1539–1565, doi:10.1256/qj.04.94, URL http://dx.doi.org/10.1256/qj.04.94.

Myriokefalitakis, S., Bergas-Massó, E., Gonçalves-Ageitos, M., Pérez García-Pando, C., van Noije, T., Le Sager, P., Ito, A., Athanasopoulou, E., Nenes, A., Kanakidou, M., Krol, M. C. and Gerasopoulos, E. (2022). Multiphase processes in the ec-earth model and their relevance to the atmospheric oxalate, sulfate, and iron cycles. *Geoscientific Model Development*, **15**(7), 3079–3120, doi:10.5194/gmd-15-3079-2022, URL https://gmd.copernicus.org/articles/15/3079/2022/.

Myriokefalitakis, S., Daskalakis, N., Gkouvousis, A., Hilboll, A., van Noije, T., Williams, J. E., Le Sager, P., Huijnen, V., Houweling, S., Bergman, T., Nüß, J. R., Vrekoussis, M., Kanakidou, M. and Krol, M. C. (2020). Description and evaluation of a detailed gas-phase chemistry scheme in the tm5-mp global chemistry transport model (r112). *Geoscientific Model Development*, **13**(11), 5507–5548, doi:10.5194/gmd-13-5507-2020, URL https://gmd.copernicus.org/articles/13/5507/2020/.

Nabat, P., Solmon, F., Mallet, M., Kok, J. and Somot, S. (2012). Dust emission size distribution impact on aerosol budget and radiative forcing over the mediterranean region: a regional climate model approach. *Atmos. Chem. Phys.*, **12**, 10545–10567.

O'Dowd, C., Smith, M., Consterdine, I. and Lowe, J. (1997). Marine aerosol, sea-salt, and the marine sulphur cycle: a short review. *Atmospheric Environment*, **31**(1), 73–80.

Ott, L. E., Pickering, K. E., Stenchikov, G. L., Allen, D. J., DeCaria, A. J., Ridley, B., Lin, R.-F., Lang, S. and Tao, W.-K. (2010). Production of lightning nox and its vertical distribution calculated from three-dimensional cloud-scale chemical transport model simulations. *Journal of Geophysical Research: Atmospheres*, **115**(D4), doi:https://doi.org/10.1029/2009JD011880, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD011880.

Pailleux, J. (1990). A global variational assimilation scheme and its application for using TOVS radiances. In *Proc. WMO International Symposium on Assimilation of Observations in Meteorology and Oceanography*, pp. 325–328, Clermont-Ferrand, France.

Parrish, D. F. and Derber, J. (1992). The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.

Peuch, V.-H., Engelen, R., Rixen, M., Dee, D., Flemming, J., Suttie, M., Ades, M., Agusti-Panareda, A., Ananasso, C., Andersson, E., Armstrong, D., Barre, J., Bousserez, N., Dominguez, J. J., Garrigues, S., Inness, A., Jones, L., Kipling, Z., Letertre-Danczak, J., Parrington, M., Razinger, M., Ribas, R., Vermoote, S., Yang, X., Simmons, A., de Marcilla, J. G. and Thepaut, J.-N. (2022). The copernicus atmosphere monitoring service: From research to operations. *Bulletin of the American Meteorological Society*, **103**(12), E2650 – E2668, doi:https://doi.org/10.1175/BAMS-D-21-0314.1, URL https://journals.ametsoc.org/view/journals/bams/103/12/BAMS-D-21-0314.1.xml.

Prignon, M., Chabrillat, S., Friedrich, M., Smale, D., Strahan, S. E., Bernath, P. F., Chipp erfield, M. P., Dhomse, S. S., Feng, W., Minganti, D., Servais, C. and Mahieu, E. (2021). Stratospheric fluorine as a tracer of circulation changes: Comparison between infrared remote-sensing observati ons and simulations with five modern reanalyses. *Journal of Geophysical Research: Atmospheres*, **126**(19), e2021JD034995, doi:https://doi.org/10.1029/2021JD034995, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021JD034995, e2021JD034995 2021JD034995.

Pye, H. O. T., Nenes, A., Alexander, B., Ault, A. P., Barth, M. C., Clegg, S. L., Collett Jr., J. L., Fahey, K. M., Hennigan, C. J., Herrmann, H., Kanakidou, M., Kelly, J. T., Ku, I.-T., McNeill, V. F., Riemer, N., Schaefer, T., Shi, G., Tilgner, A., Walker, J. T., Wang, T., Weber, R., Xing, J., Zaveri, R. A. and Zuend, A. (2020). The acidity of atmospheric particles and clouds. *Atmospheric Chemistry and Physics*, **20**(8), 4809–4888, doi:10.5194/acp-20-4809-2020, URL https://acp.copernicus.org/articles/20/4809/2020/.

Reddy, M. S., Boucher, O., Bellouin, N., Schulz, M., Balkanski, Y., Dufresne, J.-L. and Pham, M. (2005). Estimates of global multicomponent aerosol optical depth and direct radiative perturbation in the laboratoire de météorologie dynamique general circulation model. *J. Geophys. Res.*, **110**, D10S16.

Remy, S. and Anguelova, M. (2021). Improving the representation of whitecap fraction and sea salt aerosol emissions in the ecmwf ifs-aer. *Remote Sens.*, **13**, 4856, doi:10.3390/rs13234856, URL https://doi.org/10.3390/rs13234856.

Rémy, S., Kipling, Z., Flemming, J., Boucher, O., Nabat, P., Michou, M., Bozzo, A., Ades, M., Huijnen, V., Benedetti, A., Engelen, R., Peuch, V.-H. and Morcrette, J.-J. (2019). Description and evaluation of the tropospheric aerosol scheme in the european centre for medium-range weather forecasts (ecmwf) integrated forecasting system (ifs-aer, cycle 45r1). *Geoscientific Model Development*, **12**(11), 4627–4659, doi:10.5194/gmd-12-4627-2019, URL https://gmd.copernicus.org/articles/12/4627/2019/.

Rémy, S., Kipling, Z., Huijnen, V., Flemming, J., Nabat, P., Michou, M., Ades, M., Engelen, R. and Peuch, V.-H. (2022). Description and evaluation of the tropospheric aerosol scheme in the integrated forecasting system (ifs-aer, cycle 47r1) of ecmwf. *Geoscientific Model Development*, **15**(12), 4881–4912, doi:10.5194/gmd-15-4881-2022, URL https://gmd.copernicus.org/articles/15/4881/2022/.

Rémy, S., Metger, S., Huijnen, V., Williams, J. and Flemming, J. (2024). An improved representation of aerosol in the ecmwf ifs-compo cy49r1 through the integration of eqsam4climv12 - a first attempt at simulating aerosol acidity. *Geoscientific Model Development*, **17**(17), 7539–7567, doi:10.5194/gmd-17-7539-2024, URL https://gmd.copernicus.org/articles/17/7539/2024/.

Ridgwell, A. J., Marshall, S. J. and Gregson, K. (1999). Consumption of atmospheric methane by soils: A process-based model. *Global Biogeochem. Cycles*, **13**(1), 59–70.

Rödenbeck, C., Keeling, R. F., Bakker, D. C. E., Metzl, N., Olsen, A., Sabine, C. and Heimann, M. (2013). Global surface-ocean pco2 and sea-air co2 flux variability from an observation-driven ocean mixed-layer scheme. *Ocean Science*, **9**(2), 193–216, doi:10.5194/os-9-193-2013.

Ryder, C. L., Marenco, F., Brooke, J. K., Estelles, V., Cotton, R., Formenti, P., McQuaid, B., J., Price, H. C., Liu, D., Ausset, P., Rosenberg, P. D., Taylor, J. W., Choularton, T., Bower, K., Coe, H., Gallagher, M., Crosier, J., Lloyd, G., Highwood, E. J. and Murray, B. J. (2018). Coarse-mode mineral dust size distributions, composition and optical properties from aer-d aircraft measurements over the tropical eastern atlantic. *Atmos. Chem. Phys.*, **18**, 17225–17257.

Sander, R. (2015). Compilation of henry's law constants (version 4.0) for water as solvent. *Atmospheric Chemistry and Physics*, **15**(8), 4399–4981, doi:10.5194/acp-15-4399-2015, URL https://acp.copernicus.org/articles/15/4399/2015/.

Sander, S. (2000). Chemical kinetics and photochemical data for use in stratospheric modeling, supplement to evaluation 12: Update of key reactions, evaluation number 13. *JPL Publication 00-3*.

Sander, S. P., Abbatt, J., Barker, J. R., Burkholder, J. B., Friedl, R. R., Golden, D. M., Huie, R. E., Kolb, C. E., Kurylo, M. J., Moortgat, G. K., Orkin, V. L. and Wine, P. H. (2011). Chemical kinetics and photochemical data for use in atmospheric studies. Evaluation number 17. *Publication 10–6*, URL http://jpldataeval.jpl.nasa.gov/.

Sanderson, M. G. (1996). Biomass of termites and their emissions of methane and carbon dioxide: A global database. *Global Biogeochemical Cycles*, **10**(4), 543–557, doi:https://doi.org/10.1029/96GB01893, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96GB01893.

Sandu, A. and Sander, R. (2006). Technical note: Simulating chemical systems in fortran90 and matlab with the kinetic preprocessor kpp-2.1. *Atmospheric Chemistry and Physics*, **6**(1), 187–195, doi:10.5194/acp-6-187-2006, URL https://acp.copernicus.org/articles/6/187/2006/.

Schery, S. (2004). Progress on global 222rn flux maps and recommendations for future research. *1st national expert meeting on sources and measurements of natural radionucleides applied to climate and air quality studies. WMO report 155*, p. 43.

Schumann, U. and Huntrieser, H. (2007). The global lightning-induced nitrogen oxides source. *Atmospheric Chemistry and Physics*, **7**(14), 3823–3907, doi:10.5194/acp-7-3823-2007, URL https://acp.copernicus.org/articles/7/3823/2007/.

Schwartz, S. (1986). Mass transport considerations pertinent to aqueous phase reactions of gases in liqui-water clouds. In *Chemistry of Multiphase Atmospheric Systems*, chapter 4, pp. 415–471, Springer, New York.

Seinfeld, J. and Pandis, S. (1998). *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. New Jersey: John Wiley and Sons.

Seinfeld, J. and Pandis, S. (2006). *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, 2nd Edn.* New Jersey: John Wiley and Sons.

Shangguan, W., Dai, Y., Duan, Q., Liu, B. and Yuan, H. (2014). A global soil data set for earth system modeling. *J. Adv. Model. Earth Syst.*, **6**, 249–263.

Shi, Q., Jayne, J. T., Kolb, C. E. and Worsnop, D. R. (2001). Kinetic model for reaction of clono2 with h2o and hcl and hocl with hcl in sulfuric acid solutions. *J. Geophys. Res.*, **106**, 24259–24274.

Simon L. Clegg, V. S. N. F. T. D. R. C. S. K. S. A. G. H., Lynn R. Mazzoleni (2019). Modelling the hygroscopic growth factors of aerosol material containing a large water-soluble organic fraction, collected at the storm peak laboratory. *Atmospheric Environment*, **214**, doi:10.1016/j.atmosenv.2019.05.068.

Sindelarova, K., Markova, J., Simpson, D., Huszar, P., Karlicky, J., Darras, S. and Granier, C. (2022). High-resolution biogenic global emission inventory for the time period 2000–2019 for air quality modelling. *Earth System Science Data*, **14**(1), 251–270, doi:10.5194/essd-14-251-2022, URL https://essd.copernicus.org/articles/14/251/2022/.

Singh, H., Salas, L., Herlth, D., Kolyer, R., Czech, E., Viezee, W., Li, Q., Jacob, D. J., Blake, D., Sachse, G. *et al.* (2003). In situ measurements of hcn and ch3cn over the pacific ocean: Sources, sinks, and budgets. *Journal of Geophysical Research: Atmospheres*, **108**(D20).

Soulie, A., Granier, C., Darras, S., Zilbermann, N., Doumbia, T., Guevara, M., Jalkanen, J.-P., Keita, S., Liousse, C., Crippa, M., Guizzardi, D., Hoesly, R. and Smith, S. (2023). Global anthropogenic emissions (cams-glob-ant) for the copernicus atmosphere monitoring service simulations of air quality forecasts and reanalyses. *Earth System Science Data Discussions*, **2023**, 1–45, doi:10.5194/essd-2023-306, URL https://essd.copernicus.org/preprints/essd-2023-306/.

Spahni, R., Wania, R., Neef, L., van Weele, M., Pison, I., Bousquet, P., Frankenberg, C., Foster, P. N., Joos, F., Prentice, I. C. and van Velthoven, P. (2011). Constraining global methane emissions and uptake by ecosystems. *Biogeosciences*, **8**(6), 1643–1665, doi:10.5194/bg-8-1643-2011, URL https://bg.copernicus.org/articles/8/1643/2011/.

Spiro, P., Jacob, D. and Logan, J. (1992). Global inventory of sulfur emissions with 1x1 deg resolution. *J. Geophys. Res.*, (97), 6023–6036.

Stavrakou, T., Peeters, J. and Müller, J.-F. (2010). Improved global modelling of ho_x recycling in isoprene oxidation: evaluation against the gabriel and intex-a aircraft campaign measurements. *Atmospheric Chemistry and Physics*, **10**(20), 9863–9878, doi:10.5194/acp-10-9863-2010, URL https://acp.copernicus.org/articles/10/9863/2010/.

Svenningsson, B., Rissler, J., Swietlicki, E., Mircea, M., Bilde, M., Facchini, M. C., Decesari, S., Fuzzi, S., Zhou, J., Monster, J. and Rosenorn, T. (2006). Hygroscopic growth and critical supersaturations for mixed aerosol particles of inorganic and organic compounds of atmospheric relevance. *Atmos. Chem. Phys.*, **6**, 1937–1952.

Tang, I. N. and Munkelwitz, H. (1994). Water activities, densities, and refractive indices of aqueous sulfates and sodium nitrate droplets of atmospheric importance. *J. Geophys. Res.*, **99**, 801–818.

Tang, I. N., Tridico, A. C. and Fung, K. H. (1997). Thermodynamic and optical properties of sea salt aerosols. *J. Geophys. Res.*, **102**, 269–275.

Thornton, J. A., Jaeglé, L. and McNeill, V. F. (2008). Assessing known pathways for ho2 loss in aqueous atmospheric aerosols: Regional and global impacts on tropospheric oxidants. *Journal of Geophysical Research: Atmospheres*, **113**(D5), doi:https://doi.org/10.1029/2007JD009236, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD009236.

Tompkins, A. M. (2005). A revised cloud scheme to reduce the sensitivity to vertical resolution. *ECMWF Tech Memo*, **599**, 1–25.

Tsigaridis, K. and Kanakidou, M. (2003). Global modelling of secondary organic aerosol in the troposphere: a sensitivity analysis. *Atmospheric Chemistry and Physics*, **3**(5), 1849–1869, doi:10.5194/acp-3-1849-2003, URL https://acp.copernicus.org/articles/3/1849/2003/.

Tsimpidi, A. P., Karydis, V. A., Pandis, S. N., and Lelieveld, J. (2016). Global combustion sources of organic aerosols: model comparison with 84 ams factor-analysis data sets. *Atmos. Chem. Phys.*, **16**, 8939–8962.

Tsimpidi, A. P., Karydis, V. A., Pozzer, A., Pandis, S. N. and Lelieveld, J. (2014). Oracle (v1.0): module to simulate the organic aerosol composition and evolution in the atmosphere. *Geosci. Model Dev.*, **7**, 3153–3172.

Verheggen, B., Cozic, J., Weingartner, E., Bower, K., Mertes, S., Connolly, P., Gallagher, M., Flynn, M., Choularton, T. and Baltensperger, U. (2007). Aerosol partitioning between the interstitial and the condensed phase in mixed-phase clouds. *Journal of Geophysical Research: Atmospheres*, **112**(D23), doi:https://doi.org/10.1029/2007JD008714, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD008714.

von Blohn, N., Diehl, K., Mitra, S. K. and Borrmann, S. (2011). Wind tunnel experiments on the retention of trace gases during riming: nitric acid, hydrochloric acid, and hydrogen peroxide. *Atmos. Chem. Phys.*, **11**, 11569–11579.

Wesely, M. (1989). Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models. *Atmospheric Environment (1967)*, **23**(6), 1293–1304, doi:https://doi.org/10.1016/0004-6981(89)90153-4, URL https://www.sciencedirect.com/science/article/pii/0004698189901534.

Wesely, M. and Hicks, B. (1977). Some factors that affect the deposition rates of sulfur dioxide and similar gases on vegetation. *Journal of the Air Pollution Control Association*, **27**, 1110–1117.

Wexler, A. and Seinfeld, J. (1990). The distribution of ammonium salts among a size and composition dispersed aerosol. *Atmos. Environ.*, **24A**, 1231–1246.

Williams, J. E., Boersma, K. F., Le Sager, P. and Verstraeten, W. W. (2017). The high-resolution version of tm5-mp for optimized satellite retrievals: description and validation. *Geoscientific Model Development*, **10**(2), 721–750, doi:10.5194/gmd-10-721-2017, URL https://gmd.copernicus.org/articles/10/721/2017/.

Williams, J. E., Huijnen, V., Bouarar, I., Meziane, M., Schreurs, T., Pelletier, S., Marecal, V., Josse, B. and Flemming, J. (2022). Regional evaluation of the performance of the global cams chemical modeling system over the united states (ifs cycle 47r1). *Geoscientific Model Development*, **15**(12), 4657–4687, doi:10.5194/gmd-15-4657-2022, URL https://gmd.copernicus.org/articles/15/4657/2022/.

Williams, J. E., Landgraf, J., Bregman, A. and Walter, A. A. (2006). A modified band approach for the accurate calculation of online photolysis rates in stratospheric-tropospheric chemical transport models. *Atmospheric Chemistry and Physics*, **6**(6), 4137–4161, doi:10.5194/acp-6-4137-2006, URL https://acp.copernicus.org/articles/6/4137/2006/.

Williams, J. E., Rémy, S., Metzger, S., Huijnen, V. and Flemming, J. (2024). An evaluation of the regional distribution and wet deposition of secondary inorganic aerosols and their gaseous precursors in ifs-compo cycle 49r1. *in preparation*.

Williams, J. E., Strunk, A., Huijnen, V. and van Weele, M. (2012). The application of the modified band approach for the calculation of on-line photodissociation rate constants in tm5: implications for oxidative capacity. *Geoscientific Model Development*, **5**(1), 15–35, doi:10.5194/gmd-5-15-2012, URL https://gmd.copernicus.org/articles/5/15/2012/.

Williams, J. E., van Velthoven, P. F. J. and Brenninkmeijer, C. A. M. (2013). Quantifying the uncertainty in simulating global tropospheric composition due to the variability in global emission estimates of biogenic volatile organic compounds. *Atmospheric Chemistry and Physics*, **13**(5), 2857–2891, doi:10.5194/acp-13-2857-2013, URL https://acp.copernicus.org/articles/13/2857/2013/.

Wiscombe, W. J. (1980). Improved mie scattering algorithms. *Applied Optics*, **19**, 1505–1509.

Yarwood, G., Rao, S., Yocke, M. and Whitten, G. (2005). Updates to the carbon bond chemical mechanism: Cb05, final report to the US EPA. *EPA Report Number: RT-0400675, available at: http://www.camx.com (last access 03 April 2019)*.

Zender, C. S., Newman, D. and Torres, O. (2003). Spatial heterogeneity in aeolian erodibility: Uniform, topographic, geomorphic, and hydrologic hypotheses. *J. Geophys. Res*, **108 (D17)**, 4543–4555.

Zhang, L., Gong, S., Padro, J. and Barrie, L. (2001). A size-segregated particle dry deposition scheme for an atmospheric aerosol module. *Atmospheric Environment*, **35**(3), 549–560, doi:https://doi.org/10.1016/S1352-2310(00)00326-5.

Zhang, L. and He, Z. (2014). Technical note: An empirical algorithm estimating dry deposition velocity of fine, coarse and giant particles. *Atmos. Chem. Phys.*, **14**, 3729–3727.