

Response from authors

Re: Review of Harmonized Emissions Component (HEMCO) 3.0 as a versatile emissions component for atmospheric models: application in the GEOS-Chem, NASA GEOS, WRF-GC, CESM2, NOAA GEFS-Aerosol, and NOAA UFS models

August 11, 2021

We thank the two Reviewers for their helpful comments. In response, we have carefully revised the manuscript to clarify the current capabilities of HEMCO 3.0. We respond to each specific comment in detail below. The referee comments are shown in red italics. Our replies are shown in black and modified text is shown in blue. The annotated line and page numbers refer to the revised copy of the manuscript.

A copy of the manuscript's complete changes is attached at the end of the response.

Reviewer #1

I have read the paper "Harmonized Emissions Component (HEMCO) 3.0 as a versatile emissions component for atmospheric models: application in the GEOS-Chem, NASA GEOS, WRF-GC, CESM2, NOAA GEFS- Aerosol, and NOAA UFS models" by Lin et al. The paper we well written and described the HEMCO component well. HEMCO 3.0 looks to be much more flexible and useful for a wider range of models that the previous version.

A few specific comments.

One point to clarify, can scaling be spatially-, temporally, or sector-explicit?

(for example, sector-specific, or regionally-specific NO_x -> NO₂/NO₂ scaling. Can this change over time?) This seems to be implied latter in the text, but should be mentioned more explicitly.

Thank you for your comment. Yes, scaling can be spatially, temporally, and sector-explicit. We have revised the text to clarify this:

P6, Lines 168-171:

Scaling factors can be specified for each individual entry in the HEMCO configuration file, allowing for different scaling factors to be applied for different inventories, sectors, and species. HEMCO accepts scaling factors as constant numbers, temporally-explicit (diurnal, day-of-week, seasonal, interannual) numbers, or as a gridded netCDF data file.

There should be a software implementation section in the paper that gives an overview of how the software is implemented and structured. It appears the language is fortran90, but this is not mentioned anywhere in the text. It would be useful to provide a few paragraphs that give an overview of how the code is structured at the "nuts and bolts" level. How do the HEMCO components communicate with each other? Are there common blocks of data that are key to HEMCO (or have more modern software encapsulation and abstraction techniques been used?)

Thank you for your suggestion. We have added descriptions of the HEMCO 3.0 code in the manuscript.

P7, Lines 178-184:

The HEMCO Core calculates emissions with summations, masks, and scaling factors specified in the HEMCO configuration file. It includes Fortran modules that define the HEMCO state, HEMCO data types (e.g., configuration options, date and time, chemical species and their physical properties, file containers storing input data, and data containers storing data processed by HEMCO), and the driver routine that computes emissions and stores them in data containers. All data types are contained in a variable called the HEMCO state (HcoState) and passed as an argument to all HEMCO subroutines in the code. This allows multiple instances of HEMCO to operate simultaneously, as multiple copies of HEMCO state can co-exist independently. The HEMCO Core also includes an error handling and logging component.

P10, Lines 216-230:

At the beginning of the run, the Model Interface Layer provides the model species list to the HEMCO Core along with any physical properties needed for computation of state-dependent emissions (for example, Henry's law constants for ocean fluxes). It also provides information to the HEMCO Core on the model environment, such as the model clock and time step size. *This information is stored in the HEMCO state by the HEMCO Core.*

At every HEMCO time step, when HEMCO is called by the model, the HEMCO Core performs the requested calculations, loading the latest available input data *into the HEMCO state's file containers* using the Data Input Layer as necessary.

While the paper gives a useful overview of the way HEMCO is used, a reader comes away from the paper with no idea of how to actually use HEMCO? For example, is it a matter connecting to just a few key interface routines?

Thank you for your suggestion. We have added further detail regarding how to implement HEMCO 3.0 in a new model in the section regarding the Data Input Layer and the Model Interface Layer, which are used to “couple” HEMCO 3.0 to a new model.

P11, Lines 266-275:

Implementation of HEMCO in a new model can be prototyped by modifying the default Model Interface Layer to use the model's data structures. *The Model Interface Layer includes at least three subroutines: initialization, run, and finalization. These subroutines need to be called by the model and provided with information about the model environment. For initialization, the model species list and their physical properties, the HEMCO grid information, and the location of the configuration file need to be provided. For run, information about the current model time and model state variables for HEMCO extensions needs to be provided, and the computed emissions and data need to be retrieved from the HEMCO Core to be passed back to the model. If the HEMCO grid is different than the model*

grid (Sect. 2.4), then the Model Interface Layer also needs to implement re-gridding capabilities. An example is HEMCO 3.0's implementation within the CESM model, described in Sect. 3.5.

As part of such a section, it would be very useful to provide some details on how a user would add an extension. This is a potentially very powerful feature of the system, so more details on that aspect would be useful. Give an indication of how much work someone would have to do to implement a new extension. It would also be useful to point to some examples in the current code base to extensions implemented in the current system.

Thank you for your suggestion. In response, we have added further information on how existing extensions were implemented, how new extensions can be added in a similar manner, and a table including the current HEMCO extensions available in HEMCO 3.0.

P9, Lines 200-211:

Emissions dependent on model state such as dust or lightning can be computed on-line by using algorithms called HEMCO extensions supplied with the HEMCO Core. For example, the current HEMCO Core includes as default the DEAD dust emission extension implementing the algorithm from Zender et al. (2003) but users may select other available extension options (such as the Ginoux et al. (2001) algorithm) or they can add a new algorithm as an extension. Alternatively, users may pre-compute these emissions based on off-line input data and disable the HEMCO extension. Both approaches are routinely used in GEOS-Chem (Weng et al., 2020). *Table 2 lists available HEMCO extensions in HEMCO 3.0. Users may add a new algorithm as an extension by creating a new extension file within the Extensions directory in HEMCO. HEMCO extensions include subroutines for initialization, run, and finalization. At every time step, the "Run" subroutine receives HEMCO state and model state information, and returns the computed emissions array to the HEMCO Core, which can then be added to the other emissions data. New state-dependent emission algorithms can be modified to fit this structure by encapsulating the bulk of the code into the "Run" routine and*

adjusting the variable names so that model state can be read through HEMCO. Most HEMCO extensions were developed in this way.

Table 2. Emission extensions available in HEMCO 3.0 as built-in algorithms.

Species ^a	Extension name and reference
Oceanic DMS, Acetone, Acetaldehyde, Methyl nitrate, Ethyl nitrate, Methanol	SeaFlux
Ship plume NO _x , HNO ₃ , O ₃	ParaNO _x (Vinken et al., 2011)
Lightning NO _x	LightNO _x (Murray et al., 2012; Ott et al., 2010)
Soil and fertilizer NO _x	SoilNO _x (Hudman et al., 2012)
Mineral dust aerosols	DEAD (Zender et al., 2003) Ginoux (Ginoux et al., 2001)
Sea salt aerosols	SeaSalt (Chin et al., 2001; Gong, 2003; Jaeglé et al., 2011)
Biogenic VOCs	MEGAN (Guenther et al., 2012)
Biomass burning	GFED (Akagi et al., 2011; Andreae et al., 2001; Giglio et al., 2013; Randerson et al., 2012; van der Werf et al., 2010) FINN (Wiedinmyer et al., 2011)
Volcanic SO ₂	Volcano (Carn et al., 2015; Ge et al., 2016)
Inorganic iodine emissions: HOI, I ₂	Inorg_Iodine (Carpenter et al., 2013; MacDonald et al., 2014)

^a DMS = dimethyl sulfide;

Reviewer #2

This manuscript describes the HEMCOv3 system, a flexible tool for processing and adapting emission inventories for air quality modelling purposes. The strength of HEMCO is in its ability to combine multiple inventories in a very flexible and transparent way, as well as in the fact that it can be used on-line with multiple atmospheric chemistry transport models. The paper is well structured and a good contribution to GMD. Nevertheless, it is sometimes lacking details. Therefore, the paper should be revised according to the following comments before being published.

The new version of HEMCO is capable of calculating emissions at any model resolution including multiscale and unstructured grids. This is highlighted by the authors as an improved capability of the system compared to its previous version. However, no description is provided in terms of: 1) how HEMCO creates multiscale and unstructured grids, 2) how the regridding of emissions is performed for these cases (i.e., how the mass conservation is ensured) and 3) how the definition of these types of grids needs to be provided by the user. Besides these points, authors should also clarify if the new version of HEMCO can perform a remapping to working grids using projections different than regular lat-lon (e.g., lambert conformal conic, polar stereographic) as these types of projections are widely used in the modelling community.

Thank you for your suggestions. The capability of HEMCO 3.0 to support other grids and projections is dependent on the implementations of the Data Input Layer and the Model Interface Layer, which provide re-gridding capabilities between the input data, the HEMCO grid (if different than the model grid), and the model grid. The HEMCO Core operates on grid boxes and is independent of the grid being used. HEMCO out-of-the-box includes the default Data Input Layer and Model Interface Layers, which support rectilinear latitude-longitude grids only. Other implementations of HEMCO 3.0 expand on this capability. For example, the GEOS-Chem High Performance (GCHP) implementation supports the cubed-sphere grid as the HEMCO grid through the *ExtData* component in MAPL, and the CESM2 implementation supports an arbitrary model grid described by an ESMF mesh file while

keeping a rectilinear latitude-longitude HEMCO grid. Future implementations may opt to use a mix of both approaches. We have updated the manuscript to clarify this.

P11, Lines 266-275:

Implementation of HEMCO in a new model can be prototyped by modifying the default Model Interface Layer to use the model's data structures. *The Model Interface Layer includes at least three subroutines: initialization, run, and finalization. These subroutines need to be called by the model and provided with information about the model environment. For initialization, the model species list and their physical properties, the HEMCO grid information, and the location of the configuration file need to be provided. For run, information about the current model time and model state variables for HEMCO extensions needs to be provided, and the computed emissions and data need to be retrieved from the HEMCO Core to be passed back to the model. If the HEMCO grid is different than the model grid (Sect. 2.4), then the Model Interface Layer also needs to implement re-gridding capabilities. An example is HEMCO 3.0's implementation within the CESM model, described in Sect. 3.5.*

P13, Lines 320-329:

While any unstructured grid may be used as the HEMCO grid, it may be desirable to use a rectilinear latitude-longitude grid for prototyping HEMCO in new models. This is because the default Data Input Layer provided with HEMCO only supports rectilinear latitude-longitude grids, and most input data available in the HEMCO database library are also on rectilinear latitude-longitude grids. By choosing such a grid, the default Data Input Layer can be readily used for quick prototyping of a new HEMCO implementation, which may then be improved upon if another HEMCO grid is more desirable. *In cases where other grids are used as the HEMCO grid or the model grid, conservative re-gridding needs to be implemented by the Data Input Layer or the Model Interface Layer. Examples of these scenarios are described in Sect. 3.2, where the HEMCO grid is a cubed-sphere grid and ExtData from MAPL with re-gridding capability is implemented as a Data Input Layer, and Sect. 3.5, where the model grid is an arbitrary grid described by an ESMF mesh file provided by the model to HEMCO, and ESMF on-line re-gridding is implemented in the Model Interface Layer.*

P19-20, Lines 462-469:

Figure 8 shows the implementation of HEMCO 3.0 in CESM2. HEMCO serves emissions to CAM-chem, GEOS-Chem, and potentially to any representation of atmospheric chemistry in CAM. The Model Interface Layer includes routines to export data processed by HEMCO to CAM's physics buffer, a temporary storage space for model components to share data at runtime. *As CAM supports a variety of grids, during initialization, the HEMCO-CESM Model Interface Layer reads the ESMF mesh file that describes the grid used by CAM, and uses ESMF on-line re-gridding (<https://earthsystemmodeling.org/regrid/>) to re-grid data between HEMCO and CAM. By using ESMF re-gridding capabilities, HEMCO is capable of operating independently of the grid used by the model, as long as the model grid description is provided to HEMCO using an ESMF mesh file.*

A general description of the “HEMCO extensions” should be provided (i.e., a brief description of how the built-in algorithms estimate emissions from vegetation, dust, lightning and ocean, including parameters and estimation methodologies used. Specific references to these extensions should be also included if available.

Thank you for your suggestion. We have added a paragraph describing the features of HEMCO extensions, how to implement them, and have included a table of the current extensions available in HEMCO 3.0.

P9, Lines 200-211:

Emissions dependent on model state such as dust or lightning can be computed on-line by using algorithms called HEMCO extensions supplied with the HEMCO Core. For example, the current HEMCO Core includes as default the DEAD dust emission extension implementing the algorithm from Zender et al. (2003) but users may select other available extension options (such as the Ginoux et al. (2001) algorithm) or they can add a new algorithm as an extension. Alternatively, users may pre-compute these emissions based on off-line input data and disable the HEMCO extension. Both

approaches are routinely used in GEOS-Chem (Weng et al., 2020). *Table 2 lists available HEMCO extensions in HEMCO 3.0. Users may add a new algorithm as an extension by creating a new extension file within the Extensions directory in HEMCO. HEMCO extensions include subroutines for initialization, run, and finalization. At every time step, the “Run” subroutine receives HEMCO state and model state information, and returns the computed emissions array to the HEMCO Core, which can then be added to the other emissions data. New state-dependent emission algorithms can be modified to fit this structure by encapsulating the bulk of the code into the “Run” routine and adjusting the variable names so that model state can be read through HEMCO. Most HEMCO extensions were developed in this way.*

Table 2. Emission extensions available in HEMCO 3.0 as built-in algorithms.

Species ^a	Extension name and reference
Oceanic DMS, Acetone, Acetaldehyde, Methyl nitrate, Ethyl nitrate, Methanol	SeaFlux
Ship plume NO _x , HNO ₃ , O ₃	ParaNO _x (Vinken et al., 2011)
Lightning NO _x	LightNO _x (Murray et al., 2012; Ott et al., 2010)
Soil and fertilizer NO _x	SoilNO _x (Hudman et al., 2012)
Mineral dust aerosols	DEAD (Zender et al., 2003) Ginoux (Ginoux et al., 2001)
Sea salt aerosols	SeaSalt (Chin et al., 2001; Gong, 2003; Jaeglé et al., 2011)
Biogenic VOCs	MEGAN (Guenther et al., 2012)
Biomass burning	GFED (Akagi et al., 2011; Andreae et al., 2001; Giglio et al., 2013; Randerson et al., 2012; van der Werf et al., 2010) FINN (Wiedinmyer et al., 2011)
Volcanic SO ₂	Volcano (Carn et al., 2015; Ge et al., 2016)
Inorganic iodine emissions: HOI, I ₂	Inorg_Iodine (Carpenter et al., 2013; MacDonald et al., 2014)

^a DMS = dimethyl sulfide;

I think it would be useful for the reader to know which are the emission inventories that are currently considered in the database library of HEMCOv3. A summary table or link to a page where this information is specified would work.

Thank you for your suggestion. We have added Table 1, which includes the current emission inventories available in the HEMCO 3.0 library.

Page 8, Lines 189-193:

Table 1 lists the emission inventories currently in the HEMCO 3.0 default database library. Users can select from that list and can easily add new inventories. The hierarchy of emission inventories is specified in the configuration file. Masking is done by superseding lower hierarchy inventories with higher hierarchy inventories so that default inventories may be overwritten by different inventories available only for a particular region, period, or category. For example, an inventory specific for China in 2018 such as MEIC may overwrite a global default inventory.

Table 1. Emission inventories currently in the default HEMCO 3.0 library.

Coverage	Species ^a	Resolution (lat × lon)	Reference
Global	CO, NO _x , NH ₃ , SO ₂ , VOCs, BC, OC, CO ₂	0.5° × 0.5°	CEDS (Hoesly et al., 2018; McDuffie et al., 2020)
Global	CO, NO _x , NH ₃ , SO ₂ , BC, OC	0.1° × 0.1°	EDGAR v4.3.2 (Crippa et al., 2018)
Global	CO, NO, NH ₃ , SO ₂	0.1° × 0.1°	EDGAR-HTAP v2 (Janssens-Maenhout et al., 2012)
Global	Natural Sources NH ₃	2° × 2.5°	GEIA (Bouwman et al., 1997)
Global	Arctic Seabird NH ₃	0.25° × 0.25°	Croft et al. (2016)
Global	Anthropogenic C ₂ H ₅ OH	2° × 2.5°	Millet et al. (2010), Olivier et al. (2003), Granier et al. (2005)
Global	C ₂ H ₆	4° × 5°	Tzompa-Sosa et al. (2017)
Global	C ₃ H ₈	2° × 2.5°	Xiao et al. (2008)
Global	CHBr ₃ , CH ₂ Br ₂	2° × 2.5°	Liang et al. (2010)

Global	CH ₃ I, CH ₂ I ₂ , CH ₂ ICl, CH ₂ I ₂ Br	1° × 1°	Ordóñez et al. (2012), Sherwen et al. (2016)
Global	Aircraft CO, NO _x , SO ₂ , VOCs, BC, OC	1° × 1°	AEIC (Stettler et al., 2011)
Global	CH ₃ CHO, C ₂ H ₅ OH	2° × 2.5°	Millet et al. (2010)
Global	Anthropogenic PM _{2.5} Dust	2° × 2.5°	AFCID (Philip et al., 2017)
Global	Ship SO ₂ , CO, NO _x	1° × 1°	ICOADS (Wang et al., 2008)
Global	Ship SO ₂	1° × 1°	ARCTAS (Eyring et al., 2005)
Global	Ship SO ₂	1° × 1°	Corbett et al. (1999)
Global	Future RCP3PD, RCP4.5, RCP6.0, RCP8.5 CO, NO _x , NH ₃ , SO ₂ , VOCs, BC, OC	0.5° × 0.5°	Holmes et al. (2013)
Global ^a	Fire CO, NO _x , NH ₃ , SO ₂ , VOCs, BC, OC	0.25° × 0.25°	GFED v4.1 (Akagi et al., 2011; Andreae et al., 2001; Giglio et al., 2013; Randerson et al., 2012; van der Werf et al., 2010)
Global	Fire CO, NO _x , NH ₃ , SO ₂ , VOCs, BC, OC, CO ₂	0.1° × 0.1°	GFAS (Di Giuseppe et al., 2018; Rémy et al., 2017; Andela et al., 2013; Kaiser et al., 2012, Xu et al., 2010; Heil et al., 2010; Di Giuseppe et al., 2016)
Global	Fire CO, NO _x , NH ₃ , SO ₂ , VOCs, BC, OC, CO ₂	0.25° × 0.25°	QFED v2.5r1 (Darmenov et al., 2013)
Global	Fire CO, NO _x , NH ₃ , SO ₂ , VOCs, BC, OC	0.25° × 0.25°	BB4CMIP (van Merle et al., 2017)
US	CO, NO _x , NH ₃ , SO ₂ , VOCs, BC, OC	0.1° × 0.1°	EPA (https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data , last retrieved 23 Jul 2021)
Canada	CO, NO _x , NH ₃ , SO ₂ , BC, OC	0.1° × 0.1°	APEI (https://pollution-waste.canada.ca/air-emission-inventory/ , last retrieved 23 Jul 2021)
Africa	CO, NO _x , NH ₃ , SO ₂ , VOCs, BC, OC	0.1° × 0.1°	DICE-Africa (Marais et al., 2016)
China	CO, NO _x , NH ₃ , SO ₂ , VOCs, BC, OC	0.5° × 0.66°	MEIC (http://meicmodel.org/ , last retrieved 23 Jul 2021; Li et al., 2017; Zheng et al., 2018)

^a VOCs = volatile organic compounds; OC = organic carbon aerosol; BC = black carbon aerosol; * implemented as extension to distribute dry matter input data into model species.

Is there any pre-processing associated to the original emission inventories so that they can be ingested by HEMCO (e.g., renaming of pollutant names to have them homogenized across all inventories)?

Thank you for your question. The input data provided to HEMCO needs to be processed into COARDS-compliant netCDF format (http://wiki.seas.harvard.edu/geos-chem/index.php/The_COARDS_netCDF_conventions_for_earth_science_data).

The renaming of pollutant names is not necessary, as the HEMCO configuration file specifies the mapping between the netCDF variable name and the model chemical species name. This allows different models or chemical mechanisms to share the same HEMCO input files without further processing needed for each individual model. It may also be desirable, for performance reasons, to split the original inventory files by year or month, so they can be read faster during the model run. HEMCO will advance through the files automatically as the model time moves forward. We have revised the manuscript to include further detail on adding new inventories.

Page 6, Lines 143-156:

The HEMCO configuration file (example in Fig. 2) controls the operation of all HEMCO layers, fully describing the relationship between the input data read by the Data Input Layer, the processing by the HEMCO Core, and the data passed to the model by the Model Interface Layer. It is organized as individual entries for data, scaling factors, and masks. Each entry is numbered or named and includes information about the source of data (usually a netCDF file name but may be a number or mathematical expression in simple cases). *For netCDF data files, each entry specifies the netCDF variable name to be read (allowing the mapping from netCDF input species to model species), the temporal range, refresh frequency, cycling option (whether to continuously cycle the data or require an exact date match), and spatial dimension (2-D or 3-D data, with the option to specify a custom vertical distribution for 2-D data).* Also included is the model species name, the scaling factors to be applied, and the hierarchy (priority order used for masking). If a data entry does not include a species name, the entry is treated as generic data and is read into HEMCO, scaled, masked, and made

available to the model upon request. Entries may be organized in the form of “collections” enabled or disabled in bulk using switches. *HEMCO comes with a default database library of emission inventories and environmental data sets that is updated with every new GEOS-Chem version release, but users can readily add their own through processing the data into COARDS-compliant netCDF format and providing the corresponding configuration file entries. A detailed HEMCO User Guide is available on the GEOS-Chem Wiki (http://wiki.seas.harvard.edu/geos-chem/index.php/The_HEMCO_User%27s_Guide, retrieved 7 January 2021).*

Page 8-9, Lines 189-199:

Table 1 lists the emission inventories currently in the HEMCO 3.0 default database library. Users can select from that list and can easily add new inventories. The hierarchy of emission inventories is specified in the configuration file. Masking is done by superseding lower hierarchy inventories with higher hierarchy inventories so that default inventories may be overwritten by different inventories available only for a particular region, period, or category. For example, an inventory specific for China in 2018 such as MEIC may overwrite a global default inventory. A simulation for later years may retain the Chinese inventory for 2018, scale it up or down, or default to the global inventory, as specified in the configuration file. Additional emission inventories can be added to the HEMCO library in COARDS-compliant netCDF format. The HEMCO configuration file allows HEMCO to remap variable names in the inventory source file to the species name in the model, specify the region that the inventory is used for, and its precedence over the existing entries in the HEMCO library, without necessitating pre-processing of the inventory source files.

The vertical preprocessing of emissions is not required anymore in HEMCOv3, as it performs an on-line vertical regridding of the original emissions. It is however not clear to me how the user can assign sector-dependent vertical profiles to input emission inventories. This is relevant for those sectors in which emissions are originally reported in a single layer (e.g., emissions from energy sector in CEDS). Perhaps an example could be added in Figure 2.

Thank you for your suggestion. We have clarified the ability to apply scaling factors and specify custom vertical distributions in the manuscript, and have included an example in Fig. 2. Since each entry in the HEMCO configuration file is species-, sector-, and inventory-specific, it is possible to apply a custom scaling factor and vertical profile to a single species, sector, or inventory. This capability is used for distributing emissions vertically by sector in GEOS-Chem, CESM2-GC, and CAM-chem.

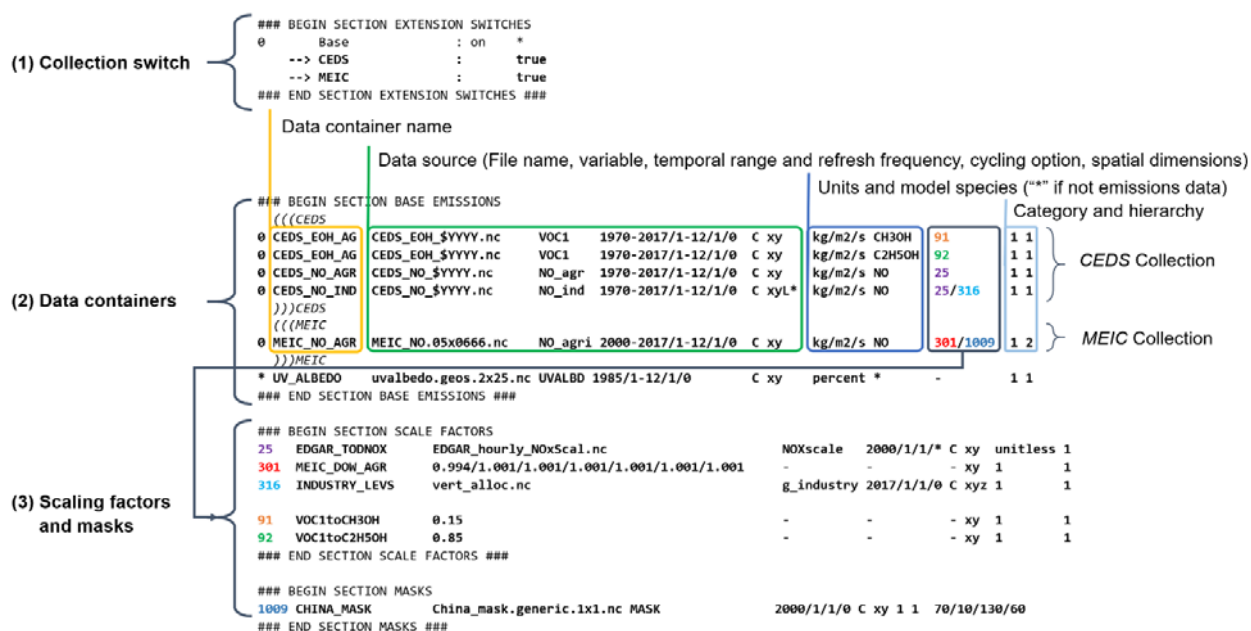


Figure 2. Sample HEMCO configuration file. The HEMCO configuration file is organized in three sections: (1) switches for “collections” of data containers, (2) data containers to be used in the model simulation, optionally organized into “collections”, and (3) scaling and masking rules to be used. Entries are organized in a similar format, including a number and/or name, the data source (netCDF file and variable name, numbers, or mathematical expressions), the temporal range and spatial dimensions, and their category and hierarchy (in the same category, data entries with higher hierarchy take precedence). For data containers, scaling factors and masks are applied by referencing the numbered scaling factor and mask entries (colored text).

Page 6, Lines 143-156:

The HEMCO configuration file (example in Fig. 2) controls the operation of all HEMCO layers, fully describing the relationship between the input data read by the Data Input Layer, the processing by the HEMCO Core, and the data passed to the model by the Model Interface Layer. It is organized as individual entries for data, scaling factors, and masks. Each entry is numbered or named and includes information about the source of data (usually a netCDF file name but may be a number or

mathematical expression in simple cases). For netCDF data files, each entry specifies the netCDF variable name to be read (allowing the mapping from netCDF input species to model species), the temporal range, refresh frequency, cycling option (whether to continuously cycle the data or require an exact date match), and spatial dimension (2-D or 3-D data, with the option to specify a custom vertical distribution for 2-D data). Also included is the model species name, the scaling factors to be applied, and the hierarchy (priority order used for masking). If a data entry does not include a species name, the entry is treated as generic data and is read into HEMCO, scaled, masked, and made available to the model upon request. Entries may be organized in the form of “collections” enabled or disabled in bulk using switches. HEMCO comes with a default database library of emission inventories and environmental data sets that is updated with every new GEOS-Chem version release, but users can readily add their own through processing the data into COARDS-compliant netCDF format and providing the corresponding configuration file entries. A detailed HEMCO User Guide is available on the GEOS-Chem Wiki (http://wiki.seas.harvard.edu/geos-chem/index.php/The_HEMCO_User%27s_Guide, retrieved 7 January 2021).

P6, Lines 168-171:

Scaling factors can be specified for each individual entry in the HEMCO configuration file, allowing for different scaling factors to be applied for different inventories, sectors, and species. HEMCO accepts scaling factors as constant numbers, temporally-explicit (diurnal, day-of-week, seasonal, interannual) numbers, or as a gridded netCDF data file.

Page 10, Lines 229-243:

HEMCO 3.0 computes vertically distributed (3-D) emissions in the same way as 2-D. HEMCO 2.0 assumed the 72-level or the 47-level GEOS grid when reading vertically distributed emissions. This required pre-processing of the emission inventory files from their original vertical grids to the supported GEOS grid. Such pre-processing is not required anymore in HEMCO 3.0, which reads vertical emission data on any sigma-pressure grid described in the input netCDF inventory file. The input data are then vertically re-gridded on-line to the model vertical grid by the Data Input Layer

using the MESSy NCREGRID package (Jöckel, 2006). This functionality is used in the WRF-GC and CESM2 models, which have user-configurable vertical grids. *HEMCO 3.0 is also capable of distributing 2-D input data on a 3-D grid, by modifying the spatial dimension field in the HEMCO configuration file. An example is shown in Fig. 2 where the 2-D NO emission field from the CEDS inventory in the industrial sector is copied vertically to all levels using the flag “xyL*”. Emissions may then be distributed using 3-D scaling factors read from a netCDF file, such as the scaling factor 316 in Fig. 2. It is additionally possible to emit all 2-D data to a particular level (e.g., to emit to level 5, using “xyL5”) or distributed among an altitude range (e.g., “xyL=1:500m” or “xyL=1:PBL”). Emission heights can additionally be read from a netCDF file. The spatial dimension parameter can be applied individually to each entry in the HEMCO configuration file, and thus can be applied to an individual inventory, sector, or species. Detailed documentation of this capability is available on the HEMCO User’s Guide (http://wiki.seas.harvard.edu/geos-chem/index.php/HEMCO_examples#Applying_2D_emissions_vertically, last retrieved 08 Aug 2021).*

HEMCOv3 can operate on a higher resolution than the model working grid for masking and scaling purposes. This feature allows achieving greater accuracy at country borders between different inventory domains. Authors illustrate this function with an example in which two emission inventories are masked at resolutions of 4x5 deg and 2x2.5 deg, the later allowing a better resolution for the US-Mexico border. This is clear to me but, should not be better to directly perform the masking and scaling operations on the original grids of the emission inventories before performing the regridding? Nowadays most of the available global and regional emission inventories are provided at very fine resolutions (i.e., 0.1x0.1 deg), which would allow having a good definition of country borders. Some inventories (e.g., EMEP) even provide the information of the emitting country per grid cell. In that case, using a country mask would not be the best solution. Is the information of emitting country used in HEMCO when given?

P10 L268: There is a limit to the resolution of the HEMCO grid. Could you specify this limit?

Thank you for your questions. We have revised the manuscript to clarify the limit and application of the HEMCO grid. It is necessary to balance the computational requirements needed to process data at a higher resolution and the desired accuracy. When multiple inventories are used, the inventories need to be re-gridded to a common HEMCO grid for processing, as the capability to perform masking by the hierarchy and category of different entries in the HEMCO configuration file has to run on a single grid. For a global model, it may be computationally infeasible to perform the masking and scaling at inventories' native resolution, but this is feasible when running a regional model or with HEMCO off-line.

Page 12-13, Lines 295-319:

Figure 3 illustrates the benefit of using a finer HEMCO grid at the boundaries between inventories. When the HEMCO grid is disabled, HEMCO runs at model resolution, and all input data, including masks, are re-gridded by the Data Input Layer to the model resolution before emissions are computed by the HEMCO Core. In the example of Fig. 3, where a national inventory for the US is to overwrite a global default inventory, this overwriting can be done only for grid cells that are fully in the US. Grid cells straddling the border must retain the global default in order to avoid under- or over-accounting, but this then loses information from the US inventory. *This is not a problem if the national inventory straddles the border and includes information on the fractional contributions from the neighboring country, but such is not the case here.* Using a finer, intermediate-resolution grid – the HEMCO grid – allows emissions at the model grid scale to more accurately blend the contributions from the two sides of the border in a single *model* grid cell. This also ensures greater consistency when using the same model simulations at different resolutions. As long as the HEMCO grid is kept at a single resolution, the calculated emissions will be consistent between simulations – no matter what model grid resolution is selected.

Another advantage of using a finer HEMCO grid is for emissions computed with extensions and dependent on both the model variables (provided on the model grid) and environmental data (provided on the HEMCO grid). If there is non-linear dependence of emissions on the environmental

data variables, then a finer HEMCO grid will produce more accurate emissions. This is the case for example in dust emission algorithms that use land type as a categorical variable.

There is a limit to the resolution of the HEMCO grid because of the need for HEMCO to store the different inventories in memory, *re-grid the data to the HEMCO grid*, and process the data at higher resolution, which may be computationally expensive. *In global simulations using GEOS-Chem Classic on a single machine at resolutions of $4^\circ \times 5^\circ$ or $2^\circ \times 2.5^\circ$ we have found that a HEMCO grid of $1^\circ \times 1.25^\circ$ is a practical limit. However, masking on the native inventory grid at much higher resolution may be desirable for regional modeling applications. One can circumvent the problem by pre-processing the emissions on their native grids using HEMCO in off-line mode. Another option is to use a regional rather than global model. For example, in nested GEOS-Chem Classic simulations at $0.25^\circ \times 0.3125^\circ$ resolution we find that HEMCO can easily handle a HEMCO grid of $0.1^\circ \times 0.1^\circ$, typical of the native resolution of inventories.*

Figure 3: “Using the 4 x finer” – Should not be 2x finer?

Thank you for pointing this out. We have removed the “4x” as the coarse resolution ($4^\circ \times 5^\circ$) and fine resolution ($2^\circ \times 2.5^\circ$) provide sufficient information.