Development of a Grid-Independent GEOS-Chem Chemical Transport Model (v9-02) as an atmospheric chemistry module for Earth System Models.

M.S. Long (mlong@seas.harvard.edu)
School of Engineering and Applied Sciences, Harvard University, Cambridge, MA USA
R. Yantosca (yantosca@seas.harvard.edu)
School of Engineering and Applied Sciences, Harvard University, Cambridge, MA USA
J. E. Nielsen (jon.e.nielsen@nasa.gov)
Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD;
Science Systems and Applications Inc., Lanham, MD
C.A. Keller (ckeller@seas.harvard.edu)
School of Engineering and Applied Sciences, Harvard University, Cambridge, MA USA
A. da Silva (arlindo.m.dasilva@nasa.gov)
Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD
M.P. Sulprizio (mpayer@seas.harvard.edu)
School of Engineering and Applied Sciences, Harvard University, Cambridge, MA USA
S. Pawson (Steven.Pawson-1@nasa.gov)
Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD
D. J. Jacob (djacob@seas.harvard.edu)
School of Engineering and Applied Sciences, Harvard University, Cambridge, MA USA

Abstract

The GEOS-Chem global chemical transport model (CTM), used by a large atmospheric chemistry research community, has been re-engineered to also serve as an atmospheric chemistry module for Earth System Models (ESMs). This was done using an Earth System Modelling Framework (ESMF) interface that operates independently of the GEOS-Chem scientific code, permitting the exact same GEOS-Chem code to be used as an ESM module or as a stand-alone CTM. In this manner, the continual stream of updates contributed by the CTM user community is automatically passed on to the ESM module, which remains state-of-science and referenced to the latest version of the standard GEOS-Chem CTM. A major step in this re-engineering was to make GEOS-Chem grid-independent, i.e., capable of using any geophysical grid specified at run time. GEOS-Chem data "sockets" were also created for communication between modules and with external ESM code. The grid-independent, ESMF-compatible GEOS-Chem is now the standard version of the GEOS-Chem CTM. It has been implemented as an atmospheric chemistry module into the NASA GEOS-5 ESM. The coupled GEOS-5/GEOS-Chem system was tested for scalability and performance with a tropospheric oxidant-aerosol simulation (120 coupled species, 66 transported tracers) using 48-240 cores and MPI distributed-memory parallelization. Numerical experiments demonstrate that the GEOS-Chem chemistry module scales efficiently for the number of cores tested, with no degradation as the number of cores increases. Although inclusion of atmospheric chemistry in ESMs is computationally expensive, the excellent scalability of the chemistry module means that the relative cost goes down with increasing number of cores in a massively parallel environment.

1. *Introduction*

Global modelling of atmospheric chemistry involves solution of the 3-D continuity equations for the concentrations of chemical species including the effects of emissions, transport, chemistry, and deposition. This is commonly done with Chemical Transport Models (CTMs) driven by input meteorological data and surface boundary conditions. CTMs are relatively simple computational tools because the chemical continuity equations are solved without coupling to atmospheric dynamics. They are adequate for many applications and play a central role in advancing knowledge of atmospheric chemistry. However, there is also increasing demand for atmospheric chemistry to be implemented as a coupled module in Earth System Models (ESMs) that represent the ensemble of processes affecting the Earth system. Here we describe a software framework through which the state-of-science GEOS-Chem CTM can be implemented seamlessly as a module in ESMs, so that the stand-alone CTM and the ESM module use exactly the same code. We describe the deployment of this capability in the NASA Goddard Earth Observing System (GEOS) developed at NASA's Global Modelling and Assimilation Office (GMAO).

GEOS-Chem (http://www.geos-chem.org) is a shared-memory parallel (OpenMP) global 3-D Eulerian CTM driven by assimilated meteorological data (Bey et al., 2001). It is used by over 100 research groups worldwide for a wide range of applications including simulation of tropospheric oxidants (Mao et al., 2013), aerosols (Fairlie et al., 2007; Jaeglé et al., 2011; Park et al., 2004; Trivitayanurak et al., 2008), carbon gases (Nassar et al., 2010; Wang et al., 2004; Wecht et al., 2014), mercury (Holmes et al., 2010; Selin et al., 2008), and stratospheric chemistry (Eastham et al., 2014; Murray et al., 2012). Development of GEOS-Chem is based on core principles of open-source code development, modular structure, nimble approach to innovation, strong version control,rigorous quality assurance (QA), extensive documentation, and user support . The large user base permits extensive model diagnosis and generates a continual stream of new developments to maintain the model at the forefront of the science. Implementation of these developments in the standard GEOS-Chem code can be done quickly and efficiently because of the simplicity of the code and the common interests of the user community. Maintaining state-of-science capability is more challenging in ESMs because of complexity of managing the central code and the need for dialogue across research communities to prioritize model development. On the other hand, CTMs such as GEOS-Chem have more difficulty staying abreast of high-performance computing (HPC) technology because of limited software engineering resources.

Here we present a re-engineered standard version of the GEOS-Chem CTM capable of serving as a flexible atmospheric chemistry module for ESMs. A key innovation is that GEOS-Chem is now grid-independent, i.e., it can be used with any geophysical grid. The same standard GEOS-Chem code can be integrated into ESMs through the Earth System Modelling Framework (ESMF, Hill et al., 2004) interface, or used as before as a stand-alone CTM driven by assimilated meteorological data. The re-

engineered grid-independent flexibility has been integrated into the standard open-code version of the GEOS-Chem CTM. The exact same scientific code in the GEOS-Chem CTM now serves as atmospheric chemistry module in the GEOS-5 ESM of the NASA Global Modeling and Assimilation Office (GMAO) (Molod et al., 2012). Scientific updates to the GEOS-Chem CTM contributed by its user community and incoroporated in the standard model following QA are automatically integrated into the GEOS-5 ESM, so that the ESM effortlessly remains state-of-science and traceable to the latest standard version of GEOS-Chem.

2. Grid-Independent GEOS-Chem Model Description

The GEOS-Chem CTM consists of four modules executing operations for chemistry and dry deposition, emissions, wet deposition, and transport (Fig. 1). GEOS-Chem solves the general Eulerian form of the coupled continuity equations for m chemical species with number density vector $\mathbf{n} = (n_1, \ldots, n_m)^{\mathrm{T}}$

$$\frac{\partial n_i}{\partial t} = -\nabla \cdot (n_i \mathbf{U}) + P_i(\mathbf{n}) - L_i(\mathbf{n}) \qquad i \in [1, m] \qquad (1)$$

Here $\mathbf{U}$ is the wind vector (including sub-grid components parameterized as turbulent diffusion and convection). and $P_i(\mathbf{n})$ and $L_i(\mathbf{n})$ are the local production and loss rates of species i including terms to describe chemical reactions, aerosol microphysics, emissions, precipitation scavenging, and dry deposition. In GEOS-Chem, as in all 3-D CTMs, equation (1) is solved by operator splitting to separately and successively apply concentration updates over finite time steps from a transport operator

$$\frac{\partial n_i}{\partial t} = -\nabla \cdot (n_i \mathbf{U}) \qquad i \in [1, m] \qquad (2)$$

and a local operator (commonly called chemical operator)

$$\frac{dn_i}{dt} = P_i(\mathbf{n}) - L_i(\boldsymbol{n}) \qquad i \in [1, m] \qquad (3)$$

The transport operator includes no coupling between species, while the chemical operator has no spatial coupling. The transport operator is further split into 1-D advection operators, a convection operator, and a boundary layer mixing operator. Operator splitting breaks down the multi-dimensionality of the coupled system (1) and enables numerical solution by finite differencing. The chemical operator in GEOS-Chem is further split into chemistry and dry deposition, emissions, and wet deposition modules for computational convenience. Gravitational settling of particles is treated as part of the chemical operator.

Wet deposition from sub-grid convective precipitation cannot be decoupled from convective transport (Balkanski et al., 1993) and is treated as part of convection in the transport operator.

The transport operators in the standard GEOS-Chem CTM are applied on fixed latitude-longitude grids (e.g. Wu et al. 2007). When integrated into an ESM, GEOS-Chem does not need to calculate its own transport; this is done separately in the ESM as part of the simulation of atmospheric dynamics, where transport of chemical species is done concurrently with transport of meteorological variables. Thus the ESM only uses GEOS-Chem to solve the chemical operator (3) over specified time steps. The GEOS-Chem chemical operator must in turn be able to accommodate any ESM grid and return concentration updates on that grid.

The chemical operator has no spatial dimensionality (0-D) and could in principle be solved independently for all grid points of the ESM. However, grouping the grid points by column is more efficient as it permits simultaneous calculation of radiative transfer, precipitation scavenging, gravitational settling, and vertically distributed emissions for all grid points within the column. Thus we take a 1-D vertical column as the minimum set of grid points to be handled by a a call to the chemical operator. Chemical operator updates for a given column can be completed without information from neighboring columns. Solving for the chemical operator column by column reduces memory overhead and facilitates scalable single program, multiple data (SPMD; Cotronis and Dongarra, 2001) parallelization in a distributed computing environment using the Message Passing Interface (MPI). It may sometimes be preferable to apply the chemical operator to ensembles of columns, grouped independent of geography, to balance the computational burden and achieve performance gains (Long et al., 2013).

Prior to this work, the horizontal grid of GEOS-Chem was defined at compile time from a limited selection of fixed latitude-longitude grids ($1/4^{o}$x$5/16^{o}$, $1/2^{o}$x$2/3^{o}$, 1^{o}x1^{o}, 2^{o}x2.5^{o}, 4^{o}x5^{o}) compatible with the advection module and offline meteorological fields. Our goal here was to re-engineer the existing GEOS-Chem code to accept any horizontal grid defined at runtime. The horizontal grid would be able to span the entire global domain, represent a single column to be calculated on a single compute node, or represent any collection of columns defined by their location. This permits use of the same scientific code for stand-alone CTM and coupled ESM applications.

2.1 *Code Modularization and Structure*

In order for the GEOS-Chem code to permit run-time horizontal grid definition, much of the FORTRAN-77 code base was updated to Fortran-90. This included extensive conversion of static to dynamically-allocatable arrays, and introduction of pointer-based derived data types. Data flow into, through, and out of GEOS-Chem's routines was reconfigured to use derived-type objects passed to routines as arguments in place of publicly-declared global-scope variables. This permitted the bundling of

data structures with similar functionality into common interfaces (data "sockets") that simplify module communication within GEOS-Chem and coupling to external components through the ESMF interface (see Section 2.2). Three sockets are defined: a meteorology and physics socket, a chemistry socket, and an input options socket. The meteorology and physics socket provides data defining geophysical state variables and arrays. This includes temperature, pressure, humidity, wind fields, and many others. The chemistry socket provides data structures for chemical species including indexing, species names, and concentrations. The input options socket provides runtime information such as calendar, grid dimensions, diagnostic definitions, and locations of offline information stored on disk. Together, these sockets incorporate all of the quantities and fields necessary for coupling to and driving modules within GEOS-Chem.

The GEOS-Chem code includes specific conditional-compilation flags to accommodate the ESMF interface and permit coupling with external data streams. These flags do not interfere with GEOS-Chem's scientific operation and are used exclusively in grid, I/O, and utility operations. There are three flags invoked as C-preprocessor statements: *ESMF_*, *EXTERNAL_GRID*, and *EXTERNAL_FORCING*. Code bounded by these flags is neither compiled nor executed unless the specific flag is enabled at compile time. The *ESMF_* flag bounds code specific for the ESMF. The *EXTERNAL_GRID* flag bounds code that allows GEOS-Chem to operate on an externally defined and initialized grid (e.g. by an ESM). The *EXTERNAL_FORCING* flag bypasses GEOS-Chem's internal, offline data I/O operations necessary for CTM operation, and replaces them with ESMF-based I/O. Users do not need to have the ESMF installed in order to run GEOS-Chem as a stand-alone CTM. The system reverts to the standard GEOS-Chem CTM code relying on the legacy module interface when compiled without these flags enabled.

The recently developed Harvard-NASA Emissions Component HEMCO (http://wiki-geos-chem.org/HEMCO/) is used for emission calculations (Keller et al., 2014). HEMCO is a Fortran-90 based, ESMF compliant, highly customizable module that uses base emissions and scale factors from a library of emission inventories to construct time-dependent emission field arrays. Emission inventories and scale factors are selected by the user in a HEMCO-specific configuration file. Emission inventories for different species and source types need not be of the same grid dimensions or domain.

The redesign of GEOS-Chem's data structures was meant to simplify coupling of GEOS-Chem with any ESM regardless of its ESMF compatibility. In the absence of an ESMF interface, users would be required to engineer a specific interface for their ESM, However, GEOS-Chem's data sockets and conditional-compilation flags facilitate this task by having all input and output data structures and associated methods conveniently located in a few specific modules.

As with all modifications to the publicly-available GEOS-Chem source code, changes made for ESM coupling and grid-independence were subject to rigorous QA by conducting prescribed 1-month and 1-

year test simulations as benchmarks (http://acmg.seas.harvard.edu/geos/geos_benchmark.html), and comparing results to the benchmarks of the previous model version. Our changes were not expected to modify any aspect of the benchmark simulation results and we verified that they did not. Results from the benchmark simulations for version 9-02k can be found here: http://wiki.seas.harvard.edu/geos-chem/index.php/GEOS-Chem_v9-02_benchmark_history#v9-02k.

2.2 *ESMF Interface*

We made GEOS-Chem ESMF-compatible for interfaceing with external ESMs. The ESMF is an open-source software application programming interface that provides a standardized high-performance software infrastructure for use in ESM design. It facilitates HPC, portability, and interoperability in Earth science applications (Collins et al., 2005).

GEOS-Chem is executed within the ESMF as a gridded component. The gridded component is the basic element of an ESMF-based program, and is defined as a set of discrete scientific and computational functions that operate on a geophysical grid. Likewise, other components of the Earth system are implemented as gridded components (e.g. atmospheric dynamics, ocean dynamics, terrestrial biogeochemistry, etc.).

Each gridded component consists of a routine establishing ESMF-specific services, and *Initialize*, *Run*, and *Finalize* operations methods for gridded component execution by the ESMF. The *Initialize* method is executed once at the beginning of the simulation and initializes component-specific runtime parameters. The *Run* method interfaces local data structures with ESMF States (see below) and executes the component code (GEOS-Chem in our case). The *Finalize* method wraps up code execution, closes any remaining open files, finalizes I/O and profiling processes, and flushes local memory.

Gridded components exchange information with each other through States. A State is an ESMF derived type that can contain multiple types of gridded and non-gridded information (Collins et al., 2005; Suarez et al., 2013). An ESMF gridded component is associated with an *Import* State and an *Export* State. The *Import* State provides access to data created by other gridded components. The *Export* State contains data that a component generates and makes available to other components. In the ESMF-enabled GEOS-Chem, data are passed into and out of the GEOS-Chem gridded component via interfacing an appropriate State with a corresponding GEOS-Chem data socket (Figure 1), making these data available within GEOS-Chem or to other ESM gridded components (see Section 2.1).

The ESMF was implemented within GEOS-Chem as an independent layer that operates on top of the CTM code. It includes code for interfacing with and executing GEOS-Chem as an ESMF gridded component. When coupling GEOS-Chem to an ESM, the GEOS-Chem transport modules are excluded

and only those modules necessary to solve Eq. (3) are used. Coupling specifically to the GEOS-5 ESM required an adaptation of GEOS-Chem's ESMF interface for the GMAO's Modelling, Analysis and Prediction Layer (MAPL) extension (Suarez et al., 2013). MAPL is otherwise not required for GEOS-Chem.

3 Implementation, Performance, and Scalability

The ESMF-enabled GEOS-Chem was embedded within the NASA GEOS-5 ESM (version Ganymed-4.0). The GEOS-5 ESM is the forward model of the GEOS-5 atmospheric data assimilation system (GEOS-DAS) (Ott et al., 2009; Rienecker et al., 2008). The system is built upon on an ESMF framework, and uses a combination of distributed memory (MPI) and, in some cases, hybrid distributed/shared memory parallelization. The dynamical core used here is based on Lin (2004), and operates on horizontal grid resolutions ranging from 2^{o}x2.5^{o} to 0.25^{o}x0.3125^{o}, with 72 vertical layers up to 0.01 hPa. Ocean surface and sea-ice boundaries are prescribed. The land and snow interfaces are based on Koster et al. (2000) and Stieglitz et al. (2001), respectively. For the coupled simulations, GEOS-5 ESM native dynamics and moist physics are applied to the GEOS-Chem chemical tracers.

All coupled GEOS-5/GEOS-Chem simulations were performed on the Discover system at the NASA Goddard Space Flight Center (http://www.nccs.nasa.gov/discover_front.html), using 12-core (dual hex-core) 2.8 GHz Intel Xeon Westmere (X5660) compute nodes equipped with 24 GB RAM, and an Infiniband DDR interconnect using the Intel compiler suite (v. 13.1.1) and MVAPICH2 (v. 1.8.1). GEOS-Chem's shared-memory (OpenMP) parallelization was disabled.

The coupled GEOS-5/GEOS-Chem system was tested on 2^{o}x2.5^{o} and 0.5^{o}x0.625^{o} grids with a standard oxidant-aerosol simulation using 120 chemical species of which 66 are transported ("chemical tracers"). Radical species with very short chemical lifetimes are not transported. The chemistry module used the RODAS-3 (4-stage, order 3(2), stiffly accurate) solver with self-adjusting internal time step (Hairer and Wanner, 1996) as part of the Kinetics Pre-processor (KPP, Eller et al., 2009; Sandu and Sander, 2006). KPP was implemented with its supplied linear algebra (BLAS Level-1) routines in place. The 2^{o}x2.5^{o} simulation used a time step of 1800 seconds for all operations. For the 0.5^{o}x0.625^{o} simulation, chemistry and system-operation time steps were both 450 seconds. Dynamics, physics, and radiation time steps were 900 seconds. For both simulations, the atmosphere used 72 vertical hybrid-sigma (pressure) levels. Simulations were run for 31 days initialized on July 1, 2006. All chemical tracers were initialized from output of a GEOS-Chem CTM (v9-02) simulation.

The 2^{o}x2.5^{o} coupled simulations were used to test scalability of the coupled system and for comparison to the GEOS-Chem CTM. We conducted simulations with 48, 96, 144, 192, and 240 total MPI processes operating on 12x4, 12x8, 12x12, 16x12, and 16x15 (lat x lon) contiguous grid point subdomains,

respectively. This represents a set of five simulations $j \in [1, 5]$. For comparison, the offline GEOS-Chem CTM (v9-02) was run on 8 shared-memory processes at 2^{o}x2.5^{o} resolution using 8-core 2.6 GHz Intel Xeon processors, reflecting a typical CTM set-up, using otherwise identical settings and initial chemical conditions as the coupled GEOS-5/GEOS-Chem simulations. Since GEOS-5 is a pure MPI application, each MPI process corresponds to a single processor core.

Figure 2 gives execution wall times for the total simulation and for the chemistry (GEOS-Chem) and dynamics gridded components. To analyze the performance and scalability results, we define the normalized scaling efficiency *S* for simulation *j* relative to simulation *j-1* as

$$S = \left(\frac{W_{x,j-1} - W_{x,j}}{W_{x,j-1}}\right)\left(\frac{N_j}{N_j - N_{j-1}}\right) \qquad (4)$$

where $W_{x,j}$ is the wall time for component *x*, and N_j is the number of cores allocated to the simulation. *S* measures how efficiently the addition of computational resources speeds up execution. For example, a value of 0.9 indicates that a doubling of computational resources decreases wall time by a factor of 1.8. A value of zero means no speed-up. A negative value means slow-down, as might result from increasing I/O. Results for 48 cores ($j = 1$) are given relative to the 8-process GEOS-Chem CTM simulation ($j = 0$), which uses different shared-memory processes and a different transport code for chemical tracers only. The two simulations are not strictly comparable but results serve to benchmark the performance of the GEOS-5/GEOS-Chem system against the GEOS-Chem CTM.

We find that the scaling efficiency for the chemistry module (GEOS-Chem) in the GEOS-5/GEOS-Chem system is 0.78 ± 0.10 for the range of cores tested. This represents excellent performance, with no decline as the number of cores increases, reflecting the independent nature of the chemistry calculation for individual columns. For that reason, we expect the excellent scalability of the chemistry module to extend to any number of cores. Scaling efficiency of the dynamics component decreases with increasing number of cores and becomes negative above 192. This reflects the small number of gridpoints allocated to individual cores increasing relative cost of communicating between processes versus operating within local memory, as well as a greater inter-nodal communication associated with additional chemical tracers. The results further suggest that the chemistry module would remain efficient for simulations beyond the range of values tested.

The 0.5^{o}x0.625^{o} resolution simulation was used to examine the performance of the GEOS-5/GEOS-Chem system when operating on a finer grid resolution than permitted by the GEOS-Chem CTM using shared-memory OpenMP parallelization. The higher resolution also increases the problem size, permitting the efficient use of more computing power. For this simulation, the horizontal grid was decomposed into

24x25 lat/lon blocks over 600 cores. The 0.5^{o}x0.625^{o} resolution simulation completed 0.35 simulation years per wall day.

About 20% of the wall time spent on chemistry in the GEOS-5/GEOS-Chem system was spent copying and flipping the vertical dimension of chemical tracer arrays between the GEOS-5 ESM and GEOS-Chem. This would be overcome to a large extent by linking GEOS-Chem tracer arrays to the ESMF using pointers, which access memory locations of preexisting variables directly. This cannot be done within the GEOS-5 ESM for two reasons: (1) GEOS-Chem stores concentrations in double-precision arrays, while the GEOS-5 system generally uses single precision. (2) GEOS-Chem indexes concentration arrays vertically from the surface of the Earth upward while the GEOS-5 system does the reverse. Such limitations are not intrinsic to GEOS-Chem and depend on the specific ESM to which GEOS-Chem is coupled; other ESMs may use different data precision and indexing. Further software engineering in GEOS-Chem could add flexibility in array definitions to accommodate different ESM configurations.

Figure 3 illustrates model results with 500 hPa O_3 mixing ratios at 12 UT on July 15, 2006 for GEOS-5/GEOS-Chem simulations at 2x2.5^{o} and 0.5^{o}x0.625^{o} resolutions, and for the GEOS-Chem CTM using GEOS-5 assimilated meteorological data at 2^{o}x2.5^{o} resolution. All three simulations are initialized from the same GEOS-Chem CTM fields at 0 UT on July 1, 2006, but have different meteorology because of differences in resolution and also because the CTM uses assimilated meteorological data while the GEOS-5/GEOS-Chem system in this implementation does not. The Figure demonstrates the fine structure of chemical transport that can be resolved with the 0.5^{o}x0.625^{o} resolution. The general patterns are roughly consistent between simulations and are reasonable compared to satellite and sonde observations (Zhang et al., 2010). A scatterplot comparing output from the different simulations (Figure 4) shows that they have comparable results. Figures 3 and 4 are intended to illustrate the GEOS-5/GEOS-Chem capability. A more thorough evaluation of GEOS-Chem's chemistry within the GEOS-5 system would require the use of the same meteorological data as the offline CTM, diagnosing the full ensemble of simulated chemical species, and investigating the effect of transport errors when using off-line meteorological fields in the CTM. This will be documented in a separate publication.

4 *Summary*

We have presented a new grid-independent version of the GEOS-Chem chemical transport model (CTM) to serve as atmospheric chemistry module within Earth system models (ESMs) through the Earth System Modelling Interface (ESMF). The new GEOS-Chem version uses any grid resolution or geometry specified at runtime. The exact same standard GEOS-Chem code (freely available from http://geos-chem.org) supports both ESM and stand-alone CTM applications. This ensures that the continual stream

of innovation from the worldwide community contributing to the stand-alone CTM is easily incorporated into the ESM version. The GEOS-Chem ESM module thus always remains state-of-science.

We implemented GEOS-Chem as an atmospheric chemistry module within the NASA GEOS-5 ESM and performed a tropospheric oxidant-aerosol simulation (120 coupled chemical species, 66 transported tracers) in that fully coupled environment. Analysis of scalability and performance for 48 to 240 cores shows that the GEOS-Chem atmospheric chemistry module scales efficiently with no degradation as the number of cores increases, reflecting the independent nature of the chemical computation for individual grid columns. Although the inclusion of detailed atmospheric chemistry in an ESM is a major computational expense, chemistry operations become relatively more efficient as the number of cores increases due to their efficient scalability.

Code Availability. GEOS-Chem source code is freely available to the public. Source code may be downloaded by following instructions found at http://wiki.seas.harverd.edu/geos-chem. At time of writing, this work used a modified version of GEOS-Chem version 9-02k as indicated in the text. All developments presented here are now included with the current GEOS-Chem version 10-01f.

Acknowledgments. This work was supported by the NASA Modelling, Analysis and Prediction (MAP) Program. The authors thank Ben Auer (NASA-GMAO) and Jack Yatteau (Harvard University) for technical assistance.

References:

Balkanski, Y. J., Jacob, D. J., Gardner, G. M., Graustein, W. C. and Turekian, K. K.: Transport and residence times of tropospheric aerosols inferred from a global three-dimensional simulation of 210Pb, J. Geophys. Res. Atmospheres, 98(D11), 20573–20586, doi:10.1029/93JD02456, 1993.

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J. and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, J. Geophys. Res., 106(D19), 23073–23, 2001.

Collins, N., Theurich, G., DeLuca, C., Suarez, M., Trayanov, A., Balaji, V., Li, P., Yang, W., Hill, C. and Silva, A. da: Design and Implementation of Components in the Earth System Modeling Framework, Int. J. High Perform. Comput. Appl., 19(3), 341–350, doi:10.1177/1094342005056120, 2005.

Cotronis, Y. and Dongarra, J.: Recent Advances in Parallel Virtual Machine and Message Passing Interface: 8th European PVM/MPI Users' Group Meeting, Santorini/Thera, Greece, September 23-26, 2001. Proceedings, Springer., 2001.

Eastham, S. D., Weisenstein, D. K. and Barrett, S. R. H.: Development and evaluation of the unified tropospheric–stratospheric chemistry extension (UCX) for the global chemistry-transport model GEOS-Chem, Atmos. Environ., 89, 52–63, doi:10.1016/j.atmosenv.2014.02.001, 2014.

Eller, P., Singh, K., Sandu, A., Bowman, K., Henze, D. K. and Lee, M.: Implementation and evaluation of an array of chemical solvers in the Global Chemical Transport Model GEOS-Chem, Geosci. Model Dev., 2(2), 89–96, 2009.

Fairlie, D. T., Jacob, D. J. and Park, R. J.: The impact of transpacific transport of mineral dust in the United States, Atmos. Environ., 41(6), 1251–1266, doi:10.1016/j.atmosenv.2006.09.048, 2007.

Hairer, E. and Wanner, G.: Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems, Springer., 1996.

Hill, C., DeLuca, C., Suarez, M., da Silva, A. and others: The architecture of the earth system modeling framework, Comput. Sci. Eng., 6(1), 18–28, 2004.

Holmes, C. D., Jacob, D. J., Corbitt, E. S., Mao, J., Yang, X., Talbot, R. and Slemr, F.: Global atmospheric model for mercury including oxidation by bromine atoms, Atmos Chem Phys, 10(24), 12037–12057, doi:10.5194/acp-10-12037-2010, 2010.

Jaeglé, L., Quinn, P. K., Bates, T. S., Alexander, B. and Lin, J.-T.: Global distribution of sea salt aerosols: new constraints from in situ and remote sensing observations, Atmospheric Chem. Phys., 11(7), 3137–3157, doi:10.5194/acp-11-3137-2011, 2011.

Keller, C. A., Long, M. S., Yantosca, R. M., Da Silva, A. M., Pawson, S. and Jacob, D. J.: HEMCO v1.0: A versatile, ESMF-compliant component for calculating emissions in atmospheric models, Geosci. Model Dev. Discuss., 7(1), 1115–1136, doi:10.5194/gmdd-7-1115-2014, 2014.

Koster, R. D., Suarez, M. J., Ducharne, A., Stieglitz, M. and Kumar, P.: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure, J. Geophys. Res. Atmospheres, 105(D20), 24809–24822, doi:10.1029/2000JD900327, 2000.

Lin, S.-J.: A "vertically Lagrangian" finite-volume dynamical core for global models, Mon. Weather Rev., 132(10), 2293–2307, 2004.

Long, M. S., Keene, W. C., Easter, R., Sander, R., Kerkweg, A., Erickson, D., Liu, X. and Ghan, S.: Implementation of the chemistry module MECCA (v2.5) in the modal aerosol version of the Community Atmosphere Model component (v3.6.33) of the Community Earth System Model, Geosci. Model Dev., 6(1), 255–262, doi:10.5194/gmd-6-255-2013, 2013.

Mao, J., Paulot, F., Jacob, D. J., Cohen, R. C., Crounse, J. D., Wennberg, P. O., Keller, C. A., Hudman, R. C., Barkley, M. P. and Horowitz, L. W.: Ozone and organic nitrates over the eastern United States: Sensitivity to isoprene chemistry, J. Geophys. Res. Atmospheres, 118(19), 2013JD020231, doi:10.1002/jgrd.50817, 2013.

Molod, A., Takacs, L., Suarez, M., Bacmeister, J., Song, I.-S. and Eichmann, A.: The GEOS-5 atmospheric general circulation model: Mean climate and development from MERRA to Fortuna. [online] Available from: http://ntrs.nasa.gov/search.jsp?R=20120011790 (Accessed 16 June 2014), 2012.

Murray, L. T., Jacob, D. J., Logan, J. A., Hudman, R. C. and Koshak, W. J.: Optimized regional and interannual variability of lightning in a global chemical transport model constrained by LIS/OTD satellite data, J. Geophys. Res. Atmospheres, 117(D20), D20307, doi:10.1029/2012JD017934, 2012.

Nassar, R., Jones, D. B. A., Suntharalingam, P., Chen, J. M., Andres, R. J., Wecht, K. J., Yantosca, R. M., Kulawik, S. S., Bowman, K. W., Worden, J. R., Machida, T. and Matsueda, H.: Modeling global atmospheric CO2 with improved emission inventories and CO2 production from the oxidation of other carbon species, Geosci Model Dev Discuss, 3(3), 889–948, doi:10.5194/gmdd-3-889-2010, 2010.

Ott, L. E., Bacmeister, J., Pawson, S., Pickering, K., Stenchikov, G., Suarez, M., Huntrieser, H., Loewenstein, M., Lopez, J. and Xueref-Remy, I.: Analysis of Convective Transport and Parameter Sensitivity in a Single Column Version of the Goddard Earth Observation System, Version 5, General Circulation Model, J. Atmospheric Sci., 66(3), 627–646, doi:10.1175/2008JAS2694.1, 2009.

Park, R. J., Jacob, D. J., Field, B. D., Yantosca, R. M. and Chin, M.: Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy, J. Geophys. Res. Atmospheres, 109(D15), D15204, doi:10.1029/2003JD004473, 2004.

Rienecker, M. M., Suarez, M. J., Todling, R., Bacmeister, J., Takacs, L., Liu, H. C., Gu, W., Sienkiewicz, M., Koster, R. D., Gelaro, R., Stajner, I. and Nielsen, J. E.: The GEOS-5 Data Assimilation System-Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. [online] Available from: http://ntrs.nasa.gov/search.jsp?R=20120011955 (Accessed 16 June 2014), 2008.

Sandu, A. and Sander, R.: Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1, Atmos Chem Phys, 6(1), 187–195, doi:10.5194/acp-6-187-2006, 2006.

Selin, N. E., Jacob, D. J., Yantosca, R. M., Strode, S., Jaeglé, L. and Sunderland, E. M.: Global 3-D land-ocean-atmosphere model for mercury: Present-day versus preindustrial cycles and anthropogenic enrichment factors for deposition, Glob. Biogeochem. Cycles, 22(2), GB2011, doi:10.1029/2007GB003040, 2008.

Stieglitz, M., Ducharne, A., Koster, R. and Suarez, M.: The Impact of Detailed Snow Physics on the Simulation of Snow Cover and Subsurface Thermodynamics at Continental Scales, J. Hydrometeorol., 2(3), 228–242, 2001.

Suarez, M., Trayanov, A., da Silva, A. and Chakraborty, P.: MAPL Manual, [online] Available from: http://geos5.org/wiki/images/f/fa/MAPL_UsersGuide.pdf, 2013.

Trivitayanurak, W., Adams, P. J., Spracklen, D. V. and Carslaw, K. S.: Tropospheric aerosol microphysics simulation with assimilated meteorology: model description and intermodel comparison, Atmos Chem Phys, 8(12), 3149–3168, doi:10.5194/acp-8-3149-2008, 2008.

Wang, J. S., Logan, J. A., McElroy, M. B., Duncan, B. N., Megretskaia, I. A. and Yantosca, R. M.: A 3-D model analysis of the slowdown and interannual variability in the methane growth rate from 1988 to 1997, Glob. Biogeochem. Cycles, 18(3), GB3011, doi:10.1029/2003GB002180, 2004.

Wecht, K. J., Jacob, D. J., Frankenberg, C., Jiang, Z. and Blake, D. R.: Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY satellite data, J. Geophys. Res. Atmospheres, 119(12), 2014JD021551, doi:10.1002/2014JD021551, 2014.

Wu, S., Mickley, L. J., Jacob, D. J., Logan, J. A., Yantosca, R. M. and Rind, D.: Why are there large differences between models in global budgets of tropospheric ozone?, J. Geophys. Res. Atmospheres, 112(D5), D05302, doi:10.1029/2006JD007801, 2007.

Figure 1. Coupling between the GEOS-Chem CTM (dashed beige box) and an ESM (blue box). The schematic shows how the coupling is managed through the ESMF, and utilizes only the GEOS-Chem components bound by the ESM box: Transport modules in the GEOS-Chem CTM are bypassed and replaced by the ESM transport modules through the atmospheric dynamics simulation .

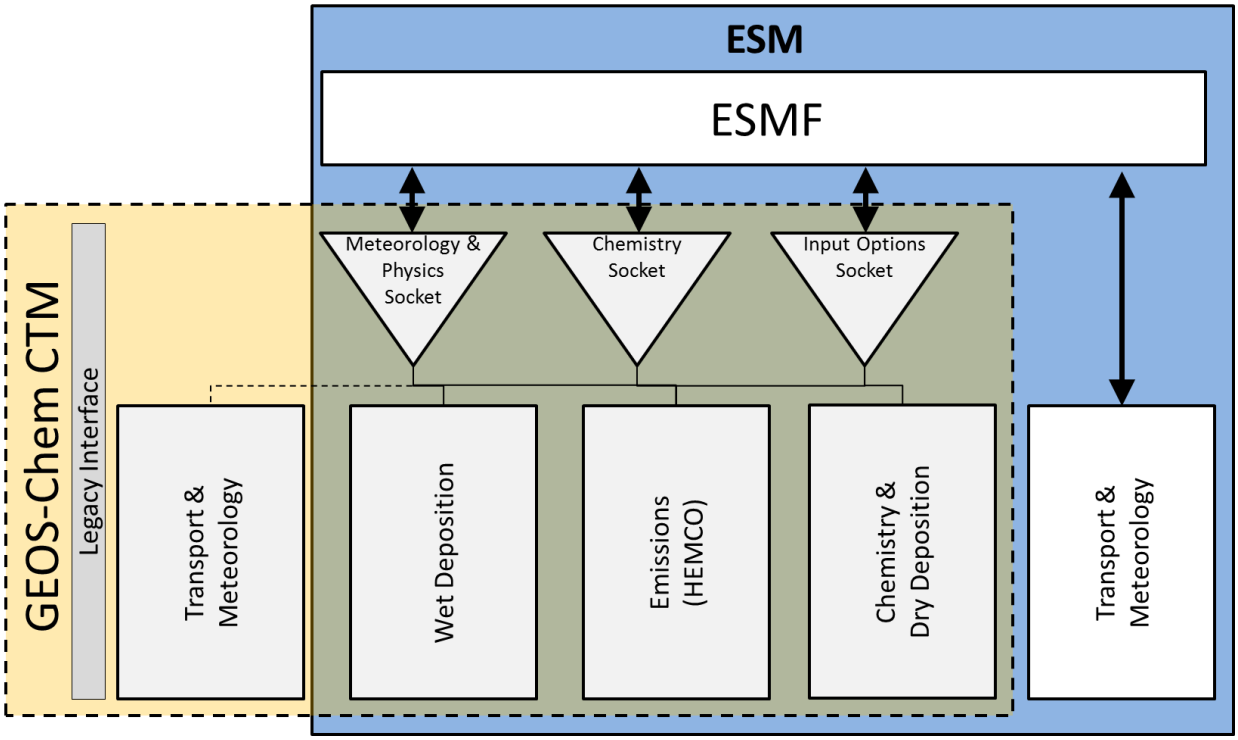

Figure 2. Performance and scalability of the GEOS-5/GEOS-Chem system for a 1-month test simulation including detailed oxidant-aerosol tropospheric chemistry at 2^{o}x2.5^{o} horizontal resolution. Top panel: total and stacked wall-times for the chemical operator (GEOS-Chem), dynamics, and other routines versus number of processor cores. Bottom panel: Scaling efficiency (Eq. 4) for chemistry, dynamics, and the full GEOS-5/GEOS-Chem system. Values shown for 48 cores are relative to the 8-process shared-memory GEOS-Chem CTM.

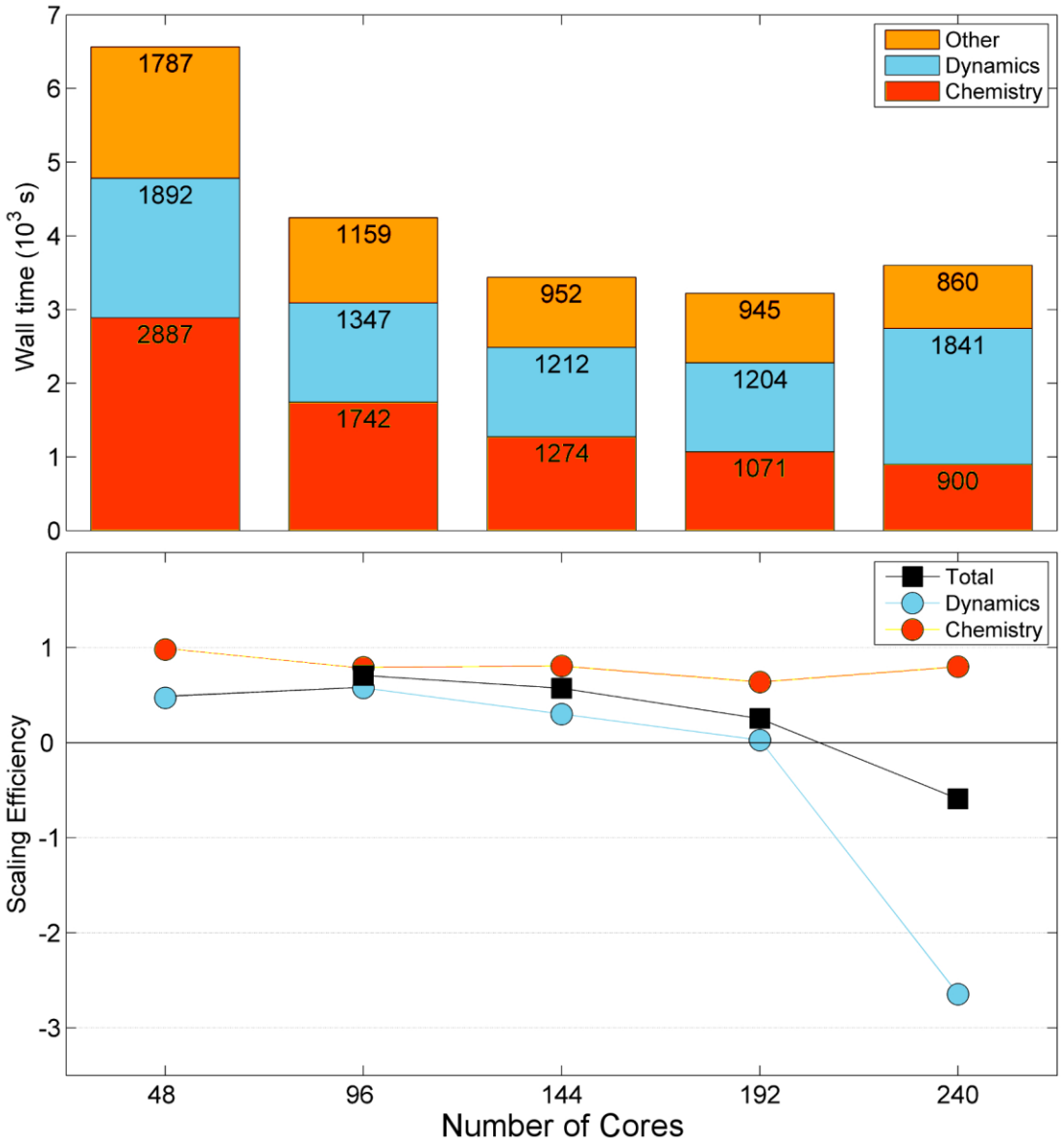

Figure 3. Instantaneous 500 hPa ozone mixing ratios (nmol mol^{-1}) at 12 UT on July 15, 2006, for CTM and ESM implementations of GEOS-Chem. Top panel: GEOS-Chem CTM at 2^{o}x2.5^{o} resolution driven by GEOS-5 assimilated meteorological data with 0.5^{o}x0.67^{o} resolution. Middle panel: GEOS-5/GEOS-Chem ESM at 2^{o}x2.5^{o} resolution. Bottom panel: GEOS-5/GEOS-Chem ESM at 0.5^{o}x0.625^{o} resolution. All three simulations are initialized with the same GEOS-Chem CTM fields at 0 UT on July 1, 2006, but the ESM as implemented here does not include meteorological data assimilation.

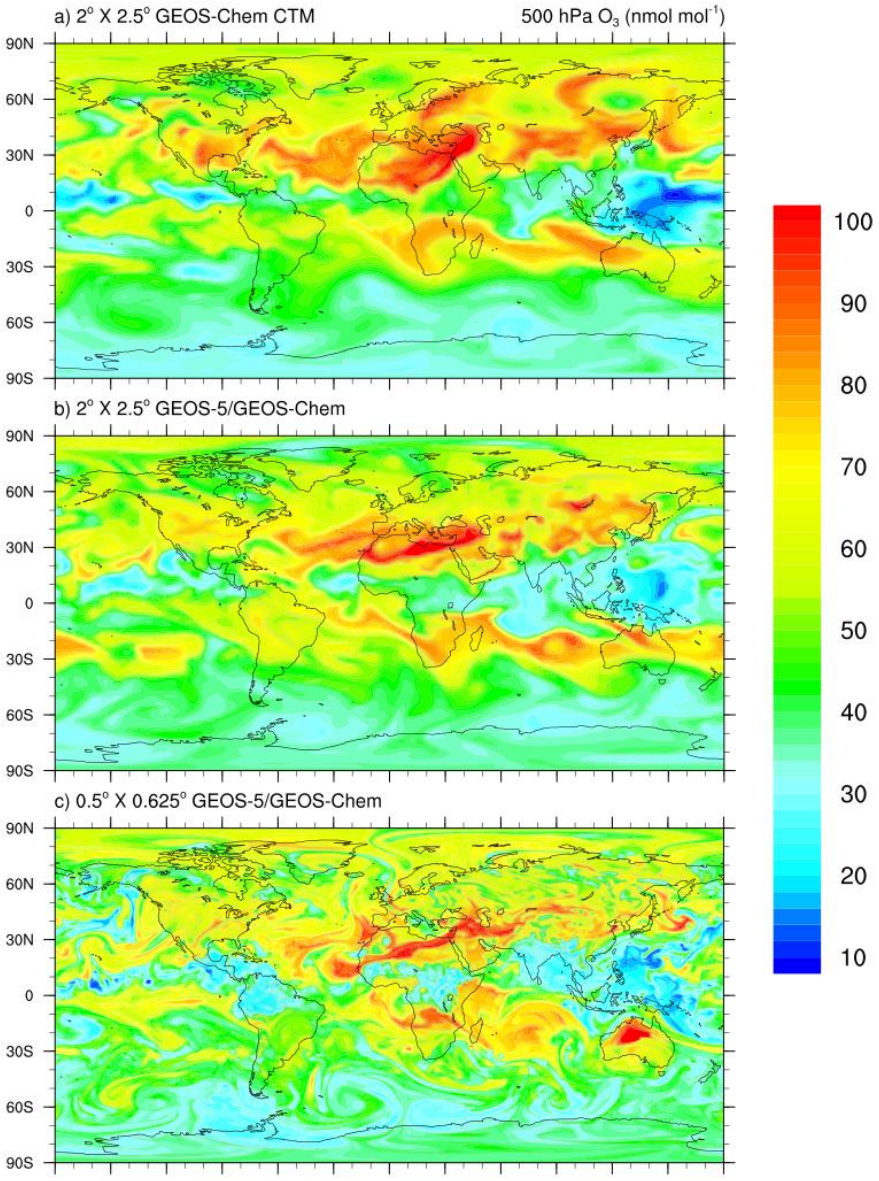

Figure 4. Comparison of instantaneous 500 hPa ozone mixing ratios (nmol mol^{-1}) at 12 UT on July 15, 2006 in the stand-alone GEOS-Chem simulation at 2^{o}x2.5^{o} horizontal resolution and the coupled GEOS-5/GEOS-Chem simulation at 2^{o}x2.5^{o} (red) and 0.5^{o}x0.625^{o} (blue) resolutions. The 0.5^{o}x0.625^{o} results are regridded to 2^{o}x2.5^{o} resolution, and each point represents a 2^{o}x2.5^{o} grid square. The reduced-major-axis regression parameters and the 1:1 line are also shown.