# MULTISCALE STRETCHED-GRIDS FOR FINE RESOLUTION SIMULATIONS OF ATMOSPHERIC COMPOSITION IN A MASSIVELY PARALLEL CHEMICAL TRANSPORT MODEL

by

Liam Bindle

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
October 2020

## Table of Contents

## List of Tables

# List of Figures

## Abstract

Modeling atmospheric composition at fine resolution globally is computationally expensive; the capability to focus on specific geographic regions using a multiscale grid is desirable. This work develops, validates, and demonstrates stretched-grids in the massively parallel version of the GEOS-Chem (GCHP) model of atmospheric composition. These multiscale grids are specified at runtime by four parameters that offer modelers nimble control of the region that is refined and the resolution of the refinement. The stretched-grid simulation is validated against global cubed-sphere simulations. The operation and flexibility of stretched-grid simulations is demonstrated with two case studies that compare simulated tropospheric $NO_2$ column densities from stretched-grid and cubed-sphere simulations to retrieved column densities from the TROPOspheric Monitoring Instrument (TROPOMI). The first case study uses a stretched-grid with a broad refinement covering the contiguous US to produce simulated columns that perform similarly to a C180 cubed-sphere simulation at less than 1/9th the computational expense. The second case study experiments with a large stretch-factor for a global stretched-grid simulation with a highly localized refinement with approximately 10 km resolution for California. The refinement improves spatial agreement with TROPOMI columns compared to a C90 cubed-sphere simulation of comparable computational demands, despite the simulation being conducted at finer resolution than parent meteorological fields. Overall this work finds that stretched-grids are a practical tool for regional fine-resolution simulations of atmospheric composition, and enables this capability for a broad scientific community. Stretched-grids are publicly available in GEOS-Chem version 13.0.0.

## List of Abbreviations and Symbols Used

$S$ Stretch factor.

$\phi'$ Output latitude after Schmidt transform.

$\phi_t$ Target latitude.

$\phi$ Input latitude to Schmidt transform.

$\theta_t$ Target longitude.

**AEIC** Aviation Emissions Inventory Code.

**AMF** Air Mass Factor.

**CEDS** Community Emissions Data System.

**CO** Carbon monoxide.

**CPU** Central Processing Unit.

**CTM** Chemical Transport Model.

**DOMINO** Dutch OMI NO2 product.

**EPA** Environmental Protection Agency.

**ESMF** Earth Systems Modeling Framework.

**FV3** GFDL Finite Volume Cubed-Sphere Dynamical Core.

**GCHP** GEOS-Chem High Performance.

**GEOS** Goddard Earth Observing System.

**GEOS-FP** GEOS Forward Processing.

**GFDL** Geophysical Fluid Dynamics Laboratory.

**GFED4** Global Fire Emissions Database version 4.

**GMAO** Global Modeling and Assimilation Office.

**HCHO** Formaldehyde.

**HEMCO** Harvard-NASA Emissions Component.

**LA** Los Angeles.

**MAD** mean absolute deviation.

**MAPL** Model Analysis and Prediction Layer.

**MB** mean bias.

**MPI** Message Passing Interface.

**NASA** National Aeronautics and Space Administration.

**NEI** National Emissions Inventory.

**NO** Nitrogen oxide.

**$NO_2$** Nitrogen dioxide.

**$NO_x$** Nitrogen oxides (NO and NO2).

**$O_3$** Ozone.

**OH** Hydroxyl radical.

**OMI** Ozone Monitoring Instrument.

**$PM_{2.5}$** Fine particulate with diameter less than 2.5 micrometers.

**QA4ECV** Quality Assurance for Essential Climate Variables.

**RMSD** root mean square deviation.

**$SO_2$** Sulfur dioxide.

**TROPOMI** TROPospheric Ozone Monitoring Instrument.

**US** United States of America.

**VOC** Volatile Organic Compounds.

## Acknowledgements

I want to express my deepest gratitude to everyone that has provided me invaluable guidance in my study. Firstly, I want to thank my supervisor Professor Randall Martin for his deeply thoughtful feedback, continual support, and his enthusiasm. His guidance has been an inspiration. I want to thank my professors in the Department of Physics and Atmospheric Science at Dalhousie University for engaging courses and their gracious help. I would also like to thank Professor Adam Bourassa and Professor Doug Degenstein from the University of Saskatchewan for introducing me to atmospheric science and for teaching me many valuable skills. I want to thank the GEOS-Chem Support Team, especially Elizabeth Lundgren from Harvard University, for their technical support in this project. I would also like to thank Thomas Clune, Benjamin Auer, and Matthew Thompson from the Global Modeling and Assimilation Office at the NASA Goddard Space Flight Center for their technical support in this project. This is a small but important subset of those that have helped me in my study; to everyone that has provided me guidance and support, I am truly grateful.

# Chapter 1

# Introduction

Atmospheric science has the critical function of understanding the two-way relationship between humans and the atmosphere. In that, an understanding of the processes and factors that control the composition of the atmosphere is imperative. Chemical transport models (CTMs) model the chemical and physical processes that govern the atmosphere's composition, and they are versatile tools for studying air quality, climate forcing, and biogeochemical cycling. Simulations of atmospheric composition are used to understand atmospheric processes and their relative importance, to interpreting satellite, aircraft, and in situ observations, and to inform evidence-based policy and design effective emission control strategies.

Global simulations of atmospheric composition are, however, computationally demanding. Chemical mechanisms in the troposphere typically involve more than 100 chemical species, emitted by anthropogenic and biogenic sources, mixing through 3-D transport on all scales. Typical global model resolutions are on the order of hundreds of kilometers and generally limited by the availability of computational resources. Massively parallel models such as GEOS-Chem in its high-performance implementation (GCHP; Eastham et al. (2018)) can run on more than 1000 cores (Zhuang et al., 2020) and is capable of 50 km resolution. The coarse resolution of global models can lead to systematic errors in applications when scales of variability finer than the model resolution are relevant, such as vertical transport and scavenging by convective updrafts (Mari et al., 2000; Li et al., 2018, 2019), nonlinear chemistry such as $NO_x$ titration (Valin et al., 2011), localized emission sources (Davis et al., 2001; Freitas et al., 2007), a priori profiles for satellite retrievals (Heckel et al., 2011; Goldberg et al., 2017; Kim et al., 2018), and simulated concentrations for population exposure estimates (Punger and West, 2013; Li et al., 2016). Multiscale grids, like nested grids, are commonly used for simulations that need to capture finer-scale modes of variability. In this work, a multiscale grid is developed, validated, and demonstrated

for the GCHP model using a technique known as grid-stretching, which is well-suited for grid-independent models like GCHP. This work enables massively parallel global multiscale simulations of atmospheric composition for a broad scientific community. These simulations are capable of spatial resolutions that are unprecedented in global models.

Several recent works set the stage for the development of stretched-grids in GCHP. Long et al. (2015) developed the grid-independent capability of GEOS-Chem. Harris et al. (2016) developed the capability for stretched-grids in the GFDL Finite-Volume Cubed-Sphere Dynamical Core (FV3), which GCHP uses to calculated advection. Eastham et al. (2018) developed the capability for GEOS-Chem to operate on cubed-sphere grids in a distributed memory framework for massive parallelization, and to use the Model Analysis and Prediction Layer (MAPL) of the NASA Global Modeling and Assimilation Office (GMAO) together with the Earth System Modeling Framework (ESMF) to couple model components. GEOS-Chem version 12.5.0 added grid-independent emissions that produce consistent emissions regardless of the model grid (The International GEOS-Chem User Community, 2019). Most recently, the Model Analysis and Prediction Layer (MAPL) version 2 (Thompson et al., 2020) of the NASA Global Modeling and Assimilation Office (GMAO) added support for stretched-grids.

Chapter 2 presents an overview of chemical transport modeling, motivates the scientific value of finer spatial resolution global CTMs, and discusses common approaches for multiscale grids. Chapter 3 presents the methodology of this work. Section 3.1 provides an overview of GCHP and its underlying gnomonic cubed-sphere grid. Section 3.2 describes grid-stretching by the Schmidt (1977) transform, following the methodology of Harris et al. (2016). Section 3.3 discusses considerations for stretched-grid simulations of atmospheric composition, and a simple procedure for choosing an appropriate stretched-factor for a given application. Section 3.5 summarizes the testing of stretched-grids in GCHP and provides a comparison of simulated oxidants and $PM_{2.5}$ concentrations from a stretched-grid simulation versus cubed-sphere simulations. Chapter 4 presents two case studies comparing stretched-grid and cubed-sphere simulations to tropospheric $NO_2$ column densities from the TROPOspheric Monitoring Instrument (TROPOMI). The first case study (Section 4.2)

considers a stretched-grid simulation with a moderate stretch-factor and a broad refinement covering the contiguous US. Simulated column densities from this computationally less-expensive stretched-grid simulation are compared to those from a fine-resolution cubed-sphere simulation. The second case study (Section 4.3) experiments with a large stretch-factor in a stretched-grid simulation with a narrow refinement covering California at 10 km resolution. Simulated concentrations from this fine-resolution stretched-grid simulation are compared to a cubed-sphere simulation with a comparable computational cost.

# Chapter 2

# Background

## 2.1 Chemical transport modeling

CTMs simulate the concentrations of atmospheric species by modeling emission, deposition, chemistry, transport, and aerosol microphysical processes. Mathematically, the rate of change of a species concentration can be written as an expression of mass conservation. For species $i$ at a location with mass density $\rho_i$, a net local source $s_i$ (sources - sinks), and wind vector $\mathbf{v}$, that is

$$\frac{\partial \rho_i}{\partial t} = \quad -\nabla \cdot (\rho_i \mathbf{v}) \quad + \quad s_i \tag{2.1}$$

This equation is known as the continuity equation. The first term on the right-hand side is commonly called the transport term and represents the ingress mass-flux of species $i$ to a volume element from advection, convection, and turbulence. The second term on the right-hand side is commonly called the local term and represents the net local source of species $i$ from emission, dry deposition, wet scavenging, chemistry, and aerosol microphysics. For model formulation, it is helpful to expand the transport term and local term into a sum of tendencies of their underlying processes. If species $i$ is a gas this gives

$$\frac{\partial \rho_i}{\partial t} = \left[\frac{\partial \rho_i}{\partial t}\right]_{adv} + \left[\frac{\partial \rho_i}{\partial t}\right]_{mix} + \left[\frac{\partial \rho_i}{\partial t}\right]_{conv} + \left[\frac{\partial \rho_i}{\partial t}\right]_{scav} + \left[\frac{\partial \rho_i}{\partial t}\right]_{chem} + \left[\frac{\partial \rho_i}{\partial t}\right]_{em} + \left[\frac{\partial \rho_i}{\partial t}\right]_{dep} \tag{2.2}$$

For an aerosol species, concentration ($\rho_i$) is parameterized by size, and additional terms for nucleation, condensation, evaporation, coagulation, and gravitational settling are required. In essence, CTMs numerically evaluate the time integral of Equation 2.2 (by operator splitting) for a coupled system of species in a Eulerian (gridded) or Lagrangian (puff) framework. This work focuses on the GEOS-Chem model, which is a Eurlerian CTM. More specifically, it focuses on GCHP which is the high-performance (massively parallel) version of GEOS-Chem.

## 2.2 The motivation for finer-resolution chemical transport models

The computational demands of modeling atmospheric composition on the global scale limit the spatial resolution of global models. Spatial resolutions of 200–400 km are typical of conventional (multithreaded) models, and spatial resolutions of 50–100 km are typical of massively parallel models. These resolutions mean global models cannot explicitly resolve modes of variability such as localized sources (e.g., ship plumes (Davis et al., 2001) and fire plumes (Freitas et al., 2007)), or convective transport and scavenging (Mari et al., 2000). In the context of nonlinear VOC-$NO_x$-$O_3$ chemistry, this a concern since grid-averaging (instantaneously diluting across an entire grid-box) a point source artificially weakens the enhancement and results in different chemistry. To treat this, global models use subgrid parameterizations of processes like convective transport (Li et al., 2018), scavenging in updrafts (Li et al., 2019), and the chemical aging of ship plumes (Vinken et al., 2011), and modelers often use multiscale grids (e.g., nested grids, described in the next section) to explicitly resolve as much fine-scale variability as possible in their region of interest.

Specifics of nonlinear atmospheric chemistry are outside the scope of this work; however, a basic understanding of the mechanisms by which systematic errors can occur by grid-averaging localized sources is useful for understanding the value of finer resolution models to studies of atmospheric composition. One example of a nonlinear response in VOC-$NO_x$-$O_3$ chemistry is $O_3$ production. The key group of species in nonlinear $O_3$ production are peroxy radicals, which are produced by the oxidation of VOCs by OH. Under very low $NO_x$ conditions, peroxy radicals can self-react, which results in no $O_3$ production. Under more moderate $NO_x$ conditions, peroxy radicals can oxidize NO to $NO_2$, resulting in $O_3$ production following the photolysis of the $NO_2$. Under high $NO_x$ conditions, the oxidation of $NO_2$ by OH sequesters the production of peroxy radicals (otherwise produced by VOCs oxidation by OH), resulting in the less efficient $O_3$ production than at lower $NO_x$ concentrations. The implications of grid resolution on $O_3$ production is well studied (Sillman et al., 1990; Jang et al., 1995; Gillani and Pleim, 1996; Kumar and Russell, 1996; Liang and Jacobson, 2000; Esler et al., 2004; Cohan et al., 2006), and estimates of the tropospheric $O_3$ burden are sensitive to model resolution at the scales typical of global models (Wild and Prather, 2006). A second example of a nonlinear response in VOC-$NO_x$-$O_3$ chemistry

is the lifetime of $NO_x$. The primary daytime sink of $NO_x$ is the oxidation of $NO_2$ by OH. And, OH concentration is strongly influenced by $NO_x$ (byway of $O_3$ production). This circular dependency in $NO_x$ establishes the nonlinear response in $NO_x$ lifetimes. Under high $NO_x$ conditions, $NO_2$ can titrate OH resulting in a long-lived plume of $NO_x$. Valin et al. (2011) found that the grid-averaging of such plumes underestimates $NO_x$ lifetime and can affect interpretations of satellite observations.

An important role of CTMs in atmospheric science is generating a priori profiles for satellite retrievals such as $NO_2$, HCHO, and $SO_2$. Accurate retrievals in polluted regions—which are of utmost interest in understanding the two-way relationship between humans and the atmosphere—need fine resolution a prioris to capture the variability of absorber profiles. Trace gas retrievals from solar backscatter measurements involve an assumption of the absorber's relative vertical distribution (shape factor) in calculating an air mass factor (AMF) to convert a slant column density to a vertical column density (Palmer et al., 2001). Shape factor assumptions are a primary contributor to AMF uncertainties which dominate $NO_2$ and HCHO retrieval uncertainties (Boersma et al., 2004; De Smedt et al., 2008; Lorente et al., 2017; Boersma et al., 2018; De Smedt et al., 2018). One reason for this is modern instruments have much finer resolution than the global CTMs that generate shape factor assumptions. For example, the NASA Standard $NO_2$ Product version 3 uses observations from the Ozone Monitoring Instrument (OMI), which has $\sim$20 km resolution (at nadir), while the product uses 100 km resolution shape factors (Krotkov et al., 2017). The Quality Assurance for Essential Climate Variables (QA4ECV) $NO_2$ and HCHO products use observations from the TROPOspheric Monitoring Instrument (TROPOMI), which has 5 km resolution, while the product uses 90 km resolution shape factors (Boersma et al., 2018). Several studies have investigated the impact of coarse resolution shape factors on $NO_2$, HCHO, and $SO_2$ retrievals in polluted regions by recalculating the AMFs with shape factors from a fine resolution regional CTM (Heckel et al., 2011; McLinden et al., 2014; Goldberg et al., 2017; Kim et al., 2018). The general result of these studies show a tendency for retrievals to be underestimated in polluted areas and overestimated in nearby cleaner areas. This tendency is because grid-averaging localized emissions results in shape factors with a weaker vertical gradient than reality in polluted areas and a stronger vertical gradient than reality in nearby cleaner areas.

These artificially weakened and strengthened vertical gradients in the shape factors cause the AMFs to be overestimated and underestimated (Palmer et al., 2001), resulting in overestimated retrievals in polluted areas and underestimated retrievals in nearby cleaner areas.

Health impact studies commonly use CTMs to simulate pollutant concentrations for population exposure estimates, since the simulated concentrations provided total spatiotemporal coverage and allow for the evaluation of various emission scenarios. For accurate estimates of population exposure, fine-resolution simulations are necessary to resolve the pollutant gradients found in densely populated areas. Simulated concentrations of primary aerosols are particularly sensitive to model resolution since their nature (direct emission, minimal influence by chemical processes) lends to fine-scale gradients that correlate with population gradients (EPA, 2007). As a result, estimates of mortality attributable to $PM_{2.5}$ can be sensitive to model resolution. Punger and West (2013) found that estimated mortality attributable to $PM_{2.5}$ exposure (in the US) based on coarse (>100 km) $PM_{2.5}$ concentrations underestimated all-cause deaths by more than 20 % compared to an estimate based on $PM_{2.5}$ concentrations at 12 km resolution. Similarly, Li et al. (2016) found that estimated mortality attributable to $PM_{2.5}$ exposure (in the US) was 8 % lower when the exposure estimates were based on simulated concentrations at 200 km resolution compared to 50 km resolution. In both studies, the lower estimates of mortality, in the estimates based on coarse-resolution simulations, were driven by differences in primary aerosols concentrations.

## 2.3 Common approaches to multiscale grids

The resolution of global CTMs is typically computationally-limited, so modelers often use multiscale grids to resolve as much fine-scale variability as possible in their region of interest. The simplest and perhaps most common multiscale grid are one-way nested grids. One-way nested simulations use a coarse global simulation to generate initial conditions and boundary conditions for finer-resolution regional simulations (Wang et al., 2004). This is similar to regional CTMs, which typically rely on initial conditions and boundary conditions from a global CTM. A drawback of one-way nesting is that the nested domains cannot feedback on the global domain. A two-way

nested simulation dynamically couples its nested domains and global domain (Krol et al., 2005). Two-way nesting can simulate intercontinental transport and improve simulations of moderately long-lived species like CO and $O_3$ (Yan et al., 2014, 2016). The main drawback of two-way nesting is the technical complexity of coupling the domains. The conventional (multithreaded) version of GEOS-Chem supports one-way nested and two-way nested simulations at 25 km and 50 km resolution. However, the high performance (massively parallel) version of GEOS-Chem (GCHP) supports neither; and, the implementation of both is non-trivial.

An alternative type of multiscale grid, which is well-established in regional climate modeling, are stretched-grids (Fox-Rabinovitz et al., 2006, 2008). Stretched-grids are deformed grids such that they have increased grid density in the region of interest, and they have grid-boxes that transition smoothly between the refinement and coarser global resolutions. Stretched-grids have the advantage of being inherently two-way coupled, and they are easy to implement in grid-independent models since the stretching operation does not alter the grid topology.

### 2.4 Massively parallel computing

A massively parallel simulation is a simulation that executes on a large number of computer processors in parallel, often using hundreds or thousands of computer cores. Modern high-performance Central Processing Units (CPUs) have up to approximately 48 cores, and therefore, massively parallel simulations need to distribute their computation across multiple computers. Because of this, inter-process communication is needed to coordinate and exchange data in the simulation. The Message Passing Interface (MPI) is a commonly used protocol for inter-process communication in massively parallel computing.

The fundamental difference between massively parallel simulations and conventional simulations is that conventional simulations are limited to a single computer's resources (its cores and memory). A massively parallel simulation leverages the computational resources of many computers. It should be noted that massively parallel simulations have code that is specially designed for coordinated execution on many processes in parallel.

# Chapter 3

# Methods

## 3.1 GEOS-Chem in its high-performance implementation (GCHP)

This work uses GEOS-Chem version 13.0.0 in its high-performance implementation (GCHP; Eastham et al. (2018)). GCHP is an offline global chemical transport model designed for massively-parallel computing systems. There are three main components in GCHP: (1) the GEOS-Chem CTM module, (2) the GFDL Finite-Volume Cubed-Sphere Dynamical Core (FV3; Putman and Lin (2007)), and (3) a component that imports archived data to drive GCHP.

GCHP uses MAPL (Suarez et al., 2007) and the Earth Systems Modeling Framework (ESMF; Hill et al. (2004)). MAPL and ESMF are abstraction layers that provide a high-level programming interface for Earth systems modeling. This high-level programming interface facilitates coupling model components (e.g., coupling the GEOS-Chem CTM module and FV3) and use of the massively-parallel computing infrastructure. Internally, MAPL and ESMF make MPI calls that perform inter-process communication, but their high-level programming interface means that model components do not have to concern themselves with low-level MPI calls or platform-specific details.

The GEOS-Chem CTM module is a grid-independent module that models gas- and aerosol-phase species concentrations in vertical columns. Being grid-independent means the module makes no assumptions of the horizontal grid, such as orthogonality or uniformity. This grid-independence means that the GEOS-Chem CTM module can be used in many different Earth systems models, regardless of their horizontal grid. The GEOS-Chem CTM module is based on the original GEOS-Chem CTM described in Bey et al. (2001), and simulates tropospheric-stratospheric chemistry by solving 3-D chemical continuity equations. The GEOS-Chem CTM module uses columnar operators (Long et al., 2015) for columnar or local calculations such as convection and chemical kinetics. Emissions are aggregated, parameterization, and computed

with the Harvard-NASA Emissions Component (HEMCO) described in (Keller et al., 2014).

GCHP couples the GEOS-Chem CTM module to the FV3 dynamical core, which calculates advection. The FV3 dynamical core uses a gnomonic cubed-sphere grid, which has several advantages over a conventional latitude-longitude grid, stemming from its more uniform grid-boxes and no polar singularities. These properties benefit the parallelization and numerical stability of transport (Putman and Lin, 2007; Eastham et al., 2018). GCHP inherits this gnomonic cubed-sphere grid. The primary driver of a GCHP simulation's computational demands is the horizontal grid resolution. A cubed-sphere grid's resolution is described with CN notation, where N is the number of equidistant segments along face edges that form the meshes.

GCHP uses offline meteorological data is from the Goddard Earth Observing System (GEOS) data assimilation system. GEOS has a 72-layer hybrid sigma-pressure grid extending from the surface to 1 Pa. GCHP inherits this vertical grid.

### 3.2 Grid-stretching

The Schmidt (1977) transform can be applied to any grid, and effectively stretches the grid to increase its density in a region. Grid-stretching in GCHP follows the methodology of Harris et al. (2016) for stretching a gnomonic cubed-sphere grid with the Schmidt (1977) transform. The procedure is two-part and starts with a standard gnomonic cubed-sphere grid. First, the refinement (the region with increased grid density) is generated over the South Pole by remapping grid-box coordinate latitudes with a modified form of the Schmidt transform

$$\phi' = \arcsin \frac{D + \sin\phi}{1 + D\sin\phi} \quad \text{where} \quad D = \frac{1 - S^2}{1 + S^2} \tag{3.1}$$

where $\phi$ is the input latitude, $\phi'$ is the output latitude, and $S$ is a parameter controlling the stretching operation's strength called the stretch-factor. This operation effectively attracts grid-box coordinates towards the South Pole along meridians (exploiting the grid's gnomonic mapping), thereby increasing the grid density. The second step is repositioning the refinement to the region of interest by rotating the grid, re-centering the refinement on a user-specified target point $\phi_t$ °N, $\theta_t$ °E. In GCHP, the rotations are

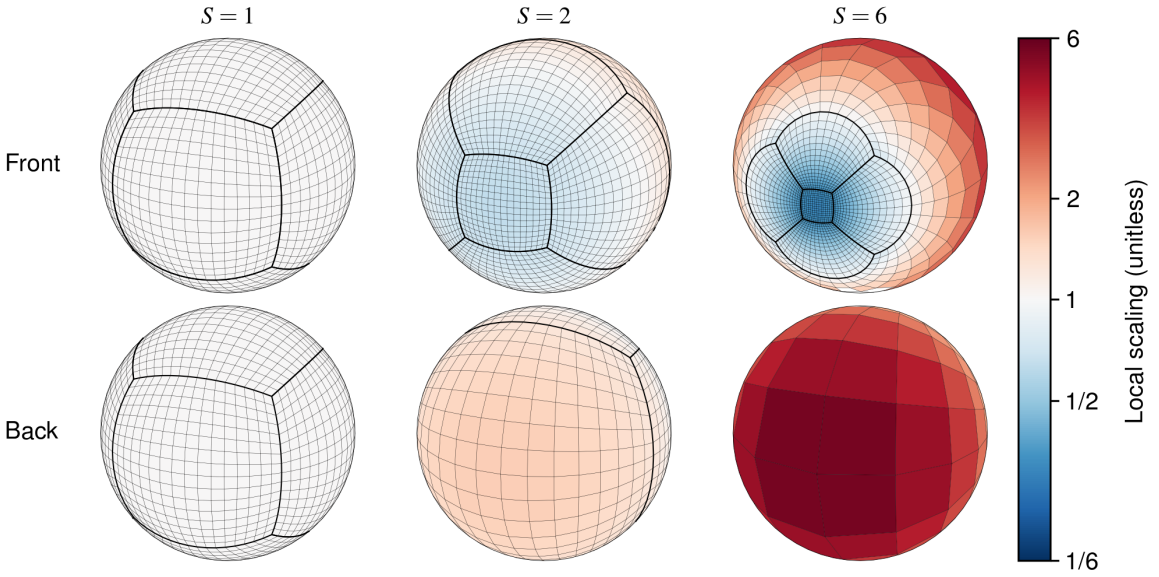


Figure 3.1: Stretched-grids with different stretch-factors ($S$). Local scaling refers to the relative change in a grid-box's edge length from the stretching operation, Eq. (3.1).

by $\phi_t + 90$ and $\theta_t + 180$ about 90 °E and 90 °N respectively. GCHP's standard cubed-sphere grid has a $-10$ °E offset, and is identical to a stretched-grid with parameters $S = 1$, $\phi_t = -90$ °N, $\theta_t = 170$ °E.

Figure 3.1 shows three stretched-grids with different stretch-factors ($S$) to illustrate how the stretching operation scales grid-box sizes (specifically edge length, which is inversely proportional to resolution). The effective resolution of the grid in the target-face (the face containing the target point) increases by approximately a factor of $S$. Two features of stretched-grids dissimilar to nested-grids are: (1) a smooth transition to the refined region, and (2) grid-boxes far from the target expand to account for the shrunken boxes (which is necessary to maintain the topology). The stretching operation is independent of the cubed-sphere grid, and therefore one can quantify the local scaling of grid-box edge lengths as a function of $S$ and distance from the target (i.e., the gradual changing of a stretched-grid's resolution relative to the grid's resolution prior to stretching). Figure 3.2 quantifies this gradual change in local scaling and the lateral distances to the target-face edges and the opposite face edges. A crucial feature to note in Figures 3.1 and 3.2 is that the extent of a stretched-grid's refinement decreases as the stretch-factor ($S$) increases. A stretched-grid with a moderate stretch-factor (e.g., 1.5–3.0) has a broad refinement (i.e., continental-scale

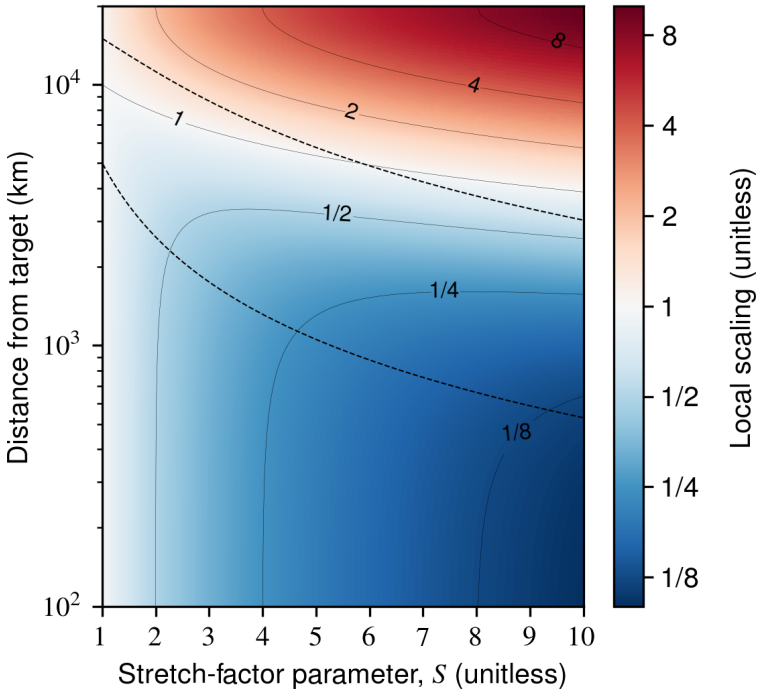


Figure 3.2: Relative change in grid-box edge length from Eq. (3.1) as a function of $S$ and distance from the target point. Dashed lines are the lateral distances to grid face interfaces (i.e., distance from the refinement center to the center of face edges). From the target point ($\phi_t$ °N, $\theta_t$ °E), the lower dashed line is the distance to (the center of) the target-face edges, and the upper dashed line is the distance to (the center of) the opposite face edges.

refinement), whereas a stretched-grid with a large stretch-factor (e.g., >5.0) has a localized refinement (i.e., regional-scale refinement).

Four scalar parameters fully describe a stretched-grid: the initial cubed-sphere resolution, the stretch-factor ($S$), the target latitude ($\phi_t$), and the target longitude ($\theta_t$). The combination of cubed-sphere resolution and stretch-factor control the resolution at the target and the minimum resolution at the target's antipode. The target latitude and longitude control the center of the refinement. Figure 3.2 shows that the relative change in grid-box edge length is approximately constant through the target-face (i.e., for a given stretch-factor, the relative change in edge length is nearly constant from the target (0 m) to the edge of the target-face (lower dashed-line)). Therefore, it is reasonable to approximate the effective resolution of the target-face as the cubed-sphere resolution times $S$. The concise parameters that define stretched-grids are an asset, as they are conceptually simple, precise, and provide the user

nimble control over the grid.

### 3.3 Choosing an appropriate stretched-factor

The stretch-factor is a seemingly simple parameter, but an appropriate value is application-specific and moderately variable. For computational efficiency, it is desirable to use a large stretch-factor, but larger stretch-factors also narrow the refinement and coarsen the resolution at the target's antipode (see Figure 3.2, resolution at the target antipode is $S^2$ times coarser than the target resolution, and $S$ times coarser than the original cubed-sphere grid resolution). For GEOS-Chem, target antipode resolutions significantly coarser than C24 warrant caution due to limited prior model evaluation at such coarse resolution. This work proposes a simple procedure to determine the maximum suitable stretch-factor for a given GCHP simulation, based on taking the lesser of two constrained upper limits for $S$:

1. **An upper limit for $S$ constrained by the required extent of the refinement.** One can determine this upper limit using Figure 3.2. Approximating the refined region as the target-face (see Section 3.2), the upper limit for $S$ is where the distance to the target-face edges (the lower dashed line) equals the required radius of refinement. For example, for a refinement with a 2000 km radius, the upper limit for $S$ is slightly less than 3.

2. **An upper limit for $S$ constrained by the maximum and minimum resolution of the stretched-grid.** The target resolution is $S^2$ times finer than the antipodal resolution; therefore, this upper limit is the square root of their ratios. For example, for a maximum and minimum resolution of C360 and C24, the upper limit for $S$ is $\sqrt{360/24} = 3.9$.

Then, take $S_{\text{max}}$ as the lesser of these two upper limits. A suitable stretch-factor is $S \lessapprox S_{\text{max}}$. GCHP requires that the cubed-sphere resolution be even (e.g., C88, C90, C92). Therefore, to follow this procedure strictly, one would take the underlying cubed-sphere resolution as the target resolution divided by $S_{\text{max}}$ rounded up to the nearest even integer, and $S$ as target resolution divided by the underlying cubed-sphere resolution (e.g., for a target resolution of C360 and $S_{\text{max}} = 3.9$, one would choose a C94 grid with $S = 3.83$).

### 3.4 Model configuration

All simulations in this work use GCHP version 13.0.0-alpha.3 with a common configuration excluding the model grid. Table 3.1 describes emissions for the US. Monthly anthropogenic emissions are based on the National Emission Inventory (NEI) for 2011 with updated annual scaling factors to account for changes in annual totals since 2011. In 2018, the scaling factor for NO was 0.64. Biomass burning emissions for 2018 are from GFED4. Aircraft emissions are from AEIC (Stettler et al., 2011). Offline grid-independent inventories for $NO_x$ emissions from soils and lightning are used (The International GEOS-Chem User Community, 2019). The offline soil $NO_x$ emissions are based on the scheme described in Hudman et al. (2012), and offline lightning $NO_x$ emissions are based on the scheme described in Murray et al. (2012). Meteorological inputs from the GEOS-FP archive, which are archived on a 0.25 °N × 0.3125 °E (25 km) grid. A 10-minute timestep is used for chemsitry and a 5-minute timestep for transport. Each simulation's initial conditions were regridded to the simulation grid, and each simulation included a 1-month spin up unless otherwiese specified.

Simulations are named according to their grid. Cubed-sphere simulations are named according to their resolution with the suffix "-global", referring to their resolution being quasi-uniform globally. For example, a C180 cubed-sphere simulation is named "C180-global". Stretched-grid simulations are named according to their refinement effective resolution, with a suffix denoting the region that is refined. For example, a stretched-grid simulation with an effective resolution of C180 in the contiguous US is named "C180e-US". "C180e" refers to the stretched-grid refinement being comparable to the resolution of a C180 cubed-sphere grid.

### 3.5 Stretched-grid development and validation

The development and testing of stretched-grids in GCHP involved multiple stages. A prototype stretched-grid simulation without chemistry and with a simplified set of emissions was developed and used to confirm stretched-grid functionality in all model components and the gridding of initial conditions. Subsequently, full chemistry and the full set of emissions were enabled, and benchmarking code was developed to

support stretched-grids, enabling the evaluation of the diagnostic output of stretched-grid simulations versus cubed-sphere simulations.

Next, a comparison of simulated oxidants and $PM_{2.5}$ from a fully developed stretched-grid simulation and cubed-sphere simulations is presented. A challenge in such comparisons is the change in grid configuration versus the reference simulation. Thus a reference cubed-sphere simulation (C96-global) is compared to a stretched-grid simulation (C96e-NA) and a second cubed-sphere simulation (C94-global) with a different gridding. The C96 and C94 cubed-sphere grids have grid-boxes with similar sizes (2.1 % difference in resolution, 4.1 % difference in areas), and the shape and orientation of their grid-boxes vary geographically in the same way. The main difference between the C96 and C94 cubed-spheres is that their grid-boxes have minimal overlap (beat frequency of 2), only overlapping near the center lines of each face. C96e-NA had an underlying C48 grid with a stretch-factor of 2.4 such that its mean grid-box area was the same as the C96 cubed-sphere in the contiguous US. C96-global, C96e-NA, and C94-global use the configuration described in Section 3.4 except they used a 3-month spin up rather than a 1-month spin up to accommodate relaxation time for CO and $O_3$. Figure 3.3 shows scatter plots of simulated mean concentrations of major oxidants and $PM_{2.5}$ from the three simulations. The comparisons of simulated means from C96-global and C94-global (right column) is included for perspective. Figure 3.3 shows that simulated means from C96e-NA are consistent with C96-global and that their differences are comparable to the difference between the two similar global cubed-sphere simulations.

Table 3.2 shows the mean bias (MB), mean absolute deviation (MAD), and root-mean-square deviation (RMSD) for C96e-NA and C94-global compared to C96-global. MBs, MADs, and RMSDs are small for both simulations. Slightly larger differences for C96e-NA are expected due to the different variability of its grid-box, and the effects of nonlinear processes propagating from the coarse resolution region to the refinement region. Nevertheless, the scores are small and or comparable to C94-global's scores.

Table 3.1: Emission sources in the contiguous US in the common model configuration.

| Source type | Inventory | Resolution | Notes |
|---|---|---|---|
| Anthropogenic | NEI-2011 | 0.1° N × 0.1° E | Monthly fluxes with updated annual totals for 2018[1] |
| Lightning | — | 0.25° N × 0.3125° E | Based on scheme described in Murray et al. (2012) |
| Soil | — | 0.25° N × 0.3125° E | Based on scheme described in Hudman et al. (2012) |
| Biomass burning | GFED4 | 0.25° N × 0.25° E | Monthly fluxes |
| Aircraft | AEIC | 1.0° N × 1.0° E | Monthly fluxes, described in Stettler et al. (2011) |
| Shipping | CEDS | 0.5° N × 0.5° E | Monthly fluxes, described in McDuffie et al. (2020) |

[1] Based on criteria pollutants National Tier 1 from https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data (Accessed May 8, 2020)

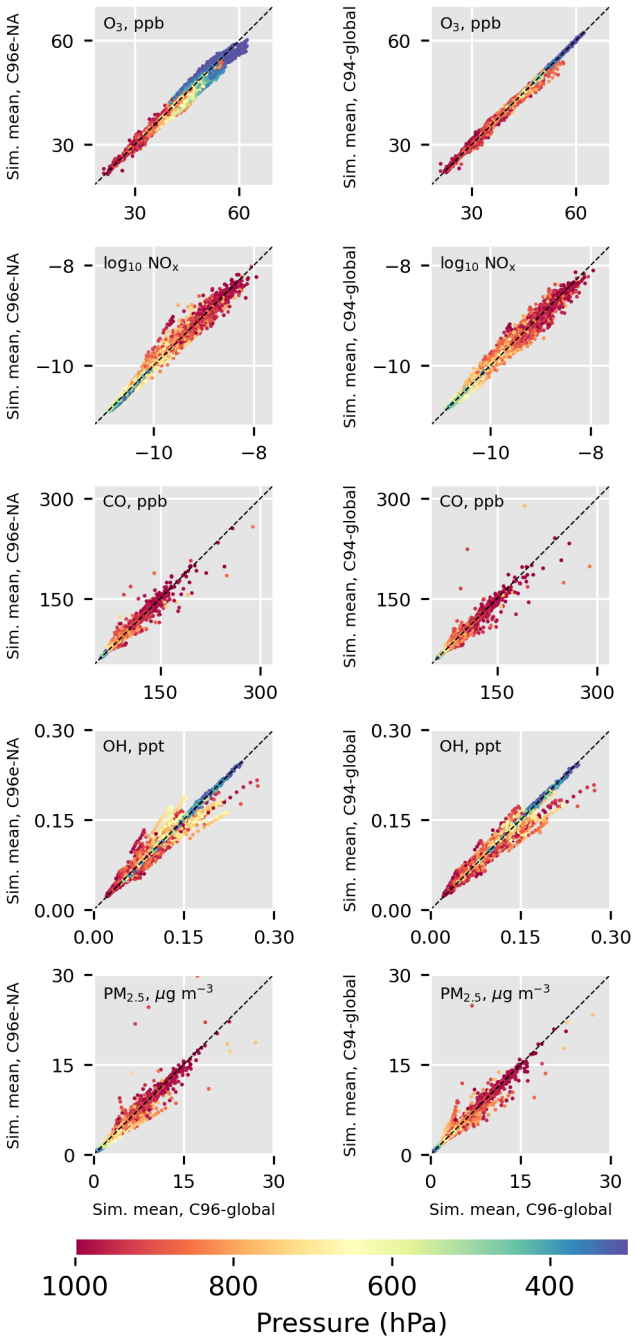


Figure 3.3: Scatter plots of mean concentrations of major oxidants and $PM_{2.5}$ from C96e-NA and C94-global verses C96-global for September 2018 for the contiguous US. The comparison of simulated concentration from C94-global and C96-global (right column) assesses the effects of different gridding that are independent of grid-stretching. Table 3.2 summarizes the data shown here.

Table 3.2: Summary of statistics for data shown in Figure 3.3.

| | | C96-global | | C96e-NA | | | | | C94-global | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ | MB | MAD | RMSD | $\mu$ | $\sigma$ | MB | MAD | RMSD |
| $O_3$ | (ppb) | 43.7 | 5.9 | 43.2 | 5.8 | -0.5 | 0.8 | 1.1 | 43.7 | 5.9 | -0.0 | 0.2 | 0.3 |
| $NO_x$ | (ppt) | 210 | 385 | 210 | 364 | -0 | 21 | 102 | 209 | 363 | -1 | 19 | 93 |
| CO | (ppb) | 81.7 | 16.7 | 82.8 | 16.7 | 1.1 | 1.4 | 2.3 | 81.6 | 16.5 | -0.1 | 0.5 | 1.9 |
| OH | (ppq) | 99 | 32 | 98 | 31 | -1 | 3 | 5 | 99 | 31 | 0 | 2 | 4 |
| $PM_{2.5}$ | ($\mu$g m$^{-3}$) | 3.83 | 2.68 | 3.64 | 2.61 | -0.19 | 0.21 | 0.42 | 3.83 | 2.66 | -0.00 | 0.08 | 0.40 |

# Chapter 4

# Stretched-Grid Case-Studies

This chapter focuses on two case studies to further demonstrate the operation of stretched-grid simulations and the flexibility of stretched-grid refinements. The first is a typical case that considers a stretched-grid simulation with a moderate stretch-factor for a broad refinement covering the contiguous US at 50 km (C180e) resolution. The second is an exploratory case that experiments with a large stretch-factor for a localized refinement covering California at 10 km (C900e) resolution. Spatial uncertainties in emissions are outside the scope of this work. The resolution of the second case study (10 km) is more than twice as fine as GEOS-Chem's typical finest resolution. This work focuses on $NO_2$ as a well measured species that is sensitive to resolution. Simulated tropospheric $NO_2$ column densities from stretched-grid and cubed-sphere simulations are compared to retrieved column densities from the TROPOMI satellite instrument.

## 4.1 Tropospheric $NO_2$ columns from TROPOMI

Simulated tropospheric $NO_2$ column densities are compared to retrieved column densities from TROPOMI. TROPOMI is an instrument onboard the Sentinel-5P satellite launched in 2018 into a sun-synchronous orbit with a local overpass time of 13:30. TROPOMI includes ultraviolet and visible band spectrometers, and the retrieved $NO_2$ column densities at $3.5\times5.5$ km$^2$ resolution use the DOMINO retrieval algorithm (Boersma et al., 2011, 2018). Observations with retrieved cloud fractions less than 10 % are included. As discussed in Section 2.2, retrieved $NO_2$ column densities are sensitive to the a priori profiles used to calculate the AMFs (Boersma et al., 2018; Lorente et al., 2017). To avoid spurious differences from the retrieval's a priori profiles when comparing simulated and retrieved $NO_2$ column densities, AMFs were recalculated with shape factors based on the mean simulated profiles for July 2018 following the approach described in Cooper et al. (2020) and Palmer et al. (2001).

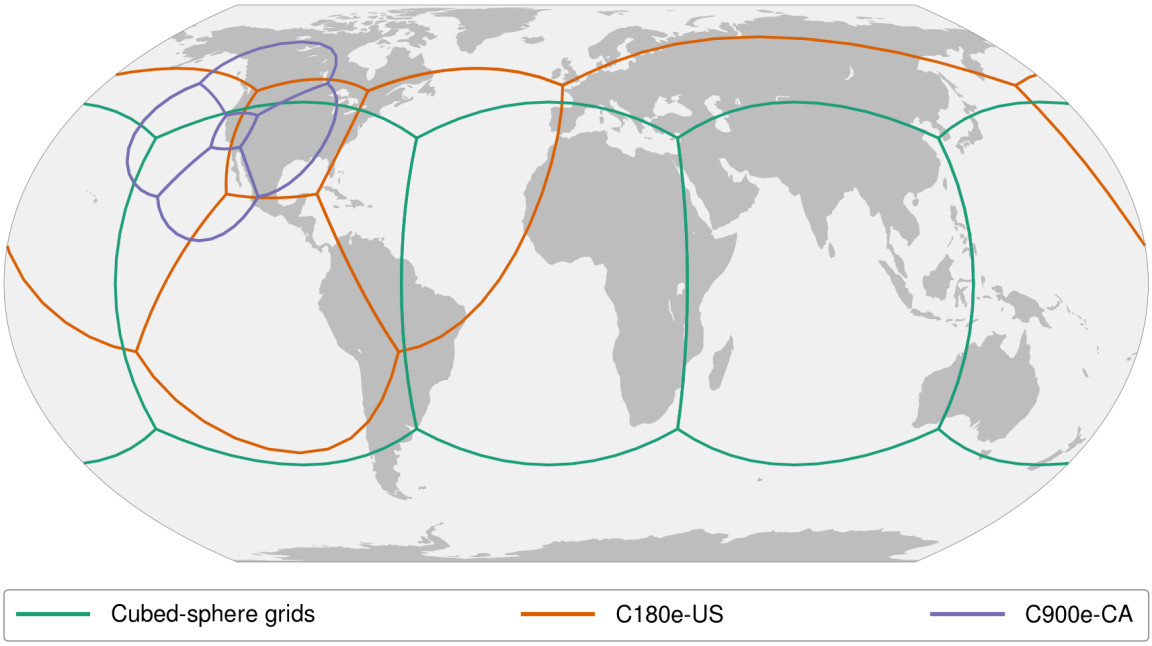


Figure 4.1: Grid faces of cubed-sphere grids and the stretched-grids used in case studies. Faces of C180-global, C180e-US, and C900e-CA simulation grids. C180e-US is a C60 stretched-grid with $S = 3$ and target 36 °N, 261 °E. C900e-CA is a C90 stretched-grid with $S = 10$ and target 37.2 °N, 240.5 °E.

## 4.2 Case study: a stretched-grid simulation with a moderate stretch-factor

Stretched-grids with moderate stretch-factors have broad refinements. This case study focuses on the contigious US and uses a C180 cubed-sphere simulation (C180-global) as a reference. For the stretched-grid simulation, a stretch-factor of 3 with an underlying C60 grid and a target point of 36 °N, 261 °E was used. This stretched-grid simulation, C180e-US, has 9 times fewer grid-boxes than C180-global, and its coarsest resolution in the southern Indian Ocean (near −36 °N, 81 °E) is approximately C20e. In addition to C180-global and C180e-US, a C60 cubed-sphere simulation (C60-global) was conducted for comparison with C180e-US. Figure 4.1 shows the grid faces of C180e-US in orange and cubed-sphere grids (C180-global and C60-global) in green.

Figure 4.2 compares tropospheric $NO_2$ column densities from TROPOMI, C180-global, C180e-US, and C60-global for the US in July 2018. The TROPOMI columns show high-concentrations over major US cities and low $NO_2$ concentrations in rural and remote areas. Simulated $NO_2$ column densities from C180-global and C180e-US

Table 4.1: Timing tests comparing the computational expense of C180-global, C180e-US, and C60-global simulations

| | C180-global[1] | C180e-US[2] | C60-global[2] |
|---|---|---|---|
| CPU time (cyr/yr)[3] | 29.3 | 1.52 | 1.33 |
| Chemistry | 24.9 | 0.99 | 1.02 |
| Dynamics | 3.5 | 0.24 | 0.24 |
| Data Input | 0.6 | 0.25 | 0.05 |
| Other | 0.3 | 0.03 | 0.02 |

[1] Two-week timing test run on 384 cores
[2] Two-week timing test run on 48 cores
[3] Core years for one simulation year (based on two-week timing tests)

are consistent throughout the domain and generally capture the plumes over cities and the low concentrations in rural and remote areas. C60-global failed to resolve local enhancements over major US cities, and comparison with C180e-US highlights the effectiveness of grid-stretching.

The computational demands of C180e-US were significantly less than C180-global and comparable to C60-global. Table 4.1 gives timing test results for C180-global, C180e-US, and C60-global. Total CPU time was reduced by nearly 20× in C180e-US relative to C180-global, resulting from fewer grid-boxes and reduced overhead. The relatively large input cost for C180e-US is suspected to be caused by a load imbalance in the online regridding of inputs. Nonetheless, the total CPU time for C180e-US was only 15 % greater than C60-global.

### 4.3 Case study: a stretched-grid simulation with a large stretch-factor

Stretched-grids with large stretch-factors have localized refinements. This case study experiments with a large stretch-factor to create a localized refinement covering California at 10 km resolution. The selected stretched-grid parameters were a stretch-factor of 10, an underlying C90 grid, and a target of 37.2 °N, 240.5 °E. This stretched-grid simulation, C900e-CA, demonstrates that computationally, stretched-grid simulations are readily capable of very fine resolutions. At the time of writing, this is the finest resolution simulation ever conducted with GEOS-Chem and comparable

to the resolution of regional models. C900e-CA is experimental. The finest resolution meteorological inputs for GEOS-Chem are from the GEOS-FP (operational product) archive on a 0.25 °N × 0.3125 °E grid (∼25 km), and therefore, C900e-CA conservatively downscales meteorological inputs (naively). Work is underway in the GCHP–GMAO community to prepare finer resolution meteorological inputs for GEOS-Chem. In the meantime, this experimental simulation leverages the resolution of anthropogenic NO emissions on a 0.1 °N × 0.1 °E (∼9 km) grid.

C900e-CA's large stretch-factor causes some grid-boxes to expand considerably. Figure 3.2 is useful for estimating how C900e-CA resolution changes with distance from its target point in central California. For example, New York is approximately 4000 km from the target. Figure 3.2 shows that for a stretch-factor of 10 and 4000 km from the target, grid-box edge lengths are scaled by ∼1. Therefore, C900e-CA's resolution in New York is approximately C90e (∼100 km). With the same procedure, one finds that at the distance of Beijing or Rome (both are nearly 10000 km from the target), C900e-CA's resolution is approximately C20e (∼460 km). C900e-CA's coarsest resolution is C9e (∼1000 km) at −37.2 °N, 60.5 °E in the southern Indian Ocean (about 2000 km southeast of Madagascar).

Figure 4.3 compares tropospheric $NO_2$ column densities from TROPOMI, C900e-CA, and a C90 cubed-sphere simulation (C90-global) for California in July 2018. The TROPOMI columns show significant fine-scale variability throughout California, high-concentrations over Los Angeles and in the San Francisco Bay Area, and smaller high-concentration features over Sacramento, Fresno, and Bakersfield. C900e-CA resolves many of the fine-scale spatial features seen in the TROPOMI columns, including the small high-concentration features over Sacramento, Fresno, and Bakersfield, despite its extreme stretch-factor and its resolution being finer than the parent meteorology. The coarse resolution of C90-global fails to resolve most spatial features seen in the TROPOMI and the C900e-CA columns and significantly underestimates high concentrations except in Los Angeles (LA). A prominent feature in the TROPOMI columns are the strong gradients along the LA Basin perimeter. There is a particularly strong gradient at the basin's northern edge, where a narrow band of mountains (the San Gabriel Mountains, ∼30 km across) separates the LA Basin and the Mojave Desert. Neither simulation captures this gradient, and LA's plume spuriously spreads

into the Mojave Desert. The affected area/grid-boxes are hatched in Figure 4.3. The basin's prevailing southwesterly wind in July, combined with the similar scale of the parent meteorology and the San Gabriel Mountains' width, suggests finer resolution meteorology could better resolve transport and orthographic boundary layer processes in the area.

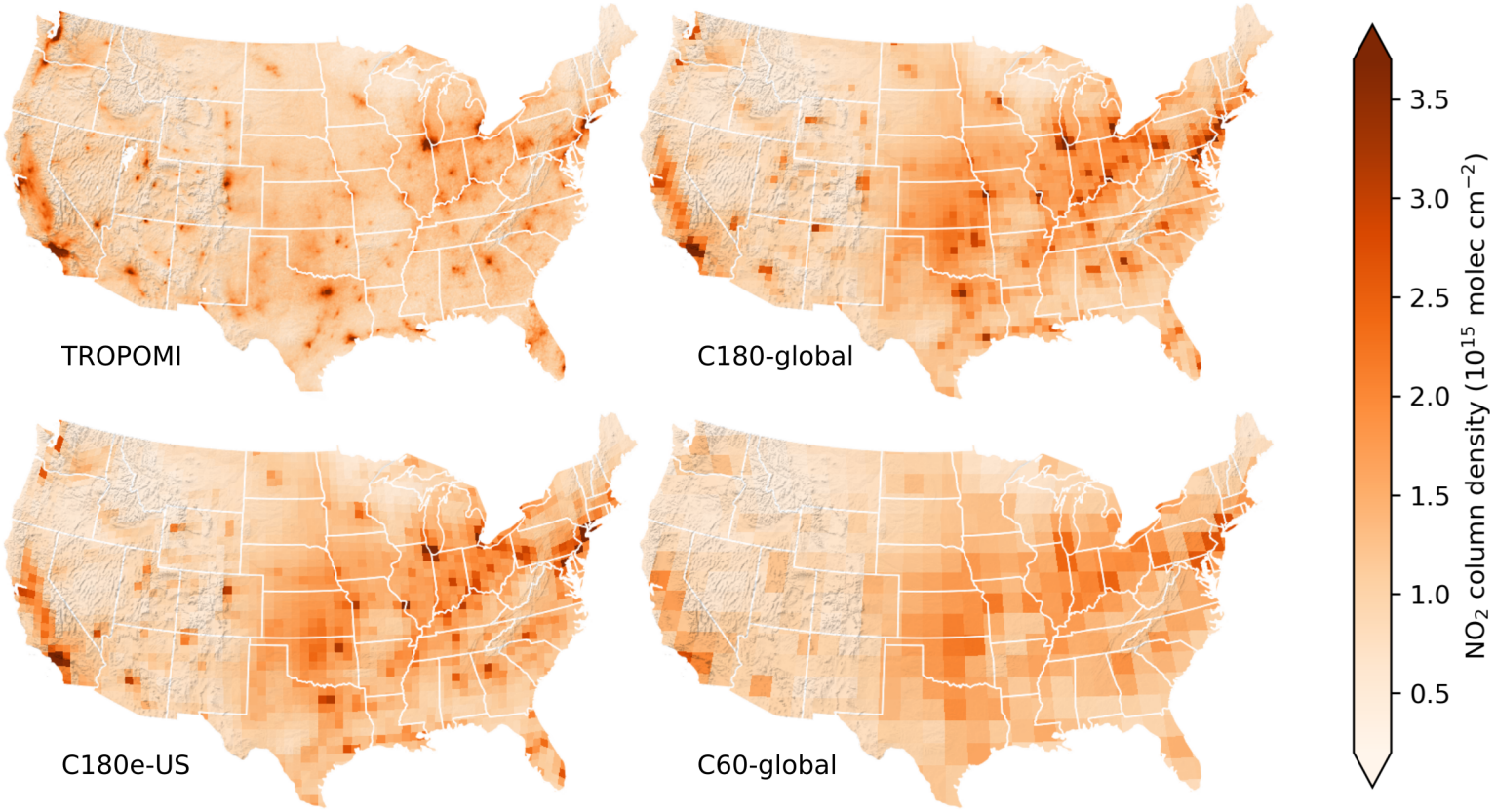


Figure 4.2: Mean tropospheric $NO_2$ column densities from C180-global and C180e-US compared to TROPOMI observations for July 2018. Simulated means only include points where TROPOMI observations were available. TROPOMI columns shown here use mean recalculated AMFs from C180-global and C180e-US.

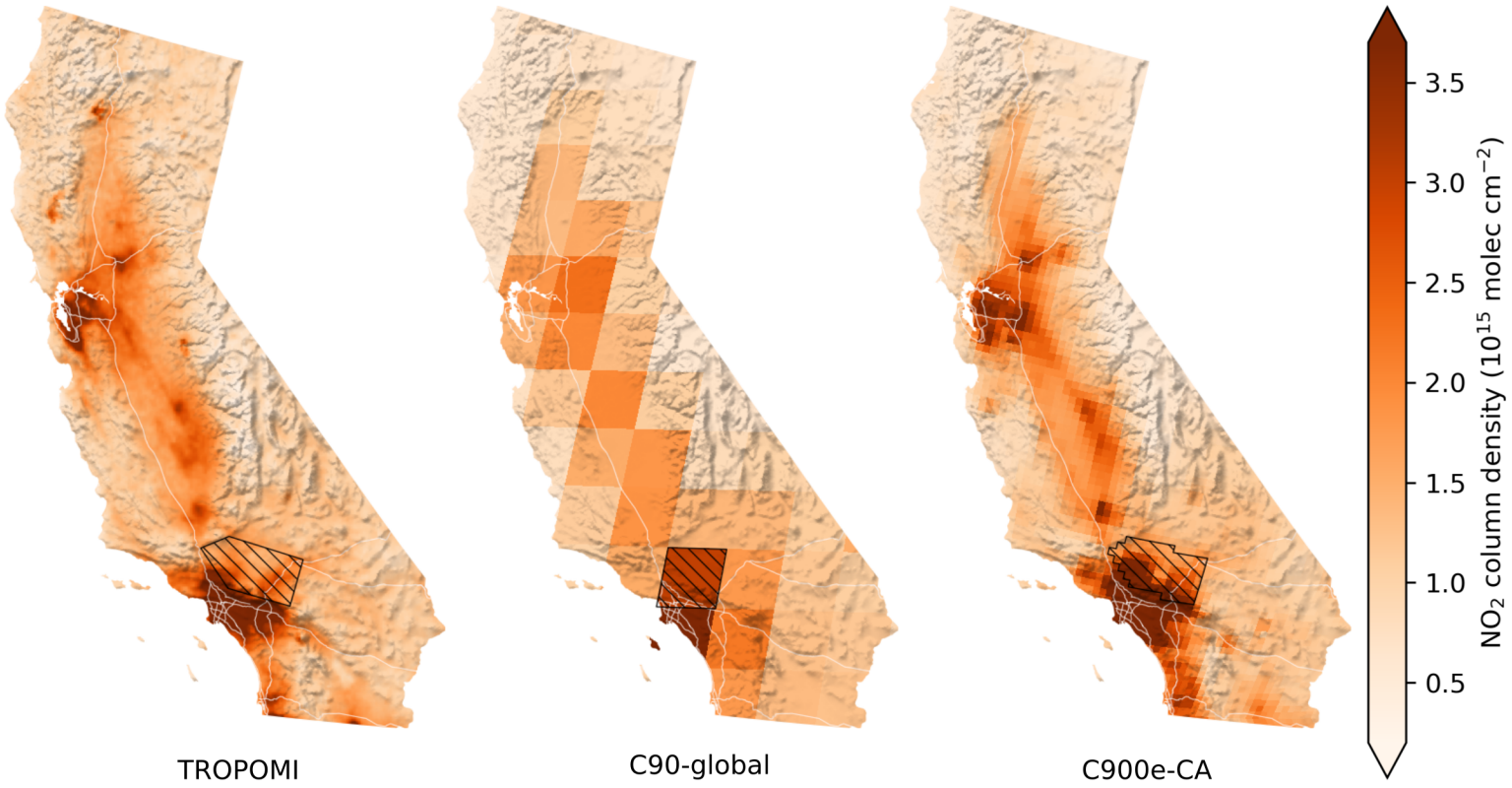


Figure 4.3: Mean tropospheric $NO_2$ column densities from C180-global and C900e-CA compared to TROPOMI observations for July 2018. Simulated means only include points where TROPOMI observations were available. TROPOMI columns shown here use recalculated AMFs from C900e-CA. Hatching shows the area/pixels affected by the spurious spreading of the Los Angeles plume.

# Chapter 5

# Conclusion

Fine resolution simulations are necessary to resolve non-linear processes, but have been impeded by computational expense. This work developed the capability for grid-stretching in the high-performance version of GEOS-Chem (GCHP), validated the capability versus global cubed-sphere simulations, and assessed its operation with two case studies compared with observations from the TROPOMI satellite instrument. The extent and position of a stretched-grid's refinement is flexible and nimbly specified by four runtime parameters. Stretched-grids offer regional- and continental-scale studies faster simulations and finer resolutions.

Generally, stretched-grids in the range of 1.5–4.0 are expected to be most applicable in GCHP because their refinement is broad, no grid-boxes outside the refinement become extremely enlarged, and the computational efficiency is well worthwhile (cubed-sphere grids with comparable resolution to the stretched-grids refinement would have 2.25–16.0 times as many grid-boxes). Larger stretch-factors have regional-scale fine-resolution applications, but warrant caution if the grid resolution becomes finer than the meteorological inputs, or coarser than C24 where prior evaluations of GEOS-Chem are limited. To help estimate an appropriate upper limit for the stretch-factor for a given GCHP application, this work proposes a simple procedure based on constraints from the required radius of refinement and the desired maximum and minimum resolution of the stretched-grid.

Computationally, stretched-grids offer unprecedented resolutions for GEOS-Chem. Improved emissions, with more accurate spatial and temporal variability, and finer resolution meteorological inputs will be needed to fully exploit of these newly achievable resolutions. Stretched-grid simulations are most efficient when the resolution at the target antipode is as coarse as possible. Studies that are insensitive to coarse resolution outside the refinement can more easily use larger stretch factors and fully

exploit the efficiency of stretched-grid simulations. Stretched-grids are publicly available and ready for scientific application in GEOS-Chem version 13.0.0.

## Bibliography


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, Journal of Geophysical Research: Atmospheres, 106, 23 073–23 095, doi:10.1029/2001JD000807, URL http://doi.wiley.com/10.1029/2001JD000807, 2001.

Boersma, K., Braak, R., and van der A, R. J.: Dutch OMI NO2 (DOMINO) data product v2. 0, Tropospheric Emissions Monitoring Internet Service on-line documentation, URL http://www.temis.nl/docs/OMI_NO2_HE5_2.0_2011.pdf, 2011.

Boersma, K. F., Eskes, H. J., and Brinksma, E. J.: Error analysis for tropospheric NO $_2$ retrieval from space: ERROR ANALYSIS FOR TROPOSPHERIC NO $_2$, Journal of Geophysical Research: Atmospheres, 109, n/a–n/a, doi:10.1029/2003JD003962, URL http://doi.wiley.com/10.1029/2003JD003962, 2004.

Boersma, K. F., Eskes, H. J., Richter, A., De Smedt, I., Lorente, A., Beirle, S., van Geffen, J. H. G. M., Zara, M., Peters, E., Van Roozendael, M., Wagner, T., Maasakkers, J. D., van der A, R. J., Nightingale, J., De Rudder, A., Irie, H., Pinardi, G., Lambert, J.-C., and Compernolle, S. C.: Improving algorithms and uncertainty estimates for satellite NO $_2$ retrievals: results from the quality assurance for the essential climate variables (QA4ECV) project, Atmospheric Measurement Techniques, 11, 6651–6678, doi:10.5194/amt-11-6651-2018, URL https://www.atmos-meas-tech.net/11/6651/2018/, 2018.

Cohan, D. S., Hu, Y., and Russell, A. G.: Dependence of ozone sensitivity analysis on grid resolution, Atmospheric Environment, 40, 126–135, doi:10.1016/j.atmosenv.2005.09.031, URL https://linkinghub.elsevier.com/retrieve/pii/S1352231005008666, 2006.

Cooper, M. J., Martin, R. V., McLinden, C. A., and Brook, J. R.: Inferring ground-level nitrogen dioxide concentrations at fine spatial resolution applied to the TROPOMI satellite instrument, Environmental Research Letters, doi:10.1088/1748-9326/aba3a5, URL https://iopscience.iop.org/article/10.1088/1748-9326/aba3a5, 2020.

Davis, D. D., Grodzinsky, G., Kasibhatla, P., Crawford, J., Chen, G., Liu, S., Bandy, A., Thornton, D., Guan, H., and Sandholm, S.: Impact of ship emissions on marine boundary layer NO $_x$ and SO $_2$ Distributions over the Pacific Basin, Geophysical Research Letters, 28, 235–238, doi:10.1029/2000GL012013, URL http://doi.wiley.com/10.1029/2000GL012013, 2001.

De Smedt, I., Müller, J.-F., Stavrakou, T., van der A, R., Eskes, H., and Van Roozendael, M.: Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, Atmospheric Chemistry and Physics, 8, 4947–4963, doi:10.5194/acp-8-4947-2008, URL https://acp.copernicus.org/articles/8/4947/2008/, 2008.

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., Van Roozendael, M., Richter, A., Hilboll, A., Peters, E., Pedergnana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., van Geffen, J., Boersma, K. F., and Veefkind, P.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project, Atmospheric Measurement Techniques, 11, 2395–2426, doi:10.5194/amt-11-2395-2018, URL https://amt.copernicus.org/articles/11/2395/2018/, 2018.

Eastham, S. D., Long, M. S., Keller, C. A., Lundgren, E., Yantosca, R. M., Zhuang, J., Li, C., Lee, C. J., Yannetti, M., Auer, B. M., Clune, T. L., Kouatchou, J., Putman, W. M., Thompson, M. A., Trayanov, A. L., Molod, A. M., Martin, R. V., and Jacob, D. J.: GEOS-Chem High Performance (GCHP v11-02c): a next-generation implementation of the GEOS-Chem chemical transport model for massively parallel applications, Geoscientific Model Development, 11, 2941–2953, doi:10.5194/gmd-11-2941-2018, URL https://www.geosci-model-dev.net/11/2941/2018/, 2018.

EPA: Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, PM2.5, and Regional Haze, 2007.

Esler, J. G., Roelofs, G. J., Köhler, M. O., and O'Connor, F. M.: A quantitative analysis of grid-related systematic errors in oxidising capacity and ozone production rates in chemistry transport models, Atmospheric Chemistry and Physics, 4, 1781–1795, doi:10.5194/acp-4-1781-2004, URL http://www.atmos-chem-phys.net/4/1781/2004/, 2004.

Fox-Rabinovitz, M., Côté, J., Dugas, B., Déqué, M., and McGregor, J. L.: Variable resolution general circulation models: Stretched-grid model intercomparison project (SGMIP), Journal of Geophysical Research, 111, D16 104, doi:10.1029/2005JD006520, URL http://doi.wiley.com/10.1029/2005JD006520, 2006.

Fox-Rabinovitz, M., Cote, J., Dugas, B., Deque, M., McGregor, J. L., and Belochitski, A.: Stretched-grid Model Intercomparison Project: decadal regional climate simulations with enhanced variable and uniform-resolution GCMs, Meteorology and Atmospheric Physics, 100, 159–178, doi:10.1007/s00703-008-0301-z, URL http://link.springer.com/10.1007/s00703-008-0301-z, 2008.

Freitas, S. R., Longo, K. M., Chatfield, R., Latham, D., Silva Dias, M. A. F., Andreae, M. O., Prins, E., Santos, J. C., Gielow, R., and Carvalho, J. A.: Including the sub-grid scale plume rise of vegetation fires in low resolution atmospheric

transport models, Atmospheric Chemistry and Physics, 7, 3385–3398, doi:10.5194/acp-7-3385-2007, URL https://acp.copernicus.org/articles/7/3385/2007/, 2007.

Gillani, N. V. and Pleim, J. E.: Sub-grid-scale features of anthropogenic emissions of NOx and VOC in the context of regional eulerian models, Atmospheric Environment, 30, 2043–2059, doi:10.1016/1352-2310(95)00201-4, URL https://linkinghub.elsevier.com/retrieve/pii/1352231095002014, 1996.

Goldberg, D. L., Lamsal, L. N., Loughner, C. P., Swartz, W. H., Lu, Z., and Streets, D. G.: A high-resolution and observationally constrained OMI NO $_2$ satellite retrieval, Atmospheric Chemistry and Physics, 17, 11 403–11 421, doi:10.5194/acp-17-11403-2017, URL https://www.atmos-chem-phys.net/17/11403/2017/, 2017.

Harris, L. M., Lin, S.-J., and Tu, C.: High-Resolution Climate Simulations Using GFDL HiRAM with a Stretched Global Grid, Journal of Climate, 29, 4293–4314, doi:10.1175/JCLI-D-15-0389.1, URL http://journals.ametsoc.org/doi/10.1175/JCLI-D-15-0389.1, 2016.

Heckel, A., Kim, S.-W., Frost, G. J., Richter, A., Trainer, M., and Burrows, J. P.: Influence of low spatial resolution a priori data on tropospheric NO $_2$ satellite retrievals, Atmospheric Measurement Techniques, 4, 1805–1820, doi:10.5194/amt-4-1805-2011, URL https://www.atmos-meas-tech.net/4/1805/2011/, 2011.

Hill, C., DeLuca, C., Balaji, Suarez, M., and Da Silva, A.: The Architecture of the Earth System Modeling Framework, Computing in Science & Engineering, 6, 18–28, doi:10.1109/MCISE.2004.1255817, URL http://ieeexplore.ieee.org/document/1255817/, 2004.

Hudman, R. C., Moore, N. E., Mebust, A. K., Martin, R. V., Russell, A. R., Valin, L. C., and Cohen, R. C.: Steps towards a mechanistic model of global soil nitric oxide emissions: implementation and space based-constraints, Atmospheric Chemistry and Physics, 12, 7779–7795, doi:10.5194/acp-12-7779-2012, URL https://acp.copernicus.org/articles/12/7779/2012/, 2012.

Jang, J.-C. C., Jeffries, H. E., Byun, D., and Pleim, J. E.: Sensitivity of ozone to model grid resolution—I. Application of high-resolution regional acid deposition model, Atmospheric Environment, 29, 3085–3100, doi:10.1016/1352-2310(95)00118-I, URL https://linkinghub.elsevier.com/retrieve/pii/135223109500118I, 1995.

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, Geoscientific Model Development, 7, 1409–1417, doi:10.5194/gmd-7-1409-2014, URL https://www.geosci-model-dev.net/7/1409/2014/, 2014.

Kim, S.-W., Natraj, V., Lee, S., Kwon, H.-A., Park, R., de Gouw, J., Frost, G., Kim, J., Stutz, J., Trainer, M., Tsai, C., and Warneke, C.: Impact of high-resolution a priori profiles on satellite-based formaldehyde retrievals, Atmospheric Chemistry and Physics, 18, 7639–7655, doi:10.5194/acp-18-7639-2018, URL https://www.atmos-chem-phys.net/18/7639/2018/, 2018.

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F., and Bergamaschi, P.: The two-way nested global chemistry-transport zoom model TM5: algorithm and applications, Atmospheric Chemistry and Physics, 5, 417–432, doi:10.5194/acp-5-417-2005, URL https://acp.copernicus.org/articles/5/417/2005/, 2005.

Krotkov, N. A., Lamsal, L. N., Celarier, E. A., Swartz, W. H., Marchenko, S. V., Bucsela, E. J., Chan, K. L., Wenig, M., and Zara, M.: The version 3 OMI NO<sub>2</sub> standard product, Atmospheric Measurement Techniques, 10, 3133–3149, doi:10.5194/amt-10-3133-2017, URL https://amt.copernicus.org/articles/10/3133/2017/, 2017.

Kumar, N. and Russell, A. G.: Multiscale air quality modeling of the Northeastern United States, Atmospheric Environment, 30, 1099–1116, doi:10.1016/1352-2310(95)00317-7, URL https://linkinghub.elsevier.com/retrieve/pii/1352231095003177, 1996.

Li, J., Wang, Y., and Qu, H.: Dependence of Summertime Surface Ozone on NO $_x$ and VOC Emissions Over the United States: Peak Time and Value, Geophysical Research Letters, 46, 3540–3550, doi:10.1029/2018GL081823, URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2018GL081823, 2019.

Li, Y., Henze, D. K., Jack, D., and Kinney, P. L.: The influence of air quality model resolution on health impact assessment for fine particulate matter and its components, Air Quality, Atmosphere & Health, 9, 51–68, doi:10.1007/s11869-015-0321-z, URL http://link.springer.com/10.1007/s11869-015-0321-z, 2016.

Li, Y., Pickering, K. E., Barth, M. C., Bela, M. M., Cummings, K. A., and Allen, D. J.: Evaluation of Parameterized Convective Transport of Trace Gases in Simulation of Storms Observed During the DC3 Field Campaign, Journal of Geophysical Research: Atmospheres, 123, 11,238–11,261, doi:10.1029/2018JD028779, URL http://doi.wiley.com/10.1029/2018JD028779, 2018.

Liang, J. and Jacobson, M. Z.: Effects of subgrid segregation on ozone production efficiency in a chemical model, Atmospheric Environment, 34, 2975–2982, doi:10.1016/S1352-2310(99)00520-8, URL https://linkinghub.elsevier.com/retrieve/pii/S1352231099005208, 2000.

Long, M. S., Yantosca, R., Nielsen, J. E., Keller, C. A., da Silva, A., Sulprizio, M. P., Pawson, S., and Jacob, D. J.: Development of a grid-independent GEOS-Chem chemical transport model (v9-02) as an atmospheric chemistry module

for Earth system models, Geoscientific Model Development, 8, 595–602, doi: 10.5194/gmd-8-595-2015, URL https://gmd.copernicus.org/articles/8/595/2015/, 2015.

Lorente, A., Folkert Boersma, K., Yu, H., Dörner, S., Hilboll, A., Richter, A., Liu, M., Lamsal, L. N., Barkley, M., De Smedt, I., Van Roozendael, M., Wang, Y., Wagner, T., Beirle, S., Lin, J.-T., Krotkov, N., Stammes, P., Wang, P., Eskes, H. J., and Krol, M.: Structural uncertainty in air mass factor calculation for NO $_2$ and HCHO satellite retrievals, Atmospheric Measurement Techniques, 10, 759–782, doi:10.5194/amt-10-759-2017, URL https://www.atmos-meas-tech.net/10/759/2017/, 2017.

Mari, C., Jacob, D. J., and Bechtold, P.: Transport and scavenging of soluble gases in a deep convective cloud, Journal of Geophysical Research: Atmospheres, 105, 22 255–22 267, doi:10.1029/2000JD900211, URL http://doi.wiley.com/10.1029/2000JD900211, 2000.

McDuffie, E., Smith, S., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E., Zheng, B., Crippa, M., Brauer, M., and Martin, R.: CEDS_GBD-MAPS: Global Anthropogenic Emission Inventory of NOx, SO2, CO, NH3, NMVOCs, BC, and OC from 1970-2017, doi:10.5281/ZENODO.3754964, URL https://zenodo.org/record/3754964, version Number: 2020_v1.0 type: dataset, 2020.

McLinden, C. A., Fioletov, V., Boersma, K. F., Kharol, S. K., Krotkov, N., Lamsal, L., Makar, P. A., Martin, R. V., Veefkind, J. P., and Yang, K.: Improved satellite retrievals of NO $_2$ and SO $_2$ over the Canadian oil sands and comparisons with surface measurements, Atmospheric Chemistry and Physics, 14, 3637–3656, doi: 10.5194/acp-14-3637-2014, URL https://www.atmos-chem-phys.net/14/3637/2014/, 2014.

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: IAV OF LIGHTNING CONSTRAINED BY LIS/OTD, Journal of Geophysical Research: Atmospheres, 117, doi:10.1029/2012JD017934, URL http://doi.wiley.com/10.1029/2012JD017934, 2012.

Palmer, P. I., Jacob, D. J., Chance, K., Martin, R. V., Spurr, R. J. D., Kurosu, T. P., Bey, I., Yantosca, R., Fiore, A., and Li, Q.: Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, Journal of Geophysical Research: Atmospheres, 106, 14 539–14 550, doi:10.1029/2000JD900772, URL http://doi.wiley.com/10.1029/2000JD900772, 2001.

Punger, E. M. and West, J. J.: The effect of grid resolution on estimates of the burden of ozone and fine particulate matter on premature mortality in the USA,

Air Quality, Atmosphere & Health, 6, 563–573, doi:10.1007/s11869-013-0197-8, URL http://link.springer.com/10.1007/s11869-013-0197-8, 2013.

Putman, W. M. and Lin, S.-J.: Finite-volume transport on various cubed-sphere grids, Journal of Computational Physics, 227, 55–78, doi:10.1016/j.jcp.2007.07.022, URL https://linkinghub.elsevier.com/retrieve/pii/S0021999107003105, 2007.

Schmidt, F.: Variable fine mesh in spectral global models, Beitr. Phys. Atmos., 50, 211–217, 1977.

Sillman, S., Logan, J. A., and Wofsy, S. C.: A regional scale model for ozone in the United States with subgrid representation of urban and power plant plumes, Journal of Geophysical Research, 95, 5731, doi:10.1029/JD095iD05p05731, URL http://doi.wiley.com/10.1029/JD095iD05p05731, 1990.

Stettler, M., Eastham, S., and Barrett, S.: Air quality and public health impacts of UK airports. Part I: Emissions, Atmospheric Environment, 45, 5415–5424, doi:10.1016/j.atmosenv.2011.07.012, URL https://linkinghub.elsevier.com/retrieve/pii/S135223101100728X, 2011.

Suarez, M., Trayanov, A., Hill, C., Schopf, P., and Vikhliaev, Y.: MAPL: a high-level programming paradigm to support more rapid and robust encoding of hierarchical trees of interacting high-performance components, in: Proceedings of the 2007 symposium on Component and framework technology in high-performance and scientific computing - CompFrame '07, p. 11, ACM Press, Montreal, Quebec, Canada, doi:10.1145/1297385.1297388, URL http://portal.acm.org/citation.cfm?doid=1297385.1297388, 2007.

The International GEOS-Chem User Community: geoschem/geos-chem: GEOS-Chem 12.5.0, doi:10.5281/ZENODO.3403111, URL https://zenodo.org/record/3403111, 2019.

Thompson, M., Auer, B., Clune, T., Trayanov, A., Jiang, W., Norris, P., Lundgren, L., Sherman, E., Jamieson, W., Bindle, L., Eastham, S. D., Putman, B., and Vikhliaev, Y.: GEOS-ESM/MAPL: MAPL 2.2 Release, doi:10.5281/ZENODO.3903435, URL https://zenodo.org/record/3903435, 2020.

Valin, L. C., Russell, A. R., Hudman, R. C., and Cohen, R. C.: Effects of model resolution on the interpretation of satellite NO $_2$ observations, Atmospheric Chemistry and Physics, 11, 11 647–11 655, doi:10.5194/acp-11-11647-2011, URL https://www.atmos-chem-phys.net/11/11647/2011/, 2011.

Vinken, G. C. M., Boersma, K. F., Jacob, D. J., and Meijer, E. W.: Accounting for non-linear chemistry of ship plumes in the GEOS-Chem global chemistry transport model, Atmospheric Chemistry and Physics, 11, 11 707–11 722, doi:10.5194/acp-11-11707-2011, URL https://acp.copernicus.org/articles/11/11707/2011/, 2011.

Wang, Y. X., McElroy, M. B., Jacob, D. J., and Yantosca, R. M.: A nested grid formulation for chemical transport over Asia: Applications to CO: NESTED GRID CO SIMULATION OVER ASIA, Journal of Geophysical Research: Atmospheres, 109, n/a–n/a, doi:10.1029/2004JD005237, URL http://doi.wiley.com/10.1029/2004JD005237, 2004.

Wild, O. and Prather, M. J.: Global tropospheric ozone modeling: Quantifying errors due to grid resolution, Journal of Geophysical Research, 111, D11 305, doi:10.1029/2005JD006605, URL http://doi.wiley.com/10.1029/2005JD006605, 2006.

Yan, Y., Lin, J., Chen, J., and Hu, L.: Improved simulation of tropospheric ozone by a global-multi-regional two-way coupling model system, Atmospheric Chemistry and Physics, 16, 2381–2400, doi:10.5194/acp-16-2381-2016, URL https://www.atmos-chem-phys.net/16/2381/2016/, 2016.

Yan, Y.-Y., Lin, J.-T., Kuang, Y., Yang, D., and Zhang, L.: Tropospheric carbon monoxide over the Pacific during HIPPO: two-way coupled simulation of GEOS-Chem and its multiple nested models, Atmospheric Chemistry and Physics, 14, 12 649–12 663, doi:10.5194/acp-14-12649-2014, URL https://www.atmos-chem-phys.net/14/12649/2014/, 2014.

Zhuang, J., Jacob, D. J., Lin, H., Lundgren, E. W., Yantosca, R. M., Gaya, J. F., Sulprizio, M. P., and Eastham, S. D.: Enabling High-Performance Cloud Computing for Earth Science Modeling on Over a Thousand Cores: Application to the GEOS-Chem Atmospheric Chemistry Model, Journal of Advances in Modeling Earth Systems, 12, doi:10.1029/2020MS002064, URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2020MS002064, 2020.