RESEARCH ARTICLE

Key Points:

- We introduce a system for simulations of regional air quality and aerosol-climate interactions with the UK Met Office Unified Model
- We evaluate the model against data from NASA's Korea-United States Air Quality campaign in 2016 over South Korea
- Despite shortcomings in the representation of some chemical and aerosol species, the system is ready for application and further development

Author Contributions:

Conceptualization: Hamish Gordon, Ken S. Carslaw, Adrian A. Hill, Paul R. Field
Data curation: Andreas Beyersdorf, Chelsea Corr-Limoges
Formal analysis: Hamish Gordon, Pratapaditya Ghosh, Xuemei Wang
Funding acquisition: Hamish Gordon
Methodology: Hamish Gordon

NUMAC: Description of the Nested Unified Model With Aerosols and Chemistry, and Evaluation With KORUS-AQ Data

Hamish Gordon[1,2], Ken S. Carslaw[2], Adrian A. Hill[3], Paul R. Field[2,3], Nathan Luke Abraham[4,5], Andreas Beyersdorf[6], Chelsea Corr-Limoges[7], Pratapaditya Ghosh[8], John Hemmings[3], Anthony C. Jones[3], Claudio Sanchez[3], Xuemei Wang[2,9], and Jonathan Wilkinson[3]

[1]Department of Chemical Engineering & Center for Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, PA, USA, [2]School of Earth and Environment, University of Leeds, Leeds, UK, [3]Met Office, Exeter, UK, [4]National Centre for Atmospheric Science, Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge, UK, [5]Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge, UK, [6]Department of Chemistry & Biochemistry, California State University, San Bernardino, CA, USA, [7]Department of Biology & Chemistry, Springfield College, Springfield, MA, USA, [8]Department of Civil and Environmental Engineering & Center for Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, PA, USA, [9]Now at Department of Chemical Engineering & Center for Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, PA, USA

Abstract We describe and evaluate a system for regional modeling of atmospheric composition with the Met Office Unified Model (UM), suitable for climate, weather forecasting and air quality applications. In this system, named NUMAC ("Nested UM with Aerosols and Chemistry"), a global model provides boundary conditions for regional models nested within it, using the Met Office's Regional Nesting Suite for multi-scale simulations. The regional models, which can run at convection-permitting or cloud-resolving scales, use the same code as the global model. The system includes double-moment prognostic aerosol microphysics with interactive chemistry of sulfur species, ozone, NO_x, and CO as in the UK Earth System Model. Double-moment prognostic cloud microphysics is optional. To test NUMAC, we compare simulations to surface and aircraft measurements from NASA's Korea-United States Air Quality campaign over South Korea. The performance of the regional model, which we run at 5 km resolution, is similar to the well-evaluated global model when the regional and global models use the same emissions. Most species such as ozone, NO_x, OH, or $PM_{2.5}$ are simulated within a factor of 2 of observations most of the time, though they are biased low compared to monitors in polluted areas (observed surface dry $PM_{2.5}$ averages 28 µgm^{-3} but we simulate 17 µgm^{-3}). Meteorology and clouds are represented satisfactorily. With higher-resolution emissions, many of the low model biases are reduced, but a tuning was required to keep NO concentrations realistic, indicating shortcomings in the chemistry scheme. We demonstrate the potential of NUMAC for studies of aerosol-cloud interactions.

Plain Language Summary Unified atmospheric simulation systems, defined as those that are routinely used to represent air quality, weather and climate with the same code, are rare. The UK Met Office Unified Model (UM) is one such system, but until recently, aerosol-cloud interactions were represented only in global climate simulations. Effects of aerosol-cloud interactions and atmospheric chemistry on Earth's radiation balance are large and must be included in global and regional climate predictions. However, partly because clouds and emissions are spatially inhomogeneous, their effects on climate are currently poorly understood. Simulations with high grid resolution can help us understand the relevant processes better. We document here a system for representing air quality, weather and climate with the UM, including chemistry-aerosol-cloud interactions, at high spatial grid resolution on regional scales. We test our system against surface and aircraft measurements of atmospheric chemical species such as ozone and particulate matter made in May 2016 in Korea. The model represents most species well, but has some shortcomings we must address in future work. Overall, however, we judge that it is ready for studies of how atmospheric chemical species and particulate matter affect air quality, weather and climate, and can be further improved as it is used.

1. Introduction

Aerosols and trace gases are important in regulating Earth's climate and are themselves strongly influenced by small-scale variability due to inhomogeneous emissions sources, complex terrain and clouds. To understand how aerosols affect weather and climate at the local scale, simulations capable of representing aerosols and clouds at

Software: Hamish Gordon, Pratapaditya Ghosh, John Hemmings, Anthony C. Jones, Claudio Sanchez
Writing – original draft: Hamish Gordon
Writing – review & editing: Hamish Gordon, Ken S. Carslaw, Adrian A. Hill, Paul R. Field, Nathan Luke Abraham, Andreas Beyersdorf, John Hemmings, Anthony C. Jones, Claudio Sanchez, Xuemei Wang

the "convection-permitting" scale of a few kilometers, or even at the "cloud-resolving" scale of a few hundred meters, are needed. Since synoptic circulations and long-range transport of moisture and chemical and aerosol species are often also of interest and can strongly affect local-scale processes, these high-resolution simulations are best embedded within a multi-scale modeling framework that is able to represent both global and regional meteorology and atmospheric composition.

State-of-the-art weather forecasting models such as the United Kingdom deterministic forecasting system (UKV, Tang et al., 2013) or the High Resolution Rapid Refresh Model (Benjamin et al., 2016) have simulated atmospheric evolution at kilometer-scale spatial resolution over large regions of Earth's surface for many years, but these models typically do not represent complex chemistry or aerosol microphysical processes. These processes are important sources of Earth system feedbacks and can also significantly affect weather prediction in certain situations, for example, in the forecasting of visibility in fog (Boutle et al., 2018; Jayakumar et al., 2021), or when high levels of pollution interact with the planetary boundary layer (Ding et al., 2021; Huang et al., 2020). Some of the models capable of representing aerosols and their interactions with clouds and climate are reviewed by Baklanov et al. (2014). Notable examples include the Weather Research and Forecasting Model with Chemistry (WRF-chem; Grell et al., 2005), COSMO-ART (Vogel et al., 2009), and RAMS (Saleeby & van den Heever, 2013).

Work to facilitate convection- or cloud-resolving modeling of atmospheric composition is motivated as follows. The land surface, clouds, convection, and emissions sources are often highly inhomogeneous, and atmospheric chemistry, aerosol and cloud processes are strongly and non-linearly coupled to these inhomogeneous properties (Bangert et al., 2011). Therefore, first, air quality predictions must account for this variability to accurately simulate exposure. Second, climate-motivated simulations of atmospheric composition in lower-resolution models are expected to incur errors even when the averages over large areas are considered. For example, Possner et al. (2016) found that cloud radiative effects from ship tracks were higher by over a factor of two in 50 km-resolution simulations compared to 1 km-resolution simulations of the same area. The authors showed that the 50 km-resolution simulations were inaccurate by attributing the higher radiative effects, in part, to the inability of the 50 km model to resolve the shipping emissions. Third, simulations at convection-permitting or cloud-resolving resolution can be evaluated in more detail against observations than lower-resolution models because the spatial scales are more similar. For example, the footprint on Earth of one pixel in a typical geostationary weather satellite instrument such as SEVIRI is of order 1 km. Thus these models can avoid by brute force the worst of the representativeness uncertainties discussed by Schutgens et al. (2017).

The Met Office Unified Model (UM; https://www.metoffice.gov.uk/research/approach/modelling-systems/unified-model, last access 23 March 2023) contains many of the pre-requisites for regional simulations with aerosols and chemistry at the kilometer scale. The UM is a seamless prediction framework that has supported global simulations across weather and climate timescales since around 1992 (Brown et al., 2012; Cullen, 1993). Furthermore, regional UM simulations at high spatial resolution, sharing code with consistent global simulations, have been supported for many years (with operational convection-permitting forecasts beginning at the Met Office in 2010). However, unlike in global climate simulations, all multi-year regional climate simulations with the UM to date (e.g., Stratton et al., 2018) have used aerosol climatologies (e.g., Stevens et al., 2017) or single-moment aerosol schemes (without prognostic aerosol number concentration) and highly simplified chemistry.

We use the term "systems" in this paper to refer to simulation setups that use the UM for different purposes. Some UM systems are very simple, involving a single simulation of the atmosphere with fixed grid resolution. Other systems are more sophisticated: they could correspond to air quality models, Earth System Models, or to research tools, such as the NUMAC (Nested Unified Model with Aerosols and Chemistry) system we present here and define below. The Rose/Cylc framework (Oliver et al., 2019) is used to manage the sequence of tasks a supercomputer needs to perform to simulate the atmosphere or Earth System for a defined period with the UM. These tasks differ between systems, and thus for technical purposes a system usually corresponds to a Rose "suite." We make a distinction between "system" and "configuration": a configuration such as GA7.1 (Walters et al., 2019) is a defined combination of science parameter settings, and multiple configurations can be, and typically are, used within one system. Most multi-scale simulations with the UM are managed by the Regional Nesting Suite defined in Rose, and use the Regional Atmosphere and Land (RAL) model configuration, version 2 of which (RAL2) was recently documented by Bush et al. (2023).

To represent aerosols, chemistry and aerosol-cloud interactions in the UM, the United Kingdom Chemistry and Aerosol (UKCA) submodel (Morgenstern et al., 2009; Mulcahy et al., 2018; O'Connor et al., 2014) can be

used, as, for example, in the global UK Earth System Model (UKESM1; Sellar et al., 2019). This submodel includes a modal double-moment aerosol microphysics scheme, a modified version of the Global Model of Aerosol Processes (GLOMAP-mode) (Mann et al., 2010; Mulcahy et al., 2020). Several chemistry schemes of varying complexity are available. A key question we examine in this paper is whether the chemistry scheme used in UKESM1 is suitable for regional high-resolution modeling in polluted environments. The Air Quality in the Unified Model (AQUM) model (Neal et al., 2017; Savage et al., 2013) is a system designed for air quality forecasting with the UM over the United Kingdom which uses one of the chemistry schemes from the UKCA sub-model, but not the aerosol microphysics scheme. The single-moment aerosol scheme it uses instead means it can only represent aerosol-cloud interactions to a limited extent (A. Jones et al., 2001). The AQUM system is also only routinely run over the UK. Prognostic double-moment aerosols were included in AQUM in initial tests (Hemmings & Savage, 2018), but they were not coupled to radiation or clouds and were not tested outside the UK.

Here, we present a system for the UM, which we call "NUMAC," designed for regional simulations that represent chemistry, aerosols and aerosol-cloud interactions at high spatial grid resolution, up to the kilometer scale. We define NUMAC as any UM simulation system based on the Regional Nesting Suite in which the full UKCA submodel (representing chemistry and aerosols) is included. NUMAC is designed to capture some of the most important atmospheric interactions and feedbacks between aerosols, clouds, chemistry, radiation and the land surface. It is not yet a full regional Earth System Model (cf., Sitz et al., 2017) because it is not yet coupled to an ocean model. However, this capability will become available in due course as aerosols and chemistry are being added to the Met Office's related Regional Coupled Suite (Castillo et al., 2022).

NUMAC does not contain new representations of atmospheric processes: indeed, we believe its consistency with other UM systems is one of its strengths. The atmosphere and land surface code is identical to the code that would be run in a convection-permitting global version of the UM (which has already been run without chemistry and aerosols (Stevens et al., 2019)). NUMAC would therefore be useful for testing affordably the potential to increase the resolution of UKESM1 (or another global UM system with prognostic aerosols) to convection-permitting scale. Like AQUM, NUMAC is not as sophisticated in its representation of aerosols as the most widely used state-of-the-art air quality models such as CMAQ, CAMx, or the more complex configurations of WRF-chem. We intend to build on the work here to increase the sophistication of the aerosol scheme in NUMAC in future. We emphasize also that NUMAC, like WRF-chem, is built on and includes a weather prediction model, and unlike chemical transport models like CMAQ or CAMx, it does not require separate driving meteorology.

Aerosol-cloud interaction is an important application of NUMAC. We therefore include the two-moment aerosol microphysics scheme GLOMAP-mode (Mann et al., 2010), as used in the global climate model (Mulcahy et al., 2020), and (optionally) the two-moment Cloud AeroSol Interacting Microphysics scheme (CASIM; Field et al., 2023; Grosvenor et al., 2017; B. J. Shipway & Hill, 2012). CASIM has similar sophistication to many of the cloud-aerosol microphysics schemes available in WRF-chem, for example, Morrison and Gettelman (2008), and is being adopted as the default microphysics scheme for the regional UM (including for operational weather prediction), starting from the RAL version 3 configuration of the model. In this paper, we build on previous studies which used some of these components (Gordon et al., 2018, 2020; Jayakumar et al., 2021; Planche et al., 2017; X. Wang et al., 2023) to combine prognostic atmospheric chemistry beyond the sulfur cycle with interactive double-moment aerosol and cloud microphysics for the first time.

In this paper we describe the NUMAC system in Section 2. We then evaluate NUMAC in detail against a case study chosen from the Korea-United States Air Quality (KORUS-AQ) field campaign in 2016. The case study and observation data sets are described in Section 3, model setup specific to this paper in Section 4 and the model evaluation in Sections 5 and 6. The aim of our model evaluation is to highlight areas where further model development is needed to ensure the model's chemistry and aerosol scheme scales to high grid resolutions. For example, we already found (Gordon et al., 2020) that the representation of aerosol activation is not fully scale invariant and plan to work on that in future. Another question we examine here is whether the chemistry scheme can handle highly inhomogeneous emissions sources that would be averaged out in a low-resolution global model. The performance of the model will most likely vary widely according to the region simulated and the meteorological conditions. However, we did not systematically tune the model for this region, we already evaluated a relatively similar model near Ascension Island (Gordon et al., 2020), and related papers in preparation will document the behavior of the model in other locations. To illustrate the versatility and the potential for a wide range of

applications of NUMAC, in Section 7 we include a demonstration of how we are applying it in two such locations, in the region of France near Paris and in the Amazon rainforest. We then discuss the advantages of the higher spatial resolution and of the higher resolution emissions inventories in Section 8, before summarizing the further developments to the model we need to make, and concluding the paper.

2. Model Description

NUMAC is defined as the Met Office's Regional Nesting Suite for multi-scale simulations with the UM, with both atmospheric chemistry and aerosol microphysics represented by the UKCA sub-model. The Regional Nesting Suite (Bush et al., 2020) allows a number of regional model domains of the atmosphere, which must be rectangular in the horizontal plane, to be defined by the user. The user specifies the location and size of these regions and the horizontal and vertical resolution. The regional models are nested inside the global UM, and multiple regional models of different resolution can be inside each other. The regional domain for our case study is shown in Figure 3. The Regional Nesting Suite (or closely related predecessors) has been tested with regional models running at spatial grid resolutions as fine as 50 m (Lean et al., 2019).

In NUMAC, both global and regional models use the same code base. This strategy has the advantage that, in principle, high-resolution regional model evaluations can be used to improve the performance of physics and chemistry schemes in the global climate model, without the disadvantages of coarse model resolution (most notably poorly resolved orography, land surface, and emissions; parameterized convection; and high representativeness uncertainties when comparing to measurements). In principle the NUMAC system does not require a specific model configuration, but for this paper, we use the global atmosphere science configuration GA7.1 (Walters et al., 2019), as used in the GC3.1 (Williams et al., 2018) and UKESM1 (Sellar et al., 2019) models. While the configuration and land surface coupling in our system are similar to the atmosphere-only UKESM global model, we did not attempt to replicate the tuning of UKESM1 (Sellar et al., 2019) or UKESM1.1 (Mulcahy et al., 2023) that occurred after the base GA7.1 configuration was established, so some settings still differ. For our regional model we use the RAL1 physical atmosphere model configuration described by Bush et al. (2020), who also outline the differences with respect to the global configuration (in dynamics, boundary layer, radiation, and sub-grid cloud). Examples of these differences are the potential for the boundary layer scheme of Lock et al. (2000) to blend with a 3D turbulent mixing scheme based on Smagorinsky (1963), which becomes active at grid resolutions finer than 1 km (Boutle et al., 2014), and the use of a different cloud parameterization for the mid-latitudes. The regional configuration is designed for regional-scale modeling at resolutions of order 1 km with no parameterized convection. For regional simulations at spatial resolutions coarser than around 10 km, it may be advisable to include the convection parameterization from the global model also in the regional model. The Joint UK Land Environment Simulator (JULES) in its GL7 configuration (Walters et al., 2019) is used in the global model, again with minor modifications (Bush et al., 2020) for the regional model.

The physics schemes that relate most to chemistry and aerosols in the most sophisticated variant of NUMAC are summarized in Figure 1. The couplings between the schemes that are most relevant to aerosol-cloud interactions are indicated on the figure and described in the following sections. In these sections, we focus our description on the regional model rather than the global model. More details are available in the original papers describing the physics schemes, which are referenced where appropriate.

2.1. Dynamics, Boundary Layer, Cloud Cover, Radiation and Land Surface

The UM uses the ENDGAME semi-Lagrangian dynamical core (Thuburn, 2016; Wood et al., 2014), which is applied to both the global and regional models, with some differences in the formulation for regional models described by Wood et al. (2014). One important difference is that the model top for the global models is usually 70 or 85 km altitude, while for regional models it is by default 40 km. Boundary layer turbulence is represented with the scheme of Lock et al. (2000) with some minor modifications (Bush et al., 2020). Radiative transfer is treated by the SOCRATES scheme (Manners et al., n.d.) based on work by Edwards and Slingo (1996). The sub-grid cloud fraction is by default that of Smith (1990) in the regional model and the PC2 (Prognostic Cloud, Prognostic Condensate) scheme (Wilson et al., 2008) in the global model. The PC2 scheme may be deployed in the regional model, as in the RAL1-T (RAL version 1-tropical) configuration, but in this paper we test only the Smith (1990) scheme in regional simulations. A more-sophisticated bimodal cloud scheme has also very recently become available for the latest UM code versions (Van Weverberg et al., 2021).
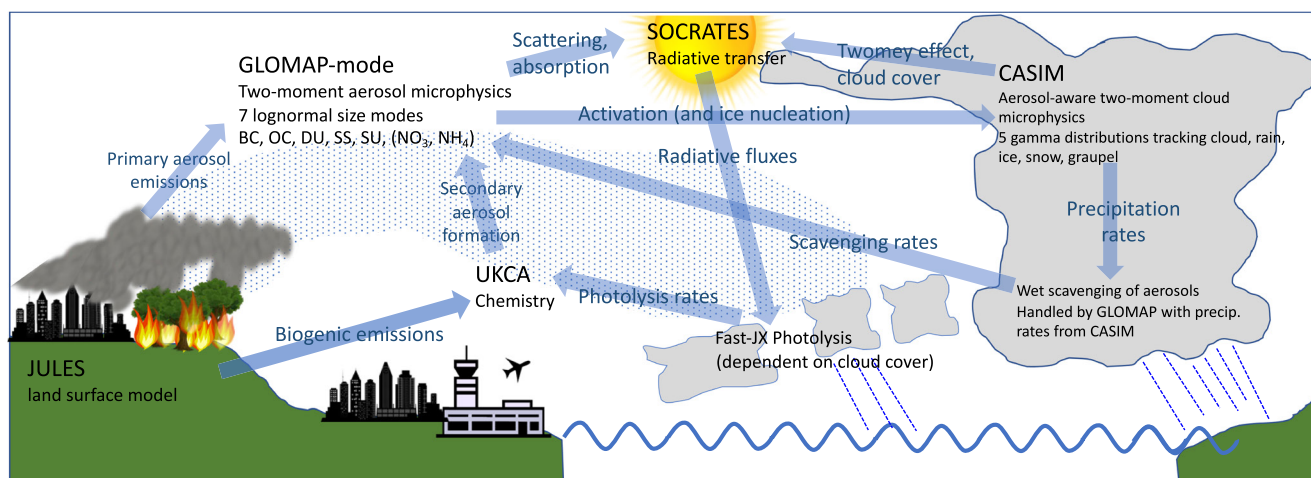
Figure 1. Schematic of the most relevant schemes for chemistry and aerosols in the Unified Model. Selected couplings between the schemes are indicated with blue arrows and blue text. A hypothetical pollution plume that interacts with clouds is shown with blue stippling. The Cloud AeroSol Interacting Microphysics (CASIM) scheme for double-moment cloud microphysics is shown, but can be substituted for the Wilson and Ballard (1999) cloud microphysics scheme with no other changes to the figure. NO_3 and NH_4 are in parentheses as they are not included in Nested Unified Model with Aerosols and Chemistry (NUMAC) yet.

In NUMAC, land surface processes are modeled using the JULES model, in much the same way as in global climate modeling systems of the UM. Emissions of monoterpenes and isoprene are parameterized in JULES based on temperature, solar radiation and plant functional type (Pacifico et al., 2011). Our JULES configuration is based on RAL1 for the regional model and GL7 for the global model, and thus in NUMAC, JULES represents nine surface types (Walters et al., 2019), while UKESM1 represents 27 types.

2.2. Initialization, User Interface, and Lateral Boundary Conditions

Regional models are typically initialized from global reanalyzes or forecasts by regridding the prognostic fields of these analyses to the regional grid. Weather forecasting analyses from the UM do not contain chemistry and aerosol fields, but they can be merged with output from climate simulations which do, using a UM application called "reconfiguration." If climate model output with these chemistry and aerosol fields is available on the case study start date and has meteorology that is consistent with the weather forecasting analyses, for example, because the climate model has been nudged to reanalysis, no further steps are needed. If not, NUMAC can be used to run a global "spin-up" simulation. The spin-up simulation can be initialized from any realistic chemistry/aerosol and meteorological output that need not represent simulations of the same date. For simulations focusing on the troposphere rather than the stratosphere, the aerosol and meteorological fields will be sufficiently consistent once this spin-up run has simulated a few months of atmospheric evolution. Then, at the start of the period of interest, or shortly before, the regional model can be initialized from this more-consistent global spin-up simulation.

It is often desirable to reinitialize the meteorological fields in NUMAC at regular intervals from operational files produced with data assimilation, to ensure the simulation closely follows observed meteorology. The user may choose to reinitialize either both the global and regional meteorological fields at the same time, or just reinitialize the global model. For example, a week-long simulation could be made up of seven global and regional simulations or "cycles," each simulating 36 hr of atmospheric evolution, and each initialized 24 hr apart. The cycles would overlap by 12 hr, to allow for adjustment time or "spin-up." This procedure is also used in the Regional Coupled Suite (Castillo et al., 2022). However, since the available operational files do not contain chemistry or aerosol fields, we added the capability to retain, or carry forward, chemistry and aerosol fields in the global and regional domains while the meteorological input fields are re-initialized at a fixed cycling frequency. In our example of the week-long simulation with seven cycles, the aerosol fields for the second 36 hr-long cycle would be initialized from the aerosol fields 24 hr into the first cycle. This procedure is frequently adopted in other models such as WRF-chem (Ha, 2022). With this capability, it would also be possible to carry forward other fields, for example, moisture tracers, from one cycle to the next if desired. The overall workflow for the first two cycles of the full multi-scale model is summarized in Figure 2, assuming that the user has chosen to reinitialize both the regional

Step	1	2	3	4	5	6	7	8	9	10
Global aerosol initialization	1									
Global initialization		1								
Global forecast			1	1	1	1				
Regional initialization			1							
Regional forecast				1	1	1				
Link to previous cycle					2					
Global initialization						2				
Global forecast							2	2	2	2
Regional initialization							2			
Regional forecast								2	2	2

Figure 2. Simulation workflow for two cycles, denoted 1 and 2. The second cycle may start before the first cycle has finished, but the previous simulation must reach the start point of the second cycle (which we assume happens for both the global and the regional model by the end of step 4, marked with double vertical lines in the table) before this can happen. The two cycles would produce overlapping simulations during steps 5, and 6 here. In principle, a third cycle could start once the second cycle reaches its start point, perhaps at the end of step 8.

and global model at regular intervals. The user would usually reinitialize both regional and global model fields, but would reinitialize the global model only if they wish the meteorology in the regional model to be constrained only by its boundary conditions, or to avoid any discontinuities in time-series of regional simulation output.

UM simulations are managed using the Rose interface to the Cylc scheduling software (Oliver et al., 2019). The Regional Nesting Suite is defined using Rose, which provides a convenient and flexible graphical user interface. With this interface, users define and simulate multiple nested regions within a global simulation, as described at the start of this section. The graphical interface also allows the user to configure the model by adding, removing or editing predefined namelist files (which can be specific to particular systems and optionally included or not). The Rose suite interface to Cylc then manages a workflow of tasks that performs the simulations, first compiling the model and preparing the input fields and then running sequential global and regional simulations for the time
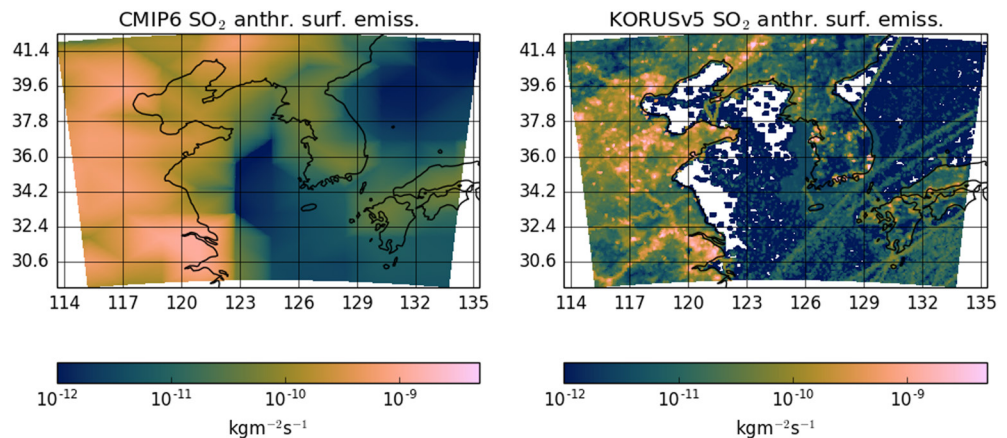
Figure 3. Comparison of our climate-model-resolution CMIP6 and 0.1°-resolution KORUSv5 SO_2 anthropogenic emissions. In both cases the emissions are plotted after regridding to the 5 km resolution of our simulations. This figure does not represent the real resolution of the CMIP6 inventory, which is 0.5°.

period the user wishes. These tasks are monitored, and may be altered during the workflow, using another graphical interface. In principle, the user need only supply enough data (initialization files and ancillary fields such as vegetation fraction or orography) to run the global simulation, and then the Regional Nesting Suite prepares input data for the regional simulations by regridding the appropriate global fields. These inputs can be selected from global input data sets at very high resolution, for example, the 100 m Shuttle Radar Tomography Mission orography (Farr et al., 2007).

Figure 1 of Bush et al. (2020) describes the regional model's boundary conditions and how tracers are advected across them via the external halo and blending zone at the edges of the regional domain. It is recommended to exclude grid cells close to the boundaries from analysis of simulation output (Gordon et al., 2018), in order to ensure the higher resolution of the regional simulation has had an effect on the prognostic fields that are advecting into the domain.

Emissions files for regional simulations with chemistry and aerosols are produced by regridding emissions data sets, either at the start of the NUMAC workflow (just before the regional and global models are initialized) or in a separate Rose suite known as the Regional Ancillary Suite. We designed NUMAC-specific scripts for these Rose suites to use either climate-resolution CMIP6 emissions or EDGAR-HTAP emissions at 0.1° spatial resolution. EDGAR-HTAP provides monthly gridded lumped emissions for 2008 and 2010, while global annual mean EDGAR v5.0 emissions are available for 1970–2015 in the same format. For this paper, we use the same code to apply the KORUSv5 inventory, although this inventory could not be used outside East Asia.

2.3. Chemistry

NUMAC can currently be configured to use two chemistry schemes with different levels of complexity within the UKCA submodel. In this paper, we use the more-complex StratTrop chemistry scheme (Archibald et al., 2020), which is also used by the UKESM1 (Sellar et al., 2019). The Archibald et al. (2020) study lists the 84 chemical species represented as prognostic tracers or in steady state and the 291 chemical reactions. The scheme is designed primarily to simulate tropospheric and stratospheric ozone, so represents the O_x, HO_x, and NO_x, cycles, some halogen chemistry, and the chemistry of isoprene and of other volatile organic compounds (VOCs) containing up to three carbon atoms. In addition, there is a simplified representation of DMS and SO_2 chemistry, and monoterpenes are included as precursors for secondary organic aerosol (SOA; as described below), but without a full chemical mechanism for their reactions. The scheme has been used extensively in global model studies over the past 5 years. The evaluation of Archibald et al. (2020) found that the model represented ozone very well on an annual mean basis, though with some seasonal biases. It underestimated NO concentrations despite predicting too much lightning in the tropics. Representation of the OH radical is mostly good except with a high bias in the tropical boundary layer.

The alternative, simpler chemistry scheme available within UKCA is known as the "offline oxidants" scheme. In this scheme, concentrations of oxidants OH, O_3, NO_3, and HO_2 are read in from files (from a previous simulation) and only the sulfur cycle and production of SOA from monoterpenes are represented. The UM with this chemistry scheme was described by Gordon et al. (2018) and, with the aerosols coupled to CASIM cloud microphysics, by Gordon et al. (2020). We do not show simulations with this oxidants scheme here, but its reduced computational expense (simulations with offline oxidants are approximately a factor two faster than those with StratTrop chemistry) will mean it is useful especially in numerical weather prediction. In future, we plan to test more complex chemistry schemes in the NUMAC system.

For regional simulations within NUMAC, the lower model top of 40 km (compared with the 85 km top of the global model) has implications for the ozone column above the model top. However, previous UM global models had a top at 39 km altitude, and a prescribed ozone column above the top of 5×10^{17} molecules cm^{-2} (compared to 6.3×10^{13} molecules cm^{-2} above the model top of 85 km in UKESM1). We therefore assume the former column ozone concentration above the model top.

To ensure consistency between global and regional models, all advected tracers in the chemistry scheme are passed through lateral boundary conditions (LBCs) by default. In future, some saving in computational expense and disk space may be obtained by not passing short-lived tracers such as the peroxy radical species through the boundary conditions, as the concentrations of these species within the domain are unlikely to be strongly influenced by the LBCs.

In the StratTrop chemistry scheme available in NUMAC, by default only convective cloud properties are passed into the lightning scheme of Price and Rind (1992). Therefore, if the convection parameterization is switched off, no NO would be produced from lightning. To address this problem in NUMAC, the lightning flash rate scheme of McCaul et al. (2009), which is already in use in the RAL1 configuration (Bush et al., 2020), is used instead. As the McCaul et al. (2009) scheme diagnoses total lightning flash rate, the simple latitude-based function of Price and Rind (1993) is used as an initial method to separate intracloud and cloud-to-ground lightning for NO production.

2.4. Aerosol Microphysics

NUMAC uses the GLOMAP-mode aerosol microphysics scheme (Mann et al., 2010) as implemented in UKCA and UKESM1 (Mulcahy et al., 2020). The simulation of aerosol-cloud interactions optionally diverges from UKESM1, otherwise the settings are the same. We include improvements by Mann et al. (2012) and the changes recommended by Mulcahy et al. (2018, 2023) to reduce the excessive aerosol radiative forcing in the GA7 and UKESM1 configurations of the global model.

Aerosols, excluding dust, are represented by five log-normal modes: nucleation, Aitken, accumulation and coarse modes, and an additional "Aitken insoluble" mode designed to hold fresh emissions of primary carbonaceous aerosols. The number of aerosol particles and the mass of each chemical component (sulfate, sea salt, black carbon, and organic carbon) in each mode are prognosed. These chemical components are thus traced separately but are assumed to be internally mixed. Sea salt and black carbon are not allowed in the nucleation mode, and sea salt is also not represented in the Aitken mode. Nitrate and ammonium aerosol are not represented in NUMAC yet. Fixed geometric standard deviations of each mode are assumed, while the median diameter of each mode is calculated from the mass and number concentration in the mode. Emissions of sea spray, sulfate and carbonaceous aerosols, described below, are allocated to modes according to the size distribution of the emitted aerosols (Mulcahy et al., 2020). Biofuel, fossil fuel, and biomass burning emissions are emitted into the Aitken insoluble mode and then transferred into the Aitken mode, a process termed "aging" by Mann et al. (2010), if 10 monolayers of sulfate or SOA condenses onto the particles. To form SOA, a generic monoterpene species is oxidized by OH, ozone, and NO_3 radicals at the reaction rates of alpha-pinene to form a non-volatile precursor species. The yield of this precursor species from the oxidation reactions is 26%, which is artificially high to compensate for the lack of SOA from isoprene or anthropogenic organic species, as discussed by Mulcahy et al. (2020). The rate of formation of SOA by this species is limited by the rate at which molecules of it collide with particles. Ongoing work will refine this treatment to include semi-volatile SOA, and SOA formation from isoprene and anthropogenic VOCs.

By default, NUMAC, in line with UKESM1, represents dust with the six-bin CLASSIC sectional scheme. Dust is emitted in nine size bins in a scheme based on that of Woodward (2001), with emissions depending on wind speed and soil moisture, and it is then redistributed to the six CLASSIC tracers. We use the same dust parameter settings as UKESM1, which lead to dust emissions being a factor of two higher than in the GC3.1 configuration (Mulcahy et al., 2020). However, we did not repeat the UKESM tuning of the bare soil fraction (Sellar et al., 2019), and the dust in NUMAC could probably be improved in a future dedicated study.

2.5. Cloud Microphysics and Aerosol-Cloud Interactions

Two cloud microphysics schemes may be used in NUMAC: the single-moment scheme of Wilson and Ballard (1999) used in the default NUMAC system or the multi-moment CASIM scheme (B. J. Shipway & Hill, 2012) as implemented in the UM by Grosvenor et al. (2017) and Field et al. (2023) and coupled to the UKCA aerosol microphysics scheme by Gordon et al. (2020). The two schemes were compared by Furtado et al. (2018) with prescribed aerosol number concentrations for a case study of organized deep convection over southern China.

The Wilson and Ballard (1999) scheme was used for the CMIP6 experiments, that is, in UKESM1 and is currently used for all Met Office operational weather forecasts, but it is intended to be replaced by CASIM in the future. It represents cloud, rain, ice, snow, and graupel masses prognostically. In its double-moment form, used in this paper, CASIM represents the number and mass concentrations of these species and also treats ice and snow separately. Both schemes represent condensation of water vapor assuming saturation adjustment.

Aerosol-cloud interactions are represented in both microphysics schemes. Both Wilson and Ballard (1999) and CASIM microphysics schemes use the Abdul-Razzak and Ghan (2000) parameterization to calculate the cloud droplet number concentration by default. If the Wilson and Ballard (1999) scheme is used, this scheme and the radiation code read in the diagnostic cloud droplet number concentration from an implementation of the Abdul-Razzak and Ghan (2000) parameterization in UKCA (West et al., 2014). This implementation assumes that the droplet concentration at all levels in a cloud is equal to that at cloud base (where aerosol activation is calculated). If the CASIM microphysics scheme is used in double-moment form, aerosol mass and number concentrations are passed to it. CASIM then calculates the hygroscopicity parameter kappa (Petters & Kreidenweis, 2007) of each aerosol mode based on its composition (Gordon et al., 2020) and activates these aerosols to cloud droplets using its own separate implementation of the Abdul-Razzak and Ghan (2000) parameterization. Alternatively, and only in CASIM, the parameterization of B. Shipway and Abel (2010) may be used, but we do not use this option in this paper. The CASIM implementation of the Abdul-Razzak and Ghan (2000) parameterization is run at each vertical level within a cloud. In the double-moment variant of the CASIM scheme, the droplet number concentration is prognostic. It is recalculated on each timestep, and then updated if it exceeds the droplet concentration that already existed in the box, with some modifications to handle sub-grid cloud fraction discussed in detail by Gordon et al. (2020).

2.6. Trace Gas and Aerosol Emissions

For both chemistry and aerosols, anthropogenic emissions (including from biomass burning) are prescribed by default from the CMIP6 inventory for global driving and regional models. All sulfur dioxide emissions are at the surface, in line with UKESM1 (Mulcahy et al., 2020). The CMIP6 emissions are read in at the ~135 km resolution of the global model. Using these low-resolution emissions in our regional model has some value in order to compare the regional and global models. For regional simulations where the comparison with the global model is not the primary motivation, higher resolution emissions can be used in the regional models, and we demonstrate an example in our evaluation below.

Natural emissions of sea spray (sea salt and primary marine organic aerosol) and dust are parameterized in NUMAC as described by Mulcahy et al. (2020). Biogenic VOC emissions are determined by JULES interactively as described earlier. The NUMAC model does not include an ocean biogeochemistry component and therefore DMS concentrations in seawater are prescribed from Lana et al. (2011) and the emission flux to the atmosphere is calculated following Liss and Merlivat (1986). The interactive fire model INFERNO (Teixeira et al., 2021) could be used to replace CMIP6 biomass burning emissions in future if needed.

3. KORUS-AQ Case Description and Observation Data Sets

The KORUS-AQ campaign (Crawford et al., 2021) was motivated by the need to understand factors driving air quality in Korea. Korea suffers from high air pollutant concentrations: for example, annual average $PM_{2.5}$ was around 25 μgm^{-3} in the most densely populated area near Seoul for much of the last decade (Y. P. Kim & Lee, 2018). A wide range of chemical species and meteorological conditions were sampled extensively during KORUS-AQ, making the campaign suitable for model evaluation.

During KORUS-AQ, the NASA DC-8 and B200 King Air aircraft were joined by the Hanseo University King Air, measuring mostly in South Korean airspace on 22 days between 2 May and 10 June 2016. Four meteorological regimes were identified: a dynamic period from 1 to 16 May, stagnation between 17 and 22 May, low-level transport and haze development between 25 and 31 May, and a blocking pattern between 1 and 7 June (Peterson et al., 2019). Of relevance to chemistry-aerosol-cloud-climate interactions, the formation of new aerosol particles was observed frequently during the campaign (Eck et al., 2020; Lee et al., 2021). In this study we focus on a model evaluation with surface measurements and with two flights of the NASA DC-8 aircraft, on 11 and 26 May 2016. We focus on individual flights when clouds were present, rather than averaging over all the flights in the campaign, in order to examine the performance of the model in cloudy conditions and to test its ability to simulate pollution plumes, which could be averaged out in an aggregate evaluation.

Model comparisons in East Asia have a long history dating back to (at least) early work by Carmichael et al. (2001). KORUS-AQ has been a focus of modeling activity in the last few years, with notable recent papers documenting a dedicated emissions inventory (Jang et al., 2019) and a WRF-chem study in which the Korean Meteorological Administration's global UM system has been used to provide boundary conditions (Ha, 2022)

as well as other evaluations of WRF-chem (Saide et al., 2020). Park et al. (2021) compared simulations of the campaign period with four variants of WRF-chem and a version of CAM-chem, along with the chemical transport models CAMx, GEOS-chem and CMAQ. These models were found to have diverse strengths and weaknesses, but common features included moderate underestimations of CO, PM_1, and ozone concentrations, with most normalized mean biases (NMBs) between 0% and −40% overall. Here and later, NMB (e.g. Emery et al., 2017) is defined as

$$NMB = \frac{\sum_i^N (S_i - O_i)}{\sum_i^N O_i} \times 100\%$$

(1)

for a set of N simulated observations S_i and real observations O_i. The ensemble of models generally did well in capturing the spatial variability of the chemical pollutants measured at the surface across the Korean peninsula, with Pearson correlation coefficients for ozone, CO, NO_2, and SO_2 exceeding 0.5. However, for surface $PM_{2.5}$, even the model ensemble could not capture the variability (the Pearson coefficient is 0.17), and individual models likely perform worse than the ensemble. The low model skill is expected as, for example, localized plumes from industrial sources are difficult to capture and many of the monitors are located close to these sources. The system we present here is similar to GEOS-chem in that we run a global low-resolution simulation and a consistent nested high-resolution simulation, but then our prognostic meteorology is more like CAM-chem or WRF-chem than GEOS-chem.

Meteorological conditions during the KORUS-AQ aircraft campaign are described by Peterson et al. (2019). We focus our study on the first three regimes: dynamic, stagnation, and haze development. We first present an evaluation of a simulation of the period from 2 to 28 May 2016 against surface measurements from the AirKorea network of the National Institute of Environmental Research of South Korea (https://www.airkorea.or.kr/eng/, last access 21 Mar 2023). Next we show an evaluation of shorter simulations of periods of interest against measurements made by a range of instruments on the NASA DC-8 aircraft, and against MODIS satellite measurements.

We simulated one flight on 11 May (local time) in the latter half of the "dynamic" regime, which featured frequent frontal passages. A front, which had recently passed over Korea, is visible outside the south-eastern corner of our simulation domain on 11 May (see later, Figure 9). We expected some influence from a strong polar jet that passes over northern North Korea at this time. However, while 3–6 May was affected by dust plumes, 11 May saw little dust. This first case study flight was also less affected than the later part of the campaign by transported air pollution from China. Thus, we also show simulations of one flight on 26 May, in the "transport and haze development" regime strongly affected by pollution advected from China. This day also saw a frontal passage. Humid conditions were favorable for secondary aerosol formation and also led to low visibility (Peterson et al., 2019).

The AirKorea network monitors, at 320 locations across Korea, measure SO_2, ozone, NO, NO_2, PM_{10}, and $PM_{2.5}$. We evaluated our simulation of all of these quantities except for PM_{10}. These monitors provide data at hourly time resolution. SO_2 mixing ratios are measured by pulse ultraviolet fluorescence, CO by an infrared sensor, NO_2 by chemiluminescence and ozone by an ultraviolet method (H. S. Kim et al., 2019). PM concentrations are measured with the beta-ray attenuation method (BAM-1020) (Travis et al., 2022).

The MODIS satellite measurements we used are from Collection 6 Level 2 retrievals, at 1 km spatial resolution for cloud properties (Platnick et al., 2015) and 3 km spatial resolution for aerosol optical depth (AOD; Levy et al., 2015). The cloud droplet concentration was calculated following Gordon et al. (2018) and Grosvenor et al. (2018).

The aircraft (Crawford et al., 2021) measured CO with an in-situ diode laser spectrometer, ozone and NO_x with a four-channel chemiluminescence instrument operated by the National Center for Atmospheric Research, and HO_x with the Airborne Tropospheric Hydrogen Oxides laser-induced fluorescence instrument operated by Penn State University (Brune et al., 2021). Ozone photolysis rates were determined by a CCD actinic flux spectrometer. SO_2 was measured by a Chemical Ionization Mass Spectrometer operated by Georgia Institute of Technology while isoprene was measured by a proton transfer reaction mass spectrometer. Aerosols were measured with the TSI Laser Aerosol Spectrometer, condensation particle counters, and a scanning mobility particle sizer; these and the cloud particle spectrometer that measured cloud droplet concentrations all form part of the NASA Langley Aerosol Research Group suite of instruments. All of the aircraft data can be accessed through the NASA archive at https://doi.org/10.5067/Suborbital/KORUSAQ/DATA01.

4. Model Setup

For our case study of the KORUS-AQ campaign we set up the UM with a global simulation at N96 ($1.875° \times 1.25°$, ~135 km) horizontal resolution to drive a regional nested grid with 5 km resolution centered at 124.5° true east and 36° true north. A rotated pole grid was used in the regional model, with the north pole at 304.5° east, 54° north. The regional model has 360 grid cells in the longitude direction and 280 in the latitude, so it covers the north-eastern part of China as well as the whole of the Korean peninsula, as shown in Figure 3. The RAL 1 configuration was used (Bush et al., 2020) for the regional model and GA7.1 for the global (Walters et al., 2019). Because emissions from China are expected to be important to Korean air quality, simulating a large part of China at high resolution means that most of the emissions that are transported to Korea should also be simulated at this high resolution, instead of being smeared out to the resolution of the global model. As we do not focus this study on aerosol-cloud interactions, which were examined with a similar model by Gordon et al. (2020) at 500 m resolution, we judge that it is unnecessary to use a finer horizontal grid spacing than 5 km here. Our global model has 85 vertical model levels extending to a model top at 85 km altitude. The regional configuration we used here has 70 vertical levels with a model top at 40 km altitude, with 61 levels below 18 km, and 16 levels below 1,000 m.

Our simulation domain includes some relatively complex terrain. The western side of Korea lies within 400 m of sea level but the eastern part contains some areas higher than 1,000 m above sea level. High ground is also found in the north-western corner of our domain, over China, and in south-western Japan.

For our long simulation of 2–28 May 2016 we reinitialized the meteorology from high-resolution ($0.25° \times 0.18°$) global UM operational analyses every 2 days starting on 2 May 2016 at 0000 UTC and ending on 28 May 2016 at 1200 UTC, and we ran each forecast simulation for 60 hr. The chemistry and aerosols from a global atmosphere-only climate model simulation were merged into the first operational analysis file, following the procedure discussed in Section 2.2. The climate simulation had the same N96 resolution as the global simulation within the NUMAC run, and was nudged to ERA-interim reanalysis temperature and horizontal winds. The nudging ensures that the meteorology in this spin-up climate simulation is sufficiently consistent with the UM analyses for the aerosols in our case study to be realistic.

In this case study, we found that reinitializing our model every 2 days from the UM operational analyses, rather than more frequently, allowed meteorology to drift to some extent. For example, for our case study of the 11 May flight, we first tried initializing the simulations on 10 May 2016 and 0000 UTC, but found a small temperature bias in the boundary layer of about +2°C during the flight which caused the cloud cover to be biased low. When we initialized the simulation from reanalysis meteorology 12 hr later, on 10 May 2016 at 1200 UTC instead, the bias in cloud cover was substantially reduced, so adopted this for our short case study instead, to avoid complicating the evaluation of chemistry and aerosols along the flight path and to allow us to study effects of the cloud cover on the chemistry. We followed the same procedure for the flight on 26 May, initializing the model at 1200 UTC on 25 May. We did not repeat our simulation of 2–28 May with a higher frequency of initializing the model in order to keep our disk usage low, as each high-resolution UM initialization file we store occupies 23 GB. However, if a long simulation with no meteorological drift was needed, the large initialization files could be separately regridded to the N96 resolution of the NUMAC global model, and the resolution of the regional model, before the simulations are run. Then only the new initialization files, around 400 MB in size for the global model and around 1 GB for our regional model, would need to be stored.

We show a set of regional simulations that use emissions downscaled from the CMIP6 inventory. We also show regional simulations where the most important emissions in the CMIP6 inventory were replaced by 0.1°-resolution emissions from the KORUSv5 inventory (Jang et al., 2019). We used KORUSv5 emissions of anthropogenic primary black and organic carbon from fossil fuel and biofuel sources, and all anthropogenic sulfur dioxide, alkanes, CO, acetone and NO_x. We note that the chemistry mechanism we used does not represent alkanes larger than propane, alkenes other than isoprene, or aromatic compounds. Biomass burning emissions, and emissions of other chemical species, were still taken from CMIP6. The CMIP6 emissions have a native resolution 0.5°, but were regridded to the N96 (~1.5°) resolution of our global model before being again regridded to our regional model grid, in order that they could be used in the global model. Thus, we could have obtained more precise results with the CMIP6 inventory in our regional model if we did not use the same regridded CMIP6 emissions as input to our regional and global models. However, had we used native resolution CMIP6 emissions, the effects on species concentrations of the resolution of the emissions inventory and the effect of the resolution of the
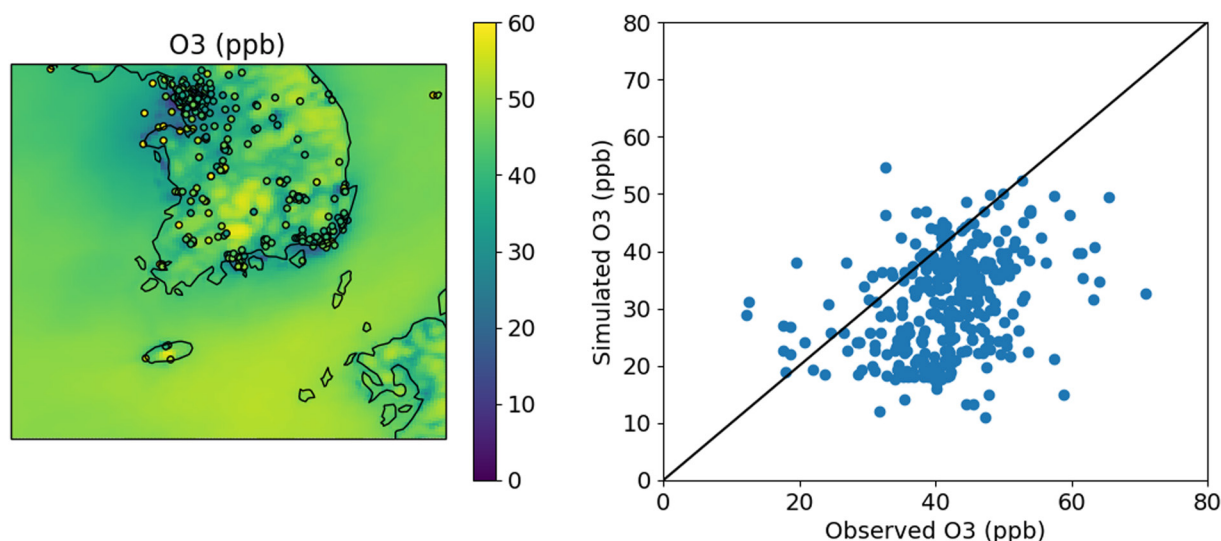
Figure 4. Simulated, time-averaged ozone spatial distribution using the KORUSv5 inventory, together with surface observations from the AirKorea monitors, and the correlation of time-averaged measured and observed ozone. The NO emissions are artificially scaled down by a factor 2.5 in the KORUSv5 inventory.

model would not be so easily separated. The KORUSv5 inventory was used in a multi-model intercomparison by Park et al. (2021). Emissions for East Asia are available at 0.1°-resolution, and specifically for South Korea at 3 km resolution. For simplicity we only used the 0.1°-resolution emissions in this study. When we used these KORUSv5 emissions, we also imposed a traffic diurnal cycle for anthropogenic (non-biomass-burning) black and organic carbon and NO_x, while in our regional and global simulations with CMIP6 emissions there is no such cycle, consistently with UKESM1. Simulation data sets presented in this paper are archived at Gordon et al. (2022).

5. Evaluation Against Surface Measurements

We evaluated simulated O_3, NO, NO_2, CO, SO_2, and $PM_{2.5}$ against surface measurements from the AirKorea network of 320 stations located as shown in Figure 4 below, using CMIP6 and KORUSv5 emissions in our regional model. The observations were compiled for the KORUS-AQ campaign into hourly means. We wrote out simulated concentrations from our model at the surface every 3 hr. While we did not simulate the entire campaign, the 26-day-long period we simulated is enough to make a reasonable comparison of our model with those studied by Park et al. (2021). We used a nearest-neighbor interpolation to produce simulated values of the air pollutants that are coincident in space and time with the observations.

Table 1 shows that the regional model produces low-biased average values of all observed pollutants when the CMIP6 emissions inventory at the resolution of our global model is used. The low bias is expected because the monitors are in general located close to emissions sources, which are completely smeared out in the CMIP6 inventory, as shown in Figure 3. The least biased vapor is ozone, which has a NMB of −14%, while the most is NO_2, with an NMB of −62%. Furthermore, there is almost no correlation between measured and predicted time-averaged values between monitors at different spatial locations, which is again not surprising given the low resolution of the emissions after regridding. However, the absolute magnitude of the emissions, while biased low, is never completely unrealistic.

With KORUSv5 emissions, the model severely overestimates surface NO concentrations, with an average NMB (calculated in the same way as the biases in Table 1) of 429%. The very high NO concentrations lead in turn to a

Table 1
Evaluation of Simulated Mean Concentrations Over 2 to 28 May 2016 Using AirKorea Observations

Pollutant	CMIP NMB (%)	CMIP R	KORUSv5 NMB (%)	KORUSv5 R
SO_2	−49	0.06	182	0.52
CO	−47	0.16	−16	0.27
O_3	−15	−0.01	−26	0.35
NO	−56	−0.01	45	0.07
NO_2	−62	0.10	−38	0.56
$PM_{2.5}$ (dry)	−39	0.06	−52	−0.05
$PM_{2.5}$ (ambient)	18	0.09	12	0.00

Note. First the time-mean of the observed and simulated data is taken, then the results are compared. CMIP and KORUSv5 columns denote regional model simulations with CMIP6 and KORUSv5 emissions inventories respectively. NMB refers to normalized mean bias, R to Pearson's correlation coefficient. This table is designed for comparison with Figure 3 of Park et al. (2021), so the R value here is a test of spatial correlation only, not temporal variability. We note that, as described in the text, the NO emissions are artificially scaled down by a factor 2.5 when the KORUSv5 inventory is used.

severe underestimate of ozone, as the simulated NO concentrations greatly exceed the ozone concentrations and therefore deplete ozone to make NO_2.

To illustrate this phenomenon explicitly, we show the timeseries of relevant gas mixing ratios in a single grid cell where the NO is overestimated, together with the median grid cell (where no anomalous behavior occurs) in Figure S13 in Supporting Information S1. The NO completely depletes the ozone concentration during the afternoon of May 5th, it remains zero during the following night, and then recovers next day when the NO concentration returns to a low value due to mixing or chemical losses. We find that if NO mixing ratios are sufficient to appreciably deplete ozone, HO_x production from ozone photolysis is suppressed (note the low OH mixing ratio in Figure S13 in Supporting Information S1) and therefore ozone production is also suppressed. In the real Korean atmosphere, as discussed in the context of KORUS-AQ by Simpson et al. (2020) and Oak et al. (2019), there are enough VOCs, and therefore peroxy radicals, to ensure this does not happen: as in the well-known HO_x cycle (e.g. Seinfeld & Pandis, 2008), they convert NO to NO_2 without ozone, resulting in ozone production when the NO_2 is photolyzed back to NO. However, in the chemical mechanism we used, the peroxy radical concentration is well known to be biased low, especially in East Asia (Archer-Nicholls et al., 2021) because there are no RO_2 species (or oxidation products that could then make RO_2) produced by anything heavier than isoprene, and so NO builds up. Schroeder et al. (2020) and Simpson et al. (2020) calculated that isoprene typically contributed only 20% of the ozone production during KORUS-AQ while aromatic species contributed almost half. We could replicate the erroneous NO build-up and ozone depletion we saw in box model simulations with the same chemical mechanism (not shown). Increasing the peroxy radical concentration in the box model by the crude expedient of adding artificially high levels of isoprene brought the NO down and prevented the ozone from being depleted. We conclude that the lack of sources of peroxy radicals in our simulation is most likely responsible for its severe overestimate of the surface NO mixing ratio.

We found that over the Korean peninsula specifically, total emissions of NO in the KORUSv5 inventory exceed those in the CMIP6 inventory by a factor of 2.5. The models presented by Park et al. (2021) do not overestimate NO when the KORUSv5 emissions are used, so the KORUSv5 emissions are unlikely to be substantially biased. However, to fix our model in the short term, we scaled down the KORUSv5 NO emissions by this factor 2.5 and found a substantial improvement in the simulated NO concentrations. Concentrations before and after this artificial adjustment are shown in Figure S11 in Supporting Information S1.

When the KORUSv5 emissions inventory is used and the NO emissions are adjusted, some positive spatial correlation between measured and simulated time-averaged concentrations at the various monitors emerges, except in the case of $PM_{2.5}$ and NO. Ozone is shown in Figure 4. However, the highest correlation, for NO_2, has an R value of only 0.56. The generally poor correlations may reflect the failure of both the 0.1°-resolution emissions inventory and the 5 km resolution model to resolve intra-urban variability, since a high fraction of the monitors are concentrated in cities. However, it may also be a symptom of the missing peroxy radicals.

When we average over all the monitors, we obtain a time series of pollutant concentrations that we can compare to observations. Ozone, SO_2, and $PM_{2.5}$ are shown in Figures 5–7 while CO, NO, and NO_2 are shown in Figures S5, S11, and S12 in Supporting Information S1. For ozone, the diurnal cycle and trend across the simulation period in concentrations are reproduced by the model well: the Pearson's R value for the correlation of the simulated and observed timeseries is 0.87 when CMIP6 emissions are used and 0.81 when KORUSv5 emissions are used. CO concentrations (Figure S5 in Supporting Information S1), like ozone, are underestimated but the diurnal and intra-month variability is realistic. Simulated NO has rather too much variability and some biases in the representation of the diurnal cycle that should be further investigated along with the initial overestimate of the concentrations before the emissions are adjusted. The variability in NO_2 (Figure S12 in Supporting Information S1) is better, at least some of the time, with the correlation between simulated and observed timeseries having an overall Pearson's R value of 0.51.

The mean SO_2 concentration is biased high in simulations with the KORUSv5 inventory, almost by a factor of 3 (Table 1). The variability between monitors is also much higher in the simulations than in the observations. However, the median SO_2 concentration (Figure 6) follows observations much more closely, suggesting that a small number of monitors, likely those close to emissions sources, are responsible for the high bias. In order to reduce biases in the climate model discussed by Mulcahy et al. (2018), all SO_2 in the model is emitted at the surface, so much of it is quite rapidly lost close to emissions sources. The variable agreement with the aircraft measurements discussed below also tentatively supports this hypothesis (SO_2 is underestimated during a flight
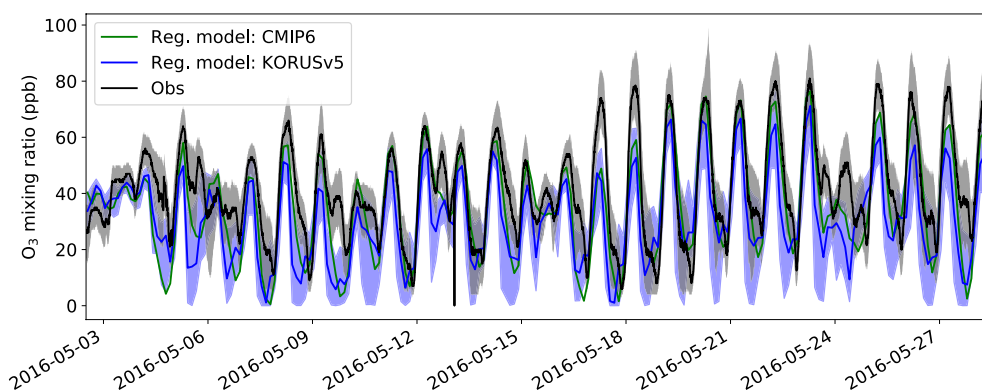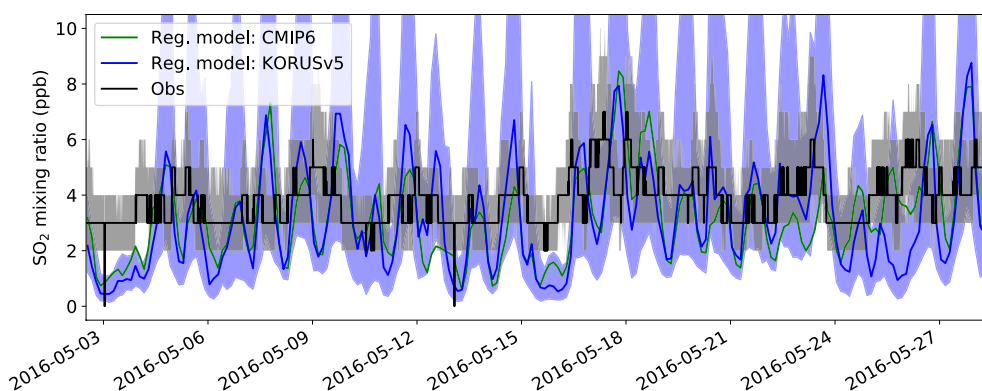
Figure 5. Observed surface ozone mixing ratios compared to ozone simulated with the CMIP6 and KORUSv5 inventories. At each time, the median simulated and observed ozone mixing ratio across the AirKorea measurement stations is plotted. Shading shows the interquartile ranges for the observations and for the simulation with the KORUSv5 inventory only. The range in the CMIP6 simulation is not shown, for clarity. The NO emissions are artificially scaled down by a factor 2.5 in the KORUSv5 inventory.

mainly over sea, on 11 May, and overestimated during a flight over land on 26 May), and so the choice of emission height should be revisited in future.

We show simulated $PM_{2.5}$ for both dry and humidified aerosols in Figure 7. As in other studies that use these data (e.g. Travis et al., 2022), we assume the observed particulate matter concentrations correspond to dried particles. The timeseries of dry $PM_{2.5}$ simulated with CMIP6 emissions, which has a Pearson's R value versus observations of 0.74 and a bias of −39%, is clearly in better agreement than the timeseries simulated with KORUSv5 emissions (Pearson's R of 0.63 and bias of −52%). However, neither simulation sustains the increase in $PM_{2.5}$ after 17 May during the stagnation period (Peterson et al., 2019), nor after 25 May during the haze development period. More work is required to improve the aerosol representation in the model, especially relating to SOA and nitrate, as we discuss later. The lack of nitrate may inhibit the accumulation of aerosol water from feedbacks discussed by Jordan et al. (2020). Our simulated humidified $PM_{2.5}$ has a stronger diurnal cycle and some interesting peaks in concentration which may be indicative of fog or haze, and would also be worth further investigation in this context.

6. Evaluation Against Aircraft and Satellite Measurements

Model results are compared to aircraft measurements by first creating 20-s averages of the observations, and then using a nearest-neighbor interpolation to find the model grid cell the aircraft was flying through at the time.

Figure 6. Observed SO_2 mixing ratios compared to simulations using the CMIP6 and KORUSv5 inventories. At each time, the median simulated and observed SO_2 mixing ratio across the AirKorea measurement stations is plotted. Shading shows the interquartile ranges for the observations and for the simulation with the KORUSv5 inventory only. The range in the CMIP6 simulation is not shown, for clarity.
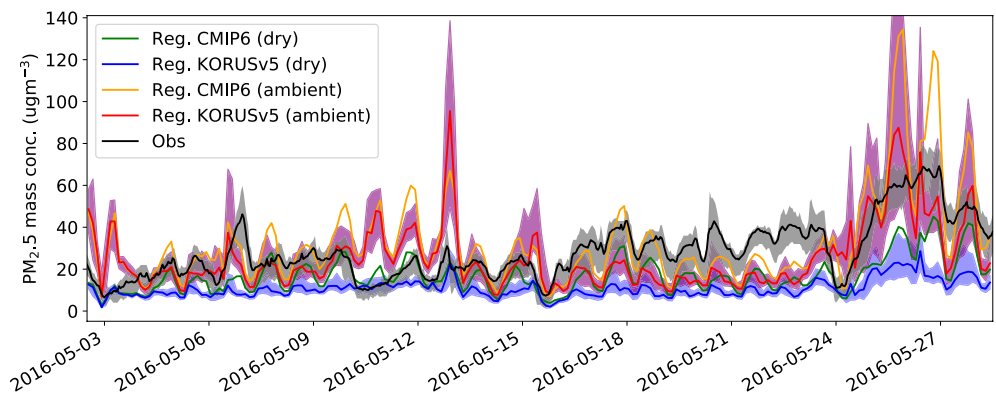
Figure 7. Observed dry and ambient (humidified) $PM_{2.5}$ mass concentrations compared to simulations using the CMIP6 and KORUSv5 inventories. At each time, the median simulated and observed $PM_{2.5}$ mass concentration across the AirKorea measurement stations is plotted. Shading shows the interquartile ranges for the observations and for the simulation with the KORUSv5 inventory only. The range in the CMIP6 simulation is not shown, for clarity.

To perform the interpolation we used Python code from the Community Intercomparison Suite (Watson-Parris et al., 2016). When flying horizontally, most likely the same regional model grid box will be sampled multiple times as the aircraft takes more than 20 s to travel 5 km, but when climbing or descending the aircraft can cross more than one grid box in the vertical (∼200 m at 2,000 m altitude) in 20 s.

6.1. Meteorology and Clouds

The temperature and relative humidity (RH) along the aircraft track agree well with the observations on 11 May, approximately 12 hr after the simulated meteorology is reinitialized, as shown in Figure 8. The RH on 26 May is biased low by around 20% between about 500 m and 2 km altitude in both the regional and global models, suggesting a minor error in the simulated meteorology. This is discussed below in the context of the simulated clouds. The results do not depend on the emissions inventory or the cloud microphysics scheme, so only our regional simulation with KORUSv5 emissions is shown.

Figure 9 shows MODIS total water path (liquid + ice), low cloud top height, and cloud droplet concentration in warm clouds only, compared to the corresponding variables from the global model, and from the regional model with both single- and two-moment microphysics schemes. The spatial pattern of the cloud is well reproduced by the model on this day, but the cloud water path to the east of Korea is overestimated by the model by at least a factor of two. While the temperature and RH along the aircraft track agree well with the observations on 11 May (Figure 8), the discrepancy between model and satellite is mostly to the east of the path of the aircraft, so it is possible that this is still due to a bias in representing these variables. Alternatively, the sub-grid cloud fraction scheme, the relatively simple scheme of Smith (1990), may be at fault. The lowest sub-figure shows that cloud-related biases are likely to affect our aircraft evaluation sporadically through the flight.

The cloud top height is reasonably well represented by the model, except near Beijing where high clouds are not simulated by the model. The generally good representation of the cloud top height of low clouds, and the evaluation of RH against the aircraft, suggests the boundary layer height is also likely to be close to observations on 11 May. The cloud droplet concentration in the simulation with the single-moment Wilson and Ballard (1999) cloud microphysics is mostly overestimated by the model compared to MODIS, by up to a factor of around three, but not so obviously overestimated compared to the aircraft measurements (lowest subfigure of Figure 9), which are admittedly rather sparse. The aircraft samples few, broken clouds, and the droplet concentration in these is not homogeneous. Either the aircraft samples only the areas where the model and observations happen to agree well, or the MODIS retrieval is unreliable.

The differences in cloud liquid water content and cloud cover between simulations with single and double-moment microphysics are relatively small, but the two-moment CASIM scheme produces a significantly lower droplet concentration, about a factor of two lower, and is in better agreement with MODIS observations on this day but worse agreement with the aircraft. The lower concentration in CASIM relative to the Wilson and Ballard (1999)
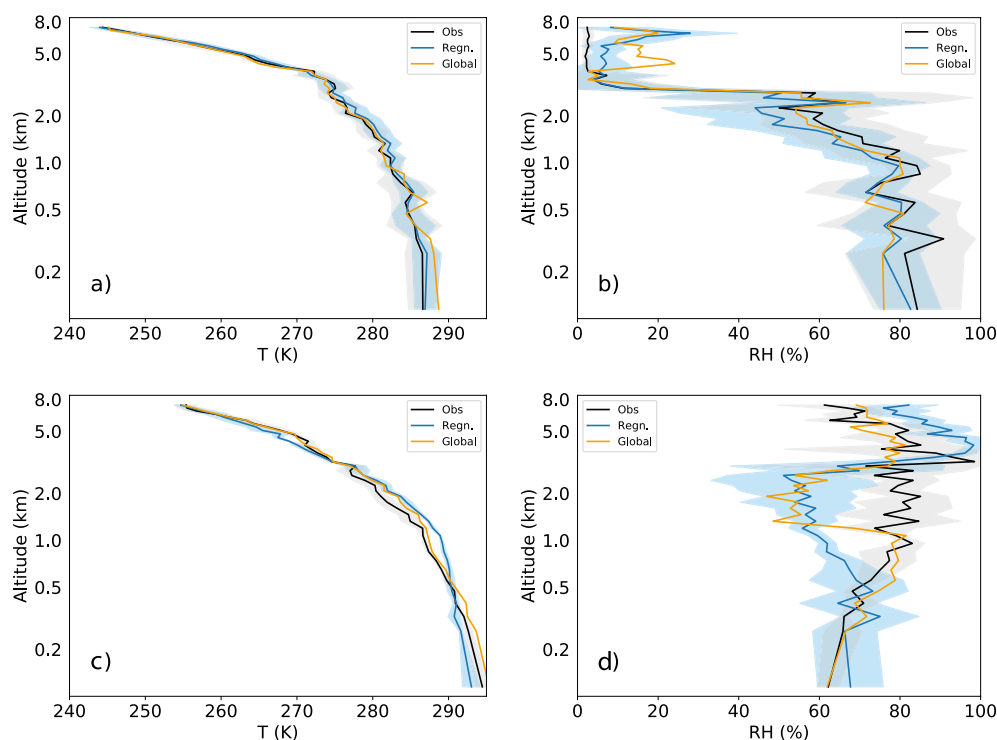
Figure 8. Vertical profiles of temperature and relative humidity (RH) compared between the model and the aircraft measurements on (a and b) 11 and (c and d) 26 May 2016. The simulated temperature and RH are interpolated onto the path of the aircraft. The means of these values and of the observations over the points sampled by the aircraft at the altitude in question are shown as solid lines and the ± 1 standard deviation interval is shaded for the observations and regional model only.

scheme is the opposite of what was observed by Gordon et al. (2020), but the relative behavior of the droplet concentration in the two schemes is known to depend on the spatial resolution of the simulations. Further dedicated efforts are required to finalize a more scale-invariant treatment of the droplet concentration; for now the resolution-dependence must be addressed by tuning following Gordon et al. (2020), which is not in the scope of this paper.

Simulated clouds on 26 May are evaluated in Figure S2 in Supporting Information S1. While the simulation of cloud water path is in generally good agreement with the satellite retrievals, there is a clear patch in simulations in the south of Korea that coincides with the path of the aircraft. The comparison of the cloud top height indicates that there is an underestimate of high cloud cover in the simulations, especially in the global model. This lack of high cloud is likely to influence simulated photolysis rates. As we focus this paper on regional simulations we do not investigate the bias in the global model further. The same trends in droplet concentration as on 11 May are apparent from the satellite image. We did not run a simulation with CASIM microphysics on this day, but we do compare simulations with CMIP6 and KORUSv5 emissions. The emissions inventory used does not strongly affect the droplet concentration.

6.2. Chemistry

A wide range of chemical species were measured by the aircraft during KORUS-AQ. We show simulations of the most important in this section, with more detailed evaluation in figures in Supporting Information S1. Table S2 in Supporting Information S1 provides summary statistics for all variables that are evaluated against aircraft measurements on both dates.

Figure 10 shows the timeseries of simulated CO interpolated onto the path of the aircraft and compared to measurements from the diode laser spectrometer on 11 May 2016. The surface simulated concentrations of CO are also compared between regional and global models at a representative time during the flight, and
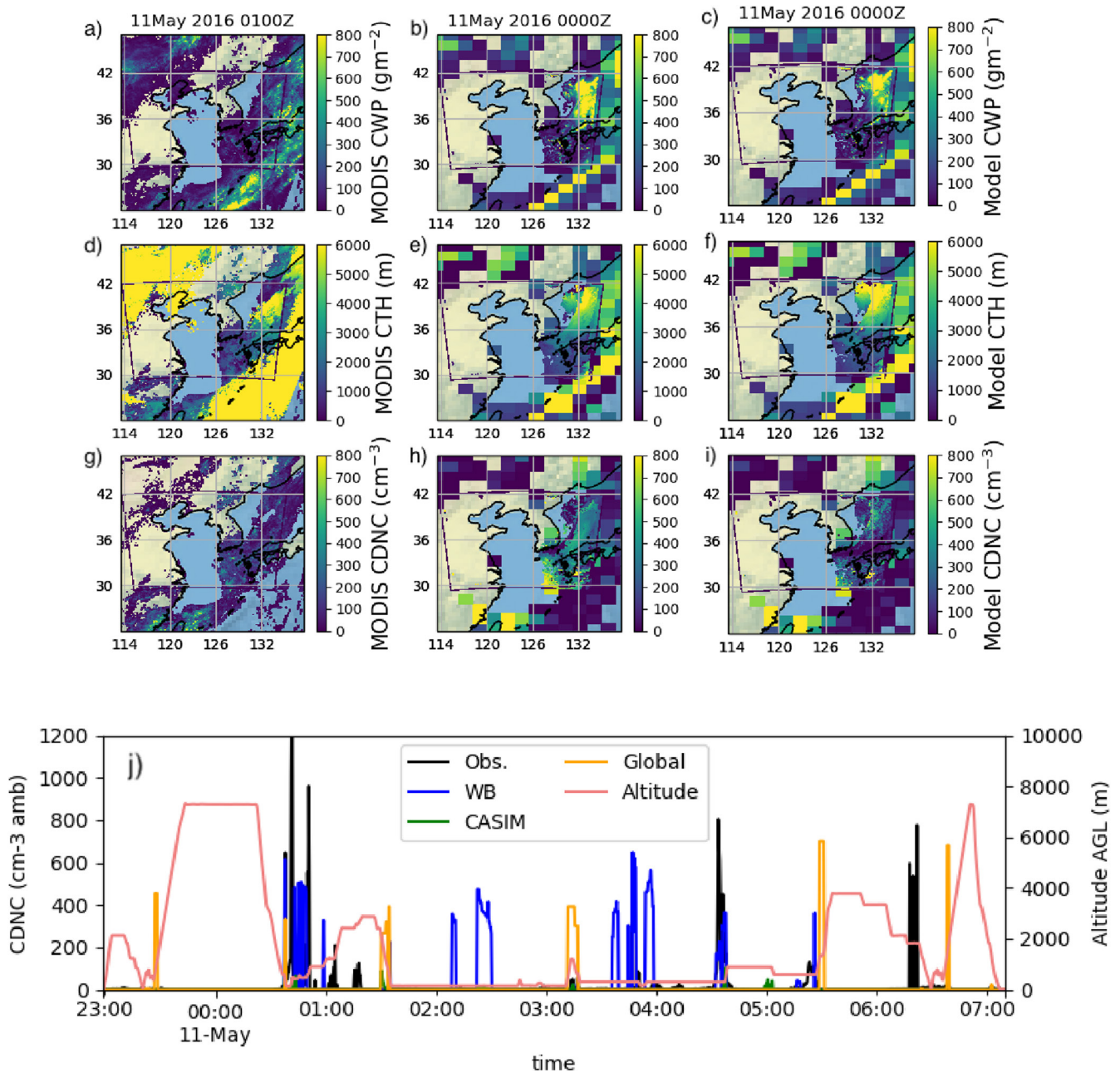
Figure 9. Cloud liquid water path (a–c), cloud top height (d–f), and droplet concentration (g–i) on 11 May 2016 at 0000 UTC (0900 local time), in MODIS satellite data (a, d, g) and the Unified Model global and regional simulations with single (b, e, h) and double-moment cloud microphysics (c, f, i). The global model grid cells are plotted underneath those of the regional model, where cloud is present. Subfigure (j) shows the timeseries of droplet concentration measured by the aircraft compared to the droplet concentration predicted by the two regional models and the global model. In the legend, WB refers to the single-moment scheme of Wilson and Ballard ([1999]). The droplet concentration predicted by the Cloud AeroSol Interacting Microphysics scheme is very low and hardly visible.

the flight path for the 11 May flight is shown on the same subfigure. Some spatial variability is evident with higher CO concentrations in the heavily urbanized Seoul region. Simulated and observed vertical profiles of these interpolated CO concentrations, averaged over this flight and the flight on 26 May, are given in Figure S6 in Supporting Information S1. The area covered by the 26 May flight is shown on Figure S9 in Supporting Information S1. The simulated CO is realistic, but underestimated by both the regional and global model by a few tens of percent most of the time, and up to around a factor of two. Overall, the NMB of the simulation with KORUSv5 emissions, relative to the observations, is −26% on 11 May and −33% on 26 May. Underestimation of CO concentrations is a common feature of models, and has been studied extensively in the context
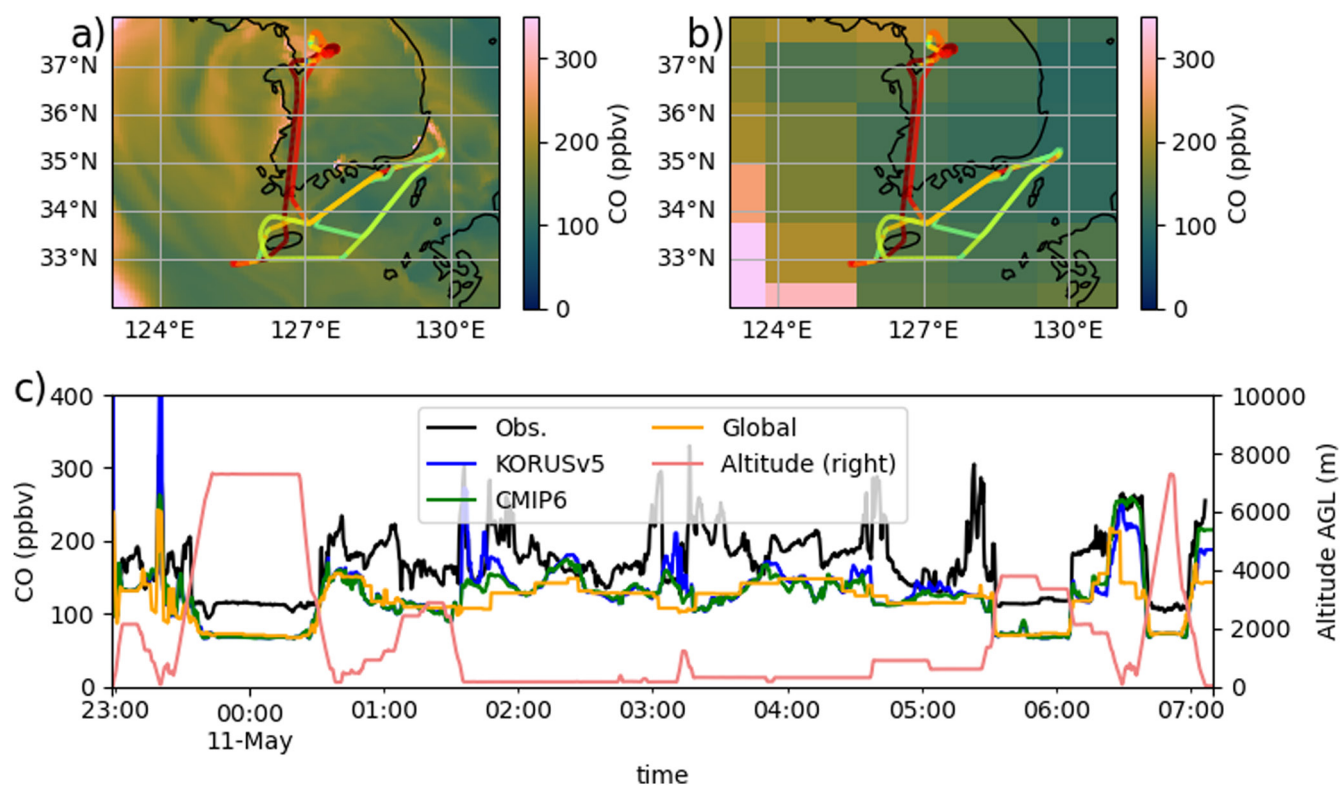
Figure 10. Simulated and observed carbon monoxide concentrations on 11 May 2016. Subfigures (a and b) show the regional model with KORUSv5 emissions and the global model approximately 50 m above the surface at 0300 UTC (1200 local time). Arrows show wind direction 50 m above the surface; their size is qualitatively proportional to the wind strength. Subfigure (c) shows the timeseries simulated by the global and regional models with the two emissions inventories (as indicated in the legend), compared to measurements from the in situ diode laser spectrometer (DACOM) on the aircraft.

of KORUS-AQ (Gaubert et al., 2020; Park et al., 2021). While we cannot draw conclusions on the origin of CO emissions from our own study, during KORUS-AQ, the underestimation of CO is thought to be mainly associated with underestimated emissions in China, as discussed by Gaubert et al. (2020) in a CAM-chem study with the same KORUSv5 emissions data set. On 11 May, the underestimate of CO concentrations may also be partly due to the lack of real-time shipping emissions in the model, since the spikes in observed CO appear to be mainly in isolated parts of the flight track over the ocean. Compared to the spatial correlation at the surface measurement sites, the Pearson correlation coefficient relating the simulation and the observations along the path of the aircraft is relatively high on both 11 and 26 May at around 0.8, which likely reflects the fact that the vertical dependence of CO concentrations is well captured by the model (Figure S6 in Supporting Information S1).

The concentrations of other species are shown in the same format as Figure S10 in Supporting Information S1. The variability in the concentrations is shown in Figure 11 for the portion of the flight on 11 May when the aircraft flew close to the surface. Summary statistics for the same period are reported in Table 2. A similar analysis for the period when the aircraft was continuously above 1,500 m altitude on 26 May is shown in Figure S1 and Table S1 in Supporting Information S1, and summary statistics for the complete flights are shown in Table S2 in Supporting Information S1.

On 11 May, generally, Table 2 indicates that the mean concentrations show mixed agreement. The global model underestimates concentrations of CO by 29%, O_3 by 23%, SO_2 by 79%, and NO by 66% while it simulates OH accurately (within 20%). Table S2 in Supporting Information S1 shows that biases in CO and O_3 over the two flights we simulated are comparable to the campaign-average biases for the models shown in Table 6 of Park et al. (2021): their ensemble mean biases are −27% for CO and −16% for ozone, while our regional simulation with KORUSv5 emissions is biased by an average of −30% for CO and −6% for ozone. The larger biases in NO and SO_2 are discussed below.
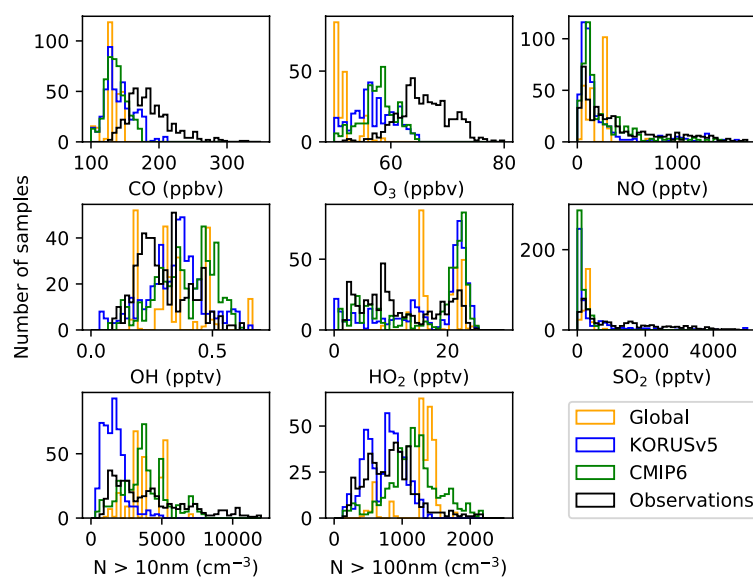
Figure 11. Frequency distributions of gas mixing ratios and aerosol concentrations on 11 May 2016 between 0145 and 0430 UTC (1045–1330 local time), when the aircraft was flying within 500 m of the surface. Global simulations, and regional simulations with the two emissions inventories (as indicated in the legend) are shown. All simulated variables are interpolated along the flight track of the aircraft at 60 s intervals (which is comparable to the time taken for the aircraft to cross a regional model grid cell in the boundary layer when flying straight and level). Because this means most model grid cells in the global model were sampled multiple times, leading to high peaks in the histogram, frequencies in the global model are scaled down by a factor 2 to ensure the regional model and observation data are visible.

The resolution of the emissions inventory is important. Variability in CO, NO, and SO_2 concentrations on 11 May is underestimated, usually by more than a factor 2, when CMIP6 emissions are used, but is much better represented when KORUSv5 emissions are used (with NO and SO_2 having simulated standard deviations within 10% of observations). Curiously, however, using the KORUSv5 inventory reduces variability in aerosol number concentrations, in poor agreement with observations. This result should be revisited in future work with a more sophisticated representation of aerosols, including nitrate following A. C. Jones et al. (2021).

On 26 May, at higher altitude, the agreement of the model with observations is generally better. The simulated concentrations of most species are within 20% of observations, except in the case of SO_2, which is greatly over-estimated (by 60% in the global model and a factor 2.3 in the regional model) when CMIP6 emissions are used. There is a corresponding bias in total aerosol number concentration likely due to excessive new particle formation (NPF). However, in contrast to our findings at the surface described in Section 5, SO_2 and particle concentrations are simulated better when KORUSv5 emissions are used (e.g., SO_2 is biased low by around 40%).

Ozone concentrations on 11 May 2016 are shown in Figure S7 in Supporting Information S1. At low altitude, the model underestimates ozone, while at high altitude it overestimates it. We speculate that biased concentrations of VOCs, and thus peroxy radicals, may be responsible for the bias in ozone. Near the surface along the western coast of Korea, ozone concentrations are anomalously low at near 30 ppbv, likely due to high NO concentrations in this area. When the NO concentration exceeds both the ozone and NO_2 concentration, NO likely depletes ozone more quickly than it is produced via NO_2 photolysis.

The OH radical concentrations are evaluated in Figures S9 and S10 in Supporting Information S1. Both the global model and regional model OH concentrations on 11 May are in good agreement with the measurements (and we note that the uncertainty in these measurements, while expected to be small compared to the model biases, is likely not negligible (Brune et al., 2021)). The regional model shows some small overestimates but reproduces the variability in the measurements better, as expected. However, both models overestimate OH on 26 May, although the regional model performs substantially better than the global model. The NMB in the global model on 26 May is 55% while that in the regional model is close to 30% for simulations with both emissions inventories. The poorer performance in the global model is likely due to biases in the cloud cover, which leads to an overestimated ozone photolysis rate at low altitudes as shown in Figure S8c in Supporting Information S1. Overall, the effect of

Table 2
Means and Standard Deviations of the Concentrations of Species Plotted in Figure 11, for the Flight on 11 May 2016 Between 0145 and 0430 UTC

	Obs. mean	Glob. mean	CMIP6 mean	KORUSv5 mean
CO (ppbv)	186.04	132.10	136.97	141.59
O_3 (ppbv)	65.98	50.67	54.06	53.06
NO (pptv)	559.76	188.85	253.85	373.08
OH (pptv)	0.30	0.35	0.38	0.34
HO_2 (pptv)	11.79	18.52	16.40	16.45
SO_2 (pptv)	1,340.09	287.64	143.00	657.66
$N > 10$ nm (cm^{-3})	4,913.16	3,880.61	3,766.33	1,605.12
$N > 100$ nm (cm^{-3})	826.93	1,268.45	1,166.16	745.03

	Obs. std	Glob. std	CMIP6 std	KORUSv5 std
CO (ppbv)	31.42	12.57	15.41	20.92
O_3 (ppbv)	4.45	2.48	7.20	7.35
NO (pptv)	769.99	99.66	276.16	744.96
OH (pptv)	0.11	0.13	0.12	0.12
HO_2 (pptv)	6.72	3.55	7.72	7.64
SO_2 (pptv)	1,477.89	116.43	133.79	1,541.37
$N > 10$ nm (cm^{-3})	3,862.30	1,217.36	1,608.02	787.77
$N > 100$ nm (cm^{-3})	360.17	285.84	382.08	270.35

Note. Comparing the means indicates the bias in the concentration while comparing the standard deviations measures the models' ability to represent variability. In the table, std refers to standard deviation, "Glob." to the global simulation, "CMIP6" to the regional simulation with CMIP6 emissions, "KORUSv5" to the regional simulation with KORUSv5 emissions, N to aerosol number concentration greater than the diameter threshold specified, and "stp" to standard temperature and pressure.

the clouds on photolysis is clearly simulated more realistically by the regional model than the global model, and this demonstrates the potential of the regional model to simulate chemistry-cloud interactions.

The sulfur dioxide concentrations are shown in Figures S3 and S4 in Supporting Information S1. Above the boundary layer on 11 May, the global models overestimates the SO_2 concentrations when CMIP6 emissions are used, while the regional models underestimate SO_2 but are in better agreement. However, in the boundary layer the agreement is poorer, leading to a low bias of around a factor 2 in the simulation with KORUSv5 emissions on 11 May apparent in Table 2. SO_2 in the boundary layer is underestimated on 26 May by the regional simulation with CMIP6 emissions (over all altitudes, the NMB is around −13%) and overestimated (NMB = +50%) with KORUSv5 emissions. These biases are likely due to imperfect shipping emissions or to biased representations of atmospheric processes. For example, the pH of cloud droplets regulates SO_2 concentrations, and in our model it is set to a global constant value of 5. This value is being revisited in other studies currently in progress, following Turnock et al. (2019).

Nitrogen monoxide concentrations are shown in Figures S14 and S15 in Supporting Information S1. Close to the surface, the NO concentrations are underestimated on 11 May, which may again be the result of poorly represented temporal variability in shipping emissions, while above the surface the concentrations are well represented. On 26 May, the mean concentration is underestimated by around 60% in simulations with CMIP6 emissions and overestimated by simulations with KORUSv5 emissions, even after these emissions are scaled down. However, all simulations are realistic. On 26 May the Pearson's R values show very high correlation between model and observations, at 0.97 for the simulation with KORUSv5 emissions, while on 11 May this correlation coefficient is lower at 0.45.

One aim of our detailed comparison of the model with data from individual flights was to test the model's ability to resolve pollution plumes. The skill of the model in this respect is mixed—for example, Figure S14 in Supporting Information S1 suggests about half of the spikes in NO mixing ratio observed on 11 May are captured—suggesting that higher time-resolution emissions or more precise diurnal cycles might be needed for this to work well.

Isoprene concentrations are evaluated in Figures S16 and S17 in Supporting Information S1. The isoprene concentrations in the flight of 26 May 2011 are very inhomogeneous because the lifetime of this gas is short and it is emitted only over land. The failure of the global model to resolve the Korean peninsula leads to spuriously high isoprene concentrations over sea in the global model, and a very high NMB of around +400% on 11 May. The regional model is in agreement with the observations close to the surface, which suggests the emissions parameterization is working well, but the concentrations of isoprene at higher altitudes are underestimated, by at least a factor 10. This bias may be at least partially because OH concentrations are slightly overestimated on this day (see Figure S9 in Supporting Information S1).

6.3. Aerosols

Number concentrations of aerosols greater than 100 nm in diameter are shown in Figures 12 and 13. Figures S18 and S19 in Supporting Information S1 show particles greater than 10 nm in diameter. The number concentration of particles with at least 100 nm diameter in the regional model is overestimated by between 30% and 55% when CMIP6 emissions are used on both 11 and 26 May. We are not aware of other model evaluations of aerosol number concentrations using KORUS-AQ data, although the study of optical properties of Saide et al. (2020) considers simulated size distributions. However, overall the performance is acceptable compared to other studies with the same aerosol microphysics scheme (Gordon et al., 2020; Ranjithkumar et al., 2021), or other models. For example, four global aerosol microphysics models compared to data from the ATom field campaign in remote regions by Williamson et al. (2019) all had biases of at least a factor 2 in their representation of aerosols of at least 60 nm diameter.

In our model with CMIP6 emissions, the number concentration of aerosols above 10 nm in diameter is simulated fairly well (over the whole flight, it is underestimated by 25% on 11 May from Table S2 in Supporting Information S1), but with KORUSv5 emissions, these number concentrations, dominated by the Aitken and nucleation modes, are more strongly underestimated on 11 May by 57%. This underestimate may be due to omission of the important contribution of ammonia, organic molecules, or amines in the NPF parameterization we use (Dunne et al., 2016; McMurry et al., 1983; R. J. Weber et al., 1998). On 26 May, the model performs better, within 11%
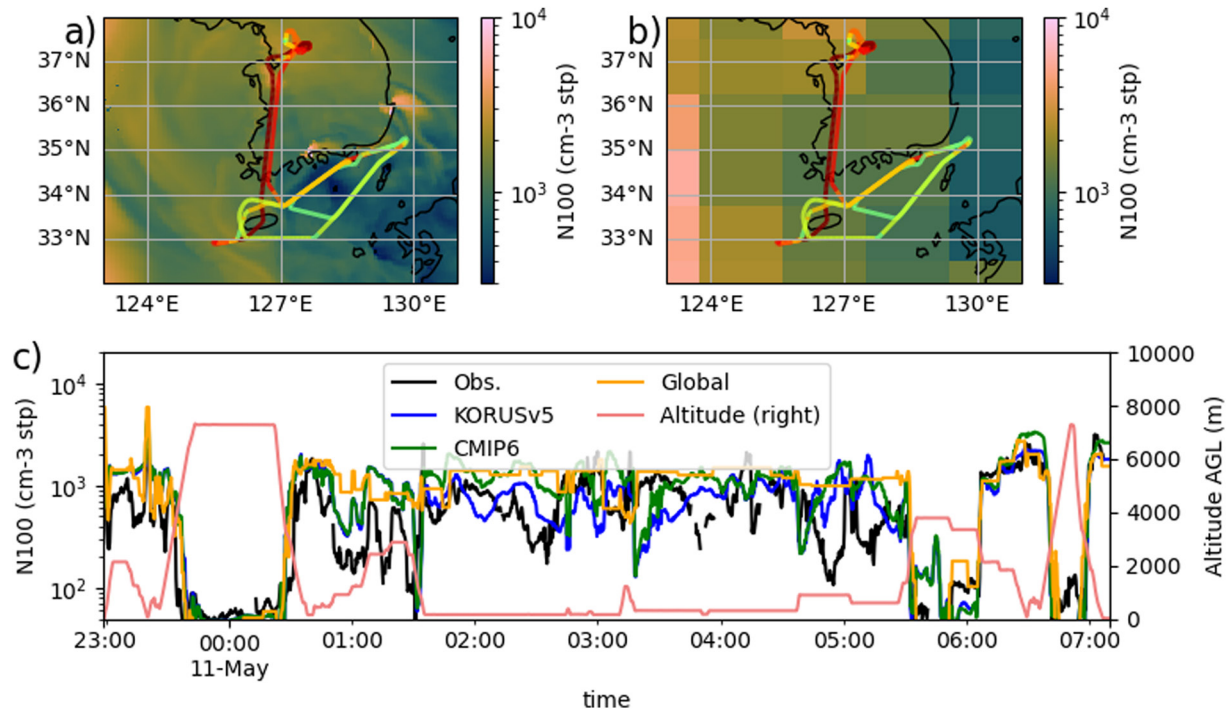
Figure 12. Particle number concentrations at standard temperature and pressure on 11 May 2016, greater than 100 nm diameter measured by the TSI 3340 Laser Aerosol Spectrometer. Subfigures (a and b) show the regional and global model approximately 50 m above the surface at 0300 UTC (1200 local time), while subfigure (c) shows the timeseries measured by the aircraft, as in Figure 10.

of observations. This improved agreement may be because high condensation sinks on this day suppress NPF. We intend to include more detailed NPF mechanisms in future studies.

The full dry aerosol size distribution is compared with observations for a period during the 11 May flight in Figure 14. During this period, the aircraft was flying at low altitude over the sea. As expected from the evaluation of number concentrations, the model underestimates the aerosol number concentration at small sizes, but performs well for larger sizes and captures the slightly bimodal size distribution via its Aitken and accumulation size modes. Number concentrations of aerosols larger than 1 µm in diameter are overestimated, primarily by dust from the six-bin treatment: the model predicts 5.1 cm^{-3} while only 1.0 cm^{-3} were observed during the period used in Figure 14. Representing dust aerosols should be revisited in future studies. Concentrations of dust are relatively low, and thus only four bins contain non-zero values during this time.
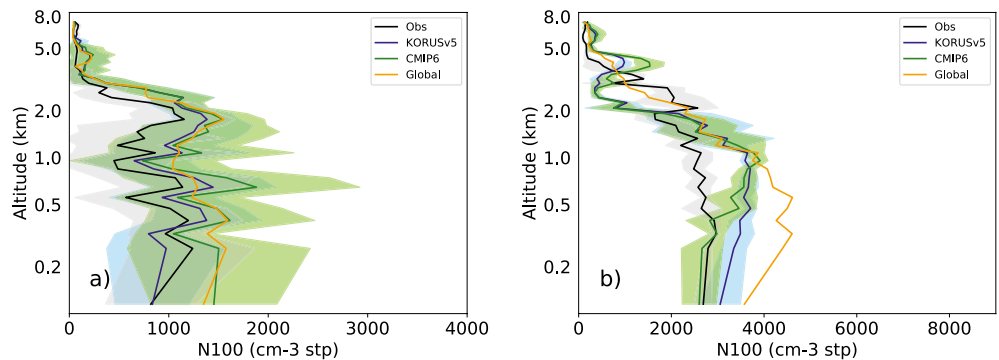
Figure 13. Vertical profiles of concentrations of dried particles greater than 100 nm in diameter on 11 May (a) and 26 (b) May 2016. The data are presented as means over the flight at standard temperature and pressure, with the interval corresponding to ±1 standard deviation shaded. Note the different scales on the x axes.
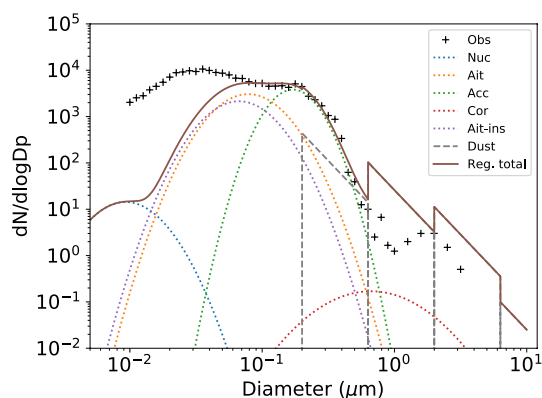
Figure 14. Dry aerosol number size distribution at standard temperature and pressure on 11 May 2016 during a representative period in the flight, 0620–0640 UTC, observed and simulated using the KORUSv5 emissions. The observations are taken from the Scanning Mobility Particle Sizer for diameters below 200 nm and the TSI 3340 Laser Aerosol Spectrometer for larger sizes. Simulated size distribution in the nucleation, Aitken, accumulation, and coarse soluble size modes, and in the dust bins, are shown with dotted or dashed lines, while the total simulated size distribution is shown with a solid line.

The AOD is compared to MODIS satellite observations (Collection 6 dark target, Level 2, 3 km pixel size, at 0.55 µm wavelength) in Figure 15. MODIS is unable to retrieve the AOD through clouds and therefore we choose two different days for this comparison: 12 and 19 May 2016. Many of the AOD trends are qualitatively well captured by the model, for example, the lower AOD over Japan on 12 May and the increase in AOD from south to north along the Chinese mainland. We regridded the AOD from MODIS so that it could be compared directly with the model; over the area plotted, where retrievals exist, the model overestimates the AOD, with a NMB of +26% on this day. The spatial patterns are captured to some extent, with a Pearson's *R* value of 0.57. On 19 May the AOD is in better agreement with observations, with a NMB of −12% and a Pearson's *R* value of 0.69. We conclude that the simulation of AOD is realistic, though with room for improvement.

## 7. Demonstration of NUMAC in Other Regional Case Studies

We are applying NUMAC in several other locations to study atmospheric composition and aerosol-cloud interactions. Figure 16 demonstrates how square and rectangular nested domains can be simulated at convection-permitting or cloud-resolving resolution for research in both pristine and polluted environments.

We are investigating how aerosols and precursor vapors are transported by resolved deep convective updrafts in the Amazon rainforest. The simulations are able to resolve the influence of the deep convective clouds on aerosol concentrations in the upper troposphere, as shown in Figures 16a and 16c. In the regional simulation (c), whose location is marked on the simulation from the global driving model (a), nucleation mode (3–10 nm diameter) aerosols at around 12 km altitude are clearly removed by the deep convective clouds, and, while the patterns are
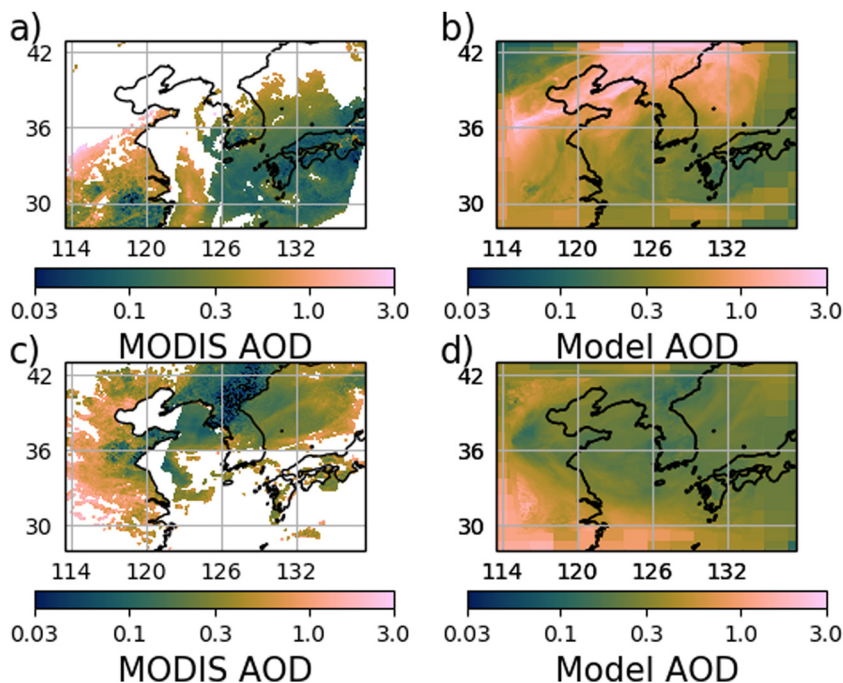
Figure 15. Aerosol optical depth on (a and b) 12 May 2016 and (c and d) 19 May 2016 in MODIS TERRA data and in the model at 0100 UTC (1000 local time). Satellite retrievals are only plotted where they exist, that is, in cloud-free areas. In subfigures (b and d), data from the regional model, with KORUSv5 emissions, is plotted on top of data from the global model.
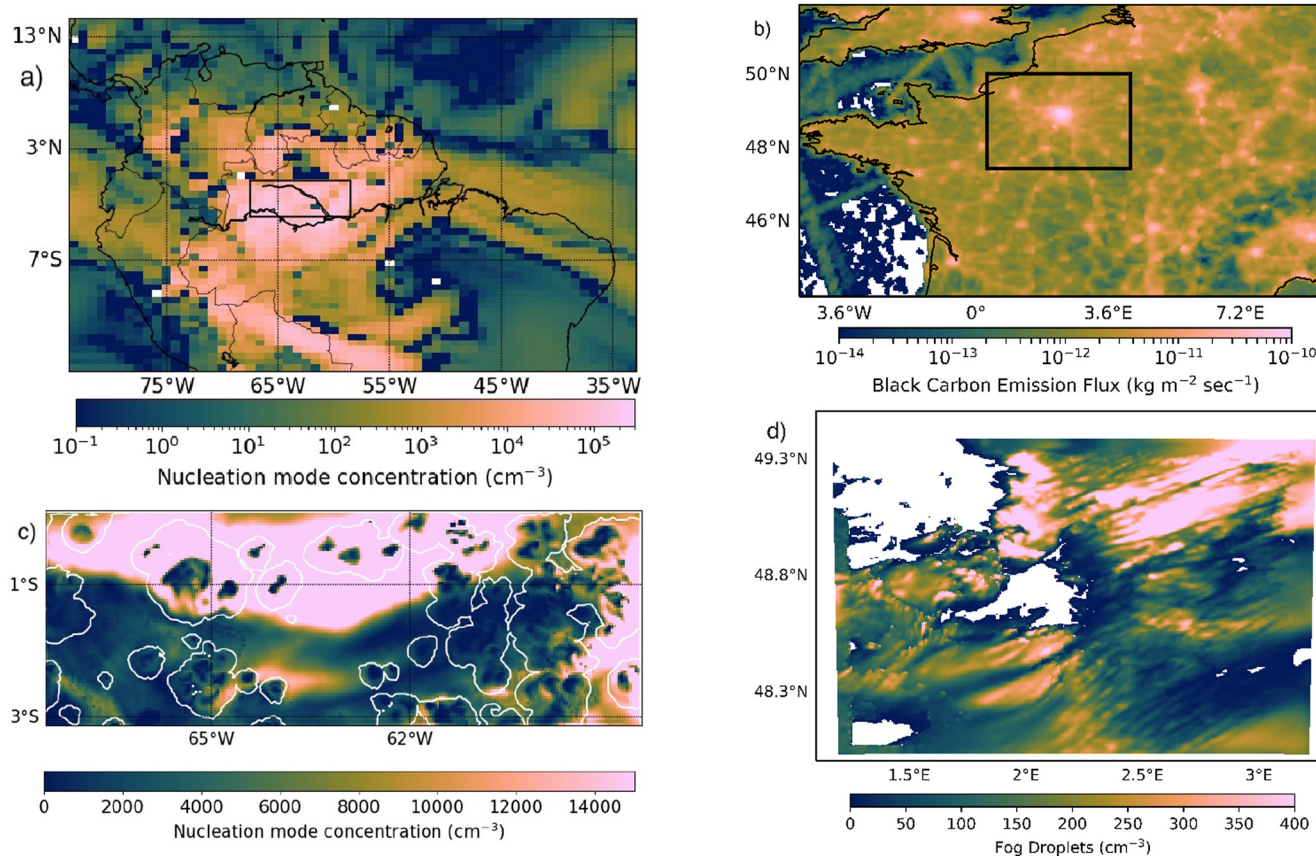
Figure 16. Example applications of the Nested Unified Model with Aerosols and Chemistry model. Subfigures (a and c) show a case study of the Amazon: (a) simulated nucleation-mode number concentrations over South America in a global simulation of 18th September 2014. The domain of the regional simulation is marked. (c) Regional simulation at 4 km resolution showing the effects of clouds on the number concentration. White contours show clouds. Subfigures (b and d) show a case study over the Paris region: (b) emission flux of black carbon from the EDGAR-HTAP inventory in a 4 km model domain over France for a simulation of ParisFog in November 2011. The domain of a 500 m-resolution regional simulation is marked. (d) Simulated fog droplet concentration at 20 m altitude in this regional simulation at 0300 UTC on 15 November 2011.

complicated, NPF may well be occurring in the cloud outflow regions (Clarke et al., 1998). Resolution of vertical transport and mixing enables an analysis of how many aerosol particles can be transported downwards directly into the boundary layer (J. Wang et al., 2016) inside and outside clouds, which is not simulated realistically in a global model with parametrized convection.

We are also working to improve the representation of aerosol-fog interactions in the model. In a 10-day case study during the ParisFog field campaign, the model is able to reproduce observed aerosol concentrations, allowing a detailed study of droplet activation in fog. Sample emissions from the 10 km-resolution EDGAR-HTAP inventory we are using, and the droplet concentrations that result in one of our simulations with 300×300 grid cells at 500 m resolution, are shown on the right of Figure 16, in subfigures (b) and (d).

8. Discussion and Conclusions

8.1. Effects of Higher Spatial Resolution Grid and Emissions

Our development of a high-resolution chemistry-climate model is motivated by the desire to simulate how atmospheric composition varies on smaller scales than can be captured by a global model. Because our simulations are on a more similar spatial scale to observations than those of a typical global climate model, we can also evaluate the simulations much more precisely (provided the parameterizations in the model work independently of the grid resolution). Our high-resolution simulations could then highlight biases in the low-resolution climate systems using the UM.

Using the same emissions in low- and high-resolution models should sometimes allow for useful direct comparisons. However, not surprisingly, if emissions with low grid resolutions of around 130 km are used in a high-resolution model, the model cannot adequately represent spatial variability within a region the size of South Korea (∼400 km): the Pearson's correlation coefficients of that simulation with observations at surface sites (Table 1) are consistently near zero for all species we considered. Furthermore, because the emissions sources are not resolved and tend to be close to the monitors, mean concentrations are biased low. Our simulation with a high resolution emissions data set is less prone to these problems, allowing problems with the model's chemistry scheme to be much more clearly identified.

Along the aircraft tracks, concentrations of NO, SO_2, isoprene and aerosol number concentrations differ markedly between regional and global simulations, and between simulations with low and high-resolution emissions, as shown in Table 2 and Table S2 in Supporting Information S1. The degree to which the Korean terrain and coastline is resolved has a large effect on the agreement of the regional and global simulations, especially for isoprene which is only emitted from land and is short-lived. Along the path of the aircraft, the mixing ratios of most other species are fairly similar (within 20%) between simulations. Especially at high altitude, this is expected for species with chemical lifetimes longer than around 1 day, as their concentrations are strongly influenced by those in the global model that provides LBCs. However, we also find that the results we quote are sensitive to the precise path of the aircraft, or to the area over which we average our model. Over the flight paths we study, surface CO concentrations in our regional model are within 20% of those in the global model. On the other hand, when we average over the whole area shown in subfigures (a) and (b) of Figure 10, surface CO concentrations are 41% higher in the regional model than the global model when the same low-resolution emissions of the global model are used in the regional model. We cannot be sure which simulation is closer to observations, as we do not have evenly distributed observations over the region. For ozone, by contrast, the discrepancy is much smaller, at around 10%. It is not surprising that the regional model produces some odd results when urban emissions from coastal cities get smeared out over regions of sea which would in reality only be influenced by shipping emissions and natural sources. CO concentrations in the regional simulation with the high-resolution KORUSv5 emissions inventory are, as expected, in somewhat better agreement with the global simulation in this area (higher than the global model by 18% rather than 41%).

Weather, air quality and climate models are steadily increasing the resolution of their spatial grids. Numerous reasons for this include the need to resolve complex terrain, emissions sources, or urban heat islands, or to avoid convection parameterizations that add uncertainty to transport and scavenging processes, or to simulate clouds more accurately. Our system is designed to facilitate this process by providing a simulation framework that can be used for development and application of both a regional and a global model. Based on our results, we speculate that as chemistry-climate models increase their spatial grid resolution and/or incorporate regionally refined or nested grids, higher resolution emissions data sets will be necessary. These data sets will allow more precise model evaluation, and sometimes will also be needed in order to avoid degradations of model performance in the high-resolution simulation compared to the lower-resolution simulation.

8.2. Future Model Development and Evaluation

In this paper we aimed to investigate whether the UKCA sub-model could be used in its entirety to represent chemistry and aerosols in high-resolution regional simulations without substantial modifications to the model's source code. Our model produces reasonable simulations of the most important variables for air quality and climate: ozone, $PM_{2.5}$ and number concentrations of cloud-forming aerosols with at least 100 nm diameter. However, there are also significant shortcomings which need to be addressed in follow-up studies. Most notably, our focused model evaluation highlighted a severe shortcoming in the chemistry mechanism. The sources of peroxy radicals in the mechanism are, most likely, insufficient to maintain ozone levels and keep NO_x concentrations realistic (see Section 5). To address this issue, we intend to adopt the more complex Common Reactive Intermediates (CRIs) chemistry scheme of Archer-Nicholls et al. (2021) and J. Weber et al. (2021), which has sources of peroxy radicals from larger VOC molecules.

We defer a detailed evaluation of aerosol composition, measured during the campaign by aerosol mass spectrometers and described by H. Kim et al. (2018) and Nault et al. (2018), to future work. We can be fairly sure that our model will exhibit substantial biases in aerosol composition in its current state, because it lacks any representation of ammonium or nitrate aerosol and has a highly simplified representation of SOA, whose

contribution tends to be severely underestimated due to the lack of any anthropogenic SOA precursors. We are actively working on including a nitrate scheme into NUMAC following its introduction into the global UM by A. C. Jones et al. (2021). The addition of the CRI chemistry scheme will also permit us to include a more sophisticated SOA representation. With these improvements, NUMAC should be tested against aerosol mass spectrometer data. Evaluation of the GEOS-chem model by Travis et al. (2022), Oak et al. (2022), and Choi et al. (2019) suggests that representing aerosol composition during KORUS-AQ will be challenging, and a stringent test of NUMAC. The nitrate-enabled NUMAC model should also lead to good simulations of ammonia concentrations, which will permit more sophisticated NPF mechanisms to be included in the model (Dunne et al., 2016).

Aerosol-cloud and cloud-aerosol interactions could also be improved by further tightening the coupling between the aerosol and cloud microphysics, in particular by introducing a more scale-aware aerosol activation scheme and by an explicit representation of "cloud-borne" aerosols. This latter functionality, representing aerosols inside and outside cloud particles with different prognostic tracers, is available in the CASIM microphysics scheme. However, as currently implemented, it does not keep track of aerosol speciation, so it cannot currently be used when CASIM is coupled to our prognostic aerosol microphysics.

8.3. Conclusions

We have introduced the prototype for a regional chemistry-climate system for the Met Office UM, capable of simulating weather, atmospheric composition and air quality at convection-permitting resolution. The system also includes a consistent global simulation to represent synoptic meteorology and long-range transport of chemical and aerosol species. We speculate based on our previous work that simulations with grid resolution as high as 333 m with the model should be possible (Gordon et al., 2020; Jayakumar et al., 2021). The motivations for modeling atmospheric composition at convection-permitting or cloud-resolving resolution are to simulate explicitly clouds, pollution plumes, and flow over complex terrain. All of these are expected to substantially impact composition in non-linear ways that will lead to errors in coarser-resolution models where these effects are necessarily averaged out.

We tested the model in a case study of the KORUS-AQ campaign in 2016. The model simulates some chemical species, such as OH concentrations, well, and its representation of ozone, CO and $PM_{2.5}$ is broadly comparable with other models that have recently been used over the same period (e.g., Park et al., 2021). However, the NO concentration is substantially biased and significant biases in SO_2 concentration, total aerosol number concentration and likely also aerosol composition also exist. The detailed model evaluation made possible by the high resolution of both the model and the emissions inventory allowed us to understand the most likely cause of the biased NO concentration: a shortage of peroxy radicals. There is also a lack of consistency in the simulated cloud droplet number concentration between different cloud microphysics schemes and likely between simulations with different grid resolutions. We plan to address these shortcomings in follow-up studies.

So far, we used a consistent modeling framework to simulate atmospheric chemistry at 135 and 5 km spatial resolutions (except that we switched off the convection parameterization for the higher-resolution model). We clearly identified situations where biases arose in the global model because of processes that are better resolved in the regional model—the clearest example being a high bias in isoprene concentrations in the global model on 10 to 11 May due to poorly resolved emissions. Further, we demonstrated that partially resolved clouds lead to inhomogeneities in photolysis rates and OH concentrations which are not captured by the global model. While OH concentrations clearly do affect non-linear processes such as NPF and the HO_x cycle, we did not yet clearly identify a bias in the global model that we could explicitly attribute to a non-linear chemical or microphysical process. However, we were able to identify key biases in both regional and global models which should be addressed in future. Despite the biases, the performance of the regional model is usually comparable to the global model, and for some metrics much better. We believe we can improve the model performance further in future while retaining mostly consistent code between horizontal grid scales. In this way NUMAC will achieve its aims: to act as a framework for simulations of air quality and atmospheric composition at high resolution, and for evaluating and refining lower-resolution global climate systems for the UM.

Data Availability Statement

We present data generated by the UK Met Office Unified Model and observations from the NASA KORUS-AQ campaign (https://doi.org/10.5067/Suborbital/KORUSAQ/DATA01), the AirKorea measurement stations and the MODIS satellite instruments. All surface and aircraft observation data is freely and publicly available at https://asdc.larc.nasa.gov/project/KORUS-AQ, last access 23 March 2023. Satellite data sets are freely and publicly available from NASA. The Terra/MODIS aerosol and cloud L2 data sets were acquired from the Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC), located in the Goddard Space Flight Center in Greenbelt, Maryland (https://ladsweb.nascom.nasa.gov/). The data generated by the Unified Model that forms the basis for all the model evaluations presented in this paper is archived at https://doi.org/10.5281/zenodo.7823962. This comprises two-dimensional time series at the surface for comparison with the AirKorea surface measurements and one-dimensional time series interpolated along the path of the aircraft, two-dimensional snapshots of the model that go with the aircraft comparison, and other data sets needed to reproduce figures in this paper. It is accessible at https://zenodo.org/record/7823962.

Software Availability Statement: Simulations presented in this paper were produced with the Met Office Unified Model (UM) version 11.6. This system includes version 5.7 of the JULES land-surface sub-model. Simulations were run using Rose version 2019.01.7 and Cylc version 7.8.11. The simulation identifiers are

- u-cg666: Global model, regional model with CMIP6 emissions, 2–28 May 2016.
- u-cj010: Global model, regional model with KORUSv5 emissions, 2–28 May 2016.
- u-cv580: Global model, regional model with CMIP6 emissions on 10–11 and 26 May.
- u-cj252: Global model, regional model with KORUSv5 emissions on 10–11 and 26 May.
- u-cg716: Global model, regional model with CASIM cloud microphysics on 10–11 May.

The source code for the UM and JULES models used in this study is free to use. However, software for this research is not publicly available due to licensing restrictions, but is available to signatories of the Met Office Software license. Full descriptions of the software, including the specific configurations used in this study, can be found in the text of this article and in articles cited therein. Software is stored in the Met Office Science Repository Service at https://code.metoffice.gov.uk/trac/home. To apply for a license, go to https://www.metoffice.gov.uk/research/collaboration/um-collaboration, and for permission to use JULES, go to the website (https://jules.jchmr.org). The Rose and Cylc software used to drive the Unified Model are public at https://github.com/metomi/rose and https://cylc.github.io/ respectively. Copies of the versions of these codes that were used for this paper are in the archival repository at https://doi.org/10.5281/zenodo.7823962.

References

Abdul-Razzak, H., & Ghan, S. J. (2000). A parameterization of aerosol activation: 2. Multiple aerosol types. *Journal of Geophysical Research*, *105*(D5), 6837–6844. https://doi.org/10.1029/1999JD901161

Archer-Nicholls, S., Abraham, N. L., Shin, Y. M., Weber, J., Russo, M. R., Lowe, D., et al. (2021). The Common Representative Intermediates Mechanism version 2 in the United Kingdom Chemistry and Aerosols Model. *Journal of Advances in Modeling Earth Systems*, *13*(5), e2020MS002420. https://doi.org/10.1029/2020MS002420

Archibald, A. T., O'Connor, F. M., Abraham, N. L., Archer-Nicholls, S., Chipperfield, M. P., Dalvi, M., et al. (2020). Description and evaluation of the UKCA stratosphere–troposphere chemistry scheme (StratTrop vn 1.0) implemented in UKESM1. *Geoscientific Model Development*, *13*(3), 1223–1266. https://doi.org/10.5194/gmd-13-1223-2020

Baklanov, A., Schlünzen, K., Suppan, P., Baldasano, J., Brunner, D., Aksoyoglu, S., et al. (2014). Online coupled regional meteorology chemistry models in Europe: Current status and prospects. *Atmospheric Chemistry and Physics*, *14*(1), 317–398. https://doi.org/10.5194/acp-14-317-2014

Bangert, M., Kottmeier, C., Vogel, B., & Vogel, H. (2011). Regional scale effects of the aerosol cloud interaction simulated with an online coupled comprehensive chemistry model. *Atmospheric Chemistry and Physics*, *11*(9), 4411–4423. https://doi.org/10.5194/acp-11-4411-2011

Benjamin, S. G., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., et al. (2016). A North American hourly assimilation and model forecast cycle: The rapid refresh. *Monthly Weather Review*, *144*(4), 1669–1694. https://doi.org/10.1175/MWR-D-15-0242.1

Boutle, I. A., Eyre, J. E. J., & Lock, A. P. (2014). Seamless stratocumulus simulation across the turbulent gray zone. *Monthly Weather Review*, *142*(4), 1655–1668. https://doi.org/10.1175/MWR-D-13-00229.1

Boutle, I. A., Price, J., Kudzotsa, I., Kokkola, H., & Romakkaniemi, S. (2018). Aerosol–fog interaction and the transition to well-mixed radiation fog. *Atmospheric Chemistry and Physics*, *18*(11), 7827–7840. https://doi.org/10.5194/acp-18-7827-2018

Brown, A., Milton, S., Cullen, M., Golding, B., Mitchell, J., & Shelly, A. (2012). Unified modeling and prediction of weather and climate: A 25-year journey. *Bulletin of the American Meteorological Society*, *93*(12), 1865–1877. https://doi.org/10.1175/BAMS-D-12-00018.1

Brune, W. H., Miller, D. O., Thames, A. B., Brosius, A. L., Barletta, B., Blake, D. R., et al. (2021). Observations of atmospheric oxidation and ozone production in South Korea. *Atmospheric Environment*, 118854. https://doi.org/10.1016/j.atmosenv.2021.118854

Bush, M., Allen, T., Bain, C., Boutle, I. A., Edwards, J., Finnenkoetter, A., et al. (2020). The first Met Office Unified Model–JULES regional atmosphere and land configuration, RAL1. *Geoscientific Model Development*, *13*(4), 1999–2029. https://doi.org/10.5194/gmd-13-1999-2020

Bush, M., Boutle, I., Edwards, J., Finnenkoetter, A., Franklin, C., Hanley, K., et al. (2023). The second Met Office Unified Model–JULES Regional Atmosphere and Land configuration, RAL2. *Geoscientific Model Development*, *16*(6), 1713–1734. https://doi.org/10.5194/gmd-16-1713-2023

Carmichael, G. R., Hayami, H., Calori, G., Uno, I., Cho, S. Y., Engardt, M., et al. (2001). Model intercomparison study of long range transport and sulfur deposition in East Asia (MICS-Asia). *Water, Air, and Soil Pollution*, *130*(1), 51–62. https://doi.org/10.1023/a:1012291200633

Castillo, J. M., Lewis, H. W., Mishra, A., Mitra, A., Polton, J., Brereton, A., et al. (2022). The Regional Coupled Suite (Rcs-Ind1): Application of a flexible regional coupled modelling framework to the Indian region at kilometre scale. *Geoscientific Model Development*, *15*(10), 4193–4223. https://doi.org/10.5194/gmd-15-4193-2022

Choi, J., Park, R. J., Lee, H.-M., Lee, S., Jo, D. S., Jeong, J. I., et al. (2019). Impacts of local vs. trans-boundary emissions from different sectors on $PM_{2.5}$ exposure in South Korea during the KORUS-AQ campaign. *Atmospheric Environment*, *203*, 196–205. https://doi.org/10.1016/j.atmosenv.2019.02.008

Clarke, A. D., Varner, J. L., Eisele, F., Mauldin, R. L., Tanner, D., & Litchy, M. (1998). Particle production in the remote marine atmosphere: Cloud outflow and subsidence during ACE 1. *Journal of Geophysical Research*, *103*(D13), 16397–16409. https://doi.org/10.1029/97JD02987

Crawford, J. H., Ahn, J.-Y., Al-Saadi, J., Chang, L., Emmons, L. K., Kim, J., et al. (2021). The Korea–United States air quality (Korus-AQ) field study. *Elementa: Science of the Anthropocene*, *9*(1), 00163. https://doi.org/10.1525/elementa.2020.00163

Cullen, M. (1993). The unified forecast/climate model. *The Meteorological Magazine*, *122*(1449), 81–94.

Ding, K., Huang, X., Ding, A., Wang, M., Su, H., Kerminen, V.-M., et al. (2021). Aerosol-boundary-layer-monsoon interactions amplify semi-direct effect of biomass smoke on low cloud formation in Southeast Asia. *Nature Communications*, *12*(1), 1–9. https://doi.org/10.1038/s41467-021-26728-4

Dunne, E. M., Gordon, H., Kãœrten, A., Almeida, J., Duplissy, J., Williamson, C., et al. (2016). Global atmospheric particle formation from CERN CLOUD measurements. *Science*, *354*(6316), 1119–1124. https://doi.org/10.1126/science.aaf2649

Eck, T., Holben, B., Kim, J., Beyersdorf, A., Choi, M., Lee, S., et al. (2020). Influence of cloud, fog, and high relative humidity during pollution transport events in South Korea: Aerosol properties and $PM_{2.5}$ variability. *Atmospheric Environment*, *232*, 117530. https://doi.org/10.1016/j.atmosenv.2020.117530

Edwards, J. M., & Slingo, A. (1996). Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model. *Quarterly Journal of the Royal Meteorological Society*, *122*(531), 689–719. https://doi.org/10.1002/qj.49712253107

Emery, C., Liu, Z., Russell, A. G., Odman, M. T., Yarwood, G., & Kumar, N. (2017). Recommendations on statistics and benchmarks to assess photochemical model performance. *Journal of the Air & Waste Management Association*, *67*(5), 582–598. https://doi.org/10.1080/10962247.2016.1265027

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, *45*(2), RG2004. https://doi.org/10.1029/2005RG000183

Field, P. R., Hill, A., Shipway, B., Furtado, K., Wilkinson, J., Miltenberger, A., et al. (2023). Implementation of a double moment cloud microphysics scheme in the UK Met Office regional numerical weather prediction model. *Quarterly Journal of the Royal Meteorological Society*, *149*(752), 703–739. https://doi.org/10.1002/qj.4414

Furtado, K., Field, P. R., Luo, Y., Liu, X., Guo, Z., Zhou, T., et al. (2018). Cloud microphysical factors affecting simulations of deep convection during the presummer rainy season in Southern China. *Journal of Geophysical Research: Atmospheres*, *123*(18), 10477–10505. https://doi.org/10.1029/2017JD028192

Gaubert, B., Emmons, L. K., Raeder, K., Tilmes, S., Miyazaki, K., Arellano, A. F., Jr., et al. (2020). Correcting model biases of co in East Asia: Impact on oxidant distributions during KORUS-AQ. *Atmospheric Chemistry and Physics*, *20*(23), 14617–14647. https://doi.org/10.5194/acp-20-14617-2020

Gordon, H., Carslaw, K. S., Hill, A. A., Field, P. R., Abraham, N. L., Beyersdorf, A., et al. (2022). NUMAC: Description of the Nested Unified Model with Aerosols and Chemistry, and evaluation with KORUS-AQ data: Supporting data [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.7278212

Gordon, H., Field, P. R., Abel, S. J., Barrett, P., Bower, K., Crawford, I., et al. (2020). Development of aerosol activation in the double-moment unified model and evaluation with clarify measurements. *Atmospheric Chemistry and Physics*, *20*(18), 10997–11024. https://doi.org/10.5194/acp-20-10997-2020

Gordon, H., Field, P. R., Abel, S. J., Dalvi, M., Grosvenor, D. P., Hill, A. A., et al. (2018). Large simulated radiative effects of smoke in the South-East Atlantic. *Atmospheric Chemistry and Physics*, *18*(20), 15261–15289. https://doi.org/10.5194/acp-18-15261-2018

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., & Eder, B. (2005). Fully coupled online chemistry within the WRF model. *Atmospheric Environment*, *39*(37), 6957–6975. https://doi.org/10.1016/j.atmosenv.2005.04.027

Grosvenor, D. P., Field, P. R., Hill, A. A., & Shipway, B. J. (2017). The relative importance of macrophysical and cloud albedo changes for aerosol-induced radiative effects in closed-cell stratocumulus: Insight from the modelling of a case study. *Atmospheric Chemistry and Physics*, *17*(8), 5155–5183. https://doi.org/10.5194/acp-17-5155-2017

Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., et al. (2018). Remote sensing of droplet number concentration in warm clouds: A review of the current state of knowledge and perspectives. *Reviews of Geophysics*, *56*(2), 409–453. https://doi.org/10.1029/2017RG000593

Ha, S. (2022). Implementation of aerosol data assimilation in WRFDA (v4.0.3) for WRF-Chem (v3.9.1) using the RACM/MADE-VBS scheme. *Geoscientific Model Development*, *15*(4), 1769–1788. https://doi.org/10.5194/gmd-15-1769-2022

Hemmings, J., & Savage, N. (2018). An initial evaluation of the GLOMAP-mode aerosol scheme for UK air-quality forecasting with AQUM. Met Office Forecasting Research Technical Report 632. Retrieved from https://www.metoffice.gov.uk/research/library-and-archive/publications/science/weather-science-technical-reports

Huang, X., Ding, A., Wang, Z., Ding, K., Gao, J., Chai, F., & Fu, C. (2020). Amplified transboundary transport of haze by aerosol–boundary layer interaction in China. *Nature Geoscience*, *13*(6), 428–434. https://doi.org/10.1038/s41561-020-0583-4

Jang, Y., Lee, Y., Kim, J., Kim, Y., & Woo, J.-H. (2019). Improvement China point source for improving bottom-up emission inventory. *Asia-Pacific Journal of Atmospheric Sciences*, *56*(1), 107–118. https://doi.org/10.1007/s13143-019-00115-y

Jayakumar, A., Gordon, H., Francis, T., Hill, A. A., Mohandas, S., Sandeepan, B. S., et al. (2021). Delhi Model with Chemistry and aerosol framework (DM-Chem) for high-resolution fog forecasting. *Quarterly Journal of the Royal Meteorological Society*, *147*(741), 3957–3978. https://doi.org/10.1002/qj.4163

Jones, A., Roberts, D. L., Woodage, M. J., & Johnson, C. E. (2001). Indirect sulphate aerosol forcing in a climate model with an interactive sulphur cycle. *Journal of Geophysical Research*, *106*(D17), 20293–20310. https://doi.org/10.1029/2000JD000089

Jones, A. C., Hill, A., Remy, S., Abraham, N. L., Dalvi, M., Hardacre, C., et al. (2021). Exploring the sensitivity of atmospheric nitrate concentrations to nitric acid uptake rate using the Met Office's Unified Model. *Atmospheric Chemistry and Physics*, *21*(20), 15901–15927. https://doi.org/10.5194/acp-21-15901-2021

Jordan, C. E., Crawford, J. H., Beyersdorf, A. J., Eck, T. F., Halliday, H. S., Nault, B. A., et al. (2020). Investigation of factors controlling $PM_{2.5}$ variability across the South Korean Peninsula during KORUS-AQ. *Elementa: Science of the Anthropocene*, *8*, 28. https://doi.org/10.1525/elementa.424

Kim, H., Zhang, Q., & Heo, J. (2018). Influence of intense secondary aerosol formation and long-range transport on aerosol chemistry and properties in the Seoul Metropolitan Area during spring time: Results from KORUS-AQ. *Atmospheric Chemistry and Physics*, *18*(10), 7149–7168. https://doi.org/10.5194/acp-18-7149-2018

Kim, H. S., Park, I., Song, C. H., Lee, K., Yun, J. W., Kim, H. K., et al. (2019). Development of a daily PM_{10} and $PM_{2.5}$ prediction system using a deep long short-term memory neural network model. *Atmospheric Chemistry and Physics*, *19*(20), 12935–12951. https://doi.org/10.5194/acp-19-12935-2019

Kim, Y. P., & Lee, G. (2018). Trend of air quality in Seoul: Policy and science. *Aerosol and Air Quality Research*, *18*(9), 2141–2156. https://doi.org/10.4209/aaqr.2018.03.0081

Lana, A., Bell, T. G., Simó, R., Vallina, S. M., Ballabrera-Poy, J., Kettle, A. J., et al. (2011). An updated climatology of surface dimethlysulfide concentrations and emission fluxes in the global ocean. *Global Biogeochemical Cycles*, *25*(1), GB1004. https://doi.org/10.1029/2010GB003850

Lean, H. W., Barlow, J. F., & Halios, C. H. (2019). The impact of spin-up and resolution on the representation of a clear convective boundary layer over London in order 100 m grid-length versions of the Met Office Unified Model. *Quarterly Journal of the Royal Meteorological Society*, *145*(721), 1674–1689. https://doi.org/10.1002/qj.3519

Lee, Y., Park, J., Kim, P., & Ghim, Y. S. (2021). New particle formation and diurnal variations in number concentrations at a rural site downwind of Seoul, Korea. *Atmospheric Pollution Research*, *12*(3), 214–223. https://doi.org/10.1016/j.apr.2021.01.014

Levy, R., Hsu, C., et al. (2015). *MODIS atmosphere L2 aerosol product*. NASA MODIS Adaptive Processing System, Goddard Space Flight Center. https://doi.org/10.5067/MODIS/MYD04_L2.006

Liss, P. S., & Merlivat, L. (1986). Air-sea gas exchange rates: Introduction and synthesis. In P. Buat-Ménard (Ed.), *The role of air-sea exchange in geochemical cycling* (pp. 113–127). Springer Netherlands. https://doi.org/10.1007/978-94-009-4738-2_5

Lock, A. P., Brown, A. R., Bush, M. R., Martin, G. M., & Smith, R. N. B. (2000). A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Monthly Weather Review*, *128*(9), 3187–3199. https://doi.org/10.1175/1520-0493(2000)128⟨3187:ANBLMS⟩2.0.CO;2

Mann, G. W., Carslaw, K. S., Ridley, D. A., Spracklen, D. V., Pringle, K. J., Merikanto, J., et al. (2012). Intercomparison of modal and sectional aerosol microphysics representations within the same 3-D global chemical transport model. *Atmospheric Chemistry and Physics*, *12*(10), 4449–4476. https://doi.org/10.5194/acp-12-4449-2012

Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., et al. (2010). Description and evaluation of GLOMAP-mode: A modal global aerosol microphysics model for the UKCA composition-climate model. *Geoscientific Model Development*, *3*(2), 519–551. https://doi.org/10.5194/gmd-3-519-2010

Manners, J., Edwards, J. M., Hill, P., & Thelen, J.-C. (n.d.). *SOCRATES (Suite Of Community RAdiative Transfer codes based on Edwards and Slingo) technical guide*. Met Office. Retrieved from https://code.metoffice.gov.uk/trac/socrates

McCaul, E. W., Goodman, S. J., LaCasse, K. M., & Cecil, D. J. (2009). Forecasting lightning threat using cloud-resolving model simulations. *Weather and Forecasting*, *24*(3), 709–729. https://doi.org/10.1175/2008WAF2222152.1

McMurry, P. H., Takano, H., & Anderson, G. R. (1983). Study of the ammonia (gas)-sulfuric acid (aerosol) reaction rate. *Environmental Science & Technology*, *17*(6), 347–352. https://doi.org/10.1021/es00112a008

Morgenstern, O., Braesicke, P., O'Connor, F. M., Bushell, A. C., Johnson, C. E., Osprey, S. M., & Pyle, J. A. (2009). Evaluation of the new UKCA climate-composition model – Part 1: The stratosphere. *Geoscientific Model Development*, *2*(1), 43–57. https://doi.org/10.5194/gmd-2-43-2009

Morrison, H., & Gettelman, A. (2008). A new two-moment bulk stratiform cloud microphysics scheme in the Community Atmosphere Model, version 3 (CAM3). Part I: Description and numerical tests. *Journal of Climate*, *21*(15), 3642–3659. https://doi.org/10.1175/2008JCLI2105.1

Mulcahy, J. P., Johnson, C., Jones, C. G., Povey, A. C., Scott, C. E., Sellar, A., et al. (2020). Description and evaluation of aerosol in UKESM1 and HadGEM3-GC3.1 CMIP6 historical simulations. *Geoscientific Model Development*, *13*(12), 6383–6423. https://doi.org/10.5194/gmd-13-6383-2020

Mulcahy, J. P., Jones, C., Sellar, A., Johnson, B., Boutle, I. A., Jones, A., et al. (2018). Improved aerosol processes and effective radiative forcing in HadGEM3 and UKESM1. *Journal of Advances in Modeling Earth Systems*, *10*(11), 2786–2805. https://doi.org/10.1029/2018MS001464

Mulcahy, J. P., Jones, C. G., Rumbold, S. T., Kuhlbrodt, T., Dittus, A. J., Blockley, E. W., et al. (2023). UKESM1.1: Development and evaluation of an updated configuration of the UK Earth System Model. *Geoscientific Model Development*, *16*(6), 1569–1600. https://doi.org/10.5194/gmd-16-1569-2023

Nault, B. A., Campuzano-Jost, P., Day, D. A., Schroder, J. C., Anderson, B., Beyersdorf, A. J., et al. (2018). Secondary organic aerosol production from local emissions dominates the organic aerosol budget over Seoul, South Korea, during KORUS-AQ. *Atmospheric Chemistry and Physics*, *18*(24), 17769–17800. https://doi.org/10.5194/acp-18-17769-2018

Neal, L. S., Dalvi, M., Folberth, G., McInnes, R. N., Agnew, P., O'Connor, F. M., et al. (2017). A description and evaluation of an air quality model nested within global and regional composition-climate models using MetUM. *Geoscientific Model Development*, *10*(11), 3941–3962. https://doi.org/10.5194/gmd-10-3941-2017

Oak, Y. J., Park, R. J., Jo, D. S., Hodzic, A., Jimenez, J. L., Campuzano-Jost, P., et al. (2022). Evaluation of secondary organic aerosol (SOA) simulations for Seoul, Korea. *Journal of Advances in Modeling Earth Systems*, *14*(2), e2021MS002760. https://doi.org/10.1029/2021MS002760

Oak, Y. J., Park, R. J., Schroeder, J. R., Crawford, J. H., Blake, D. R., Weinheimer, A. J., et al. (2019). Evaluation of simulated O_3 production efficiency during the KORUS-AQ campaign: Implications for anthropogenic NO_x emissions in Korea. *Elementa: Science of the Anthropocene*, *7*, 56. https://doi.org/10.1525/elementa.394

O'Connor, F. M., Johnson, C. E., Morgenstern, O., Abraham, N. L., Braesicke, P., Dalvi, M., et al. (2014). Evaluation of the new UKCA climate-composition model – Part 2: The Troposphere. *Geoscientific Model Development*, *7*(1), 41–91. https://doi.org/10.5194/gmd-7-41-2014

Oliver, H., Shin, M., Matthews, D., Sanders, O., Bartholomew, S., Clark, A., et al. (2019). Workflow automation for cycling systems. *Computing in Science & Engineering*, *21*(4), 7–21. https://doi.org/10.1109/mcse.2019.2906593

Pacifico, F., Harrison, S. P., Jones, C. D., Arneth, A., Sitch, S., Weedon, G. P., et al. (2011). Evaluation of a photosynthesis-based biogenic isoprene emission scheme in JULES and simulation of isoprene emissions under present-day climate conditions. *Atmospheric Chemistry and Physics*, *11*(9), 4371–4389. https://doi.org/10.5194/acp-11-4371-2011

Park, R. J., Oak, Y. J., Emmons, L. K., Kim, C.-H., Pfister, G. G., Carmichael, G. R., et al. (2021). Multi-model intercomparisons of air quality simulations for the KORUS-AQ campaign. *Elementa: Science of the Anthropocene*, *9*(1), 00139. https://doi.org/10.1525/elementa.2021.00139

Peterson, D. A., Hyer, E. J., Han, S.-O., Crawford, J. H., Park, R. J., Holz, R., et al. (2019). Meteorology influencing springtime air quality, pollution transport, and visibility in Korea. *Elementa: Science of the Anthropocene*, *7*, 57. https://doi.org/10.1525/elementa.395

Petters, M. D., & Kreidenweis, S. M. (2007). A single parameter representation of hygroscopic growth and cloud condensation nucleus activity. *Atmospheric Chemistry and Physics*, *7*(8), 1961–1971. https://doi.org/10.5194/acp-7-1961-2007

Planche, C., Mann, G. W., Carslaw, K. S., Dalvi, M., Marsham, J. H., & Field, P. R. (2017). Spatial and temporal CCN variations in convection-permitting aerosol microphysics simulations in an idealised marine tropical domain. *Atmospheric Chemistry and Physics*, *17*(5), 3371–3384. https://doi.org/10.5194/acp-17-3371-2017

Platnick, S., Ackerman, S., King, M., Meyer, K., Menzel, W. P., Holz, R. E., et al. (2015). *MODIS atmosphere L2 cloud product (06_l2)*. NASA MODIS Adaptive Processing System, Goddard Space Flight Center. https://doi.org/10.5067/MODIS/MYD06_L2.006

Possner, A., Zubler, E., Lohmann, U., & Schär, C. (2016). The resolution dependence of cloud effects and ship-induced aerosol-cloud interactions in marine stratocumulus. *Journal of Geophysical Research: Atmospheres*, *121*(9), 4810–4829. https://doi.org/10.1002/2015JD024685

Price, C., & Rind, D. (1992). A simple lightning parameterization for calculating global lightning distributions. *Journal of Geophysical Research*, *97*(D9), 9919–9933. https://doi.org/10.1029/92JD00719

Price, C., & Rind, D. (1993). What determines the cloud-to-ground lightning fraction in thunderstorms? *Geophysical Research Letters*, *20*(6), 463–466. https://doi.org/10.1029/93GL00226

Ranjithkumar, A., Gordon, H., Williamson, C., Rollins, A., Pringle, K., Kupc, A., et al. (2021). Constraints on global aerosol number concentration, SO_2 and condensation sink in UKESM1 using ATom measurements. *Atmospheric Chemistry and Physics*, *21*(6), 4979–5014. https://doi.org/10.5194/acp-21-4979-2021

Saide, P. E., Gao, M., Lu, Z., Goldberg, D. L., Streets, D. G., Woo, J.-H., et al. (2020). Understanding and improving model representation of aerosol optical properties for a Chinese haze event measured during KORUS-AQ. *Atmospheric Chemistry and Physics*, *20*(11), 6455–6478. https://doi.org/10.5194/acp-20-6455-2020

Saleeby, S. M., & van den Heever, S. C. (2013). Developments in the CSU-RAMS aerosol model: Emissions, nucleation, regeneration, deposition, and radiation. *Journal of Applied Meteorology and Climatology*, *52*(12), 2601–2622. https://doi.org/10.1175/JAMC-D-12-0312.1

Savage, N. H., Agnew, P., Davis, L. S., Ordóñez, C., Thorpe, R., Johnson, C. E., et al. (2013). Air quality modelling using the Met Office Unified Model (AQUM OS24-26): Model description and initial evaluation. *Geoscientific Model Development*, *6*(2), 353–372. https://doi.org/10.5194/gmd-6-353-2013

Schroeder, J. R., Crawford, J. H., Ahn, J.-Y., Chang, L., Fried, A., Walega, J., et al. (2020). Observation-based modeling of ozone chemistry in the Seoul metropolitan area during the Korea-United States Air Quality Study (KORUS-AQ). *Elementa: Science of the Anthropocene*, *8*, 3. https://doi.org/10.1525/elementa.400

Schutgens, N., Tsyro, S., Gryspeerdt, E., Goto, D., Weigum, N., Schulz, M., & Stier, P. (2017). On the spatio-temporal representativeness of observations. *Atmospheric Chemistry and Physics*, *17*(16), 9761–9780. https://doi.org/10.5194/acp-17-9761-2017

Seinfeld, J., & Pandis, S. (2008). *Atmospheric chemistry and physics*. John Wiley & Sons.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., et al. (2019). UKESM1: Description and evaluation of the U.K. Earth System Model. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4513–4558. https://doi.org/10.1029/2019MS001739

Shipway, B., & Abel, S. (2010). Analytical estimation of cloud droplet nucleation based on an underlying aerosol population. *Atmospheric Research*, *96*(2), 344–355. https://doi.org/10.1016/j.atmosres.2009.10.005

Shipway, B. J., & Hill, A. A. (2012). Diagnosis of systematic differences between multiple parametrizations of warm rain microphysics using a kinematic framework. *Quarterly Journal of the Royal Meteorological Society*, *138*(669), 2196–2211. https://doi.org/10.1002/qj.1913

Simpson, I. J., Blake, D. R., Blake, N. J., Meinardi, S., Barletta, B., Hughes, S. C., et al. (2020). Characterization, sources and reactivity of volatile organic compounds (VOCs) in Seoul and surrounding regions during KORUS-AQ. *Elementa: Science of the Anthropocene*, *8*, 37. https://doi.org/10.1525/elementa.434

Sitz, L. E., Di Sante, F., Farneti, R., Fuentes-Franco, R., Coppola, E., Mariotti, L., et al. (2017). Description and evaluation of the Earth System Regional Climate Model (RegCM-ES). *Journal of Advances in Modeling Earth Systems*, *9*(4), 1863–1886. https://doi.org/10.1002/2017MS000933

Smagorinsky, J. (1963). General circulation experiments with the primitive equations: I. The basic experiment. *Monthly Weather Review*, *91*(3), 99–164. https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2

Smith, R. N. B. (1990). A scheme for predicting layer clouds and their water content in a general circulation model. *Quarterly Journal of the Royal Meteorological Society*, *116*(492), 435–460. https://doi.org/10.1002/qj.49711649210

Stevens, B., Fiedler, S., Kinne, S., Peters, K., Rast, S., Müsse, J., et al. (2017). MACv2-SP: A parameterization of anthropogenic aerosol optical properties and an associated Twomey effect for use in CMIP6. *Geoscientific Model Development*, *10*(1), 433–452. https://doi.org/10.5194/gmd-10-433-2017

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., et al. (2019). DYAMOND: The dynamics of the atmospheric general circulation modeled on non-hydrostatic domains. *Progress in Earth and Planetary Science*, *6*(1), 1–17. https://doi.org/10.1186/s40645-019-0304-z

Stratton, R. A., Senior, C. A., Vosper, S. B., Folwell, S. S., Boutle, I. A., Earnshaw, P. D., et al. (2018). A Pan-African convection-permitting regional climate simulation with the Met Office Unified Model: CP4-Africa. *Journal of Climate*, *31*(9), 3485–3508. https://doi.org/10.1175/JCLI-D-17-0503.1

Tang, Y., Lean, H. W., & Bornemann, J. (2013). The benefits of the Met Office variable resolution NWP model for forecasting convection. *Meteorological Applications*, *20*(4), 417–426. https://doi.org/10.1002/met.1300

Teixeira, J. C., Folberth, G. A., O'Connor, F. M., Unger, N., & Voulgarakis, A. (2021). Coupling interactive fire with atmospheric composition and climate in the UK Earth System Model. *Geoscientific Model Development*, *14*(10), 6515–6539. https://doi.org/10.5194/gmd-14-6515-2021

Thuburn, J. (2016). ENDGame: The new dynamical core of the Met Office weather and climate prediction model. In P. J. Aston, A. J. Mulholland, & K. M. Tant (Eds.), *UK success stories in industrial mathematics* (pp. 27–33). Springer International Publishing. https://doi.org/10.1007/978-3-319-25454-8_4

Travis, K. R., Crawford, J. H., Chen, G., Jordan, C. E., Nault, B. A., Kim, H., et al. (2022). Limitations in representation of physical processes prevent successful simulation of $PM_{2.5}$ during KORUS-AQ. *Atmospheric Chemistry and Physics*, *22*(12), 7933–7958. https://doi.org/10.5194/acp-22-7933-2022

Turnock, S. T., Mann, G. W., Woodhouse, M. T., Dalvi, M., O'Connor, F. M., Carslaw, K. S., & Spracklen, D. V. (2019). The impact of changes in cloud water pH on aerosol radiative forcing. *Geophysical Research Letters*, *46*(7), 4039–4048. https://doi.org/10.1029/2019GL082067

Van Weverberg, K., Morcrette, C. J., Boutle, I. A., Furtado, K., & Field, P. R. (2021). A bimodal diagnostic cloud fraction parameterization. Part I: Motivating analysis and scheme description. *Monthly Weather Review*, *149*(3), 841–857. https://doi.org/10.1175/MWR-D-20-0224.1

Vogel, B., Vogel, H., Bäumer, D., Bangert, M., Lundgren, K., Rinke, R., & Stanelle, T. (2009). The comprehensive model system COSMO-ART: Radiative impact of aerosol on the state of the atmosphere on the regional scale. *Atmospheric Chemistry and Physics*, *9*(22), 8661–8680. https://doi.org/10.5194/acp-9-8661-2009

Walters, D., Baran, A. J., Boutle, I. A., Brooks, M., Earnshaw, P., Edwards, J., et al. (2019). The Met Office Unified Model global atmosphere 7.0/7.1 and JULES global land 7.0 configurations. *Geoscientific Model Development*, *12*(5), 1909–1963. https://doi.org/10.5194/gmd-12-1909-2019

Wang, J., Krejci, R., Giangrande, S., Kuang, C., Barbosa, H. M., Brito, J., et al. (2016). Amazon boundary layer aerosol concentration sustained by vertical transport during rainfall. *Nature*, *539*(7629), 416–419. https://doi.org/10.1038/nature19819

Wang, X., Gordon, H., Grosvenor, D. P., Andreae, M. O., & Carslaw, K. S. (2023). Contribution of regional aerosol nucleation to low-level CCN in an amazonian deep convective environment: Results from a regionally nested global model. *Atmospheric Chemistry and Physics*, *23*(7), 4431–4461. https://doi.org/10.5194/acp-23-4431-2023

Watson-Parris, D., Schutgens, N., Cook, N., Kipling, Z., Kershaw, P., Gryspeerdt, E., et al. (2016). Community Intercomparison Suite (CIS) v1.4.0: A tool for intercomparing models and observations. *Geoscientific Model Development*, *9*(9), 3093–3110. https://doi.org/10.5194/gmd-9-3093-2016

Weber, J., Archer-Nicholls, S., Abraham, N. L., Shin, Y. M., Bannan, T. J., Percival, C. J., et al. (2021). Improvements to the representation of BVOC chemistry-climate interactions in UKCA (vn11.5) with the CRI-Strat 2 mechanism: Incorporation and evaluation. *Geoscientific Model Development Discussions*, 1–52. https://doi.org/10.5194/gmd-2021-119

Weber, R. J., McMurry, P. H., Mauldin, L., Tanner, D. J., Eisele, F. L., Brechtel, F. J., et al. (1998). A study of new particle formation and growth involving biogenic and trace gas species measured during ACE 1. *Journal of Geophysical Research*, *103*(D13), 16385–16396. https://doi.org/10.1029/97JD02465

West, R. E. L., Stier, P., Jones, A., Johnson, C. E., Mann, G. W., Bellouin, N., et al. (2014). The importance of vertical velocity variability for estimates of the indirect aerosol effects. *Atmospheric Chemistry and Physics*, *14*(12), 6369–6393. https://doi.org/10.5194/acp-14-6369-2014

Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., et al. (2018). The Met Office global coupled model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *Journal of Advances in Modeling Earth Systems*, *10*(2), 357–380. https://doi.org/10.1002/2017MS001115

Williamson, C. J., Kupc, A., Axisa, D., Bilsback, K. R., Bui, T., Campuzano-Jost, P., et al. (2019). A large source of cloud condensation nuclei from new particle formation in the tropics. *Nature*, *574*(7778), 399–403. https://doi.org/10.1038/s41586-019-1638-9

Wilson, D. R., & Ballard, S. P. (1999). A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. *Quarterly Journal of the Royal Meteorological Society*, *125*(557), 1607–1636. https://doi.org/10.1002/qj.49712555707

Wilson, D. R., Bushell, A. C., Kerr-Munslow, A. M., Price, J. D., & Morcrette, C. J. (2008). PC2: A prognostic cloud fraction and condensation scheme. I: Scheme description. *Quarterly Journal of the Royal Meteorological Society*, *134*(637), 2093–2107. https://doi.org/10.1002/qj.333

Wood, N., Staniforth, A., White, A., Allen, T., Diamantakis, M., Gross, M., et al. (2014). An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quarterly Journal of the Royal Meteorological Society*, *140*(682), 1505–1520. https://doi.org/10.1002/qj.2235

Woodward, S. (2001). Modeling the atmospheric life cycle and radiative impact of mineral dust in the Hadley Centre climate model. *Journal of Geophysical Research*, *106*(D16), 18155–18166. https://doi.org/10.1029/2000JD900795