

Response to the Interactive Comments:

Using a Virtual Machine environment for developing, testing, and training for the UM-UKCA Composition-Climate Model, using Unified Model version 10.9 and above

5 Nathan Luke Abraham^{1,2}, Alexander T. Archibald^{1,2}, Paul Cresswell³, Sam Cusworth³, Mohit Dalvi³, David Matthews³, Steven Wardle³, and Stuart Whitehouse³

1. Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, UK

2. National Centre for Atmospheric Science, UK

3. Met Office, FitzRoy Road, Exeter, EX1 3PB, UK

10 *Correspondence to:* N. Luke Abraham (luke.abraham@atm.ch.cam.ac.uk)

We would like to thank both reviewers for taking the time to read the manuscript. We include their full comments below in ***bold italics***, with our responses interspaced below in normal text. Changes to the manuscript text are given in *italics*.

Response to Anonymous Referee #1

General Comments:

15 ***This manuscript describes a virtual machine (VM) for low-resolution testing and code development of the Met Office Unified Model, focusing on the chemistry and aerosol sub-module. The use of VM's and containers remains relatively new in the geophysical model development community, and the exploration presented here is a worthwhile contribution. The contribution is generally well-written and complete, subject to some***
20 ***clarifications requested below.***

We would like to thank the reviewer for their positive helpful comments.

Specific Comments:

25 ***While the manuscript is generally well written, I found Section 4.2 somewhat confusing and possibly unnecessary. The comparison between two HPC platforms and a 1x2 decomposition in the VM did not lead to any conclusions that I could discern. I can't see how the statement that the "compile time is comparable on the VM" is supported by the evidence presented. I am also unable to find a description of where the VM is***
30 ***deployed in these tests.***

The relevant issue here is the overhead of the VM, which we can expect will be large compared to e.g. a container. A 1x2 HPC test, then a 1x2 test with the VM deployed on the same HPC, would directly address that question.

- 5 In Section 2 we state that “For this study we use an Ubuntu 16.04LTS host with 32 Intel Xeon 2.0GHz cores and 64GB of memory, and an Ubuntu 16.04LTS guest configured with 2, 4, 8, or 16 cores and up to 20GB of memory as required.” and this is the case for VM integrations throughout the manuscript. The text has been altered to highlight this:

10 *Table 2 shows very approximate timings on these three different environments (ARCHER, the XCS-C, and the VM set-up described in Section 2) for the compile (including code extract), reconfiguration, and atmosphere tasks for the N48L38 UKCA configuration described in Section 4.*

- 15 In Section 4.2 we wish to highlight how a Virtual Machine environment can be useful for model development by comparing how a similar model set-up, that has designed to be run quickly, can be used on the different platforms available. For this we used the optimum configuration on each of the different systems, rather than using a 1x2 configuration everywhere, as we wanted to highlight the differences during a typical development cycle. It is for these reasons that only very rough timings are given. As the Reviewer highlights there is additional overhead when using the VM compared to using a container, but we did not wish to perform a detailed timing comparison as the configuration presented here is designed for model development and testing, and not for production simulations where model performance would need to be optimised. The additional text has been added to emphasize this point:

Here we are not considering a detailed comparison of identical N48L38 UKCA configurations on different computing systems but are instead interested in the typical configuration that would be used on each system when making new code developments.

- 25 On the VM, GNU compiles took around 8-10 minutes for initial builds, with subsequent incremental builds taking less than a minute. On NEXCS GNU compiles took around 4 minutes for an initial compile and 2-3 minutes for subsequent incremental compiles. Times for the Intel and Cray Compiler Environment compiles were longer than both of these. We have amended the text to reduce the emphasis on the compile time:

- 30 *While the model run-time is fastest on the two supercomputers, when queue times are considered the VM may be a better option for model development jobs over ARCHER. More cores can also be provided to the VM if available, further reducing run-times (see Table 4).*

We would also be unlikely to be able to deploy the VM on to either HPC environments due to security concerns.

A second comment is that while the results are likely reproducible with a few limitation of hardware (maybe), the general lack of access to the code due to restricted access does diminish the value of this contribution.

5 As the Known Good Output is internally generated within the VM system, any changes in model evolution between different VM instances due to the differences in the Host hardware are accounted for during the rose-stem tests. As this system is designed to be a self-contained platform for model development, testing, and training, the lack of reproducibility of the results between different Hosts should not be an issue for users.

The nature of the UM licence means that its use is restricted. However, for those with access we believe that this will be a
10 useful tool, and there has already been an uptake in this VM system by UKCA and UM users.

We also wish to highlight how useful small systems, such as the VM environment described here, are for users and developers, without the need to use HPC resources.

Technical corrections:

15 ***P1L22: Break up the sentence beginning with “The UM: : :”***

The text now reads:

*The UM comprises of over 965,000 lines of computer code, mainly written in Fortran 90/95. A comprehensive test harness is
20 in place at the Met Office and some partner sites such as Australia, New Zealand, South Korea, Poland, and elsewhere, to aid in development and to ensure that the code is well-tested, which is comprehensive but labour intensive to maintain.*

P2L15: Please clarify the discussion of Rose and Cylc. As written they both appear to be workflow management systems of some kind.

25

Cylc is a workflow engine, used to schedule the various tasks required, whereas Rose is used to configure the UM and edit the input namelists. We have amended the text in the manuscript to clarify this:

*Rose is a system for creating, editing, and running application configurations and it is used as the graphical user interface for
30 the UM to configure the input namelists.*

P4L1: How does the time step affect memory? I can't follow that sentence.

It is the reduction in model levels that reduced the memory, not the change in the model timestep. The reduction in the number of levels then allows the timestep to be increased, which is then beneficial for model run time. The sentence has been changed to read:

- 5 *The number of vertical levels can be reduced further to 38 (with a model top at 40km), as is used in the HadGEM2-ES configuration of the UM (Collins et al., 2011; O'Connor et al., 2014), which allows the model to only need 7.5GB of memory and allows the dynamical timestep to be increased from 20 to 30 minutes, also reducing the overall run-time.*

P10L8: I suggest a new paragraph at “The VM”

10

This change has been made. The text now appears:

However, for the VM this maintenance, while possible, is labour intensive and usually unnecessary.

- 15 *The VM is not designed as a system that should be maintained through UM versions, it is designed to be used for simple code development against a single, stable UM release.*

P11L22: “mule-cumf” needs explanation

- 20 The mule utility is included as part of the UM install and contains a number of useful tools for examining UM files. One of these tools allows for the comparison of UM files, called “mule-cumf”, which allows for a bit-wise comparison of all fields contained in UM output files to be performed. However, in the context of this discussion these details are unnecessary, and so the reference to mule-cumf has been removed. The text now reads:

The output from these two jobs is then compared in the same way as for the KGO comparison.

- 25 **Response to Anonymous Referee #2**

- Abraham et al. (2018) describe how a global atmospheric chemistry-climate model (UM-UKCA) was modified for use in a virtual machine with a low memory footprint, making it easily portable to most modern desktop or laptop computers with the appropriate host software installed. The resolution of the model was significantly reduced,*
- 30 *and a number of other features were disabled, rendering the model configuration unsuitable for production use. The point of the exercise however was to provide an environment suitable for certain kinds of model development and testing, and for training*

new users of the system.

The modifications to the base model are described, and the different use cases for this system are discussed. This paper should be interesting and useful for users of the virtual machine environment, and other workers who are also considering packaging such a model in a virtual environment. I recommend publication in GMD after a few minor revisions.

We would like to thank the reviewer for their positive and helpful comments.

10 *In your "Code Availability" section, you describe the restrictions and availability of the UM, but you do not mention anything about obtaining a VM image, except for a cryptic reference to "UMDP X10", which is not described any further. It might seem a bit pedantic, but since the whole paper is about this virtual machine, it would be nice to see some more text about its availability.*

15

Under the *Code Availability* section, we have included a comment regarding the availability of the Virtual Machine configuration settings, which are freely available via GitHub. As mentioned in this section, all Unified Model Documentation Papers, including UMDP X10, are available under the UM licence. This paper is not intended to be a replacement for UMDP X10, and the citations for Cresswell (2018) and all other UMDPs include the URLs of these document on the Met Office Science Repository Service (U.K. Met Office, 2018). We have therefore clarified the text in the *Code Availability* section to:

20

The Met Office Virtual Machine can be obtained from <https://github.com/metomi/metomi-vm> (Matthews, 2018). Due to intellectual property right restrictions, we cannot provide either the source code or documentation papers for the UM.

25 *Obtaining the UM. The Met Office Unified Model is available for use under licence. The functionality discussed here is fully available in the UM trunk from version 11.0, with detailed instructions on using how to install the UM on the VM provided by UMDP X10 (Cresswell, 2018), and the UM version 10.9 UKCA Tutorials are described in Abraham et al. (2017). A number of research organisations and national meteorological services use the UM in collaboration with the Met Office to undertake basic atmospheric process research, produce forecasts, develop the UM code and build and evaluate Earth system models.*

30 *For further information on how to apply for a licence see <http://www.metoffice.gov.uk/research/modelling-systems/unified-model>.*

The UMDP citations appear as:

Abraham, N. L., Archibald, A. T., Bellouin, N., Boucher, O., Braesicke, P., Bushell, A., Carslaw, K., Collins, B., Dalvi, M., Dennison, F., Emmerson, K., Folberth, G., Haywood, J., Hewitt, A., Johnson, C., Kipling, Z., Macintyre, H., Mann, G., Telford, P., Merikanto, J., Morgenstern, O., O'Connor, F., Ordonez, C., Osprey, S., Pringle, K., Pyle, J., Rae, J., Reddington, C., Savage, N., Sellar, A., Spracklen, D., Stier, P., West, R., Mulcahy, J., Woodward, S., Boutle, I. and Woodhouse, M. T.: UMDP 084: United Kingdom Chemistry and Aerosol (UKCA) Technical Description. [online] Available from: https://code.metoffice.gov.uk/doc/um/vn11.0/papers/umdp_084.pdf (Accessed 23 February 2018), 2018.

Cresswell, P.: Unified Model Documentation Paper X10: Unified Model Virtual Machine Guide. [online] Available from: https://code.metoffice.gov.uk/doc/um/vn11.0/papers/umdp_X10.pdf (Accessed 6 March 2018), 2018.

Whitehouse, S.: Unified Model Documentation Paper X09: The UM rose-stem Suite for External Users. [online] Available from: https://code.metoffice.gov.uk/doc/um/vn11.0/papers/umdp_X09.pdf (Accessed 16 February 2018), 2018.

Page 4, line 2: Which diagnostics were turned off, and why are these "unnecessary"?

The Unified Model has a comprehensive diagnostics package and can output a large number of diagnostic fields from many different UM code sections. Many UM climate configurations contain a standard set of diagnostics that are included in long 20-year simulations that are then processed to produce a series of standard plots to aid in model evaluation. The original N48L70 GA7+StratTrop configuration contained 695 separate diagnostic requests, 43 of which were required for UKCA coupling. These 652 additional diagnostics are the “unnecessary” ones, as it is possible to run the model without any diagnostic requests at all (except for those required for UKCA coupling). All fields that are required for UKCA coupling can be found in Tables 39 and 40 of Appendix C in UMDP 84 (Abraham *et al.*, 2018).

The final GA7+StratTrop configuration includes 19 diagnostics from UKCA chemistry and aerosols that are used in the rose-stem tests, listed below. We did not include this information in the manuscript as we believed that these specifics were unnecessary as users can add or remove diagnostics from their own simulations as required.

Diagnostic
GLOMAP-mode accumulation mode (soluble) optical depth
Ozone (O ₃) mass-mixing ratio (kg kg ⁻¹)
GLOMAP-mode Aitken mode (soluble) organic matter mass-mixing ratio (kg kg ⁻¹)
Age of air (s)
Cloud droplet number concentration (m ⁻³)
Nitrogen dioxide (NO ₂) mass-mixing ratio from UKCA (kg kg ⁻¹)
Primary emissions flux of sulphuric acid aerosol mass to GLOMAP-mode Aitken mode (soluble) (mol s ⁻¹)
Geometric (number) mean dry diameter of particles in GLOMAP-mode accumulation mode (soluble) (m)
Vertically integrated plume scavenging of Cl in the GLOMAP-mode coarse mode (soluble) (mol s ⁻¹)
Flux through NO + RO ₂ reactions (mol s ⁻¹)
Ozone dry deposition flux (mol s ⁻¹)
NO _x wet deposition flux (mol s ⁻¹)
NO _x emissions from lightning (mol s ⁻¹)
NO surface emissions (kg m ⁻² s ⁻¹)
NO aircraft emissions (kg m ⁻² s ⁻¹)
Photolysis rate JO1D (s ⁻¹)
Ozone (O ₃) mass-mixing ratio on pressure levels (kg kg ⁻¹)
Heaviside function on pressure levels
Flux through NO + RO ₂ reactions on pressure levels (mol s ⁻¹)

We have amended the text of the manuscript to read:

5 *The memory requirements can be lowered even more by turning off diagnostics and diagnostic sections that are not required for UKCA coupling to reduce the number of large three-dimensions arrays that need to be allocated, leaving a model configuration that only needs 4GB of memory.*

10 *Page 4, line 9: A short discussion of the differences in the ozone simulated by the two configurations would be appropriate here. To put the changes into context, perhaps these differences could be compared with the multi-model spread in any of the recent model intercomparison exercises.*

We do not believe that such a quantitative or qualitative comparison would be useful in the context of the purpose of this configuration and could in fact be misleading by implying that the fidelity of the output from the N48L38 configuration is

suitable for scientific studies. The changes to the model dynamics brought about by lowering of model resolution and model top mean that this configuration will rapidly deviate from a higher-resolution control. This low-resolution configuration is only able to be run for a few model hours, rather than the longer times that would be required for a meaningful study into model resolution, such as Stock *et al.* (2014). These results will also be highly biased by the initial conditions used.

5

Table 1: I found the "Configuration details" of the black and purple curves slightly confusing.

Instead of giving the number of timesteps in one model hour, it seems more natural to give the length of the timestep (eg. 20 minutes and 30 minutes).

10 We have amended the table to:

Table 1: Description of the curves presented in Figure 1.

Curve in Figure 1	UM resolution	Configuration details	Peak memory usage (GiB) 1×1 MPI processes without OpenMP
Green curve	N48L70 3.75°×2.5° with 70 vertical levels up to 80km	GA6 configuration with 27 dynamical timesteps (20-minute dynamical timestep).	1.62
Black curve	N48L70 3.75°×2.5° with 70 vertical levels up to 80km	GA7+StratTrop configuration with 3 dynamical timesteps and 1 chemical timestep (20-minute dynamical timestep, 1-hour chemical timestep).	12.13
Purple curve	N48L38 3.75°×2.5° with 38 vertical levels up to 40km	GA7+StratTrop configuration with 2 dynamical timesteps and 1 chemical timestep (30-minute dynamical timestep, 1-hour chemical timestep).	7.43
Blue curve	N48L38 3.75°×2.5° with 38 vertical levels up to 40km	GA7+StratTrop configuration with 2 dynamical timesteps and 1 chemical timestep (30-minute dynamical timestep, 1-hour chemical timestep). Only minimal diagnostic output is included.	6.84
Red curve	N48L38 3.75°×2.5° with 38 vertical levels up to 40km	GA7+StratTrop configuration with 2 dynamical timesteps and 1 chemical timestep (30-minute dynamical timestep, 1-hour chemical timestep). Only minimal diagnostic output is included and the CFMIP Observation Simulator Package (COSIP) has been disabled.	3.99

5 **Figure 2: By plotting the results of the full and reduced configurations next to each other, the reader is required to eyeball the two figures to get a sense of the differences between the simulations. For panels (b) and (d), I would prefer to see the plot of the full configuration as-is, and then a plot of the difference between the reduced and full**

versions. Also, what does "3-hour mean" mean in this context? Is this a single 3-hour period from one particular day? More information is needed here.

As mentioned, the changes in model dynamics mean that a quantitative difference between these two configurations is likely to be misleading. Due to computational limitations, the low-resolution N48L38 model is only designed to be run for 3 model hours, with all model output being given as 3-hour means. For comparison, the N96L85 configuration was sampled in the same way so that a 3-hour mean could be plotted for both.

These configurations should never be compared in a quantitative way – the low resolution configuration on the VM is designed only for model development, testing, and training, and should not be used for meaningful scientific studies. To emphasize the differences in resolution Figures 2b and 2d have been replaced with surface plots over Europe, and the caption has been updated. We have added the following text at the end of Section 4.1:

Here both models were run for 3 model hours, with output provided as a mean over this period.

Figure 2 is now:

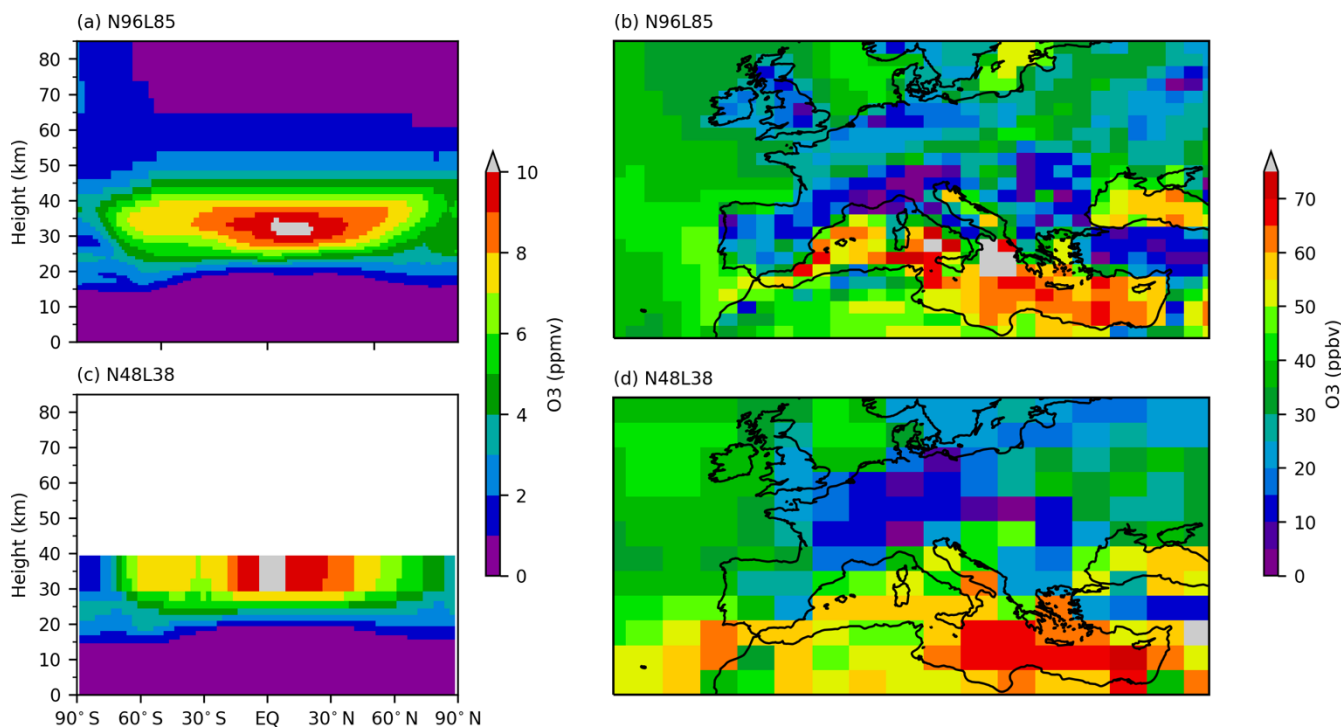


Figure 2: Comparison of modelled ozone from N96L85 and N48L38 UKCA configurations, highlighting the differences in model resolution and vertical extent. All plots are 3-hour means from the first 3 model hours of simulation. (a, b) N96L85 UM-UKCA configuration. (c, d) N48L38 UM-UKCA configuration.

Page 6, line 10: The first "to" should be "do".

The text now reads:

- 5 *When developing a change often what is most important is the speed of compilation and the over-all turn-around from “what changes do I need to implement?” to “the model runs successfully”.*

Page 11, line 22: What is "mule-cumf"?

- 10 As discussed in the reply to Reviewer #1, the “mule-cumf” tool allows for a bit-wise comparison of all fields contained in UM output files to be performed. However, in the context of this discussion these details are unnecessary, and so the reference to mule-cumf has been removed. The text now reads:

The output from these two jobs is then compared in the same way as for the KGO comparison.

15

Page 11, line 24: "of is desirable". I think you left out some words here.

Indeed, the text “MPI processes” is missing. The text now reads:

- 20 *Scientifically, maintaining results when changing the number (or decomposition) of MPI processes is desirable as it allows for the resources to be changed during long climate runs, allowing for more or less processes to be used without affecting the model output.*

- 25 ***Page 13, line 1: Can you provide some kind of link to the online tutorials? Ideally in a permanently-available archive? Or could you zip them and add them as a supplement to the paper?***

- All the references for the UKCA tutorials have DOIs that point to archived PDF files of the tutorial webpages that are held online in the University of Cambridge Apollo Repository (Abraham, 2013; Abraham et al., 2017; Abraham and Mann, 2014,
30 2016). The citation style for these references has been updated to include the live URL as well as the DOI, and these now read as:

Abraham, N. L.: UKCA & UMUI Tutorials for UM8.2, Online Learning Materials, doi:10.17863/CAM.22149 [online] Available from: http://www.ukca.ac.uk/wiki/index.php/UKCA_&_UMUI_Tutorials (Accessed 26 March 2018), 2013.

Abraham, N. L. and Mann, G. W.: UKCA Chemistry and Aerosol Tutorials for UM8.4, Online Learning Materials, doi:10.17863/CAM.22151 [online] Available from:

http://www.ukca.ac.uk/wiki/index.php/UKCA_Chemistry_and_Aerosol_Tutorials (Accessed 26 March 2018), 2014.

Abraham, N. L. and Mann, G. W.: UKCA Chemistry and Aerosol Tutorials at vn10.4 using Rose & Cylc, Online Learning Materials, doi:10.17863/CAM.22152 [online] Available from:

5 http://www.ukca.ac.uk/wiki/index.php/UKCA_Chemistry_and_Aerosol_Tutorials_at_vn10.4 (Accessed 26 March 2018), 2016.

Abraham, N. L., Bellouin, N. and Schmidt, A.: UKCA Chemistry and Aerosol Tutorials at vn10.9 using Rose & Cylc, Online Learning Materials, doi:10.17863/CAM.22153 [online] Available from:

10 http://www.ukca.ac.uk/wiki/index.php/UKCA_Chemistry_and_Aerosol_Tutorials_at_vn10.9 (Accessed 26 March 2018), 2017.

Additional Changes

In addition to the above changes, the following changes have also been made:

- 15
- The reference for RCMES has been updated from NASA Jet Propulsion Laboratory (2018) to Lee et al. (2018):

Lee, H., Goodman, A., McGibbney, L., Waliser, D., Kim, J., Loikith, P., Gibson, P. and Massoud, E.: Regional Climate Model Evaluation System powered by Apache Open Climate Workbench v1.3.0: an enabling tool for facilitating regional climate studies, Geosci. Model Dev. Discuss., 1–23, doi:10.5194/gmd-2018-113, 2018.

20

- The in-text citation for the version 10.9 tutorials has been changed slightly within the text of the Code Availability section, although the reference has remained the same

Abraham et al. (2017)

25

- The following has been added to the acknowledgments:

The work of MD was supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra.

References

30 Abraham, N. L.: UKCA & UMUI Tutorials for UM8.2, Online Learning Materials, doi:10.17863/CAM.22149 [online]

- Available from: http://www.ukca.ac.uk/wiki/index.php/UKCA_&_UMUI_Tutorials (Accessed 26 March 2018), 2013.
- Abraham, N. L. and Mann, G. W.: UKCA Chemistry and Aerosol Tutorials for UM8.4, Online Learning Materials, doi:10.17863/CAM.22151 [online] Available from: http://www.ukca.ac.uk/wiki/index.php/UKCA_Chemistry_and_Aerosol_Tutorials (Accessed 26 March 2018), 2014.
- 5 Abraham, N. L. and Mann, G. W.: UKCA Chemistry and Aerosol Tutorials at vn10.4 using Rose & Cylc, Online Learning Materials, doi:10.17863/CAM.22152 [online] Available from: http://www.ukca.ac.uk/wiki/index.php/UKCA_Chemistry_and_Aerosol_Tutorials_at_vn10.4 (Accessed 26 March 2018), 2016.
- Abraham, N. L., Bellouin, N. and Schmidt, A.: UKCA Chemistry and Aerosol Tutorials at vn10.9 using Rose & Cylc, Online Learning Materials, doi:10.17863/CAM.22153 [online] Available from: http://www.ukca.ac.uk/wiki/index.php/UKCA_Chemistry_and_Aerosol_Tutorials_at_vn10.9 (Accessed 26 March 2018), 2017.
- 10 Abraham, N. L., Archibald, A. T., Bellouin, N., Boucher, O., Braesicke, P., Bushell, A., Carslaw, K., Collins, B., Dalvi, M., Dennison, F., Emmerson, K., Folberth, G., Haywood, J., Hewitt, A., Johnson, C., Kipling, Z., Macintyre, H., Mann, G., Telford, P., Merikanto, J., Morgenstern, O., O'Connor, F., Ordóñez, C., Osprey, S., Pringle, K., Pyle, J., Rae, J., Reddington, C., Savage, N., Sellar, A., Spracklen, D., Stier, P., West, R., Mulcahy, J., Woodward, S., Boutle, I. and Woodhouse, M. T.: UMDP 084: United Kingdom Chemistry and Aerosol (UKCA) Technical Description. [online] Available from: https://code.metoffice.gov.uk/doc/um/vn11.0/papers/umdp_084.pdf (Accessed 23 February 2018), 2018.
- 15 Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A. and Woodward, S.: Development and evaluation of an Earth-System model – HadGEM2, *Geosci. Model Dev.*, 4(4), 1051–1075, doi:10.5194/gmd-4-1051-2011, 2011.
- 20 Cresswell, P.: Unified Model Documentation Paper X10: Unified Model Virtual Machine Guide. [online] Available from: https://code.metoffice.gov.uk/doc/um/vn11.0/papers/umdp_X10.pdf (Accessed 6 March 2018), 2018.
- 25 Lee, H., Goodman, A., McGibbney, L., Waliser, D., Kim, J., Loikith, P., Gibson, P. and Massoud, E.: Regional Climate Model Evaluation System powered by Apache Open Climate Workbench v1.3.0: an enabling tool for facilitating regional climate studies, *Geosci. Model Dev. Discuss.*, 1–23, doi:10.5194/gmd-2018-113, 2018.
- Matthews, D. P.: Vagrant virtual machines with FCM + Rose + Cylc installed, [online] Available from: <https://github.com/metomi/metomi-vms>, 2018.
- 30 NASA Jet Propulsion Laboratory: RCMES: Regional Climate Model Evaluation System, [online] Available from: <https://rcmes.jpl.nasa.gov/> (Accessed 27 March 2018), 2018.
- O'Connor, F. M., Johnson, C. E., Morgenstern, O., Abraham, N. L., Braesicke, P., Dalvi, M., Folberth, G. A., Sanderson, M. G., Telford, P. J., Voulgarakis, A., Young, P. J., Zeng, G., Collins, W. J. and Pyle, J. A.: Evaluation of the new UKCA climate-composition model-Part 2: The troposphere, *Geosci. Model Dev.*, 7(1), doi:10.5194/gmd-7-41-2014, 2014.

Stock, Z. S., Russo, M. R. and Pyle, J. A.: Representing ozone extremes in European megacities: the importance of resolution in a global chemistry climate model, *Atmos. Chem. Phys.*, 14(8), 3899–3912, doi:10.5194/acp-14-3899-2014, 2014.

U.K. Met Office: Met Office Science Repository Service, [online] Available from: <https://code.metoffice.gov.uk/trac/home/> (Accessed 10 April 2018), 2018.

Using a Virtual Machine environment for developing, testing, and training for the UM-UKCA Composition-Climate Model, using Unified Model version 10.9 and above

Nathan Luke Abraham^{1,2}, Alexander T. Archibald^{1,2}, Paul Cresswell³, Sam Cusworth³, Mohit Dalvi³, David Matthews³, Steven Wardle³, and Stuart Whitehouse³

1. Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, UK

2. National Centre for Atmospheric Science, UK

3. Met Office, FitzRoy Road, Exeter, EX1 3PB, UK

Correspondence to: N. Luke Abraham (luke.abraham@atm.ch.cam.ac.uk)

10 **Abstract.** The Met Office Unified Model (UM) is a state-of-the-art weather and climate model that is used operationally worldwide. UKCA is the chemistry and aerosol sub model of the UM that enables interactive composition and physical atmosphere interactions, but which adds an additional 120,000 lines of code to the model. Ensuring that the UM code and UM-UKCA (the UM running with interactive chemistry and aerosols) is well tested is therefore essential. While a comprehensive test harness is in place at the Met Office and partner sites to aid in development, this is not available to many UM users. Recently the Met Office have made available a Virtual Machine environment that can be used to run the UM on a desktop or laptop PC. Here we describe the development of a UM-UKCA configuration that is able to run within this Virtual Machine while only needing 6GB of memory, before discussing the applications of this system for model development, testing, and training.

1 Introduction

20 The Met Office Unified Model (UM; Cullen, 1993) is a state of the art general circulation model, that is used operationally worldwide for weather forecasts and climate projections. The UM comprises of over 965,000 lines of computer code, mainly written in Fortran 90/95, and a comprehensive test harness is in place at the Met Office and some partner sites such as Australia, New Zealand, South Korea, Poland, and elsewhere, to aid in development and to ensure that the code is well-tested, which is comprehensive but labour intensive to maintain. There are over 450 UM developers, and this testing system is not available to many of those that do not work at the Met Office. There are three UM releases per year, and code must be tested as part of the standard code development process. The UM code itself is hosted at the Met Office Science Repository Service (U.K. Met Office, 2018c), which is accessible to all users.

Recently the Met Office have provided a Virtual Machine (VM) configuration that can be used to run the UM on a desktop or laptop PC (Matthews, 2018). A VM is essentially an emulation of a complete and self-contained computer (or *guest*), running as software within another computer (or *host*) which virtualises the hardware that the guest operating system is installed on.

A similar system is a *container*, which will only include the software stack required to execute a program that is then portable between different computers without needing a complex compile and install process. These virtual environments are useful when using large and complex computer models as they are standardised and consistent between different host platforms, although changes to the underlying hardware means that results may not be identical between simulations on two different hosts using otherwise identical guest machines. As the environments are consistent and repeatable, these are often useful for training.

The Weather Research and Forecasting model (WRF; Skamarock et al., 2008) has been run on both VMs and containers (Hacker et al., 2017). Due to the standardised environment that they provide, VMs are also useful to enable analysis of model data by providing a consistent platform for analysis tools that is available to all researchers, such as that provided by the Regional Climate Model Evaluation System (RCMES), which has also been used for training users of this system ([NASA Jet Propulsion Laboratory, 2018](#))([Lee et al., 2018](#)). Here we describe the implementation of the United Kingdom Chemistry and Aerosols (UKCA) composition-climate model within the Met Office VM to better facilitate development and testing of new model code, and to provide a standardised environment for training new users in how to use UM-UKCA.

2 The Met Office Virtual Machine

The Met Office Virtual Machine makes use of VirtualBox (Oracle, 2018) and Vagrant (HashiCorp, 2018) to provision the VM and will automatically install the required packages that are needed to use the UM, namely FCM (U.K. Met Office, 2018a), Rose (U.K. Met Office, 2018d), and Cylc (NIWA, 2018). FCM is a build and version control system built around Subversion. [Rose is a system for creating, editing, and running application configurations](#) [Rose is a collection of tools that aim to improve the use of scientific computer models, both for research and operational use](#), and it is used as the graphical user interface for the UM to configure the input namelists. Cylc is a workflow tool that is used to schedule the various tasks needed to run an instance of the UM in the correct sequence. With these installed it is then possible to install the UM on the VM, and this process is documented in detail in UM documentation paper X10 (Cresswell, 2018). It is also possible to install the Met Office Iris Python library (U.K. Met Office, 2018b) on the VM, which is used to read and process UM output files. The time taken to install everything and get the UM up and running is dependent on host machine and internet connection speed, but can be as quick as 30 minutes.

Some simple UM configurations are provided that only require 2 cores and 3GB of memory in total, based around low resolution versions of the Global Atmosphere 6 (Walters et al., 2017a), limited area, or the single-column model configuration of the UM (Cullen, 1993). However, configurations using UKCA chemistry and aerosol schemes usually require much more memory due to the large number of three-dimensional fields used.

For this study we use an Ubuntu 16.04LTS host with 32 Intel Xeon 2.0GHz cores and 64GB of memory, and an Ubuntu 16.04LTS guest configured with 2, 4, 8, or 16 cores and up to 20GB of memory as required. The Met Office VM can run a number of different guest GNU/Linux operating systems, although Ubuntu 16.04LTS is the current recommended distribution.

It can also be used with GNU/Linux, macOS, and Windows hosts and will take up a maximum of 30GB of hard-disk space. There is only a single user on the VM (named *vagrant*), and for simplicity the UM is installed into this user's home directory, rather than to a central location on the VM's filesystem (as would need to be done on a HPC system). Due to license restrictions, only freely available Fortran and C compilers can be used on the VM. The GNU gfortran and gcc compilers are used as standard as they come with the guest Ubuntu distribution used by this system and have adequate feature support to compile the UM.

3 The UM-UKCA model

The United Kingdom Chemistry and Aerosols model is part of the Met Office Unified Model and provides a framework for adding aerosol and chemistry schemes to the UM. Aerosol microphysics is simulated by the GLOMAP-mode model (Mann et al., 2010), and this can be run with prescribed oxidant fields as in Global Atmosphere 7 (GA7; Walters et al., 2017b), or coupled interactively to the different UKCA chemistry schemes provided. The chemistry scheme used here is *StratTrop* (also known as *CheST*), which is a combination of the stratospheric chemistry scheme of Morgenstern et al., 2009 and the tropospheric chemistry scheme of O'Connor et al., 2014. This scheme has previously been used in standalone studies (Banerjee et al., 2014; Esenturk et al., 2018; Ferracci et al., 2018; Finney et al., 2016; Hardiman et al., 2017) and for the Chemistry Climate Model Initiative (Morgenstern et al., 2017), and will be used in the upcoming UK Earth system model, UKESM1. A full technical description of UKCA can be found in UM documentation paper 84 (Abraham et al., 2018). UKCA makes up over 12.5% of the UM code, at just over 121,000 lines.

Here we will use the *StratTrop* chemistry added to a GA7 configuration, which will require an additional 169 transported chemical and aerosol tracers and other 3D fields over the GA6 configuration currently provided on the VM, along with extra diagnostics calculated at run-time. Typical climate configurations of UKCA use the so-called N96L85 model resolution, $1.875^\circ \times 1.25^\circ$ with 85 vertical levels up to 85km. This configuration requires around 120GB of memory, and so running on a standard desktop is not feasible. However, it is possible to use a lower resolution, and turn-off high-memory sections of code, to allow a model configuration to fit within the specifications of a standard desktop, as described in Section 4.

4 Developing a low memory UKCA configuration

4.1 Changes to the model configuration

Figure 1 shows the development of a low-memory UKCA configuration that is suitable to be used on the VM, with further details provided in Table 1. Taking a N96L85 climate configuration based around the Global Atmosphere 7 configuration of the UM with UKCA *StratTrop* chemistry included, and re-gridding this to N48L70 ($3.75^\circ \times 2.5^\circ$ with 70 vertical levels up to 80km) allowed the required memory to be reduced to only 12GB. While this is a great improvement, it is still too much to allow the set-up to be used easily in a VM on a personal computer. The number of vertical levels can be reduced further to 38

(with a model top at 40km), as is used in the HadGEM2-ES configuration of the UM (Collins et al., 2011; O'Connor et al., 2014), which allows the model to only need 7.5GB of memory and allows the dynamical timestep to be increased from 20 to 30 minutes, also reducing the overall run-time which allows the model to only need 7.5GB. The memory requirements can be lowered even more by turning off unnecessary diagnostics and diagnostic sections that are not required for UKCA coupling to reduce the number of large three-dimensional arrays that need to be allocated, leaving a model configuration that only needs 4GB of memory. The other sharp peaks that are consistent between configurations are from sections such as dynamics, convection, or unavoidable calls to the diagnostics routines, and so cannot be turned off to reduce the memory requirements further.

When combined with the VM's operating system, using this 4GB configuration means that it is possible to run UM-UKCA with only 6GB required in total. This total memory requirement increases to 8GB if stricter compiler checks are used (see Section 5). Figure 2 compares UKCA ozone from a N96L85 climate configuration and from the low-memory N48L38 configuration that can be used on the VM. Here both models were run for 3 model hours, with output provided as a mean over this period.

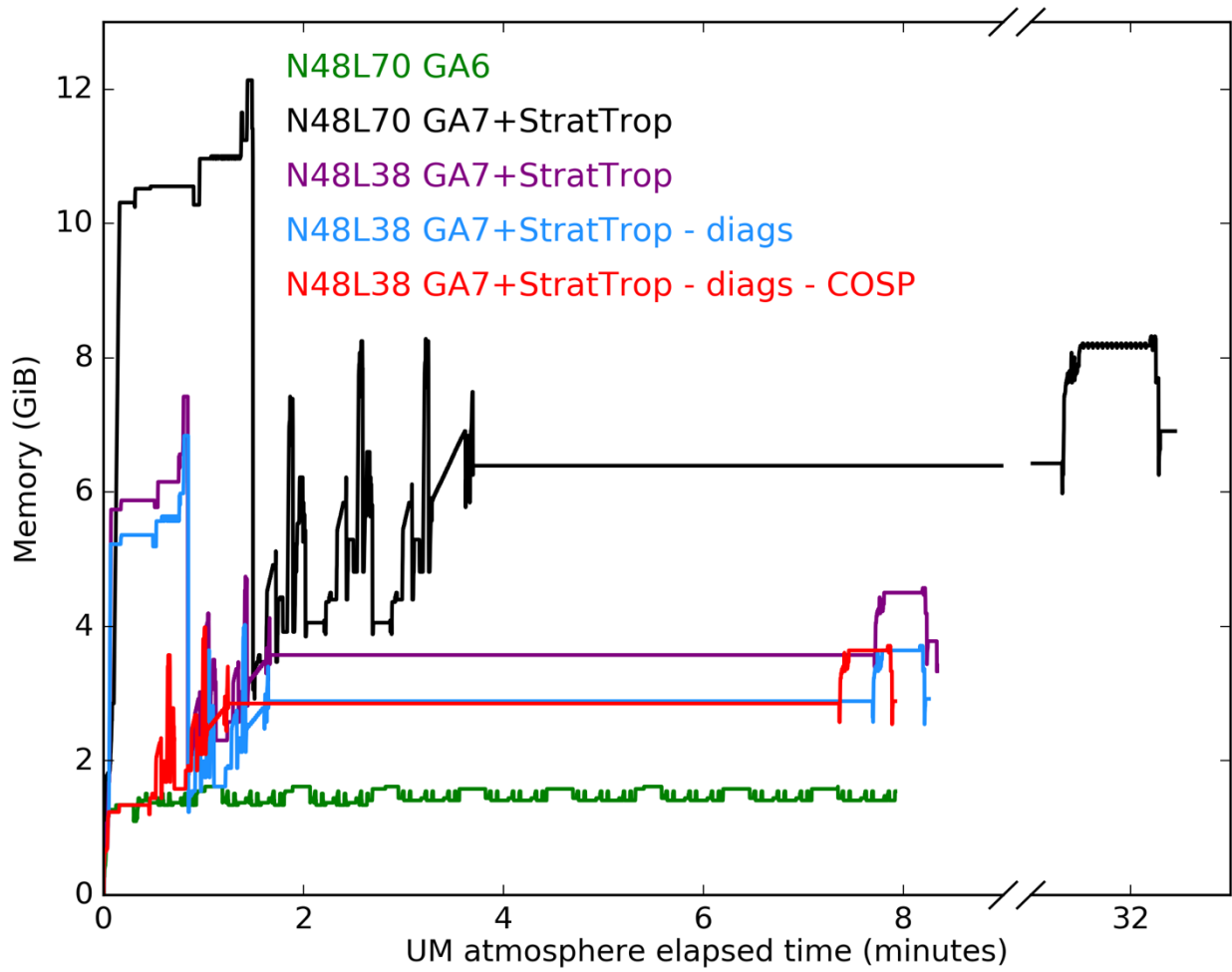


Figure 1: Memory usage of various UM configurations on the VM. Each simulation was performed using 1 MPI process without using OpenMP parallelisation. Further details of these different configurations are given in Table 1.

Table 1: Description of the curves presented in Figure 1.

Curve in Figure 1	UM resolution	Configuration details	Peak memory usage (GiB) 1×1 MPI processes without OpenMP
Green curve	N48L70 3.75°×2.5° with 70 vertical levels up to 80km	GA6 configuration with 27 dynamical timesteps (<u>9—model—hours20-minute dynamical timestep</u>).	1.62
Black curve	N48L70 3.75°×2.5° with 70 vertical levels up to 80km	GA7+StratTrop configuration with 3 dynamical timesteps and 1 chemical timestep (<u>1—model—hour20-minute dynamical timestep, 1-hour chemical timestep</u>).	12.13
Purple curve	N48L38 3.75°×2.5° with 38 vertical levels up to 40km	GA7+StratTrop configuration with 2 dynamical timesteps and 1 chemical timestep (<u>1—model—hour30-minute dynamical timestep, 1-hour chemical timestep</u>).	7.43
Blue curve	N48L38 3.75°×2.5° with 38 vertical levels up to 40km	GA7+StratTrop configuration with 2 dynamical timesteps and 1 chemical timestep (<u>1—model—hour30-minute dynamical timestep, 1-hour chemical timestep</u>). Only minimal diagnostic output is included.	6.84
Red curve	N48L38 3.75°×2.5° with 38 vertical levels up to 40km	GA7+StratTrop configuration with 2 dynamical timesteps and 1 chemical timestep (<u>1—model—hour30-minute dynamical timestep, 1-hour chemical timestep</u>). Only minimal diagnostic output is included and the CFMIP Observation Simulator Package (COSIP) has been disabled.	3.99

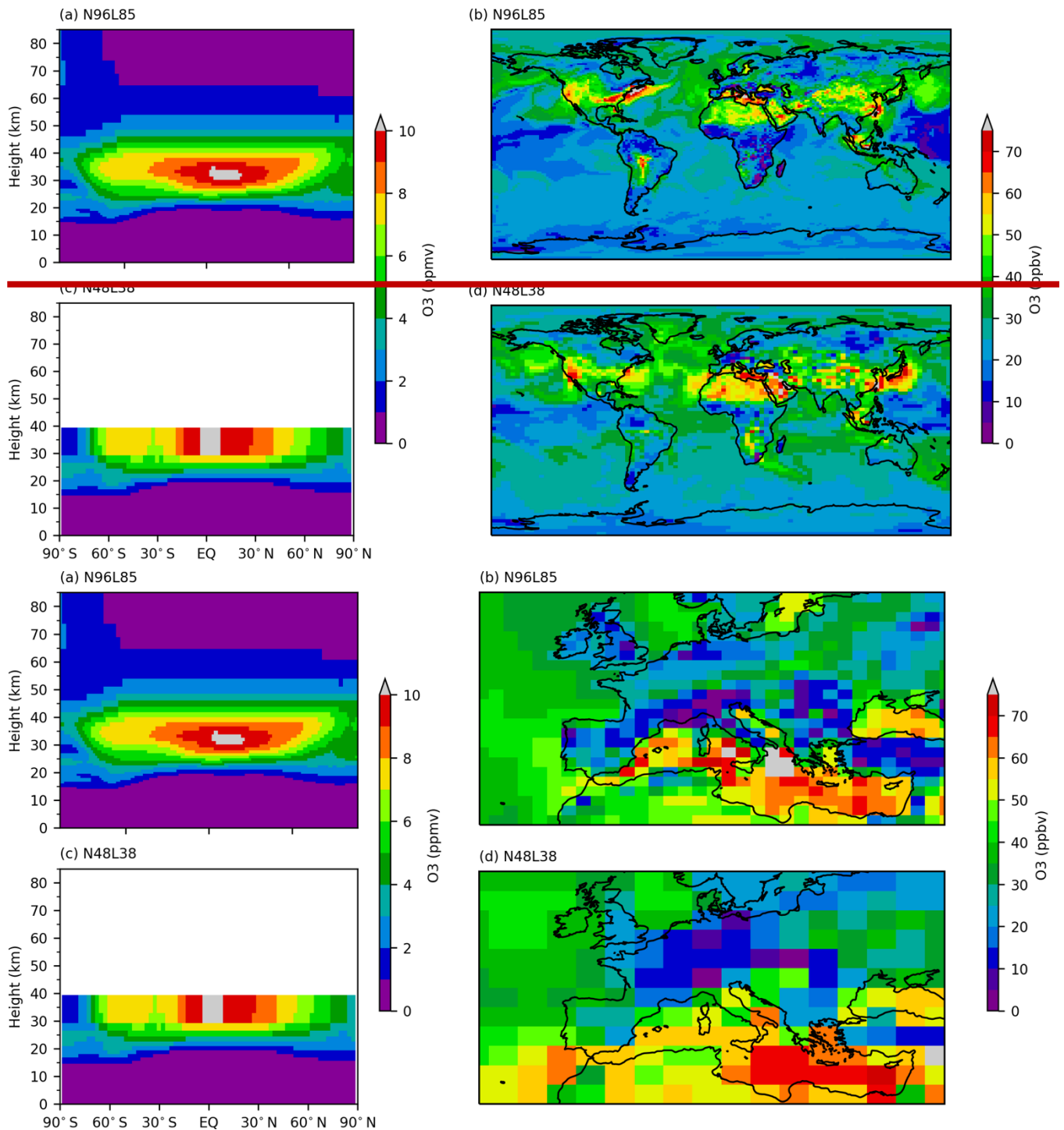


Figure 2: Comparison of modelled ozone from N96L85 and N48L38 UKCA configurations, highlighting the differences in model resolution and vertical extent. All plots are 3-hour means from the first 3 model hours of simulation. (a, b) 3-hour mean ozone from N96L85 UM-UKCA configuration. (c, d) 3-hour mean ozone from N48L38 UM-UKCA configuration.

4.2 Comparison to HPC implementation

Within the U.K., Academic researchers who use the UM have access to at least two supercomputers, ARCHER (a Cray XC30), and the Met Office XCS-C (a Cray XC40). Combined these have 57168 cores and 179136 GB of memory available for running the UM, and are both suitable for large production runs, such as for the CMIP6 historical and DECK experiments (Eyring et al., 2016), or for AerChemMIP (Collins et al., 2017).

However, not all work involves long and large simulations. When developing a change often what is most important is the speed of compilation and the over-all turn-around from “what changes tdo I need to implement?” to “the model runs successfully”.

Table 2 shows very approximate timings on these three different environments (ARCHER, the XCS-C, and the VM set-up described in Section 2) for the compile (including code extract), reconfiguration, and atmosphere tasks for the N48L38 UKCA configuration described in Section 4. Here we are not considering a detailed comparison of identical N48L38 UKCA configurations on different computing systems but are instead interested in the typical configuration that would be used on each system when making new code developments. Care should be taken with these results, as they are dependent on many factors that are impossible to isolate. The compile, reconfiguration, and atmosphere tasks may each use a different number of MPI processes on the three platforms, with different clock speeds and using different filesystems. Different compilers, compiler versions, and flags are used between the machines. Here the run-length has been increased to 3 model hours (6 dynamical timesteps, 3 chemical timesteps).

While the model run-time is fastest on the two supercomputers, the compile time is comparable on the VM. When queue times are considered the VM will may be a better option for model development jobs over ARCHER, especially as mMore cores can also be provided to the VM if available, further reducing run-times (see Table 4).

Throughput on the XCS-C is broadly better than on ARCHER and the VM, although both of these supercomputers are large and complex, and will always have delays due to planned (and unplanned) downtime for example. As the VM is a *single-user* environment that can be installed on a personal computer, the whole system is simpler and less error prone as there is no waiting, no queues, no filesystem I/O contention etc. This means that, for initial developments at least, the VM environment is likely to be the easiest and fastest to use due to the immediate turnaround.

5 Testing new model developments

Each working copy of the UM includes a complete Rose suite that can be used to test the UM or install it to a fixed location. This suite is used with the *rose-stem* utility, which is used to run such test suites; the test suite itself is known as *rose-stem* (Whitehouse, 2018) and contains a number of different applications, known as *apps*, which perform different tasks that can be automated. One of the functions of this suite is to install the UM itself, but it can also be used to run a series of known UM configurations and test them. As well as confirming that the tasks complete successfully, *rose-stem* is also able to compare the output from these configurations to already created standard output (also called *Known Good Output*, or KGO). This forms

part of the standard UM development process at the Met Office and several partner sites, as each site has its own rose-stem configuration, including the Virtual Machine. When run with a new development, rose-stem will enable a user to test that their code compiles and runs, and also whether it changes the results of known tests. Rose-stem is usually called from within a copy of the UM code; for installation this will be a copy of the UM trunk at a particular version (e.g. vn10.9), but for testing developments this will need to be a local copy containing the desired changes. It should be noted that if model output is no longer identical to the current KGO, this does not mean that the new output is not scientifically valid, just that the changes made to the code have also changed the results. Further testing and validation will be required to determine the scientific impact of these code changes.

Table 2: Rough timings for running UM version 10.9 N48L38 UKCA suite on ARCHER, the XCS-C, and the VM without using OpenMP. The UM could not be used with GNU gfortran on ARCHER as some dependencies as not been made available with this compiler.

Step	ARCHER (XC30)	XCS-C (XC40)	Virtual Machine
Cray <i>cce</i> initial compile	34 minutes	15 minutes	-
Cray <i>cce</i> incremental compile	5-7 minutes	3 minutes	-
Intel <i>ifort</i> initial compile	19 minutes	9 minutes	-
Intel <i>ifort</i> incremental compile	6-7 minutes	1 minute	-
GNU <i>gfortran</i> initial compile	-	4 minutes	8-10 minutes
GNU <i>gfortran</i> incremental compile	-	2-3 minutes	45 seconds
Reconfiguration task, used to produce the initial conditions file.	3-4 minutes (Intel, 6×4 processes)	15-30 seconds (GNU, 4×9 processes)	25-30 seconds (GNU, 1×2 processes)
Atmosphere task	40 seconds (Intel, 6×4 processes)	45 seconds (GNU, 4×9 processes)	12 minutes (GNU, 1×2 processes)

At UM version 11.0 there are 109 different configurations of the UM tested at the Met Office with rose-stem, with more at partner sites, which test different science, compiler, and machine settings. On the Virtual Machine there are only 12 UM tests, with 3 of these being UKCA configurations. On the VM the UKCA rose-stem jobs provide the following checks:

15

1. Test that new code compiles without errors when using both *safe* and *rigorous* compiler settings (see Table 3).
2. Tests that new code runs without errors when using both *safe* and *rigorous* compiler settings (see Table 3).
3. Tests whether new code changes results (this may be expected, depending on the change made).
4. Tests whether new code changes maintain results when using different processor decompositions.
5. Tests whether OpenMP parallelisation has been affected by new code changes, which then change results.

20

Table 3: UKCA rose-stem jobs available on the VM, showing the differences between Fortran compiler flags used for safe and rigorous compile options. C flags also differ. Both safe jobs point to the same KGO files.

Rose-stem job name within the ukca group	Compile type	GNU gfortran compiler flags	Number of UM OpenMP threads	Total VM memory required (GB)
vm_n48_ukca_eg_noomp	safe	-O2 -Werror	0	6
vm_n48_ukca_eg_omp_noios	safe	-O2 -Werror -fopenmp	2	6
vm_n48_ukca_eg_omp_noios_comp_check	rigorous	-O0 -Wall -ffpe-trap=invalid,zero -fbounds-check -Warray-bounds -fcheck-array-temporaries -finit-real=nan -fimplicit-none -fopenmp	2	8

A listing of the UKCA rose-stem jobs on the VM can be found in Table 3. Each of these is constructed using a *graph*, which links the different tasks of the job together by their dependencies. A simplified graph of these UKCA rose-stem tests at UM version 11.0 is shown in Figure 3. Additionally, equivalent GNU N48L38 tests to those on the VM have been implemented at the Met Office and are performed daily against the UM trunk. These test if any new code changes are likely to break the tests on the VM, effectively reducing the need for frequent VM rose-stem testing against the current and most recent UM code.

Table 4: Approximate timings for different UKCA configurations on the VM at UM version 11.0. The compiler settings are listed in detail in Table 3.

Compiler settings	Safe	Safe	Rigorous
Number of UM OpenMP threads	0	2	2
Approximate run-time on 2-core VM (1x2) (minutes)	8	11	29
Approximate run-time on 4-core VM (1x2) (minutes)	8	4	17
Approximate run-time on 4-core VM (1x4) (minutes)	5	8	22
Approximate run-time on 8-core VM (1x4) (minutes)	5	3	14
Approximate run-time on 8-core VM (1x8) (minutes)	3	6	26
Approximate run-time on 16-core VM (1x8) (minutes)	3	3	22

5.1 Known Good Output tests

In order to test against already produced *Known Good Output* (KGO), this KGO must be produced and installed into a standard location. At the Met Office, this step is performed manually when new code is committed to the UM trunk, and so a growing and continual series of KGO is maintained and available; this allows any unexpected change of results to be detected as changes are made to the trunk and facilitates the testing of changes to the UM at or near the latest trunk revision. However, for the VM this maintenance, while possible, is labour intensive and usually unnecessary. The VM is not designed as a system that should be maintained through UM versions, it is designed to be used for simple code development against a single, stable UM release. This is also important due to other possible changes to the UM, Rose, and Cylc systems that would also be time-consuming to maintain. On the VM the KGO for each rose-stem job must be generated and installed when the UM is installed as it is machine dependent. This can be done by using the command

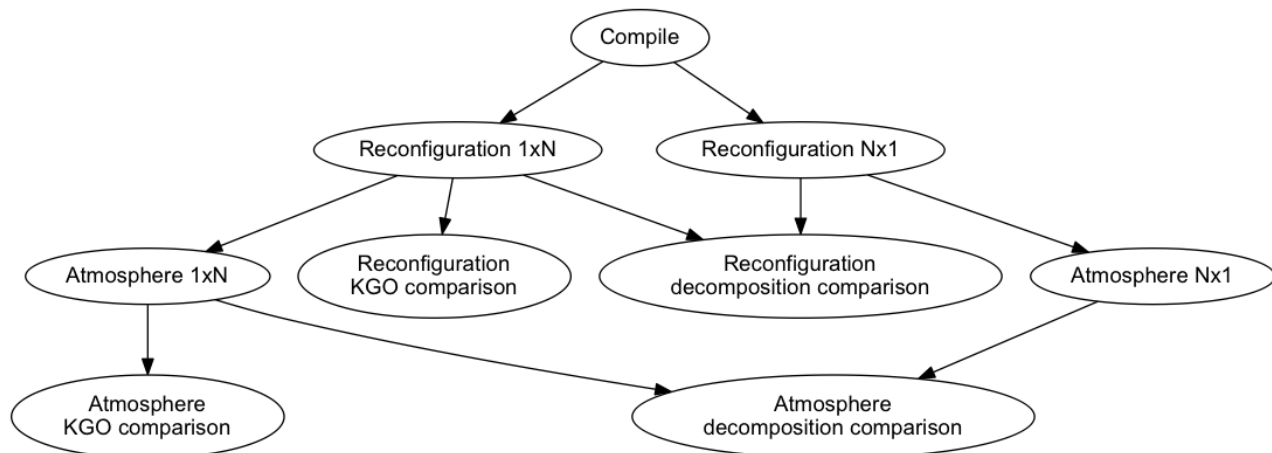


Figure 3: Simplified example rose-stem graph, showing the different tasks performed during a single job.

```
rose stem --group=ukca -S GENERATE_KGO=true
```

when running `rose-stem` on a copy of the UM trunk at the required version. The `ukca` term refers to the group of all the UKCA `rose-stem` jobs available on the VM (listed in Table 3).

5

Once the KGO for the UKCA jobs is in place, it is then possible to run the `rose-stem` tests on a local copy of the UM containing a code change by using the command

```
rose stem --group=ukca
```

10

within the top-level directory of the local copy of the UM code. It is also possible to run separate jobs individually by replacing the `ukca` above by the `rose-stem` job name (see Table 3). Once the various tasks are complete, `rose-stem` will automatically perform a bit-wise comparison between the UM output files generated by the reconfiguration and UM atmosphere steps and will highlight any differences between the KGO files and the files produced from the `rose-stem` jobs. Often major science changes or bugfixes will result in differences, but small changes or new schemes that are turned off should usually preserve the model evolution and hence the KGO tests should pass. If these KGO tests do not pass when they would be expected to do so, this may point to an error in the implementation.

15

5.2 Processor decomposition tests

From UM version 11.0, when the flag `-S INTEGRATION_TESTING=true` is passed to `rose-stem`, the UKCA jobs can be run in two different configurations to enable processor decomposition testing. Here the UM will use either $1 \times N$ processes or $N \times 1$ processes. By default, $N=2$, but this can be changed by adding `-S MPI_TASKS=N` on the call to `rose-stem`. The output from these two jobs is then compared [using mule-cumf](#) in the same way as for the KGO comparison. By increasing the value of N it is possible to reduce the run-time of these configurations, as seen in Table 4.

20

Scientifically, maintaining results when changing the number (or decomposition) of [MPI processes](#) is desirable as it allows for the resources to be changed during long climate runs, allowing for more or less [MPI](#) processes to be used without affecting the model output. Technically, it helps test for code correctness. There may, for example, be bounds errors or race conditions that would never be encountered if only one decomposition is routinely tested, but that may show up if the run conditions are varied. By ensuring several decompositions give the same answer we reduce the risk of bugs in the code; something that changes answers in a second decomposition may cause a crash in third or produce nonsense.

25

Additionally, any impacts on OpenMP parallelisation can be tested by running job `vm_n48_ukca_eg_omp_noios`, as the KGO is produced using job `vm_n48_ukca_eg_noomp` and should be identical for both.

30

6 Training new users

When training new users of UM-UKCA how to use the model, we begin with the premise “*what are the most common things that a new PhD student or researcher would need to know how to do in order to use UKCA?*”. When starting with UM-UKCA, often new users will wish to perform studies that will involve a combination of the following:

5

- Creating or modifying emissions
- Adding new chemical species
- Adding new reactions
- Adding new deposition processes

10

- Adding new diagnostics

as well as needing to learn how to output and process UM and UKCA diagnostics, and also wanting to become more familiar with the model’s user-interface.

We can in fact cover all these points by imagining two new chemical species, ALICE and BOB, and the chemical equation

15



where *Secondary Organic Species* will condense to form secondary organic aerosol (Mann et al., 2010), ALICE is dry deposited, and BOB is wet deposited. The specifics of the emission field, reaction and deposition rates are unimportant, so long as they have reasonable values.

20

The steps to include the toy chemistry above to UM-UKCA can then be broken down into the following tasks:

1. Add ALICE and BOB as transported tracers to the UM and UKCA
2. Add surface emissions of ALICE, comprising:
 - a. Re-gridding of a provided $1^\circ \times 1^\circ$ dataset to N48 resolution
 - b. Saving this re-gridded emissions field to NetCDF with the correct metadata required by UKCA
 - c. Including this new NetCDF emissions file in the UKCA namelist within Rose
3. Add the bimolecular reaction R1 to the UKCA StratTrop chemical mechanism
4. Add the dry-deposition of ALICE to UKCA
5. Add the wet-deposition of BOB to UKCA
6. Add new diagnostic reaction fluxes to UM-UKCA to output the fluxes through steps introduced in 3, 4, and 5 above
7. Process UM output to calculate aerosol optical depth

25

30

To aid new users of UM-UKCA a series of online tutorials have been developed that go through the various (and complex) steps above required to make changes to the model. Detailed instructions are available for UM versions 8.2 (Abraham, 2013), 8.4 (Abraham and Mann, 2014), 10.4 (Abraham and Mann, 2016), and 10.9 (Abraham et al., 2017) using ARCHER. Additionally, at UM version 10.9 this UKCA training can be performed exclusively on the VM using a configuration based on the N48L38 model described in Section 4. Figure 4 shows typical output from the above tasks after they have been completed.

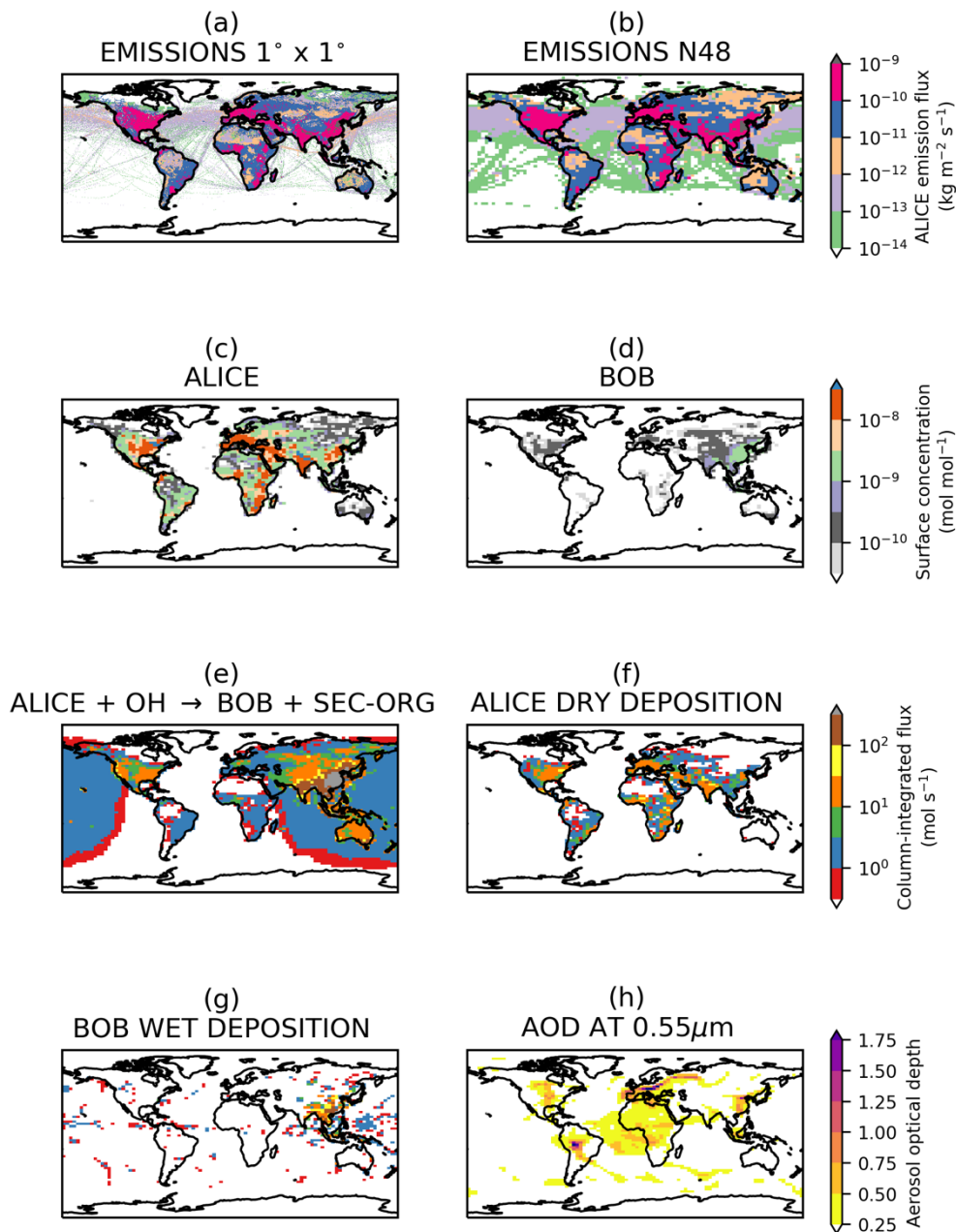


Figure 4: Showing output from the various UKCA tutorial tasks. (a) $1^\circ \times 1^\circ$ supplied September emissions. (b) Emissions re-gridded to N48. (c) 3-hour mean ALICE surface mole fraction after the tutorials have been completed. (d) as for (c), but for BOB. (e) 3-hour mean column-integrated flux through reaction R1. (f) 3-hour mean column-integrated ALICE dry deposition flux. (g) 3-hour mean column-integrated BOB dry deposition flux. (h) 3-hour mean total Aerosol Optical Depth from GLOMAP-mode at $0.55\mu\text{m}$. ALICE and BOB were both initialised to $1.0 \times 10^{22} \text{ kg kg}^{-1}$.

7 Conclusions

As can be seen in Table 2, the UM-UKCA configurations on the VM require longer runtimes than equivalent configurations on HPC systems. Using 2 MPI processes the N48L38 configuration would take at least 32 hours to simulate one model month, although using a server as a host and adding more processes reduces this as seen in Table 4.

5 The decrease in horizontal resolution will be somewhat detrimental to the representation of the model dynamics and other physical or chemical processes (Stock et al., 2014; Strachan et al., 2013), although many UKCA studies have been performed using N48 resolution (Banerjee et al., 2014; Bednarz et al., 2016; Keeble et al., 2017; Nowack et al., 2014). However, the reduction in the number of vertical levels (see Figure 2) is a significant issue as the chemistry scheme used, StratTrop, is designed to simulate both stratospheric and tropospheric chemistry, and without a full model stratosphere it is computationally
10 inefficient to include all the reactions in this scheme. Also, due to this low top, the look-up table photolysis scheme, used in the region above the Fast-JX photolysis scheme (Telford et al., 2013), will not be called.

The chemistry scheme used in HadGEM2-ES (O'Connor et al., 2014) would be appropriate for this level structure, but this would then not test the equivalent settings to the N96L85 GA7+StratTrop configuration, and so the usefulness of the tests listed in Table 4 would be reduced. By keeping as many settings as possible the same between these configurations it allows
15 the VM tests to be traceable to N96L85 tests performed on the Met Office systems.

While the VM is ideal for initial development and testing of new code, once long science integrations need to be done it is necessary to switch to a standard N96L85 climate configuration on a HPC system. Simulations of around 20-years in length are usually required to see if any new science introduces a significant change to stratospheric composition, and these are not feasible to be performed on the VM. The VM could however be used for stand-alone scientific studies making use of the
20 UKCA Box Model (Esenturk et al., 2018) or UKCA called from the Single Column Model, although further work would be required to allow these configurations to work on the VM.

While the toy chemistry discussed in Section 6 does not cover all reaction types covered and may be considered simpler than most changes new users may wish to make, it is sufficiently detailed enough to give users experience with many different parts of the UKCA code. While using the VM will give users experience of using and editing the UM's Rose graphical user interface,
25 it will not provide experience in using UM-UKCA on a HPC system. However, this is also an advantage, as it means that it is possible to gain experience in using and developing UM-UKCA before committing supercomputer resources.

Despite these limitations, the VM is an easy-to-use system that is available to all current and potential users, and creates a consistent environment for model development, testing, and training. This system is complementary to, rather than being a replacement for, using UM-UKCA on HPC systems.

30 Code Availability

[The Met Office Virtual Machine can be obtained from https://github.com/metomi/metomi-vms \(Matthews, 2018\).](https://github.com/metomi/metomi-vms) Due to intellectual property right restrictions, we cannot provide either the source code or documentation papers for the UM.

Obtaining the UM. The Met Office Unified Model is available for use under licence. The functionality discussed here is fully available in the UM trunk from version 11.0, with [detailed](#) instructions on [using how to install the UM on the VM](#) provided by UMDP X10 (Cresswell, 2018), and the UM version 10.9 UKCA Tutorials are described in [Abraham et al., 2018a](#) [Abraham et al. \(2017\)](#). A number of research organisations and national meteorological services use the UM in collaboration with the Met Office to undertake basic atmospheric process research, produce forecasts, develop the UM code and build and evaluate Earth system models. For further information on how to apply for a licence see <http://www.metoffice.gov.uk/research/modelling-systems/unified-model>.

Acknowledgements

10 This work used the ARCHER UK National Supercomputing Service (<http://www.archer.ac.uk>). This work used Monsoon2, a collaborative High-Performance Computing facility funded by the Met Office and the Natural Environment Research Council. This work used the NEXCS High Performance Computing facility funded by the Natural Environment Research Council and delivered by the Met Office.

UKCA training has been supported by the NERC ACITIES atmospheric chemistry modelling network, grant number 15 NE/K001280/1 and NERC Advanced Training Short Courses scheme, grant numbers NE/M006220/1, NE/N000129/1, NE/N019091/1, and NE/P020089/1. We would like to thank the UM Systems Team at the Met Office. NLA would like to thank John Pyle, Bryan Lawrence, and NCAS Computational Modelling Services for supporting this work. [The work of MD was supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra.](#)

20 *Copyright statement.* The works published in this journal are distributed under the Creative Commons Attribution 4.0 License. This licence does not affect the Crown copyright work, which is re-usable under the Open Government Licence (OGL). The Creative Commons Attribution 4.0 License and the OGL are interoperable and do not conflict with, reduce or limit each other.

© Crown copyright 2018

25 References

Abraham, N. L.: UKCA & UMUI Tutorials for UM8.2, Online Learning Materials, doi:10.17863/CAM.22149 [online] Available from: http://www.ukca.ac.uk/wiki/index.php/UKCA_&_UMUI_Tutorials (Accessed 26 March 2018), 2013.

Abraham, N. L. and Mann, G. W.: UKCA Chemistry and Aerosol Tutorials for UM8.4, Online Learning Materials, doi:10.17863/CAM.22151 [online] Available from: http://www.ukca.ac.uk/wiki/index.php/UKCA_Chemistry_and_Aerosol_Tutorials (Accessed 26 March 2018), 2014.

- Abraham, N. L. and Mann, G. W.: UKCA Chemistry and Aerosol Tutorials at vn10.4 using Rose & Cylc, Online Learning Materials, doi:10.17863/CAM.22152 [online] Available from: http://www.ukca.ac.uk/wiki/index.php/UKCA_Chemistry_and_Aerosol_Tutorials_at_vn10.4 (Accessed 26 March 2018), 2016.
- 5 Abraham, N. L., Bellouin, N. and Schmidt, A.: UKCA Chemistry and Aerosol Tutorials at vn10.9 using Rose & Cylc, Online Learning Materials, doi:10.17863/CAM.22153 [online] Available from: http://www.ukca.ac.uk/wiki/index.php/UKCA_Chemistry_and_Aerosol_Tutorials_at_vn10.9 (Accessed 26 March 2018), 2017.
- Abraham, N. L., Archibald, A. T., Bellouin, N., Boucher, O., Braesicke, P., Bushell, A., Carslaw, K., Collins, B., Dalvi, M.,
- 10 Dennison, F., Emmerson, K., Folberth, G., Haywood, J., Hewitt, A., Johnson, C., Kipling, Z., Macintyre, H., Mann, G., Telford, P., Merikanto, J., Morgenstern, O., O'Connor, F., Ordóñez, C., Osprey, S., Pringle, K., Pyle, J., Rae, J., Reddington, C., Savage, N., Sellar, A., Spracklen, D., Stier, P., West, R., Mulcahy, J., Woodward, S., Boutle, I. and Woodhouse, M. T.: UMDP 084: United Kingdom Chemistry and Aerosol (UKCA) Technical Description. [online] Available from: https://code.metoffice.gov.uk/doc/um/vn11.0/papers/umdp_084.pdf (Accessed 23 February 2018), 2018.
- 15 Banerjee, A., Archibald, A. T., Maycock, A. C., Telford, P., Abraham, N. L., Yang, X., Braesicke, P. and Pyle, J. A.: Lightning NO_x, a key chemistry-climate interaction: Impacts of future climate change and consequences for tropospheric oxidising capacity, *Atmos. Chem. Phys.*, 14(18), doi:10.5194/acp-14-9871-2014, 2014.
- Bednarz, E. M., Maycock, A. C., Abraham, N. L., Braesicke, P., Dessens, O. and Pyle, J. A.: Future Arctic ozone recovery: the importance of chemistry and dynamics, *Atmos. Chem. Phys.*, 16(18), 12159–12176, doi:10.5194/acp-16-12159-2016,
- 20 2016.
- Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A. and Woodward, S.: Development and evaluation of an Earth-System model – HadGEM2, *Geosci. Model Dev.*, 4(4), 1051–1075, doi:10.5194/gmd-4-1051-2011, 2011.
- 25 Collins, W. J., Lamarque, J.-F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., Shindell, D. and Smith, S. J.: AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6, *Geosci. Model Dev.*, 10(2), 585–607, doi:10.5194/gmd-10-585-2017, 2017.
- Cresswell, P.: Unified Model Documentation Paper X10: Unified Model Virtual Machine Guide. [online] Available from: https://code.metoffice.gov.uk/doc/um/vn11.0/papers/umdp_X10.pdf (Accessed 6 March 2018), 2018.
- 30 Cullen, M. J. P.: The unified forecast/climate model, *Meteorol. Mag.*, 122(1449), 81–94, 1993.
- Esenturk, E., Abraham, L., Archer-Nicholls, S., Mitsakou, C., Griffiths, P., Archibald, A. and Pyle, J.: Quasi-Newton Methods for Atmospheric Chemistry Simulations: Implementation in UKCA UM Vn10.8, *Geosci. Model Dev. Discuss.*, 1–31, doi:10.5194/gmd-2018-32, 2018.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J. and Taylor, K. E.: Overview of the Coupled

- Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev.*, 9(5), 1937–1958, doi:10.5194/gmd-9-1937-2016, 2016.
- Ferracci, V., Heimann, I., Abraham, N. L., Pyle, J. A. and Archibald, A. T.: Global modelling of the total OH reactivity: investigations on the “missing” OH sink and its atmospheric implications, *Atmos. Chem. Phys. Discuss.*, 1–33, doi:10.5194/acp-2018-12, 2018.
- 5 Finney, D. L., Doherty, R. M., Wild, O. and Abraham, N. L.: The impact of lightning on tropospheric ozone chemistry using a new global lightning parametrisation, *Atmos. Chem. Phys.*, 16(12), doi:10.5194/acp-16-7507-2016, 2016.
- Hacker, J. P., Exby, J., Gill, D., Jimenez, I., Maltzahn, C., See, T., Mullendore, G. and Fossell, K.: A Containerized Mesoscale Model and Analysis Toolkit to Accelerate Classroom Learning, Collaborative Research, and Uncertainty Quantification, *Bull. Am. Meteorol. Soc.*, 98(6), 1129–1138, doi:10.1175/BAMS-D-15-00255.1, 2017.
- 10 Hardiman, S. C., Butchart, N., O’Connor, F. M. and Rumbold, S. T.: The Met Office HadGEM3-ES chemistry–climate model: evaluation of stratospheric dynamics and its impact on ozone, *Geosci. Model Dev.*, 10(3), 1209–1232, doi:10.5194/gmd-10-1209-2017, 2017.
- HashiCorp: Vagrant, [online] Available from: <https://www.vagrantup.com/> (Accessed 27 March 2018), 2018.
- 15 Keeble, J., Brown, H., Abraham, N. L., Harris, N. R. P. and Pyle, J. A.: On ozone trend detection: using coupled chemistry–climate simulations to investigate early signs of total column ozone recovery, *Atmos. Chem. Phys. Discuss.*, 1–17, doi:10.5194/acp-2017-1206, 2017.
- Lee, H., Goodman, A., McGibbney, L., Waliser, D., Kim, J., Loikith, P., Gibson, P. and Massoud, E.: Regional Climate Model Evaluation System powered by Apache Open Climate Workbench v1.3.0: an enabling tool for facilitating regional climate studies, *Geosci. Model Dev. Discuss.*, 1–23, doi:10.5194/gmd-2018-113, 2018.
- 20 Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., Pickering, S. J. and Johnson, C. E.: Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition–climate model, *Geosci. Model Dev.*, 3(2), 519–551, doi:10.5194/gmd-3-519-2010, 2010.
- Matthews, D. P.: Vagrant virtual machines with FCM + Rose + Cylc installed, [online] Available from: <https://github.com/metomi/metomi-vms>, 2018.
- 25 Morgenstern, O., Braesicke, P., O’Connor, F. M., Bushell, A. C., Johnson, C. E., Osprey, S. M. and Pyle, J. A.: Evaluation of the new UKCA climate–composition model – Part 1: The stratosphere, *Geosci. Model Dev.*, 2(1), 43–57, doi:10.5194/gmd-2-43-2009, 2009.
- Morgenstern, O., Hegglin, M. I., Rozanov, E., O’Connor, F. M., Abraham, N. L., Akiyoshi, H., Archibald, A. T., Bekki, S., Butchart, N., Chipperfield, M. P., Deushi, M., Dhomse, S. S., Garcia, R. R., Hardiman, S. C., Horowitz, L. W., Jöckel, P., Josse, B., Kinnison, D., Lin, M., Mancini, E., Manyin, M. E., Marchand, M., Marécal, V., Michou, M., Oman, L. D., Pitari, G., Plummer, D. A., Revell, L. E., Saint-Martin, D., Schofield, R., Stenke, A., Stone, K., Sudo, K., Tanaka, T. Y., Tilmes, S., Yamashita, Y., Yoshida, K. and Zeng, G.: Review of the global models used within phase 1 of the Chemistry–Climate Model Initiative (CCMI), *Geosci. Model Dev.*, 10(2), 639–671, doi:10.5194/gmd-10-639-2017, 2017.

- NIWA: Cylc - a workflow engine, [online] Available from: <http://cylc.github.io/cylc/> (Accessed 27 March 2018), 2018.
- Nowack, P. J., Luke Abraham, N., Maycock, A. C., Braesicke, P., Gregory, J. M., Joshi, M. M., Osprey, A. and Pyle, J. A.: A large ozone-circulation feedback and its implications for global warming assessments, *Nat. Clim. Chang.*, 5, 41, 2014.
- O'Connor, F. M., Johnson, C. E., Morgenstern, O., Abraham, N. L., Braesicke, P., Dalvi, M., Folberth, G. A., Sanderson, M. G., Telford, P. J., Voulgarakis, A., Young, P. J., Zeng, G., Collins, W. J. and Pyle, J. A.: Evaluation of the new UKCA climate-composition model-Part 2: The troposphere, *Geosci. Model Dev.*, 7(1), doi:10.5194/gmd-7-41-2014, 2014.
- Oracle: VirtualBox, [online] Available from: <https://www.virtualbox.org/> (Accessed 27 March 2018), 2018.
- Skamarock, C., Klemp, B., Dudhia, J., Gill, O., Barker, D., Duda, G., Huang, X., Wang, W. and Powers, G.: A Description of the Advanced Research WRF Version 3, , doi:10.5065/D68S4MVH, 2008.
- 10 Stock, Z. S., Russo, M. R. and Pyle, J. A.: Representing ozone extremes in European megacities: the importance of resolution in a global chemistry climate model, *Atmos. Chem. Phys.*, 14(8), 3899–3912, doi:10.5194/acp-14-3899-2014, 2014.
- Strachan, J., Vidale, P. L., Hodges, K., Roberts, M. and Demory, M.-E.: Investigating global tropical cyclone activity with a hierarchy of AGCMs: the role of model resolution, *J. Clim.*, 26(1), 133–152, 2013.
- Telford, P. J., Abraham, N. L., Archibald, A. T., Braesicke, P., Dalvi, M., Morgenstern, O., O'Connor, F. M., Richards, N. A. D. and Pyle, J. A.: Implementation of the Fast-JX Photolysis scheme (v6.4) into the UKCA component of the MetUM chemistry-climate model (v7.3), *Geosci. Model Dev.*, 6(1), 161–177, doi:10.5194/gmd-6-161-2013, 2013.
- 15 U.K. Met Office: FCM: Flexible Configuration Management, [online] Available from: <http://metomi.github.io/fcm/doc/> (Accessed 27 March 2018a), 2018.
- U.K. Met Office: Iris: A Python library for Meteorology and Climatology, [online] Available from: <http://scitools.org.uk/iris/index.html> (Accessed 27 March 2018b), 2018.
- 20 U.K. Met Office: Met Office Science Repository Service, [online] Available from: <https://code.metoffice.gov.uk/trac/home/> (Accessed 10 April 2018c), 2018.
- U.K. Met Office: Rose: a framework for managing and running meteorological suites, [online] Available from: <http://metomi.github.io/rose/doc/rose.html> (Accessed 27 March 2018d), 2018.
- 25 Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S., Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P. L., Wilkinson, J., Willett, M., Woolnough, S. and Xavier, P.: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations, *Geosci. Model Dev.*, 10(4), 1487–1520, doi:10.5194/gmd-10-1487-2017, 2017a.
- 30 Walters, D., Baran, A., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whittall, M., Williams, K. and Zerroukat, M.: The Met Office Unified Model

Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, *Geosci. Model Dev. Discuss.*, 1–78, doi:10.5194/gmd-2017-291, 2017b.

Whitehouse, S.: Unified Model Documentation Paper X09: The UM rose-stem Suite for External Users. [online] Available from: https://code.metoffice.gov.uk/doc/um/vn11.0/papers/umdp_X09.pdf (Accessed 16 February 2018), 2018.