Sprint Documentation #09

LAGOOn[i] Sprint

Catrin Meyer[1], Felix Plöger[2], Lars Hoffmann[1], Jan Clemens[2], Nicole Thomas[2], Rolf Müller[2]

[1]Jülich Supercomputing Center, Forschungszentrum Jülich GmbH, Jülich, Germany
[2]Institute of Climate and Energy Systems (ICE-4), Stratosphere, Forschungszentrum Jülich GmbH, Jülich, Germany

Contact: info@nat-esm.de

Published on 05.12.2024 on https://www.nat-esm.de/services/accepted-sprint

1 Summary

The LAGOOn sprint focused on combining two Lagrangian trajectory models, CLaMS-traj and MPTRAC, to enhance the computational efficiency of CLaMS (the Chemical Lagrangian Model of the Stratosphere) by utilizing MPTRAC's (MPTRAC = Massive-Parallel Trajectory Calculations) GPU-ready capabilities. The responsible RSE, in interaction with the sprint team, developed a wrapper to enable the use of MPTRAC's C functionality within CLaMS, which is written in Fortran. It will also allow CLaMS to run on GPU-based systems with the OpenACC implementation of MPTRAC. Moreover, the RSE with the team created a standalone Fortran application, tested across various systems, and successfully integrated into the MPTRAC repository with added test cases and documentation (https://github.com/slcs-jsc/mptrac). This sprint laid the groundwork for future GPU optimization of CLaMS and serves as a blueprint for combining legacy Fortran code with modern C-based components. Both models were added to the natESM system as optional components. The result is a significant step forward towards improving performance and interoperability between older and newer code generations, with future modules now able to be ported to GPUs.

2 General Information

Start and end date:	16.10.2023 – 30.06.2024
Intended period:	6 FTE months (9 months with part time (2/3) FTE)
Responsible RSE:	Catrin Meyer (JSC)
Responsible scientist:	Felix Plöger (ICE-4)

Lagrangian transport models track individual air parcels as they move through the atmosphere, considering turbulence, wind patterns, and other factors influencing the motion of atmospheric particles. These models generally simulate the behavior of large numbers of particles and can provide a detailed picture of how trace gases, aerosol particles, and clouds move and disperse in the atmosphere. Lagrangian transport models are used extensively in atmospheric science to study the

[i] CLaMS and MPTRAC: A Lagrangian toolkit for the natESM strategy (LAGOOn)

transport, dispersion, and transformation of atmospheric constituents, such as pollutants, trace gases, and aerosol particles. The models are used for various applications, including studies of atmospheric chemistry, air pollution, cloud microphysics, volcanic ash dispersion, and long-range transport processes. In particular regarding the representation of small-scale transport processes and structures (filaments), Lagrangian transport models are advantageous compared to Eulerian models because they can better resolve the processes and structures.

In the German community, several Lagrangian codes are in frequent use: the Lagrangian transport model EMAC–ATTILA (ATTILA; Atmospheric Tracer Transport in a LAgrangian model), (Brinkop and Jöckel, 2019), the model MPTRAC (Hoffmann et al., 2016, 2022*)* and the Chemical Lagrangian model of the Stratosphere (CLaMS, McKenna et al., 2002a,b; Konopka et al., 2007; Pommrich et al., 2014). Further Lagrangian models used in the community are ATLAS (Wohltmann et al., 2010), which resembles CLaMS in many respects, LAMETTA, FLEXPART (https://imgw.univie.ac.at/forschung/atmosphaerische-transportprozesse/transportmodell-flexpart/) and HYSPLIT (Stein et al. 2015). The focus of this sprint was on MPTRAC and CLaMS-traj; in particular on the coupling of the numerically efficient and GPU-ready Lagrangian advection of MPTRAC into the CLaMS model framework.

The Chemical Lagrangian Model of the Stratosphere (CLaMS) and Massive-Parallel Trajectory Calculations (MPTRAC) are two Lagrangian trajectory codes for simulating atmospheric transport. The CLaMS-traj code, written in Fortran, has been expanded since its development in the 1990s to include dedicated modules for chemistry, atmospheric dynamics, and cloud physics. The development of MPTRAC started in the 2010s. It is written in C and special emphasis was put on the highly parallel implementation of the algorithms, so that it is capable of fully exploiting the capabilities of modern GPU-based HPC systems by utilizing the OpenACC directive-based programming approach. To start porting the CLaMS model system to GPUs, we explored the possibility of replacing the CLaMS advection (trajectory) core with the MPTRAC advection scheme. During this sprint, both codes were analyzed, and the sprint team developed a concept. As part of the concept, the RSE implemented a wrapper using ISO_C_BINDINGS to call the MPTRAC C-routines from Fortran. The RSE developed a Fortran standalone application and tested it successfully with multiple compilers and multiple systems (Laptop and HPC systems). A test suite and documentation were added by the RSE with input from the sprint team. This establishes a foundation for porting CLaMS to GPUs, which can now proceed module by module. Moreover, this example is a good demonstrator how to combine modules from different code generations together for efficient future use.

The Chemical Lagrangian Model of the Stratosphere (CLaMS) is a modular model system which can be applied to pure trajectory and to 3-dimensional global simulations (e.g., McKenna et al, 2002a,b; Pommrich et al, 2014, Ploeger et al., 2024). CLaMS includes packages for chemistry, ice cloud formation, and sedimentation. By combining a sophisticated mixing algorithm with a diabatic vertical coordinate (Konopka et al., 2007), CLaMS is well suited to simulate transport around transport barriers, such as the tropopause, the polar vortex or the edge of the Asian monsoon anticyclone (Ploeger et al.,2017; Sonnabend et al., 2024; Vogel et al., 2019). In recent years the CLaMS model has been applied to problems ranging from polar ozone loss, stratospheric age-of-air studies, and trends in stratospheric water vapor (e.g., Grooß and Müller, 2021; Poshyvailo-Strube et al., 2022; Konopka et al., 2022). Further, the CLaMS transport algorithm has been used in the frame of climate models (Charlesworth et al., 2023; Ploeger et al., 2024). The source code of CLaMS is written in Fortran90, and an MPI-based parallelization is applied for the core programs, such as the advective trajectory calculation.

The Lagrangian transport model Massive-Parallel Trajectory Calculations (MPTRAC) was introduced by Hoffmann et al. (2016), discussing Lagrangian transport simulations for volcanic eruptions with different meteorological data sets. In recent years MPTRAC was used in several case studies of simulating long-range transport processes of volcanic SO_2 (Heng et al., 2016; Wu et al. 2017, 2018; Liu et al., 2020; Cai et al. 2021). In addition to calculating the trajectories of air

parcels, MPTRAC includes modules and parameterizations for turbulent diffusion and subgrid-scale wind fluctuations, convection, sedimentation, dry and wet deposition, hydroxyl chemistry, and more. MPTRAC is designed for large-scale atmospheric simulations on current HPC and GPU systems through a flexible MPI-OpenMP-OpenACC hybrid parallelization (Liu et al., 2020; Hoffmann et al., 2022, 2024).

3 Sprint Objectives

The existing Lagrangian models, such as CLaMS-traj and MPTRAC, each offer distinct strengths and weaknesses depending on the application. Combining the strengths of these models (e.g., CLaMS's chemistry and representation of small-scale mixing and MPTRAC's flexible hybrid parallelization and computational efficiency) opens up new opportunities for building the next generation of Lagrangian models. The primary goal of this sprint was to integrate the computational kernels for advection of these models into a shared library for Lagrangian transport modeling. This sprint was also designed to lay the basis and develop a concept of combining legacy code written in Fortran with recently developed code written in C.

4 Procedure and Insights

4.1 Technical Approach / Procedure

The sprint started with an in-depth review of the key concepts underlying the MPTRAC and CLaMS models. For MPTRAC, this involved following the documentation (https://github.com/slcs-jsc/mptrac), executing the provided instructions, and running the example case. Subsequently, a detailed examination of the source code was conducted to gain further insight. The primary focus of this sprint was on the trac.c module, which handles Lagrangian particle dispersion. Some general functions of MPTRAC were already included in a library mptrac.c and mptrac.h. During the sprint the RSE transformed the existing code base of MPTRAC into a library. The MPTRAC code is written in C and, when coupled to CLaMS, its functions need to be called from Fortran. The RSE therefore developed a Fortran standalone application to test the functionality of the developed wrapper. An essential step in the sprint was to integrate multi language programming with Fortran and C in a single program. Since Fortran 2003, a standardized approach for combining Fortran and C code exists through the ISO_C_BINDINGS (https://gcc.gnu.org/onlinedocs/gfortran/ISO_005fC_005fBINDING.html), which allows for the generation of interoperable procedures, derived-type declarations, and global variables between Fortran and C. The compiler is informed with the BIND(C) attribute that a symbol shall be interoperable with C. The RSE tested the exchange between Fortran and C in small steps, starting with exchanging a scalar value, followed by the exchange of arrays, and finally with a structure composed of multiple data types. Additionally, the tests were performed with different compilers (GNU, NVIDIA, Intel) on the supercomputer JUWELS and on a laptop (under Linux). Finally, the RSE developed a standalone application for trajectory calculations. The application is written in Fortran with interfaces to the used C-structures and the used C-functions. This application and the associated interfaces are included in the source code repository of MPTRAC (which is open). Additionally, the RSE developed a test case, which calls this application and compares the results to a reference data set. This test is also included in the MPTRAC repository and will be checked regularly to find inconsistencies with updated code versions. As a side effect the documentation of MPTRAC was updated by the sprint team during the sprint to get more detailed information about the individual functions for new users and developers.

At the start of the work with CLaMS, the RSE reviewed the documentation available at http://clams.icg.kfa-juelich.de/CLaMS/traj, followed by conducting an example simulation. At the start of the sprint the main program for calculating the trajectory calculation was included in the subroutine TIMSTP with global definitions. It turned out that further modularization of the CLaMS trajectory code would be helpful, similar to MPTRAC. We achieved this modularization

during the sprint, and it is now available. Additionally, we improved the documentation. To enable the usage of CLaMS functions, which are written in Fortran, from a C program the RSE tested the general exchange in small test programs, and the exchange of scalars and arrays with different compilers on the supercomputer JUWELS and on a laptop (under Linux). For the other direction, namely calling CLaMS Fortran functions from MPTRAC C programs, we generated a concept of calling Fortran functions from C. This concept will be the basis for the coupling in this direction.

4.2 General Insights

When integrating C and Fortran code, the ISO_C_BINDINGS module is the standard tool that facilitates interoperability between the two languages. Introduced in Fortran 2003, this standardized interface allows Fortran programs to call C functions and vice versa, ensuring smooth and reliable communication and data exchange. This capability is particularly valuable in scientific computing and high-performance applications, where both languages are often employed to leverage their respective strengths: in our case to combine the scientific capabilities of CLaMS with the HPC capabilities of MPTRAC.

For the exchange of structures between Fortran and C, we found it to be crucial to ensure that the order and array sizes in the Fortran interface match those in the original C structure. The structures in MPTRAC consist of extensive variable lists (up to 309 for the control parameter list) and data types that may also depend on self-defined structures (struct of struct). Since MPTRAC's development is ongoing, the content of the structures may change in future versions. Therefore, the RSE developed a test case to compare the original structure with the structure in the interface. In this way, incorrect ordering and missing variables can be quickly detected and fixed. The test script is now integrated in the testing workflow used in MPTRAC.

5 Results

The result of the sprint was the successful implementation of a standalone Fortran application calling MPTRAC C functions for calculating trajectories. Moreover, the interoperability of the MPTRAC trajectory module and the CLaMS model system was established. To achieve this, we developed a concept for combining Fortan and C codes. The ISO_C_BINDINGS provide a framework to generate procedures, derived-type declarations, and global variables in Fortran that are interoperable with C. This framework was used by the RSE to develop a Fortran standalone application with interfaces for including the C-structures and C-functions from MPTRAC. The RSE integrated the standalone application into a test suite, which calls the MPTRAC applications from Fortran and compares the results to a reference data set. The test suite was compiled using different compilers (GNU, NVIDIA, Intel) on the JUWELS supercomputer and on a (Linux) Laptop. Comparisons with the reference dataset revealed that there are no discrepancies. In the future, the test suite will be executed on a regular basis to identify inconsistencies after code updates. Additionally, we expanded the documentation of MPTRAC to enable getting more detailed information about the individual functions.

Figure 1 shows trajectory calculations, performed by the RSE with the standalone Fortran application calling MPTRAC C functions. Shown is the trajectory of volcanic ash from the eruption of the Puyehue-Cordón Caulle volcano, Chile, in June 2011.

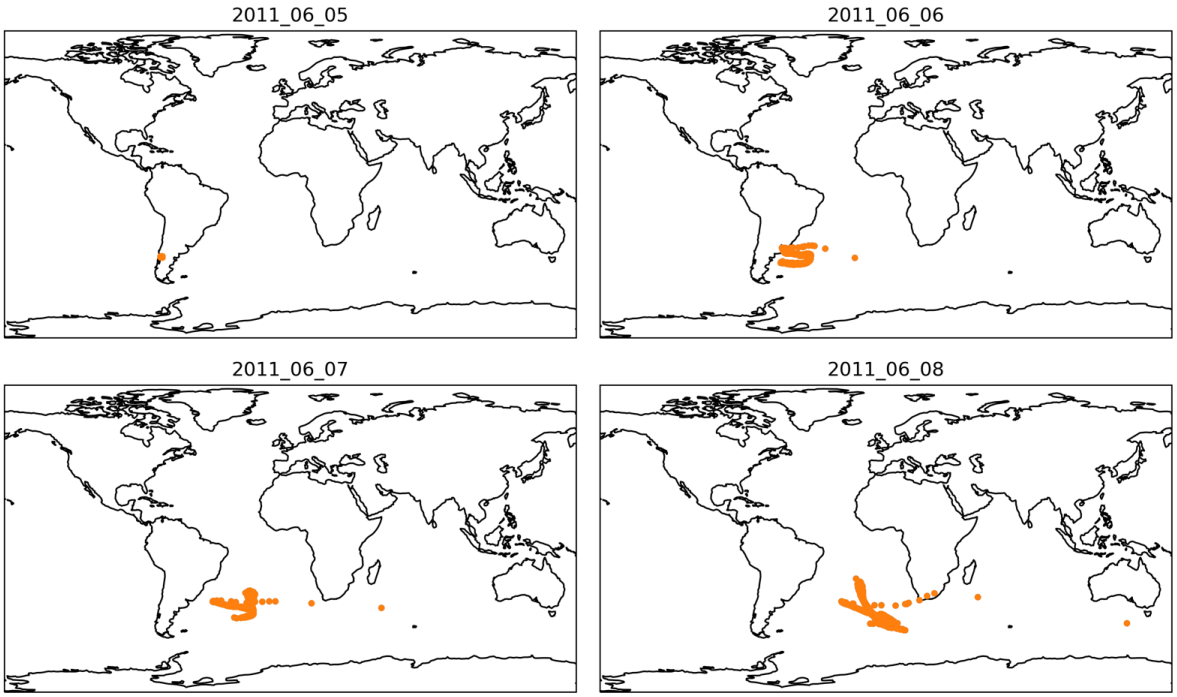

Figure 1: Trajectory calculations with standalone Fortran application calling MPTRAC C functions, offline-driven by ERA5 reanalysis data. Trajectory of volcanic ash from the eruption of the Puyehue-Cordón Caulle volcano, Chile, in June 2011.

During this sprint we modularized the CLaMS trajectory code further and improved the documentation. The RSE tested the exchange between Fortran functions called from a C program with values and arrays on the JUWELS supercomputer and on a (Linux) Laptop with different compilers. This constitutes the basis of developing a standalone C application calling Fortran functions from CLaMS. The sprint is a significant step towards developing a common Lagrangian modeling toolkit. Such a common library would allow optimization of Lagrangian modeling methods in different model systems as well as future usage on different hybrid hardware systems.

6 Conclusions and Outlook

The successful implementation of the standalone Fortran application, which calls MPTRAC C functions for improved trajectory calculations, is a significant step towards developing a common Lagrangian modeling toolkit. Such a toolkit will be highly valuable for the community. Following the developed concept of using ISO_C_BINDINGS, the calling of MPTRAC C functions can now be directly integrated in CLaMS. The first step of porting CLaMS to GPU is achieved and can now be continued module-wise. The implementation of a standalone C application calling CLaMS Fortran functions and at the end MPTRAC using CLaMS functions will be followed up on the basis of the developed concept of the sprint. The Lagrangian toolkit can be used standalone and within the MESSy framework, and the insights gained during the sprint are useful for ongoing and planned developments (e.g., in WarmWorld Better). Overall, the sprint was an important contribution to providing Lagrangian modeling competence to natESM. Additionally MPTRAC (https://nat-esm-system.dkrz.de/Optional%20Components/MPTRAC.html) and the CLaMS system (https://nat-esm-system.dkrz.de/Optional%20Components/CLaMS.html) were added to the natESM model system as optional components during the sprint.

7 References

- GitHub Repository MPTRAC: https://github.com/slcs-jsc/mptrac
- Documentation CLaMS Trajectory code: http://clams.icg.kfa-juelich.de/CLaMS/traj

- GitLab Repository CLaMS: https://jugit.fz-juelich.de/clams/clams-git

Brinkop, S. and Joeckel, P.: ATTILA 4.0: Lagrangian advective and convective transport of passive tracers within the ECHAM5/MESSy (2.53.0) chemistry–climate model, Geosci. Model Dev., 12, 1991–2008, https://doi.org/10.5194/gmd-12-1991-2019, 2019.

Cai, Z., Griessbach, S., and Hoffmann, L.: Improved estimation of volcanic SO2 injections from satellite retrievals and Lagrangian transport simulations: the 2019 Raikoke eruption, Atmos. Chem. Phys., 22, 6787–6809, https://doi.org/10.5194/acp-22-6787-2022, 2022.

Charlesworth, E., Ploeger, F., Birner, T., Baikhadzhaev, R., Abalos, M., Abraham, L., Akiyoshi, H., Bekki, S., Dennison, F., Jöckel, P., Keeble, J., Kinnison, D., Morgenstern, O., Plummer, D., Rozanov, E., Strode, S., Zeng, G., and Riese, M.: Stratospheric water vapor affecting atmospheric circulation, Nat. Commun., 14, 3925, https://doi.org/10.1038/s41467-023-39559-2, 2023.

Grooß, J.-U. and Müller, R.: Simulation of record Arctic stratospheric ozone depletion in 2020, J. Geophys. Res., 126,e2020JD033339, https://doi.org/10.1029/2020JD033339, 2021.

Heng, Y., Hoffmann, L., Griessbach, S., Roessler, T., and Stein, O.: Inverse transport modeling of volcanic sulfur dioxide emissions using large-scale simulations, Geosci. Model Dev., 9, 1627–1645, https://doi.org/10.5194/gmd-9-1627-2016, 2016.

Hoffmann, L., Roessler, T., Griessbach, S., Heng, Y., and Stein, O.: Lagrangian transport simulations of volcanic sulfur dioxide emissions: Impact of meteorological data products, J. Geophys. Res., 121, 4651–4673, https://doi.org/10.1002/2015JD023749, 2016.

Hoffmann, L., Baumeister, P. F., Cai, Z., Clemens, J., Griessbach, S., Günther, G., Heng, Y., Liu, M., Haghighi Mood, K., Stein, O., Thomas, N., Vogel, B., Wu, X., and Zou, L.: Massive-Parallel Trajectory Calculations version 2.2 (MPTRAC-2.2): Lagrangian transport simulations on graphics processing units (GPUs), Geosci. Model Dev., 15, 2731–2762, https://doi.org/10.5194/gmd-15-2731-2022, 2022.

Hoffmann, L., Haghighi Mood, K., Herten, A., Hrywniak, M., Kraus, J., Clemens, J., and Liu, M.: Accelerating Lagrangian transport simulations on graphics processing units: performance optimizations of Massive-Parallel Trajectory Calculations (MPTRAC) v2.6, Geosci. Model Dev., 17, 4077–4094, https://doi.org/10.5194/gmd-17-4077-2024, 2024.

Konopka, P., Günther, G., Müller, R., dos Santos, F. H. S., Schiller, C., Ravegnani, F., Ulanovsky, A., Schlager, H., Volk, C. M., Viciani, S., Pan, L. L., McKenna, D.-S., and Riese, M.: Contribution of mixing to upward transport across the tropical tropopause layer (TTL), Atmos. Chem. Phys., 7, 3285–3308, 2007.

Konopka, P., Tao, M., Ploeger, F., Hurst, D. F., Santee, M. L., Wright, J. S., and Riese, M.: Stratospheric moisteningafter 2000, Geophys. Res. Lett., 49, e2021GL097609, https://doi.org/https://doi.org/10.1029/2021GL097609, 2022.

Liu, M., Huang, Y., Hoffmann, L., Huang, C., Chen, P., and Heng, Y.: High-Resolution Source Estimation of Volcanic Sulfur Dioxide Emissions Using Large-Scale Transport Simulations, in: Computational Science – ICCS 2020, edited by Krzhizhanovskaya, V. V., Závodszky, G., Lees, M. H., Dongarra, J. J., Sloot, P. M. A., Brissos, S., and Teixeira, J., pp. 60–73, Springer International Publishing, https://doi.org/10.1007/978-3-030-50420-5_5, 2020.

McKenna, D. S., Grooß, J.-U., Günther, G., Konopka, P., Müller, R., Carver, G., and Sasano, Y.: A new Chemical Lagrangian Model of the Stratosphere (CLaMS): 2. Formulation of chemistry scheme and initialization, J. Geophys. Res., 107, 4256, https://doi.org/10.1029/2000JD000113, 2002a.

McKenna, D. S., Konopka, P., Grooß, J.-U., Günther, G., Müller, R., Spang, R., Offermann, D., and Orsolini, Y.: A new Chemical Lagrangian Model of the Stratosphere (CLaMS): 1. Formulation of advection and mixing, J. Geophys. Res., 107, 4309, https://doi.org/10.1029/2000JD000114, 2002b.

Ploeger, F., Konopka, P., Walker, K., and Riese, M.: Quantifying pollution transport from the Asian monsoon anticyclone into the lower stratosphere, Atmos. Chem. Phys., 17, 7055–7066, https://doi.org/10.5194/acp-17-7055-2017, 2017.

Ploeger, F., Birner, T., Charlesworth, E., Konopka, P., and Müller, R.: Moist bias in the Pacific upper troposphere and lower stratosphere (UTLS) in climate models affects regional circulation patterns, Atmos. Chem. Phys., 24, 2033–2043, https://doi.org/10.5194/acp-24-2033-2024, 2024.

Pommrich, R., Müller, R., Grooß, J.-U., Konopka, P., Ploeger, F., Vogel, B., Tao, M., Hoppe, C. M., Günther, G., Spelten, N., Hoffmann, L., Pumphrey, H.-C., Viciani, S., D'Amato, F., Volk, C. M., Hoor, P., Schlager, H., and Riese, M.: Tropical troposphere to stratosphere transport of carbon monoxide and long-lived trace species in the Chemical Lagrangian Model of the Stratosphere (CLaMS), Geosci. Model Dev., 7, 2895–2916, https://doi.org/10.5194/gmd-7-2895-2014, 2014.

Poshyvailo-Strube, L., Müller, R., Fueglistaler, S., Hegglin, M. I., Laube, J. C., Volk, C. M., and Ploeger, F.: How can Brewer–Dobson circulation trends be estimated from changes in stratospheric water vapour and methane?, Atmos. Chem. Phys., 22, 9895–9914, https://doi.org/10.5194/acp-22-9895-2022, 2022.

Stein, A. F., R. R. Draxler, G. D. Rolph, B. J. B. Stunder, M. D. Cohen, and F. Ngan. "NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System", Bulletin of the American Meteorological Society 96, 12, 2059-2077, https://doi.org/10.1175/BAMS-D-14-00110.1, 2015.

Sonnabend, J., Grooß, J.-U., Ploeger, F., Hoffmann, L., Jöckel, P., Kern, B., and Müller, R.: Lagrangian transport based on the winds of the icosahedral nonhydrostatic model (ICON), Meteorol. Z., 33, 229–242, https://doi.org/10.1127/metz/2024/1207, 2024.

Vogel, B., Müller, R., Günther, G., Spang, R., Hanumanthu, S., Li, D., Riese, M., and Stiller, G. P.: Lagrangian simulations of the transport of young air masses to the top of the Asian monsoon anticyclone and into the tropical pipe, Atmos. Chem. Phys., 19, 6007–6034, https://doi.org/10.5194/acp-19-6007-2019, 2019.

Wohltmann, I., Lehmann, R., and Rex, M.: The Lagrangian chemistry and transport model ATLAS: simulation and validation of stratospheric chemistry and ozone loss in the winter 1999/2000, Geosci. Model Dev., 3, 585–601, https://doi.org/10.5194/gmd-3-585-2010, 2010.

Wu, X., Griessbach, S., and Hoffmann, L.: Equatorward dispersion of a high-latitude volcanic plume and its relation to the Asian summer monsoon: a case study of the Sarychev eruption in 2009, Atmos. Chem. Phys., 17, 13 439–13 455, https://doi.org/10.5194/acp-17-13439-2017, 2017.

Wu, X., Griessbach, S., and Hoffmann, L.: Long-range transport of volcanic aerosol from the 2010 Merapi tropical eruption to Antarctica, Atmos. Chem. Phys., 18, 15 859–15 877, https://doi.org/10.5194/acp-18-15859-2018, 2018.