

CMAS



The Community Modeling and Analysis System Center 22 Years of Serving a Global User Community

by Saravanan Arunachalam

An overview of the operations at the CMAS Center, including the distribution and support of the CMAQ modeling system.

The Community Modeling and Analysis System (CMAS) Center was established by the U.S. Environmental Protection Agency (EPA) in 2001 as a mechanism for transferring primary responsibility for the development, maintenance, and applications support of the Models-3 air quality modeling system from EPA to a separate organization serving the needs of the air quality modeling community. The center was designed to leverage the community's intellectual capital to continue refining air quality modeling tools to improve and support their applications for critical problem analysis and decision-making.

From research to application to outreach, the goal of the CMAS Center is to advance the community modeling paradigm through the establishment of a centralized resource to serve the environmental modeling community. The main objectives of the center were to serve as a bridge between segments of the air quality modeling community, foster growth of developer and user communities, and serve as a hub for model education and training. At the core of CMAS operations is the distribution and support of the Community Multiscale Air Quality (CMAQ) modeling system and the Sparse Matrix Operator Kernel Emissions (SMOKE) model. The modeling support expanded, over the years, to include additional modeling tools that are discussed below.

The CMAS Center was initially hosted by MCNC Environmental Programs in Research Triangle Park, NC. In January 2003, the CMAS Center transitioned to the University of North Carolina (UNC) at Chapel Hill, where it remains operational today. The host organization for the CMAS Center is awarded by a periodic competitive proposal process, and the current term for the CMAS Center runs through July 2026.¹

Prior to the 1990s, advancements in air quality model technology could not be easily shared between modelers because of technical incompatibilities in algorithms, software, operating systems, and so forth. By standardizing with open-source, advanced modeling systems, the CMAS Center enables collaborative development and linking of models for meteorology, emissions, air quality, hydrology, and environmental and health effects. Initial design focused on using a “modular” approach, where developers separated individual atmospheric processes into independent pieces, which then act like building blocks that execute one aspect of functionality. The CMAS Center-supported models, such as CMAQ and SMOKE, use the concept of “modularity” to empower developers to upgrade existing atmospheric processes or add new ones, thus ensuring rapid evolution of software to reflect advances in scientific knowledge and to meet changing needs of the environmental modeling community.

To address the challenges of the 1990 U.S. Clean Air Act Amendments and to study the complex relationship between multiple pollutants, EPA developed the CMAQ modeling system,^{2,3} a third-generation, multi-scale, multi-pollutant modeling system. With a community model approach, a

need was identified to promote community air quality modeling and its connections with human/ecological health and climate/air quality assessments. Working toward this goal, the CMAS Center is designed to leverage the modeling community's intellectual capital in air quality modeling and analyses and to continually refine the tools available for addressing air quality issues through collaboration.

The CMAS Center has four primary functions to lead and coordinate the activities: Outreach, Research, Software Development, and Applications and User Support. In addition, the CMAS Center has an External Advisory Committee (EAC) to advise and guide its operations. Nominated by the CMAS Center Director, the EAC is currently composed of 13 members and includes ex officio members from EPA, members from state agencies, industry, academia, consultants, regional planning organizations, international users, and expert users—each having a voting capacity, and have a rotating term limit of two years.

Software Support

The CMAS Center supports 15 different software products,⁴ including CMAQ, SMOKE, the Input/Output Applications Programming Interface (I/O API), the Atmospheric Model Evaluation Tool (AMET), the Community LINE source (C-LINE) model, the Community model for near-PORT (C-PORT) applications, the Control Strategy Tool (CoST), the Emissions Modeling Framework (EMF), the Fertilizer Emission Scenario Tool for CMAQ (FEST-C), the Meteorology-Chemistry Interface Processor (MCIP), the Research LINE Source (R-LINE) model, the Spatial Allocator, the Speciation Tool, the Surrogate Tools, and the Visual Environment for Rich Data Interpretation (VERDI). Since January 2018, over 6,200 users from across the globe (approximately 37% from the United States) have downloaded software from the CMAS Center. (Note that some of the CMAS Center products are also available through other hosting venues and that the download count does not reflect accesses from other venues.)

Until 2017, CMAQ source code distribution was maintained through Concurrent Version System (CVS). Since then, source code is maintained and distributed via a GitHub repository.⁵ All development happens through version control on GitHub internally within EPA, and periodically, the EPA team shares release-ready versions of the code with the CMAS Center staff. The CMAS Center provides final testing and benchmarking of updates to CMAQ prior to public release and in consultation with EPA staff.

Support for users to install, run, and develop model applications are key CMAS Center activities. The CMAS Center has a multifaceted approach to supporting the software it hosts, at no cost to the user. First, the center develops hands-on self-guided tutorials and user guides, in collaboration with EPA, and makes them available. Since 2005, the CMAS Center has facilitated interactions between software users and software developers to provide user support, identify software

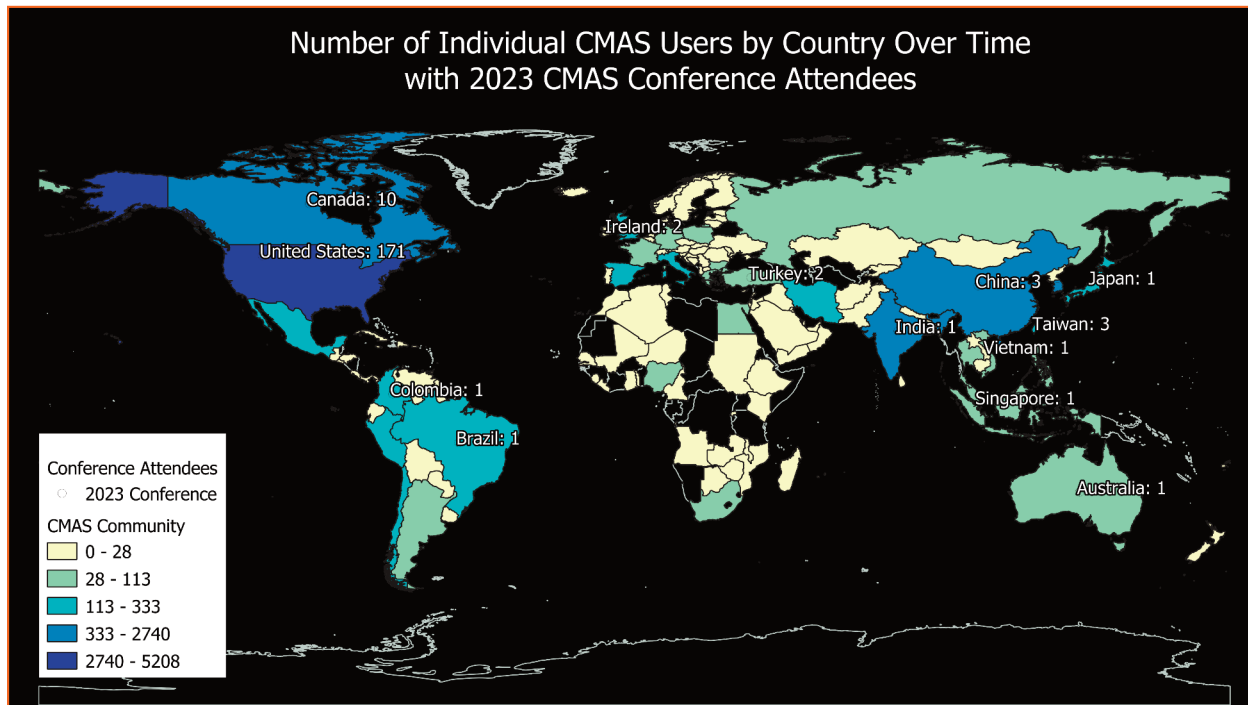


Figure 1. Global footprint of the CMAS Center user community, showing current active users and attendees at the 22nd Annual Conference held in 2023.

deficiencies (“bugs”), and target areas that require improvement in the models and tools hosted by the Center.

Since 2018, the CMAS Center has hosted an online forum through Discourse⁶ with topical headings that provide structure, documentation of recurring issues, and a searchable capability to minimize duplication. Over 25 specific topics are predefined, and the user can post a question under any topic. While the CMAS center staff has primary responsibility to monitor the forum for questions from the user community, often solutions to forum posts are provided by other community members or by developers of the specific software product from EPA or UNC. All discussion is archived, which provides a searchable repository of information for future users who may face the same problem. On average, this forum receives about 50 new issues each month across all models and tools, from 934 current registered users.

The CMAS Center also hosts an online data warehouse for input/output model datasets, air quality observations, and a tool-suite for data processing and model evaluation. This data warehouse is targeted to support easy online access to model datasets for the user community, and to avoid serving them one at a time through the now archaic method of copying to external hard drives and shipping. While the data warehouse was initially hosted on a File Transfer Server (FTP), the CMAS Center started hosting model datasets on Google Drive for various model years between 2002 and 2017. In 2022, the CMAS Center started to transition to Amazon’s Open Data Sponsorship Program (AWS ODP),⁷ as part of Amazon’s Sustainability Data Initiative for hosting the data warehouse.

The CMAS Center currently hosts complete data for comprehensive benchmarking of new CMAQ releases, and data from the EPA Air Quality Time Series (EQUATES) study with consistent treatment of emissions, meteorology, and model configuration for the years 2002–2019.⁸ Reproducing the benchmark on a user’s individual computing environment ensures that the model is installed accurately. The EQUATES dataset was developed by EPA to provide a consistent set of long-term emissions inputs for human health, ecological, and evaluation applications of CMAQ or other air quality models. Currently, the Google Drive and AWS ODP collectively host more than 150 Terabytes of data from the CMAS Center, and plans are afoot to grow the warehouse on AWS ODP.

Recent contributions from the CMAS Center are new guidance and hands-on tutorials on cloud computing and running CMAQ on two of the popular cloud vendors’ resources: Amazon Web Services (AWS) and Microsoft Azure. Running the model on a cloud-based server is a new paradigm that the regional-to-global atmospheric modeling community, in general, is exploring, to avoid maintaining computer servers within one’s organization. Additional details are in Adams et al. found elsewhere in this issue.⁹

CMAS Training, Workshops, and Outreach

The CMAS Center offers training, workshops, and conferences to the community at large, and is funded through registration fees, which are independent of EPA’s support for operating the Center. The CMAS Center Training focuses on two main models—the CMAQ model and the SMOKE modeling system—using a combination of lectures and hands-on

training. Starting in 2005, training was offered four times a year in-person at the CMAS Center. Since 2015, the CMAS Center transitioned to offering SMOKE and CMAQ trainings twice per year virtually through UNC's Sakai online classroom instructional system and AWS-based instances. The transition to virtual training has significantly expanded the access to the CMAS Center training to the global audience across multiple time zones. The CMAS Center has also offered the training onsite in multiple countries in Europe, Asia, and South America starting in 2006. Beyond introductory trainings on SMOKE and CMAQ, advanced training has also been offered on special topics, such as EPA's Environmental Benefits Mapping and Analyses Program (BenMAP) and Interpretive Model Analyses and Evaluation on an on-demand basis. To date, the CMAS Center has trained over 2,100 users across the world.

A key outreach activity of the CMAS Center is hosting an annual conference. The 22nd annual conference was held in Chapel Hill, NC, October 16–18, 2023. The three-day conference is comprised of plenary keynote speeches on current/emerging topics followed by 10 technical sessions on topics in atmospheric chemistry modeling and applications. About 100 oral and poster presentations are made each year, and attendance has grown from a little less than 200 attendees in the year 2005 to as high as 350 in the year 2020. The annual conference provides a venue for the CMAS Center user community to gather, exchange information on the latest developments in models, available datasets, and lessons learned from model applications, and informal conversations that often lead to collaborative research projects.

Over the past 22 years, the CMAS Center has continued to work toward the objectives set forth by EPA. The CMAS Center has been instrumental in growing the air quality modeling

community, especially outside the United States (see Figure 1). More importantly, the CMAS Center supported suite of models is being used to study air pollution and guide air quality management within and outside the United States. As an example, through the conference and training conducted by the CMAS Center in South America, the local authorities have learned that air pollution is not local to one city or state, and a broad collaboration amongst regional and local entities is needed to mitigate air pollution issues and protect public health. On the positive side, there has been increased collaboration among agencies in recent years. However, there are still areas for improvement in the operations of the CMAS Center, such as acquiring sustained and increased resources for outreach and user support activities, and more explicit involvement from the EAC to enable the CMAS Center to become self-sustaining.

Looking Forward

The goals for the CMAS Center in the next decade are to (a) have some of the CMAS Center products such as CMAQ and SMOKE installed as a software package for easy replication by users on one of the cloud vendors' marketplaces; (b) have an interactive web-based analysis of model predictions versus observations (both in-situ and remotely sensed); (c) have emissions, meteorology, and air quality data available for each case study in an integrated manner along with the observations for in-depth diagnostic analytical capabilities; and (d) expand outreach activities to additional parts of the world such as Africa where air pollution issues are rapidly emerging with several observational efforts underway. The CMAS Center remains committed to facilitating air quality management by continuing to grow a global user community that can easily develop air quality model applications for their custom needs, focus efforts on understanding and enhancing model science, and perform interpretive analyses. **em**

Saravanan Arunachalam is Director of the CMAS Center and Deputy Director of the Institute for the Environment at The University of North Carolina at Chapel Hill. Email: sarav@email.unc.edu.

Disclaimer: The views expressed in this article are those of the author and do not necessarily represent the views or policies of EPA.

Acknowledgment: The U.S. Environmental Protection Agency (EPA), through its Office of Research and Development and Office of Air and Radiation, partially funded and collaborated in the research described here under EP W 16 014 to the University of North Carolina at Chapel Hill. This article is dedicated to Dr. Adel Hanna, CMAS Center Director, 2003–2019.

References

1. CMAS Center. See <https://cmascenter.org/index.cfm>.
2. Appel, K.W.; Bash, J.O.; Fahey, K.M.; Foley, K.M.; Gilliam, R.C.; Hogrefe, C.; Hutzell, W.T.; Kang, D.; Mathur, R.; Murphy, B.N.; Napelenok, S.L.; Nolte, C.G.; Pleim, J.E.; Pouliot, G.A.; Pye, H.O.T.; Ran, L.; Roselle, S.J.; Sarwar, G.; Schwede, D.B.; Sidi, F.I.; Spero, T.L.; Wong, D.C. The Community Multiscale Air Quality (CMAQ) model versions 5.3 and 5.3.1: System updates and evaluation; *Geosci. Model Dev.* 2021, *14*, 2867–2897; <https://doi.org/10.5194/gmd-14-2867-2021>.
3. Byun, D.; Schere, K.L. Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system; *Appl. Mech. Rev.* 2006, *59*, 51–77.
4. CMAS Center Release Calendar. See https://cmascenter.org/download/release_calendar.cfm.
5. GitHub repository. See <https://github.com/USEPA/CMAQ>.
6. CMAS Center Forum. See <https://forum.cmascenter.org>.
7. AWS Open Data Portal. See https://aws.amazon.com/marketplace/pp/prodview-7uloskv5wkgmc?sr=0-3&ref_=beagle&applicationId=AWSMPContessa.
8. Foley, K.M.; Pouliot, G.A.; Eyth, A.; Aldridge, M.F.; Allen, C.; Appel, K.W.; Bash, J.O.; Beardsley, M.; Beidler, J.; Choi, D.; Farkas, C.; Gilliam, R.C.; Godfrey, J.; Henderson, B.H.; Hogrefe, C.; Koplitz, S. N.; Mason, R.; Mathur, R.; Misenis, C.; Possiel, N.; Pye, H. O. T.; Reynolds, L.; Roark, M.; Roberts, S.; Schwede, D.B.; Seltzer, K.M.; Sonntag, D.; Talgo, K.; Toro, C.; Vukovich, J.; Xing, J.; Adams, E. 2002–2017 Anthropogenic Emissions Data for Air Quality Modeling over the United States; *J. Data in Brief* 2023, 109022; <https://doi.org/10.1016/j.dib.2023.109022>.
9. Adams, E.; Efstathiou, C.; Adelman, Z.; Foley, K.; Sidi, F.; Hao, W.; Arunachalam, S. Air Quality Modeling on the Cloud: Pairing Data Availability and High-Performance Clusters; *EM* January 2024.