MUSICA Tutorial Series 2021-2022:
1. How to use MUSICAv0 output

MUSICA: MUlti-Scale Infrastructure for Chemistry and Aerosols

12 November 2021
Here we value respectful dialogue, please...

Offer Constructive Feedback
Consider New Ideas
Show Appreciation
Encourage Innovation
Share The Air
Acknowledge Teamwork
Overview of Tutorial Series

Overall goal is to introduce various components of MUSICA:
Variable resolution global modeling (MUSICAv0)
Box modeling (MusicBox)
Model evaluation (MELODIES)

- November 12: How to use MUSICAv0 output (unstructured grids)
- November 19: Office hours
- December 10: How to run MUSICAv0 (CAM-chem with CONUS grid)
- January 14: How to create your own variable resolution grid
- February 11: How to run MusicBox
- March 11: How to use MELODIES
- April 8: Additional topics, user requests
You are a group with diverse interests and experience

**Topics:** Air quality, smoke transport, data assimilation, PM2.5 trends, aerosols, aerosol-cloud interactions, SOA, stratospheric aerosol injections, aviation NOx, volcanoes, dust storm and biogenic forecast, impact on health & crop yield, Asian monsoon, water vapor isotopes

**Regions:** Brazil, Latin America & Caribbean, Mexico, southern Africa, Himalayan region, Asia, Europe, Michigan, Thailand and SE Asia, UTLS

**Career stage:** Grad students, postdocs, professors, research scientists, ...

We hope to make this accessible and interesting for all!
Agenda for Today

- Introduction and overview of MUSICA (Louisa Emmons)
- Overview of unstructured grids (Louisa Emmons)
- Description of the Community CONUS grid simulation, format and contents of output files (Louisa Emmons)
- Demonstration of plotting and analysis tools (Duseong Jo)
- Hands-on exercises (everyone)

Goal for the Day:
- Introduce the format of MUSICAv0 output - for users of Community simulation, those who want to run MUSICAv0
- Introduce some python plotting notebooks to handle unstructured grids
Introduction and Overview of MUSICA
MUSICA: MUlti-Scale Infrastructure for Chemistry & Aerosols

A new model-independent infrastructure, which will enable chemistry and aerosols to be simulated at different resolutions in a coherent fashion

Will facilitate use of a variety of chemistry schemes, physics parameterizations and atmospheric models

Coupled to other earth system component models (land, ocean, sea ice, etc.)

Whole atmosphere framework: troposphere to thermosphere

https://www2.acom.ucar.edu/sections/multi-scale-chemistry-modeling-musica

MUSICA Vision paper published in BAMS (Pfister et al., 2020: https://doi.org/10.1175/BAMS-D-19-0331.1)
Model-Independent Chemistry Module (MICM)

Database of chemical mechanisms and data needed for solving chemistry

Allows easily changing the chemical mechanism

Will allow use of the same chemistry in different atmosphere models and

MusicBox: MICM in a box model:  [https://github.com/NCAR/music-box](https://github.com/NCAR/music-box)

Available with command-line control or browser interface

Allows for easy:

- Modification of chemical mechanism
- Specification of initial and time-varying environment

Browser interface plots results, allows comparison of 2 mechanisms
Choices for variable resolution atmosphere models

Spectral Element (SE - cubed sphere)

Currently running in CAM

Model for Prediction Across Scales (MPAS - hexagonal mesh)

Non-hydrostatic allowing for finer scales
MUSICAv0 is a configuration of the Community Earth System Model (CESM):

**CAM-chem** (Community Atmosphere Model with Chemistry)

With Spectral Element (SE) dynamical core and Regional Refinement (RR) [CAM-chem-SE-RR]

At finer resolution, emissions and chemistry are more accurately represented

Simulated pollutants on human exposure-relevant scales

Global feedbacks are directly included

Most of the grid points are in refined region, so no additional cost to simulate the whole globe

Users can create their own grids
MUSICAv0 – Results

Impact of higher resolution on chemistry versus increased chemical complexity

Becky Schwantes (NOAA), Forrest Lacey (NCAR/ACOM+RAL)
Papers in prep.

Improving representation of fire emissions

Analysis of Korea air quality and impact of model resolution

Wenfu Tang, NCAR/ACOM+ASP

Duseong Jo, NCAR/ACOM+ASP
Simulations in support of ACCLIP
Asian Summer Monsoon Chemical and Climate Impact Project

The regionally-refined MUSICA v0 Asia grid enhances the convective transport of pollutants by the Asian monsoon into the UTLS when compared to WACCM.

Sampling these air masses is a key objective of the ACCLIP field phase in summer 2022.

Ren Smith, NCAR/ACOM
MELODIES for MUSICA: A modular framework to compare model results and observations of atmospheric chemistry

MELODIES: Model EvaLuation using Observations, Diagnostics and Experiments Software

- Modular framework
- User-friendly interface
- User Guides will be produced
- Tutorial for community, targeting students and postdocs

MELODIES is being built on MONET
https://github.com/noaa-oar-arl/menet
MUSICA Goals

• To be developed collaboratively with university and government researchers
• To become the next-generation community infrastructure for atmospheric chemistry & aerosol research
• To deepen existing, and establish new, working relations of the research community with a variety of users ranging from the research community to stakeholders
• To contribute to both advancing the science and to providing relevant and actionable information for the development of mitigation policies or warning systems
Community Involvement Welcome

We invite the community to participate in development, evaluation and application of MUSICA:
https://www2.acom.ucar.edu/sections/multi-scale-chemistry-modeling-musica

Working groups:
• Model Architecture
• Emissions and Deposition
• Chemical Schemes
• Aerosols
• Physics, Transport, sub-scale Processes
• Whole Atmosphere
• Evaluation and Data Assimilation

Visit MUSICA website to join working groups
Implementation plans are being developed

Also:
• Join email list to receive MUSICA Newsletters and updates
• Access existing MUSICAv0 simulations
• Contribute to python library
• Contribute to MELODIES
Overview of unstructured grids
Description of Spectral Element in CAM

https://doi.org/10.1029/2017MS001257

**CAM-SE: (Lauritzen et al., 2018)**

CAM-SE uses a continuous Galerkin finite element method (Taylor et al., 1997) referred to as Spectral Elements (SE):

- Physical domain: Tile the sphere with quadrilaterals using the gnomonic cubed-sphere projection
- Computational domain: Mapped local Cartesian domain
- Each element operates with a Gauss-Lobatto-Legendre (GLL) quadrature grid
  - Gaussian quadrature using the GLL grid will integrate a polynomial of degree $2N - 1$ exactly, where $N$ is degree of polynomial
- Elementwise the solution is projected onto a tensor product of 1D Legendre basis functions
  - by multiplying the equations of motion by test functions, weak Galerkin formation
  - all derivatives inside each element can be computed analytically!

CESM Tutorial: https://www.cesm.ucar.edu/events/tutorials/2019/files/Lecture2-lauritzen.pdf
The spectral element (SE) method has been considered as a numerical method for the fluid flow solver in global weather/climate models. The main motivations were the SE methods' near-perfect scalability, GPU acceleration, high-order accuracy for smooth problems, and mesh refinement capabilities. For some time CESM has supported a SE dynamical core option in CAM discretized on a cubed-sphere grid.

[Laurizen et al., 2018; Fig 1]
The Spectral Element model output is on unstructured grids.

Model output provides the lat, lon of the center of each model grid, but the edges and vertices are not.

This information is available in “SCRIP” files for each grid.

The SCRIP files are read in the plotting examples to allow plotting maps to show the native grid.

The left panel shows the Gauss-Lobatto-Legendre (GLL) grid with $N_p \times N_p$ quadrature points defined on a standard element $[-1,1]^2$, where $N_p=4$. The right panel shows the cubed-sphere grid system tiled with spectral elements $\Omega_e$, where $N_e$ is the number of elements in each coordinate direction on a panel (in this case $N_e=5$). Each element $\Omega_e$ on has the GLL grid structure.

[Laurizen et al., 2018; Fig 2]
Pause to start JupyterHub

To be ready for the hands-on portion, start your JupyterHub server now - it may be slow if everyone connects at once…

https://www2.acom.ucar.edu/sites/default/files/workshop/MUSICA_Tutorial_2021-11-12_jupyterhub.pdf

Just first 3 steps (slides 2, 3, 4)
Description of the Community CONUS grid simulation
Format and contents of output files
Description of the Community MUSICAv0 simulation

CESM2.2 CAM-chem with Spectral Element, Regional Refinement (1/8 degree) over CONUS
Chemistry: MOZART-TS1 (standard trop-strat scheme in CESM2)
Meteorology: nudged to MERRA2 reanalyses
Simulated years: 2012-2013 (so far)
Emissions:
- Anthropogenic: CAMS-v5.1
- Fires: QFED (FINN EFs applied to QFED CO2)
- Biogenic: MEGANv2.1 in CLM
- Ocean: OASISS for DMS

CO tags: anthropogenic for various regions (CO01 - CO12), biomass burning (CO13)
Community MUSICAv0 Simulation Output

Output available on NCAR DASH Repository: https://doi.org/10.5065/tgbj-yv18

Output:

- Each file has a single variable (e.g., T, O3, etc.) for a timeseries
- Monthly averages are grouped by year
  - Available on native variable resolution grid and interpolated to 0.9x1.25 regular grid
- Daily averages are grouped by month (on variable resolution grid)
- Hourly averages will be available soon

A comparable simulation at uniform ne30 (1-degree) resolution for 2010-2020 is also available. This provides monthly initial conditions files that can be regridded for variable resolution simulations.
Community MUSICAv0 CONUS Simulations

Version: 1.0

The Community MUSICAv0 CONUS simulations have been produced using the Community
Version 2.2.0 with comprehensive tropospheric and stratospheric TSI chemistry, using the
dynamical core, with regional refinement over the contiguous U.S., having a grid mesh with
horizontal grid spacing over most of the globe that refines to 1/8 degree (~14 km) over the
simulation without the regional refinement for comparison, link to follow.

The file names are as follows:
*.cam.h0*.nc files are monthly mean output fields on the 0.9x1.25 degrees regular horizontal
*.cam.h1*.nc files are monthly mean output fields on the native grid;
*.cam.h2*.nc files are daily mean output fields on the native grid.

DOI
https://doi.org/10.5065/tgbj-yv18

Download Data and Documentation (2658 Files, 193.87 GB Total)

- Individual Files - View, select and download individual files from this Dataset.
- Zip File - Download a ZIP file containing all files.
- Wget shell script - Download all files using Wget, preferred for Linux.
- Curl shell script - Download all files via Curl, preferred for MacOS.

Temporal Resolution
1.0 month
1.0 day
6.0 hour

Spatial Resolution
1.0 degrees Latitude
1.0 degrees Longitude
0.125 degrees Latitude
0.125 degrees Longitude

Related Links
MUSICA
Use search box to refine list to a manageable size
Dimensions and grids

Standard CESM (finite volume) regular grids:
dimensions:
- \( \text{lat} = 192 \)
- \( \text{lon} = 288 \)
- \( \text{time} = \text{UNLIMITED} \); // (1 currently)
- \( \text{lev} = 32 \)

- \( \text{lat(lat)} \)
- \( \text{lon(lon)} \)
- \( \text{lev(lev)} \)
- \( \text{O3(time, lev, lat, lon)} \)

Spectral element output:
dimensions:
- \( \text{ncol} = 174098 \)
- \( \text{time} = \text{UNLIMITED} \); // (1 currently)
- \( \text{lev} = 32 \)

- \( \text{lat(ncol)} \)
- \( \text{lon(ncol)} \)
- \( \text{lev(lev)} \)
- \( \text{area(ncol)} \)
- \( \text{O3(time, lev, ncol)} \)
Additional grid information files

The CESM output files do not currently contain all the information about unstructured grids.

SCRIP file: contains centers *and corners* of each grid cell.
Vertical dimension

The dimension `lev` indexes the model layers:
- index 0 = top of the model
- index 31 = surface layer (if number of lev = 32) [in python can use index = -1]

The variable `lev` is the global mean pressure for each model layer (mid-level), units = hPa

The variable `PMID` is the pressure for each grid box:
PMID(time, lev, ncol) ;
    PMID:units = "Pa" ;
    PMID:long_name = "Pressure at layer midpoints" ;

`Z3(time, lev, ncol)` - Geopotential Height (above sea level) of each grid box; units = "m"
`PS(time, ncol)` - surface pressure; units = "Pa"
`PDELDRY(time, lev, ncol)` - layer thickness; units = "Pa"
Output variables available

Chemical species: find explanations of the chemistry and the species in Emmons et al., JAMES, 2020: https://doi.org/10.1029/2019MS001882

Meteorological & dynamics variables, for example:

- T: temperature (K); U,V: wind speeds (m/s); PRECC, PRECL: precipitation
- CLOUD, CLDTOT: cloud fraction, integrated column cloud fraction
- PBLH: boundary layer height (m)

Emissions diagnostics:

- Total surface emissions: SF{species}
- Biogenic emissions: MEG_{species}
- integrated column of vertical emissions: {species}_CLXF
- Lightning emissions: LNO_COL_PROD (2D), LNO_PROD (3D)

Deposition:

- Dry deposition velocity, flux: DV_{species}, DF_{gas-species}, {aerosol}DDF
- Wet dep, integrated flux: WD_{species}

Various output streams available

/glade/campaign/acom/acom-climate/tlmes/CO_CONUS/
f.e22.FCcotagsNudged.ne0CONUSne30x8.cesm220.2012-01/atm/

$casename = f.e22.FCcotagsNudged.ne0CONUSne30x8.cesm220.2012-01

./hist/ {many variables for small number of timesteps in each file}
$casename.cam.h0.YYYY-MM.nc : Monthly mean files, interpolated to 0.9x1.25 regular grid
$casename.cam.h1.YYYY-MM.nc : Monthly means, on native grid
$casename.cam.h2.YYYY-MM-DD-00000.nc : Daily averages, on native grid
$casename.cam.h3.YYYY-MM-DD-00000.nc : 6-hr output, interpolated to 0.9x1.25 regular grid
$casename.cam.h4.YYYY-MM-DD-00000.nc : 1-hr output, on native grid

./proc/tseries/month_1/ {one variable for long time series in each file}
$casename.cam.h0.{species}.YYYY01-YYYY12.nc
$casename.cam.h1.{species}.YYYY01-YYYY12.nc

./proc/tseries/day_1/
$casename.cam.h2.{species}.YYYYMM01-YYYYMM31 : one species, daily averages for 1 month

Use the ‘ncdump’ command on any file from the UNIX command line to get a list of the variables, their dimensions, long names, and units
Community Simulation: h0 files

f.e22.FCcotagsNudged.ne0CONUSne30x8.cesm220.2012-01.cam.h0.YYYY-MM.nc

- Monthly mean files interpolated to 0.9° x 1.25° (lat x lon) regular grid
- Many variables, but they are **interpolated (not conservatively regridded)** to the regular grid, so not suitable for budgets (e.g., emissions totals, etc.)
- Helpful for quick looks using tools you have for regular grids
Community Simulation: h1 files

f.e22.FCcotagsNudged.ne0CONUSne30x8.cesm220.2012-01.cam.h1.YYYY-MM.nc

- Monthly means on SE variable resolution grid *(fewer variables than h0 because of file size)*
- Gas-phase Species: CH4, CO, H2O, O3, O3S {stratospheric ozone}, OH
- Aerosols: PM25_SRF (PM2.5 at surface only), bc_a1, bc_a4, dst*, ncl* num*, pom*, so4*, soa*, CCN3
- Total surface emissions: SFCO, SFDMS, SFNO, SFSO2, OCN_FLUX_DMS
- Biogenic emissions from MEGAN: MEG_ISOP, MEG_MTERP
- Lightning NO emissions: LNO_COL_PROD(time, ncol)
- Meteorology: T, U, V, Q, CLOUD(time, lev, ncol), CLDTOT(time, ncol), PRECC, PRECL, PBLH, PS {surface pressure}, CAPE
Community Simulation: h2, h3, h4 files

f.e22.FCcotagsNudged.ne0CONUSne30x8.cesm220.2012-01.cam.h2.YYYY-MM-DD-00000.nc
  • Daily averages on native grid

f.e22.FCcotagsNudged.ne0CONUSne30x8.cesm220.2012-01.cam.h3.YYYY-MM-DD-00000.nc
  • Interpolated to 0.9x1.25 regular grid, 6-hr output
  • Many species suitable for regional model boundary conditions

f.e22.FCcotagsNudged.ne0CONUSne30x8.cesm220.2012-01.cam.h4.YYYY-MM-DD-00000.nc
  • On native grid, hourly output
  • Some species only available for surface layer: {species}_SRF
Community Simulation: Timeseries files

Directory:
/glade/campaign/acom/acm-climate/times/C0_CONUS/
f.e22.FCcotagsNudged.ne0CONUSNe30x8.cesm220.2012-01/atm/proc/tseries/

• Each file has 1 variable for a month (hourly output) or a year (monthly averages)

./month_1/
• $casename.cam.h0.{species}.YYYY01-YYYY12.nc (interpolated to regular grid)
• $casename.cam.h1.{species}.YYYY01-YYYY12.nc (on native grid)

./day_1/
• $casename.cam.h2.{species}.YYYYMM01-YYYYMM31 : one species, daily averages for 1 month
Demonstration of python plotting examples
Python Plotting Examples

Tutorial example notebooks:  https://github.com/NCAR/musica-tutorial

Examples of a variety of plots:  
https://ncar.github.io/CAM-chem/index.html

Contributions welcome!
We are going to learn

• Read MUSICAv0 output file(s)
• View file structure
• Basic array manipulation
• Various 2D map applications with an unstructured grid output (global, regional, custom colorbar, overlaying observations, log-scale plot, multi-panel, gridlines)
• Calculate global/regional emission total
• Calculate and plot vertical profile
• Simple regridding and cross section calculation
• Find an index of a specific location on the unstructured MUSICAv0 grid
• Timeseries plot
• Regrid unstructured grid data to structured grid data and save as a new NetCDF file