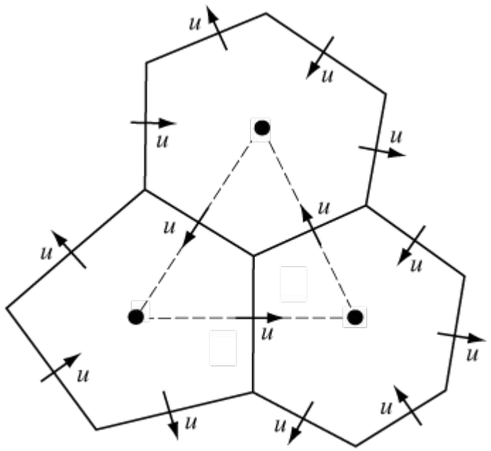
MPASSIT: A scalable tool for MPAS data post-processing

Larissa Reames
Cooperative Institute for Severe and High-Impact Weather Research and Operations (CIWRO),
University of Oklahoma, Norman, OK
NOAA/OAR National Severe Storms Laboratory, Norman, OK

Motivation

- MPAS data can be difficult to visualize or analyze
- Rudimentary lat-lon conversion provided with MPAS is limited
- NSSL decided late 2022 to run CONUS MPAS runs – how to post-process?
- Solution: MPASSIT, based on code and methods of chgres_cube, a UFS pre-processing utility



MPASSIT Overview

- Interpolates data from the irregular MPAS mesh to any regular grid
  - regional or global
  - output grid specified w/ wrfout or grid parameters like WPS
- Runs in parallel and supports bilinear, conservative, and nearest neighbor interpolation
- Variables processed and target grid properties (like WPS parameters) are configurable at runtime
- Output in netCDF format, optionally WRF-identical and UPP-compatible
- https://github.com/LarissaReames-NOAA/MPASSIT

Compiling MPASSIT

1. Load necessary libraries
   a. Necessary libraries
      i. basic: cmake, intel/gnu, mpi
      ii. specific: netcdf (?hdf5), esmf
   b. NOAA RDHPCS platforms Hera/Jet: modulefiles are included; no need to load separately
   c. Other platforms: create a `build.[target]` file in modulefiles folder that sets appropriate environmental variables
2. Build command:

```
./build.sh [target=target compiler=compiler debug=true/false]
```

Running MPASSIT

- Run directory must contain:
  - histlist_2d, histlist_3d, histlst_soil (if processing soil fields), diaglist; these tell MPASSIT which variables to process
  - namelist
- Ensure that you're using enough memory based on how large your input (MPAS)/target(output) grids are and how many variables you're requesting (esp. 3d variables)
- Example run command from a slurm submission script to go from ~3km CONUS MPAS to 3km CONUS target grid and only a few 3d arrays processed:

```
srun -n 120 mpassit namelist.input
```

Sample namelist – ‘lambert’ target grid

```
&config
  grid_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.init.nc_0325"
  hist_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.history.2024-03-25_09.00.00.nc"
  diag_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.diag.2024-03-25_09.00.00.nc"
  block_decomp_file="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.graph.info.part.120"
  output_file="/lfs4/NAGAPE/hpc-wof1/lreames/mpassit_out.nc"
  interp_diag=.true.
  interp_hist=.true.
  wrf_mod_vars=.true.
  esmf_log=.true.
  nx = 1800
  ny = 1060
  dx = 3000.0
  dy = 3000.0
  ref_lat = 38.50
  ref_lon = -97.50
  truelat1 = 38.5
  truelat2 = 38.5
  stand_lon = -97.5
  target_grid_type = 'lambert'
/
```

Files to process; grid file can be either grid or init file

namelist option not shown: set **is_regional** = .false. for a global grid (default is .true.)

Sample namelist – ‘lambert’ target grid

```
&config
  grid_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.init.nc_0325"
  hist_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.history.2024-03-25_09.00.00.nc"
  diag_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.diag.2024-03-25_09.00.00.nc"
  block_decomp_file="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.graph.info.part.120"
  output_file="/lfs4/NAGAPE/hpc-wof1/lreames/mpassit_out.nc"
  interp_diag=.true.
  interp_hist=.true.
  wrf_mod_vars=.true.
  esmf_log=.true.
  nx = 1800
  ny = 1060
  dx = 3000.0
  dy = 3000.0
  ref_lat = 38.50
  ref_lon = -97.50
  truelat1 = 38.5
  truelat2 = 38.5
  stand_lon = -97.5
  target_grid_type = 'lambert'
/
```

block decomposition file from gpmetis; does not have to be the same one you use for forecasts, just make sure you run MPASSIT with the same # of cores as the block decomp file is configured for; OPTIONAL but highly recommended for larger (CONUS) grids

namelist option not shown: set **is_regional** = .false. for a global grid (default is .true.)

Sample namelist – ‘lambert’ target grid

```
&config
  grid_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.init.nc_0325"
  hist_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.history.2024-03-25_09.00.00.nc"
  diag_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.diag.2024-03-25_09.00.00.nc"
  block_decomp_file="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.graph.info.part.120"
  output_file="/lfs4/NAGAPE/hpc-wof1/lreames/mpassit_out.nc"
  interp_diag=.true.
  interp_hist=.true.
  wrf_mod_vars=.true.
  esmf_log=.true.
  nx = 1800
  ny = 1060
  dx = 3000.0
  dy = 3000.0
  ref_lat = 38.50
  ref_lon = -97.50
  truelat1 = 38.5
  truelat2 = 38.5
  stand_lon = -97.5
  target_grid_type = 'lambert'
/
```

Full pathname of desired output file

namelist option not shown: set **is_regional** = .false. for a global grid (default is .true.)

Sample namelist – ‘lambert’ target grid

```
&config
  grid_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.init.nc_0325"
  hist_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.history.2024-03-25_09.00.00.nc"
  diag_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.diag.2024-03-25_09.00.00.nc"
  block_decomp_file="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.graph.info.part.120"
  output_file="/lfs4/NAGAPE/hpc-wof1/lreames/mpassit_out.nc"
  interp_diag=.true.
  interp_hist=.true.
  wrf_mod_vars=.true.
  esmf_log=.true.
  nx = 1800
  ny = 1060
  dx = 3000.0
  dy = 3000.0
  ref_lat = 38.50
  ref_lon = -97.50
  truelat1 = 38.5
  truelat2 = 38.5
  stand_lon = -97.5
  target_grid_type = 'lambert'
/
```

T/F whether to process hist and/or diag files

namelist option not shown: set **is_regional** = .false. for a global grid (default is .true.)

Sample namelist – ‘lambert’ target grid

```
&config
  grid_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.init.nc_0325"
  hist_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.history.2024-03-25_09.00.00.nc"
  diag_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.diag.2024-03-25_09.00.00.nc"
  block_decomp_file="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.graph.info.part.120"
  output_file="/lfs4/NAGAPE/hpc-wof1/lreames/mpassit_out.nc"
  interp_diag=.true.
  interp_hist=.true.
  wrf_mod_vars=.true.
  esmf_log=.true.
  nx = 1800
  ny = 1060
  dx = 3000.0
  dy = 3000.0
  ref_lat = 38.50
  ref_lon = -97.50
  truelat1 = 38.5
  truelat2 = 38.5
  stand_lon = -97.5
  target_grid_type = 'lambert'
/
```

Whether to make files identical to WRF output (metadata, additional variables, staggered winds). Use .true. if you plan to process through UPP

namelist option not shown: set **is_regional** = .false. for a global grid (default is .true.)

Sample namelist – ‘lambert’ target grid

```
&config
  grid_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.init.nc_0325"
  hist_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.history.2024-03-25_09.00.00.nc"
  diag_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.diag.2024-03-25_09.00.00.nc"
  block_decomp_file="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.graph.info.part.120"
  output_file="/lfs4/NAGAPE/hpc-wof1/lreames/mpassit_out.nc"
  interp_diag=.true.
  interp_hist=.true.
  wrf_mod_vars=.true.
  esmf_log=.true.
  nx = 1800
  ny = 1060
  dx = 3000.0
  dy = 3000.0
  ref_lat = 38.50
  ref_lon = -97.50
  truelat1 = 38.5
  truelat2 = 38.5
  stand_lon = -97.5
  target_grid_type = 'lambert'
/
```

T/F Create ESMF log files. Useful for debugging but can make your directory messy if using large # processes.

namelist option not shown: set **is_regional** = .false. for a global grid (default is .true.)

Sample namelist – ‘lambert’ target grid

```
&config
  grid_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.init.nc_0325"
  hist_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.history.2024-03-25_09.00.00.nc"
  diag_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.diag.2024-03-25_09.00.00.nc"
  block_decomp_file="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.graph.info.part.120"
  output_file="/lfs4/NAGAPE/hpc-wof1/lreames/mpassit_out.nc"
  interp_diag=.true.
  interp_hist=.true.
  wrf_mod_vars=.true.
  esmf_log=.true.
  nx = 1800
  ny = 1060
  dx = 3000.0
  dy = 3000.0
  ref_lat = 38.50
  ref_lon = -97.50
  truelat1 = 38.5
  truelat2 = 38.5
  stand_lon = -97.5
  target_grid_type = 'lambert'
/
```

Target grid specifications. Set these exactly as you would in WPS.

namelist option not shown: set **is_regional** = .false. for a global grid (default is .true.)

Sample namelist – ‘lambert’ target grid

```
&config
  grid_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.init.nc_0325"
  hist_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.history.2024-03-25_09.00.00.nc"
  diag_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.diag.2024-03-25_09.00.00.nc"
  block_decomp_file="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.graph.info.part.120"
  output_file="/lfs4/NAGAPE/hpc-wof1/lreames/mpassit_out.nc"
  interp_diag=.true.
  interp_hist=.true.
  wrf_mod_vars=.true.
  esmf_log=.true.
  nx = 1800
  ny = 1060
  dx = 3000.0
  dy = 3000.0
  ref_lat = 38.50
  ref_lon = -97.50
  truelat1 = 38.5
  truelat2 = 38.5
  stand_lon = -97.5
  target_grid_type = 'lambert'
/
```

Grid type. For parameter-specified grids, options of ‘lambert’ and ‘lat-lon’ are available (‘mercator’ and ‘polar’ untested)

namelist option not shown: set **is_regional** = .false. for a global grid (default is .true.)

Sample namelist – ‘file’ target grid

```
&config
  grid_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.init.nc_0325"
  hist_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.history.2024-03-25_09.00.00.nc"
  diag_file_input_grid="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.diag.2024-03-25_09.00.00.nc"
  block_decomp_file="/lfs4/NAGAPE/hpc-wof1/lreames/MPASSIT/run/mpas.graph.info.part.120"
  output_file="/lfs4/NAGAPE/hpc-wof1/lreames/out_mpassit.nc"
  interp_diag=.true.
  interp_hist=.true.
  wrf_mod_vars=.true.
  esmf_log=.true.
  target_grid_type = 'file'
  file_target_grid = "/lfs4/NAGAPE/hpc-wof1/lreames/wrfinput_d01"
/
```

Remove all grid specification details and add new “file_target_grid” option and set target_grid_type=’file’

namelist option not shown: set **is_regional** = .false. for a global grid (default is .true.)

Sample histlist_3d, histlist_2d

histlist_3d – 3-D vars from history files

zgrid	PHB
w	W
theta	T
uReconstructZonal	U
uReconstructMeridional	V
qv	QVAPOR
qc	QCLOUD
qr	QRAIN
qi	QICE
qs	QSNOW
qg	QGRAUP
ni	QNICE
nr	QNRAIN
pressure	P_HYD
rho	MUB

variable name in MPAS history file

name you'd like in the output

histlist_2d – 2-D vars from history files

surface_pressure	PSFC
xland	XLAND
skintemp	TSK
snow	SNOW
snowh	SNOWH
sst	SST

diaglist (variables from diag files) and **histlist_soil** (soil temp/moisture variables from history files) will look similar

Limitations

- Default of **bilinear interpolation** for any variable not hardcoded in list of conservative (‘snow’, ‘snowh’) or nearest neighbor ('ivgtyp', 'isltyp', 'xland', 'landmask'). You could modify these yourself (in **init_input_hist_fields** in **input_data.F90**) if you’d like to change that behavior.
- All 3d history variables are assumed to be of vertical dimension **nz** except **zgrid** and **w**. Again, you could modify this list if you need to.
- A note on target grid choice : You can choose whatever target grid you’d like so long as part of it overlaps with the MPAS grid. It can be much smaller, or larger. If larger, outlying regions will be zero-filled (see planned improvements for future changes).

Future Improvements

- Make variable-specific regridding options user-configurable in the histlist/diaglist files
- User-specified fill values for out-of-bounds grids

Final Thoughts

- https://github.com/LarissaReames-NOAA/MPASSIT
- Thank you to everyone who's tested the code or contributed changes, including Craig Schwartz and Yunheng Wang
- Please reach out! larissa.reames@noaa.gov or make a github issue if you find a bug or would like to contribute a new feature

Questions?