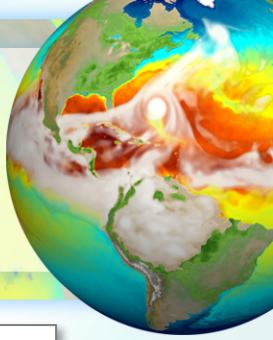
An Introduction to MPAS-Analysis



Main developers:

Xylar Asay-Davis

Milena Veneziani

Contributors:

Sterling Baldwin

Riley X. Brady

Darin Comeau

Charles Doutriaux

Jeremy Fyke

Matthew Hoffman

Joseph Kennedy

Mark Petersen

Stephen Price

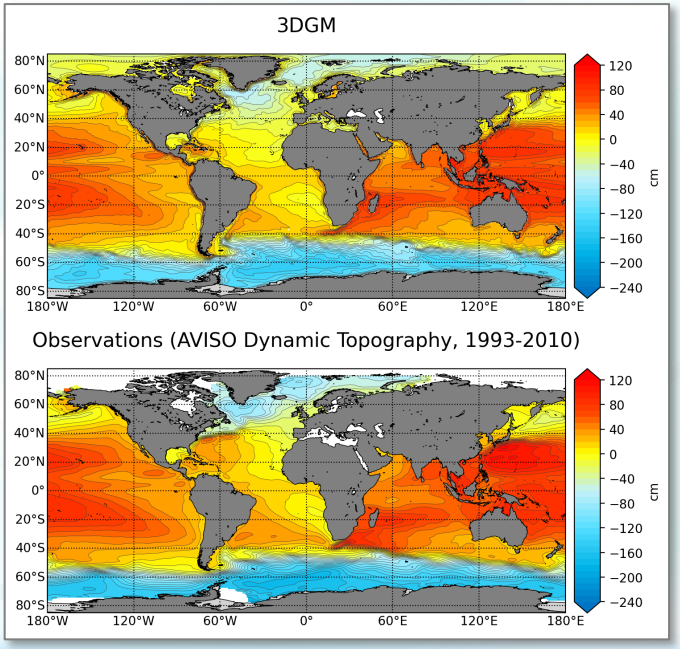
Kevin Rosa

Greg Streletz

Adrian Turner

Luke Van Roekel

Phillip J. Wolfram



Sea surface height

E3SM Energy Exascale Earth System Model

U.S. DEPARTMENT OF ENERGY

Outline

- Setting Up and Running MPAS-Analysis
- Example Results
- Most Common Config Options
- Under the Hood
- Future Plans

Outline

- Setting Up and Running MPAS-Analysis
- Example Results
- Most Common Config Options
- Under the Hood
- Future Plans

Installing MPAS-Analysis

- Details:
  - MPAS-Analysis Documentation
  - MPAS-Analysis Tutorial
- The gist:
  - Install Miniconda3

```
conda config --add channels conda-forge
conda create -n mpas-analysis mpas-analysis
conda activate mpas-analysis
```

  - Download
    - observations
    - mapping and mask files for standard meshes

```
download_analysis_data -o /path/to/mpas_analysis/diagnostics
```

E3SM Results for Input

- E3SM simulation directory:

```
baseDirectory = /global/cscratch1/sd/sprice/e3sm_scratch/cori-
knl/20200610.A_WCYCL1850-DIB-ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM
```

- A subdirectory with an MPAS-Ocean restart file:

```
runSubdirectory = run
```

- Subdirectories with ocean and sea-ice monthly averaged data:

```
oceanHistorySubdirectory = archive/ocn/hist
seaIceHistorySubdirectory = archive/ice/hist
```

- And namelists and “streams” files describing MPAS parameters and output:

```
oceanNamelistFileName = run/mpaso_in
oceanStreamsFileName = run/streams.ocean
seaIceNamelistFileName = run/mpassi_in
seaIceStreamsFileName = run/streams.seaice
```

Configuring MPAS-Analysis

- Configuration is with Python cfg (also called ini) files:

```
[runs]
# mainRunName is a name that identifies the simulation being analyzed.
mainRunName = runName

[execute]
# the number of parallel tasks (1 means tasks run in serial, the default)
parallelTaskCount = 1

# the parallelism mode in ncclimo ("serial" or "bck")
ncclimoParallelMode = serial
```

- The default config file contains over 1,000 config options
  - Lots of flexibility
  - A bit overwhelming
- Override defaults with one or more custom config files
  - We'll go over some common config options later in the presentation

Run the code

- Run:

```
$ mpas_analysis 20200610.A_WCYCL1850.ne30_ECwISC30to60E1r2.cori-knl.cfg
```

- (Better yet, run a batch job)
- Typical output:

```
Running tasks:   8% |###                                          | ETA:   0:02:34

Log files for executed tasks can be found in
/media/xylar/bbyates/analysis/output/GMPAS-QU240wLI/native_transects_mpas_tools/logs
Total setup time: 0:00:07.21
Total run time: 0:01:19.78
Generating webpage for viewing results...
Done.
```

- Copy and/or chmod the resulting web output so you can view it in a web portal

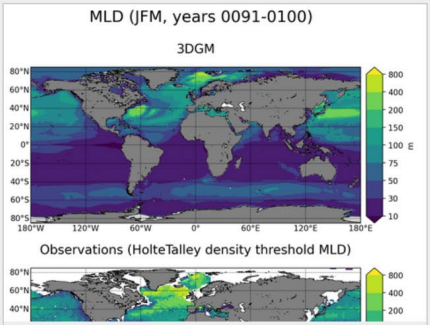
Web Interface

MPAS-Analysis Diagnostics

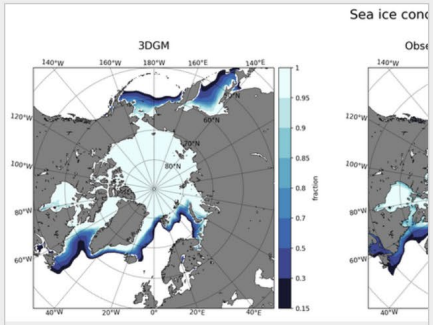
Run: 3DGM

MPAS
Model for Prediction Across Scales

Components



Ocean



Sea Ice

Web Interface: Provenance

Provenance

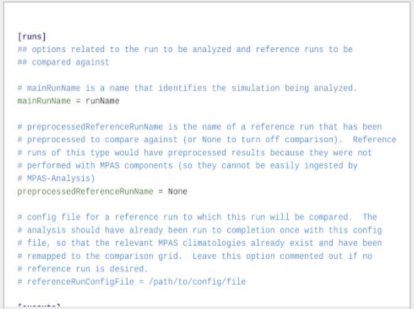
MPAS-Analysis version: 1.2.6

Git Hash: abcf191

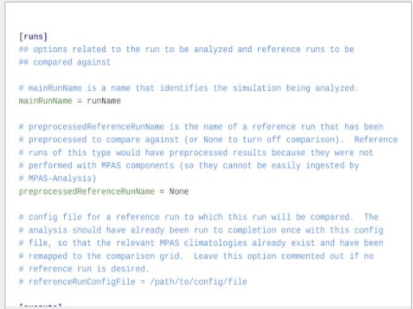
Command line:

```
/global/project/projectdirs/m3412/sprice/analysis/MPAS-Analysis/mpas_analysis/__main__.py configs/polarRegions.conf config.curren
```

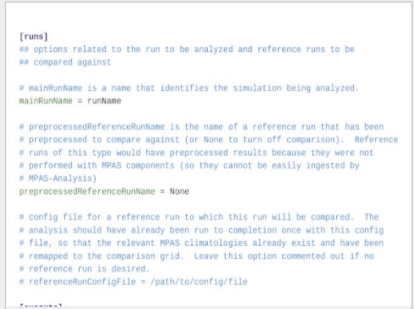
MPAS-Analysis Configuration Files



polarRegions.conf



config.current2



Complete Configuration File

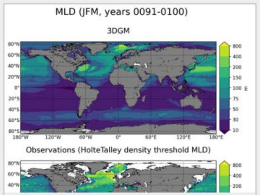
Web Interface: Ocean

MPAS-Analysis Diagnostics: Ocean

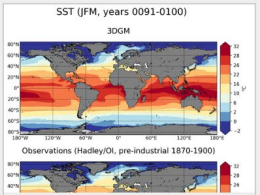
Run: 3DGM

MPAS
Model for Prediction Across Scales

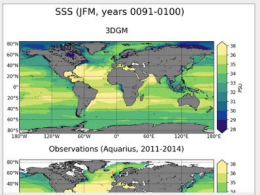
Quick Links



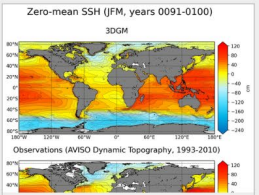
Global Mixed-Layer Depth



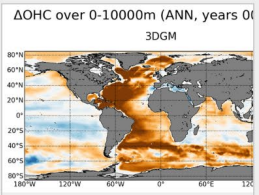
Global Sea Surface Temperature



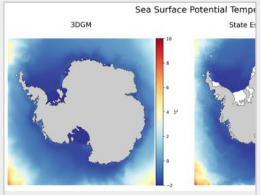
Global Sea Surface Salinity



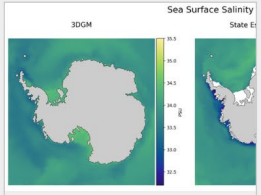
Global Sea Surface Height



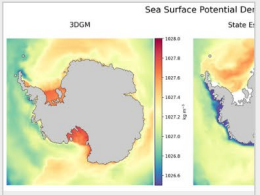
Global OHC Anomaly



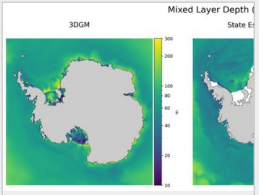
Antarctic Potential Temperature



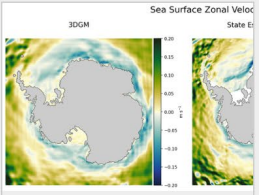
Antarctic Salinity



Antarctic Potential Density



Antarctic Mixed Layer Depth



Antarctic Zonal Velocity

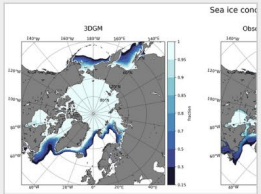
Web Interface: Sea Ice

MPAS-Analysis Diagnostics: Sea Ice

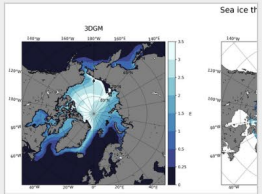
Run: 3DGM

MPAS
Model for Prediction Across Scales

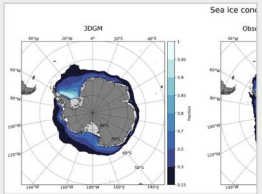
Quick Links



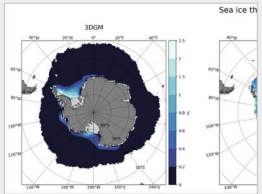
Northern-Hemisphere Sea-Ice Concentration



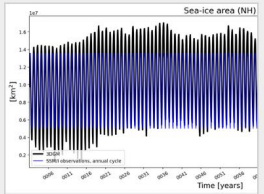
Northern-Hemisphere Sea-Ice Thickness



Southern-Hemisphere Sea-Ice Concentration



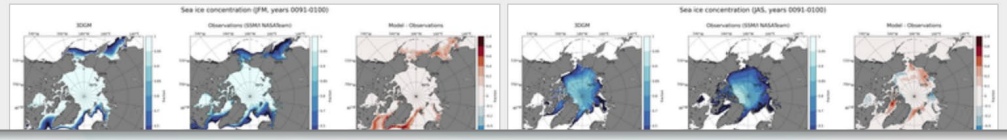
Southern-Hemisphere Sea-Ice Thickness



Time Series

Northern-Hemisphere Sea-Ice Concentration

Observations: SSM/I NASATeam



Log Files

- Recall from run output:

```
Log files for executed tasks can be found in /media/xylar/bbyates/analysis/output/GMPAS-QU240wLI/native_transects_mpas_tools/logs
```

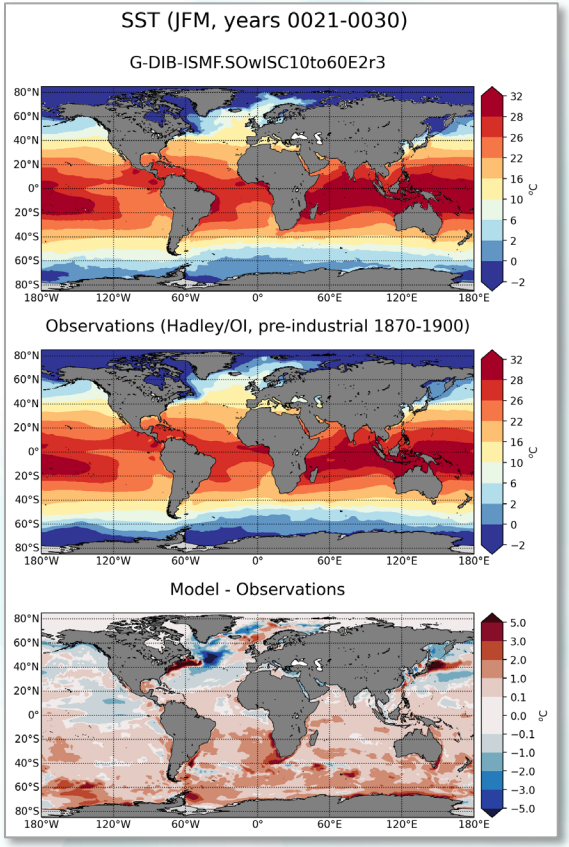
- One log file per analysis task (typically several hundred)
- Useful for diagnosing errors:

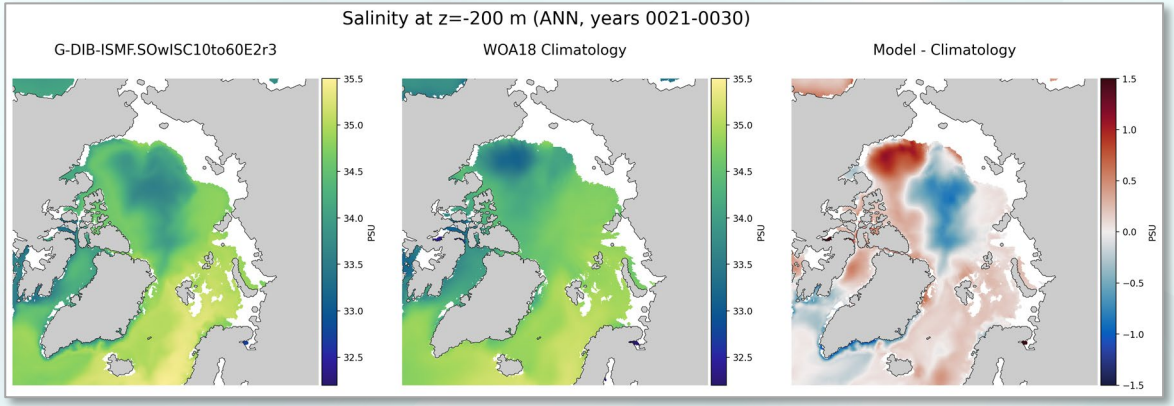
```
$ cd /media/xylar/bbyates/analysis/output/GMPAS-QU240wLI/native_transects_mpas_tools/logs
$ grep Error *.log
```

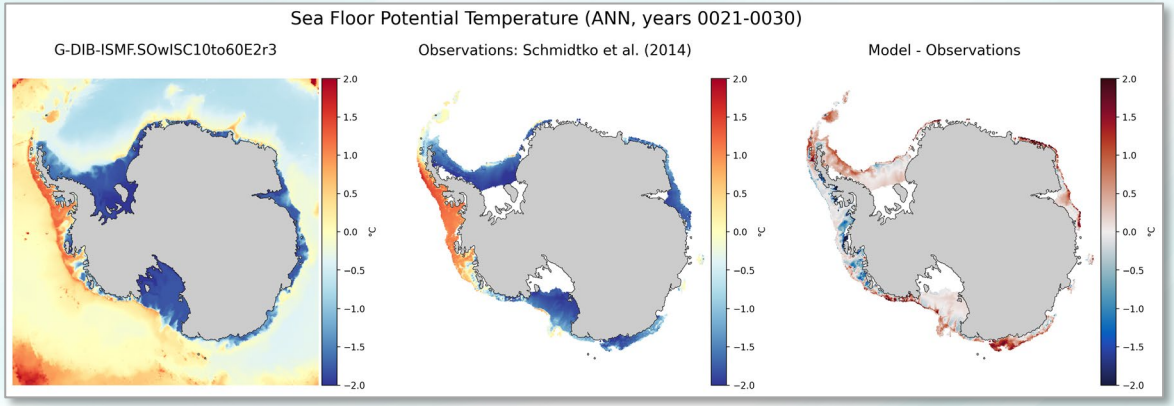
Outline

- Setting Up and Running MPAS-Analysis
- Example Results
- Most Common Config Options
- Under the Hood
- Future Plans

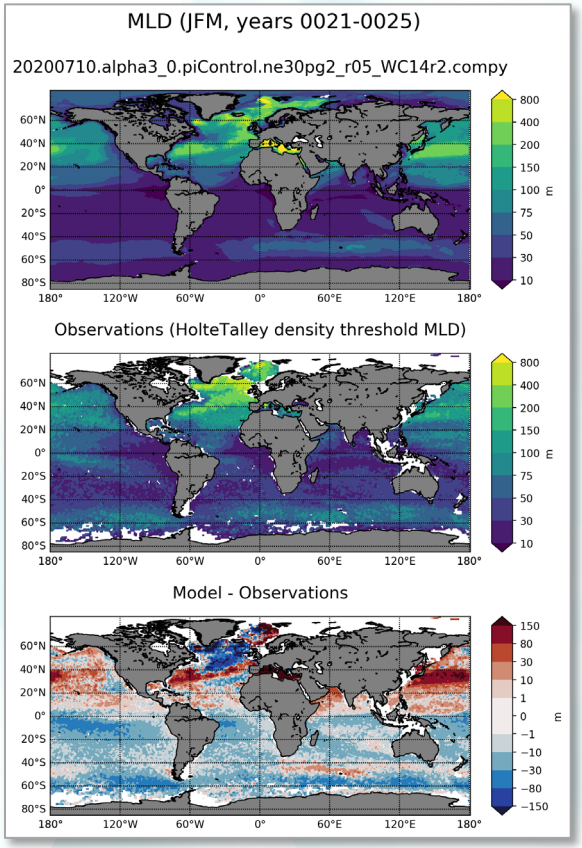
Temperature and Salinity: Var. Depths & Regions

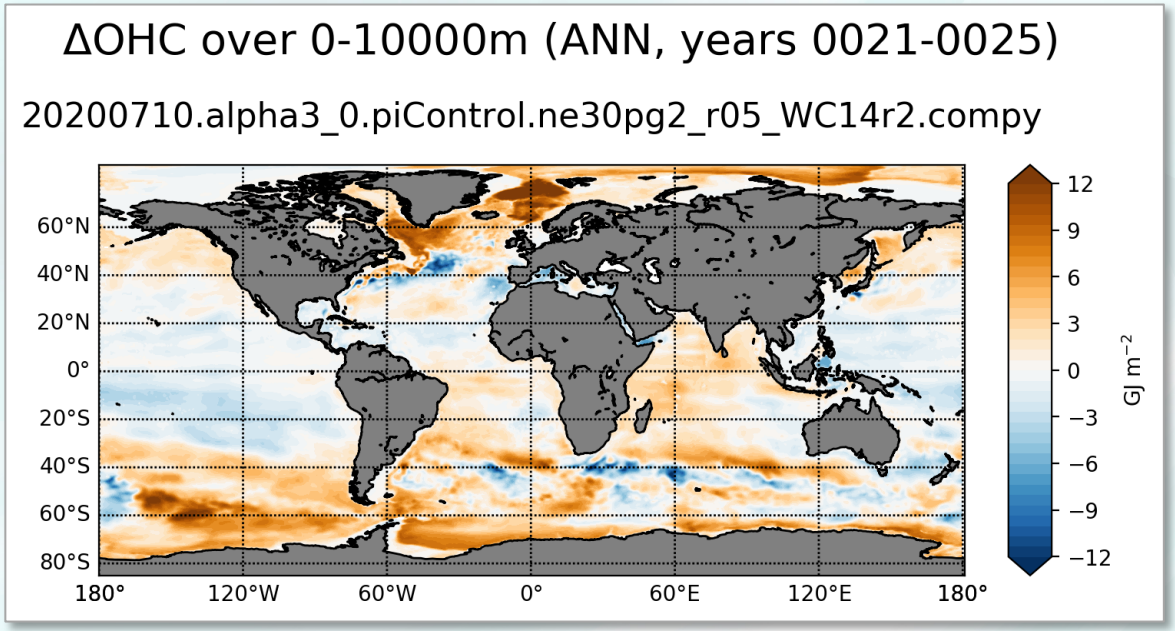




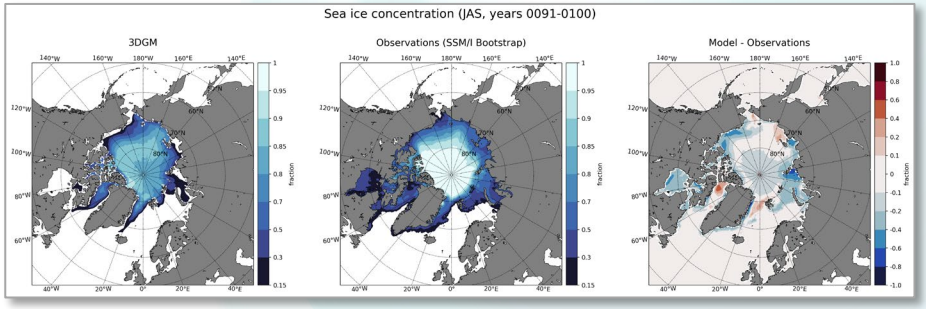


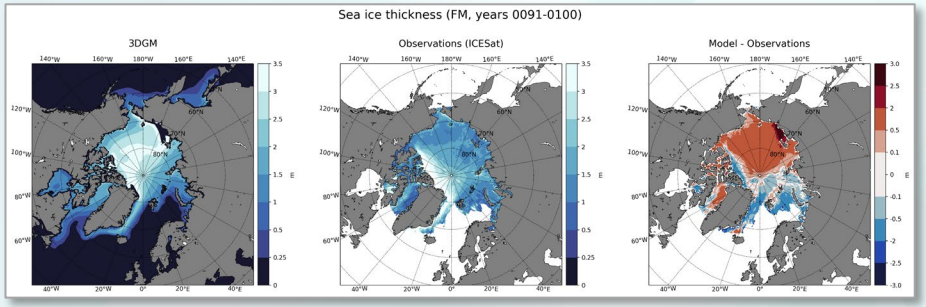
Mixed Layer Depth and Ocean Heat Content

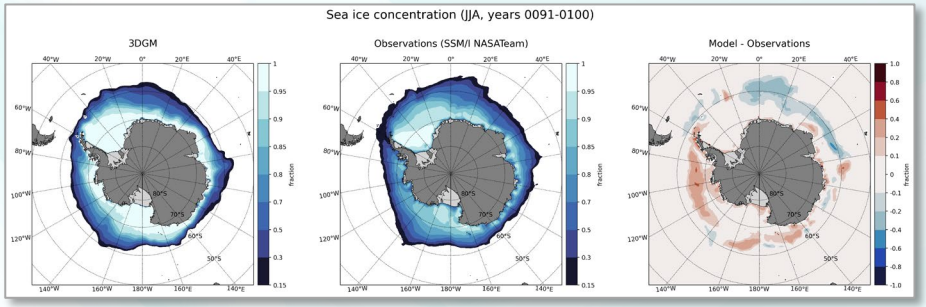


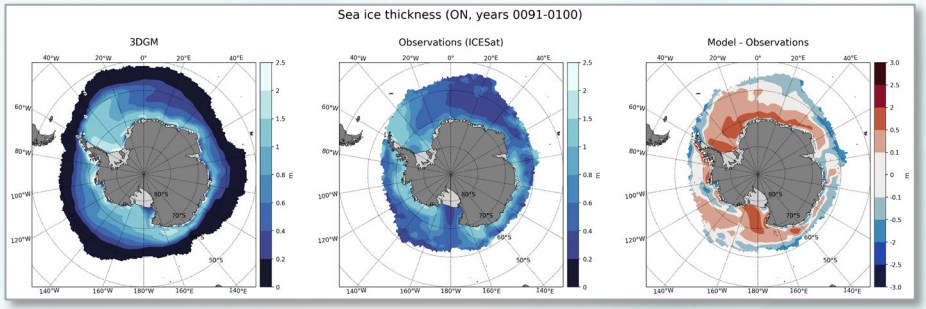


Sea Ice Concentration and Thickness

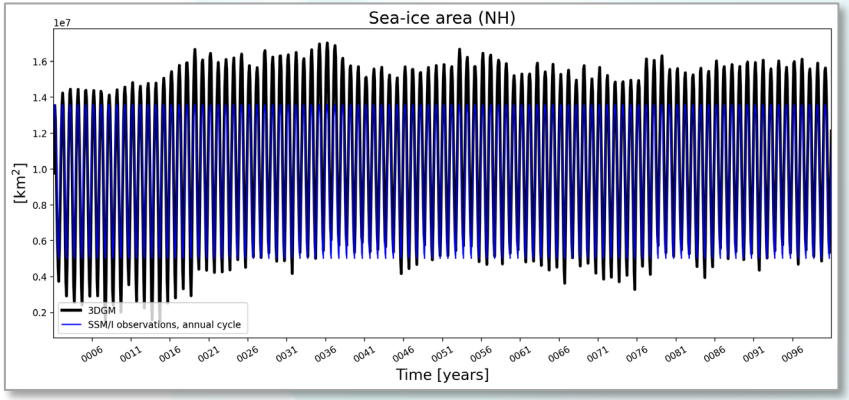


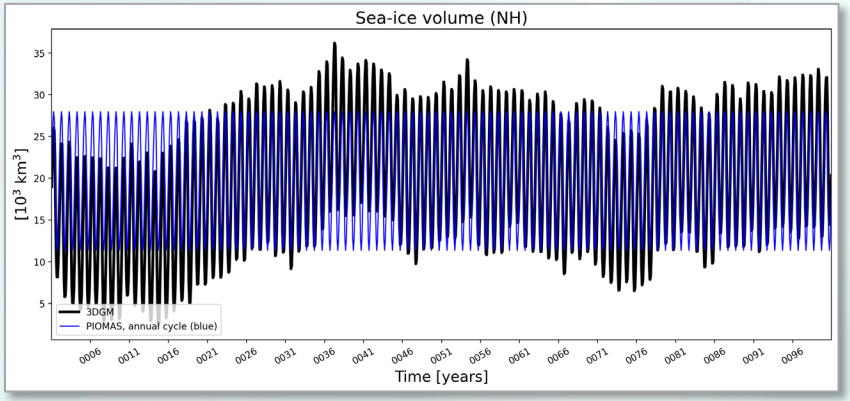


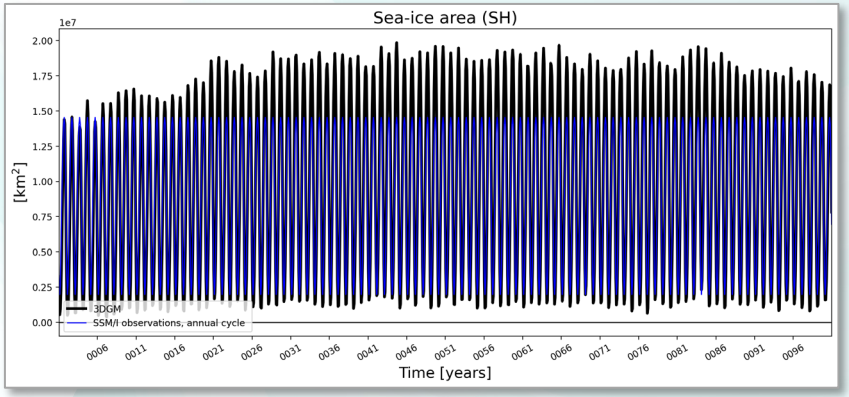


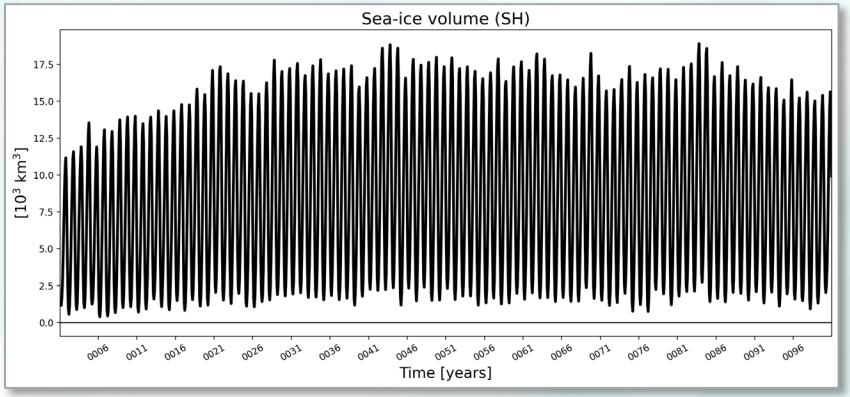


Sea Ice Area and Volume

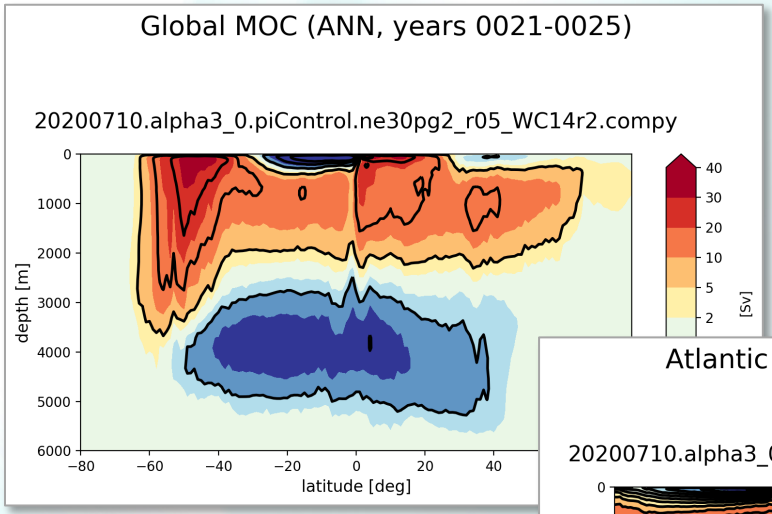


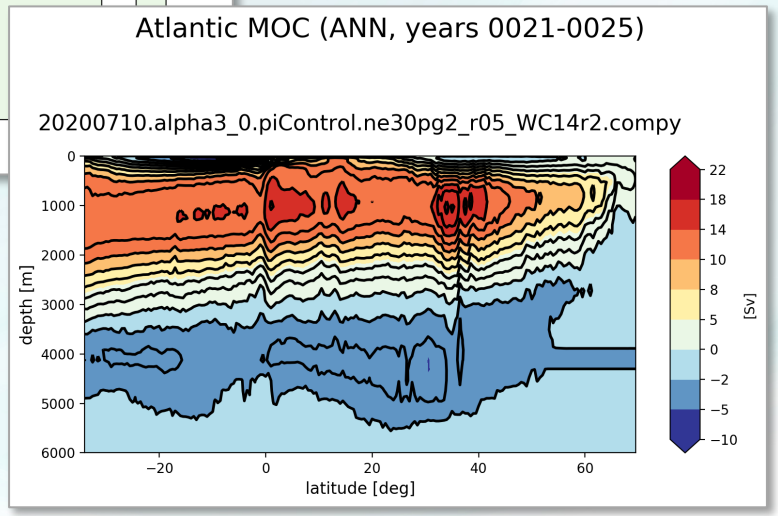


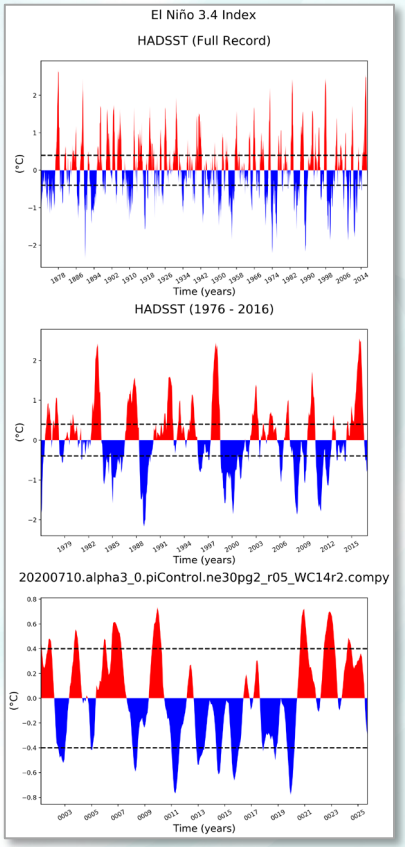




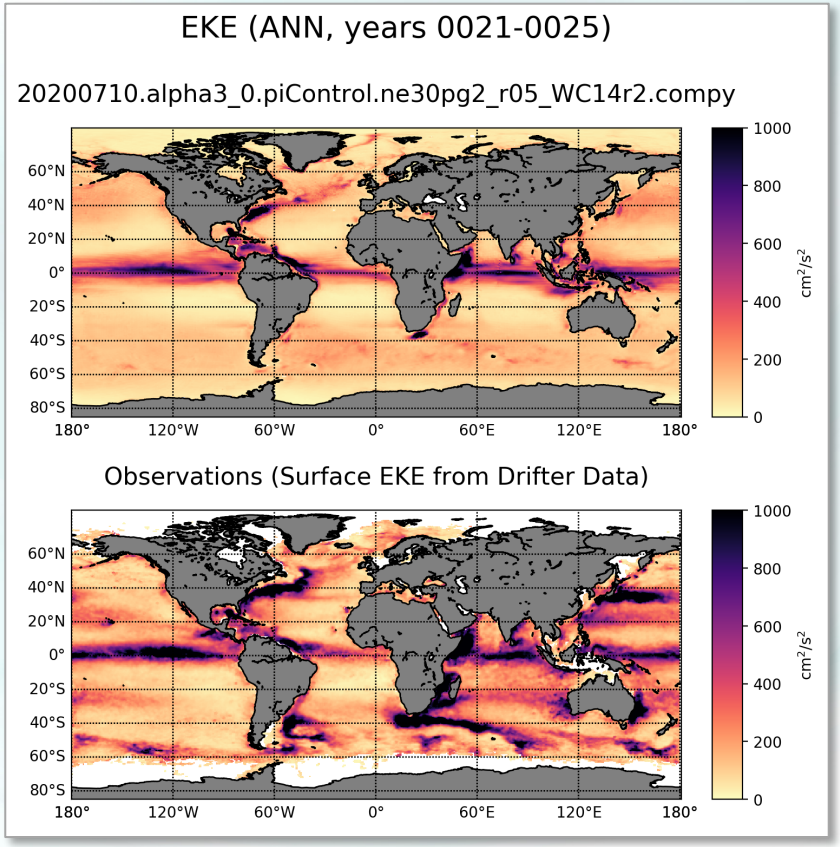
Meridional Overturning Circulation and El Niño



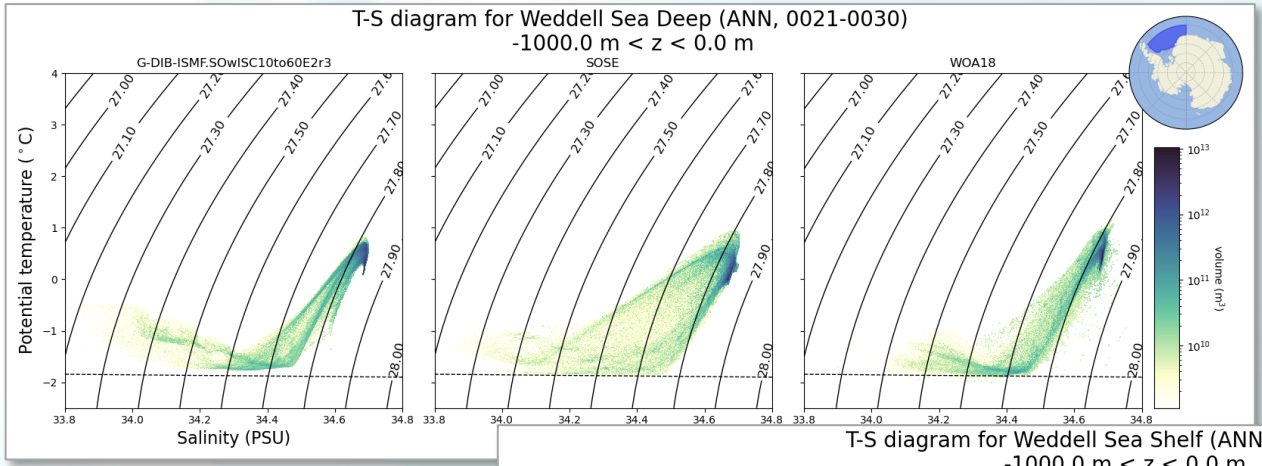


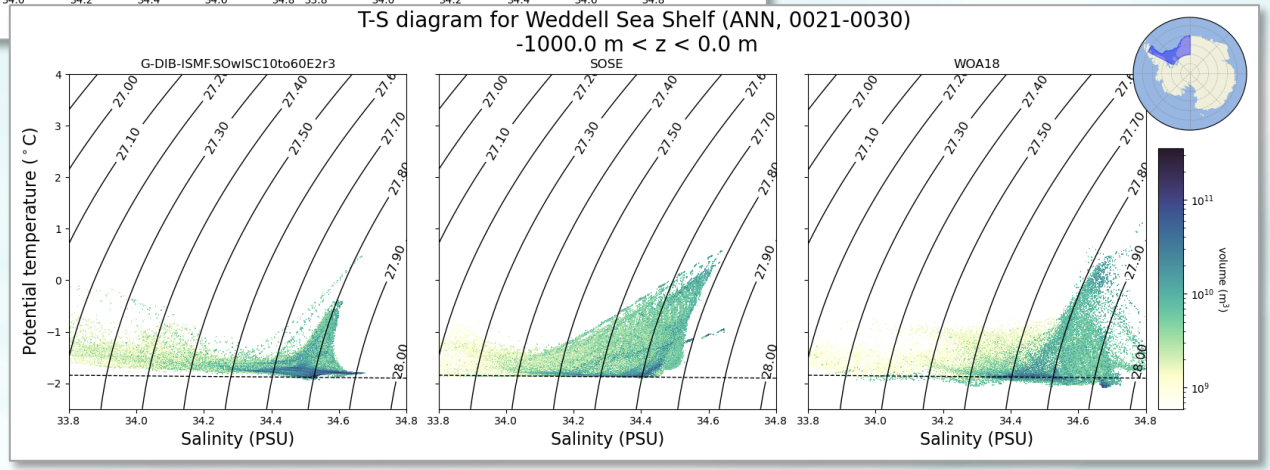


Eddy Kinetic Energy

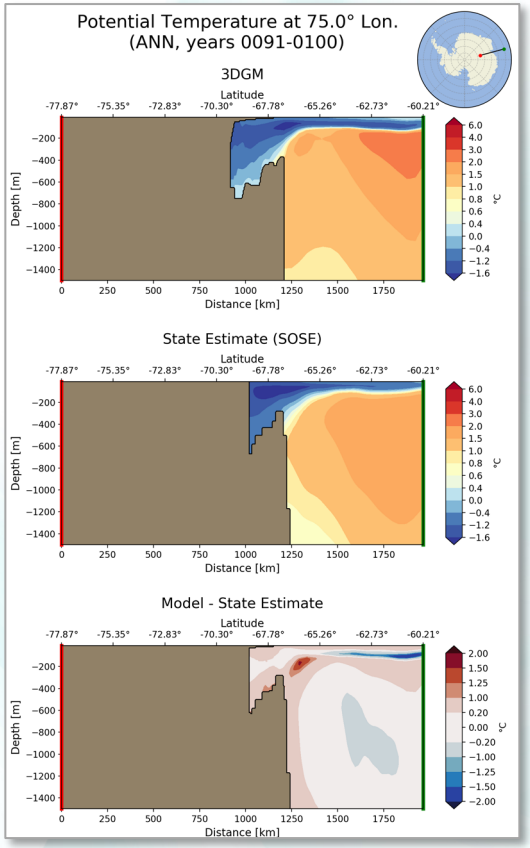


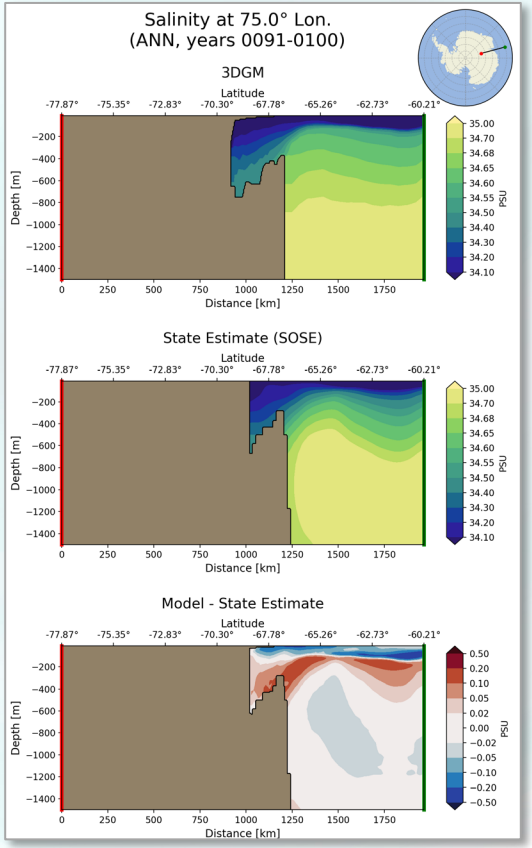
Temperature/Salinity Diagrams

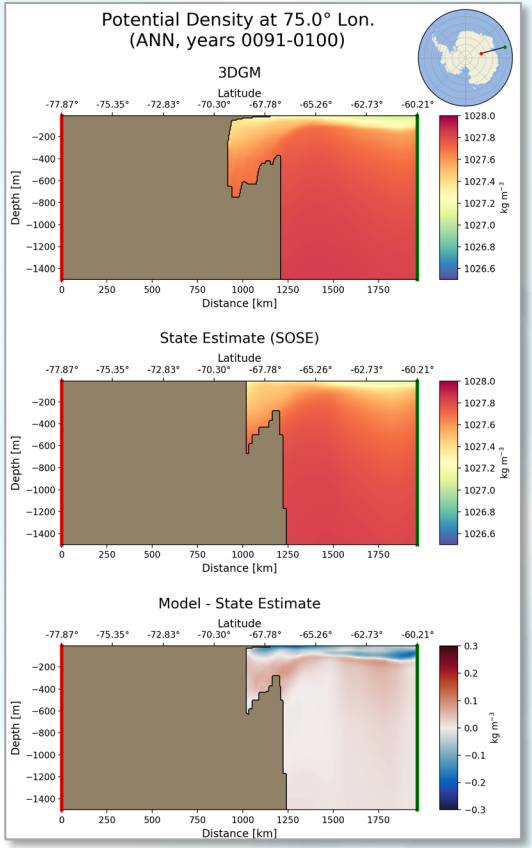




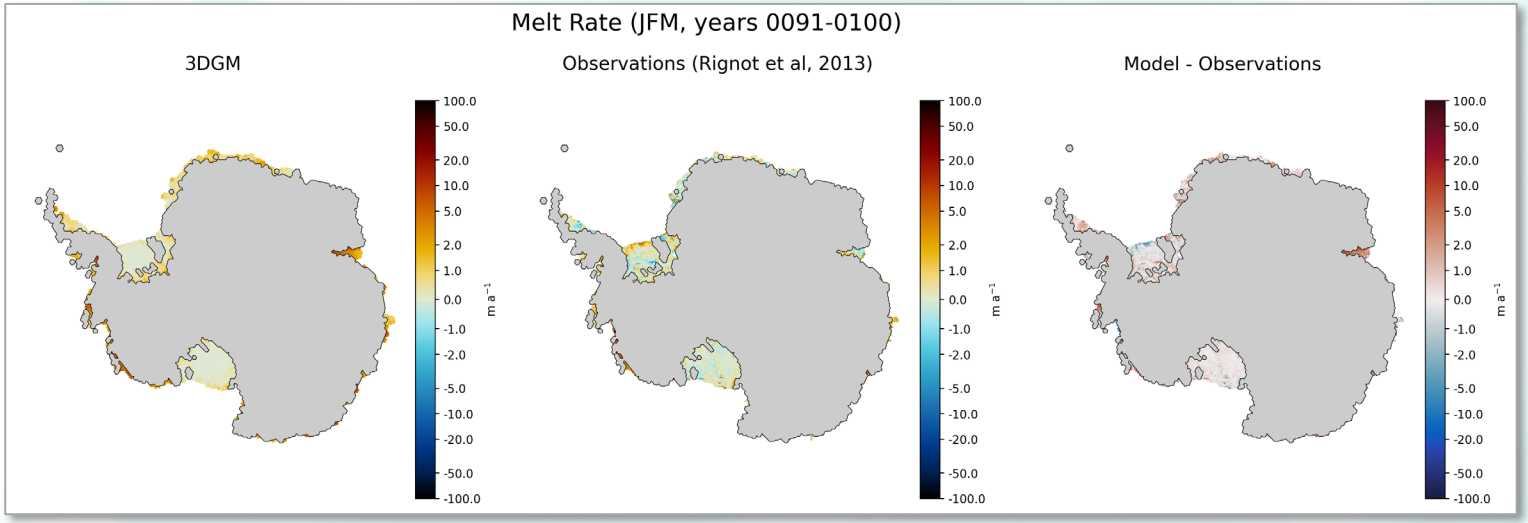
Ocean Transects

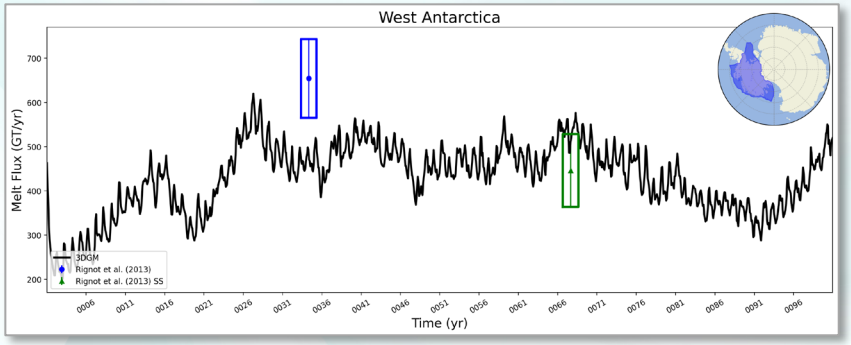


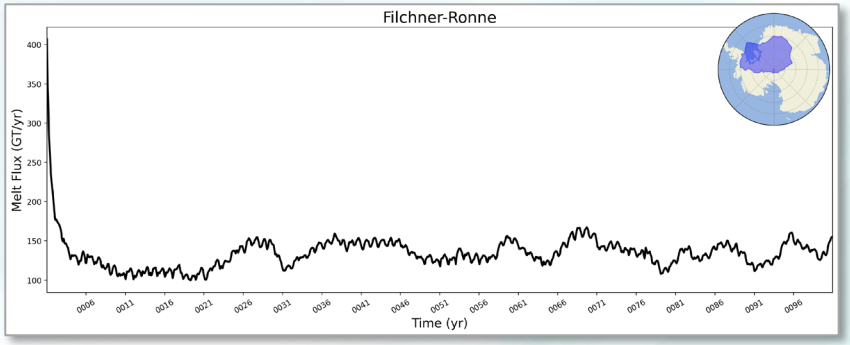




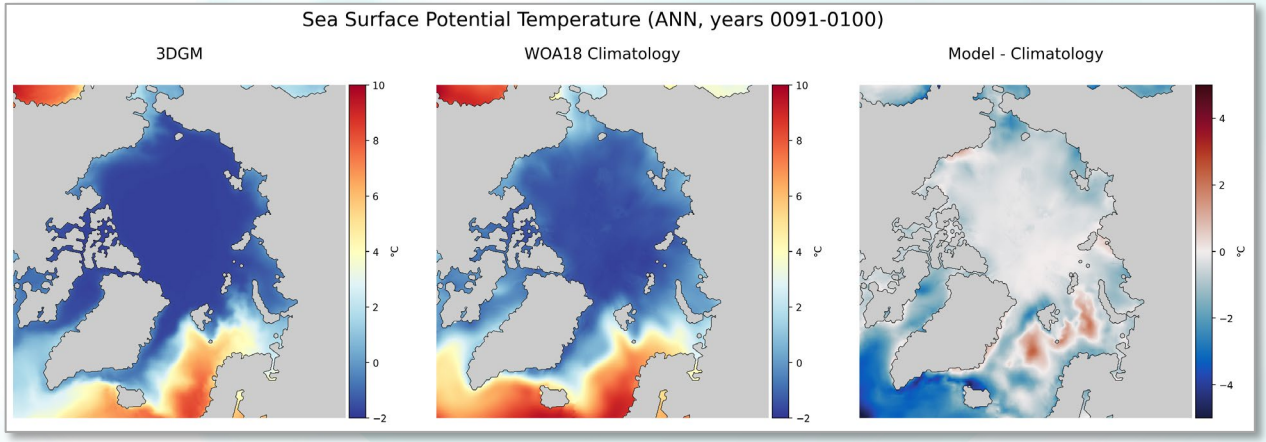
Melt Rates Below Ice Shelves

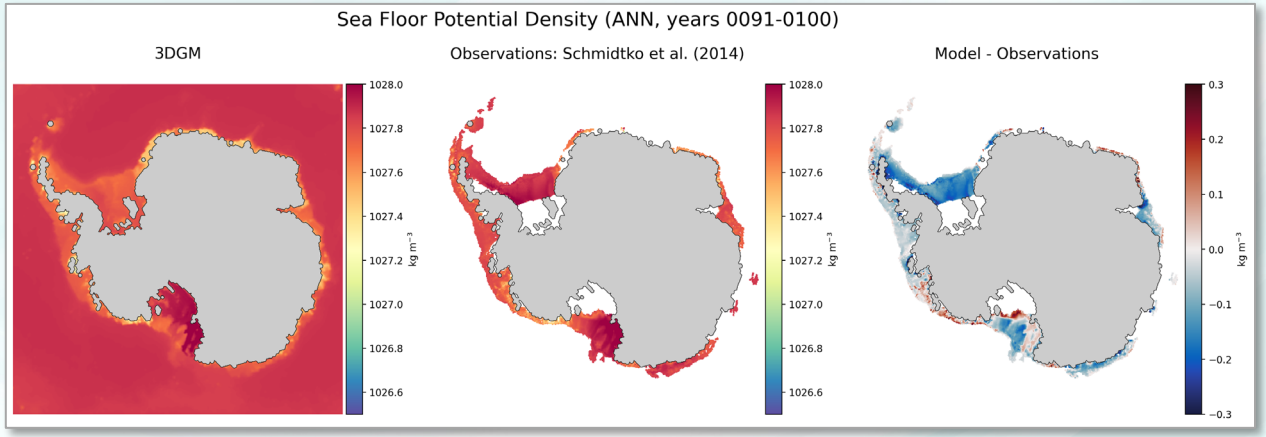




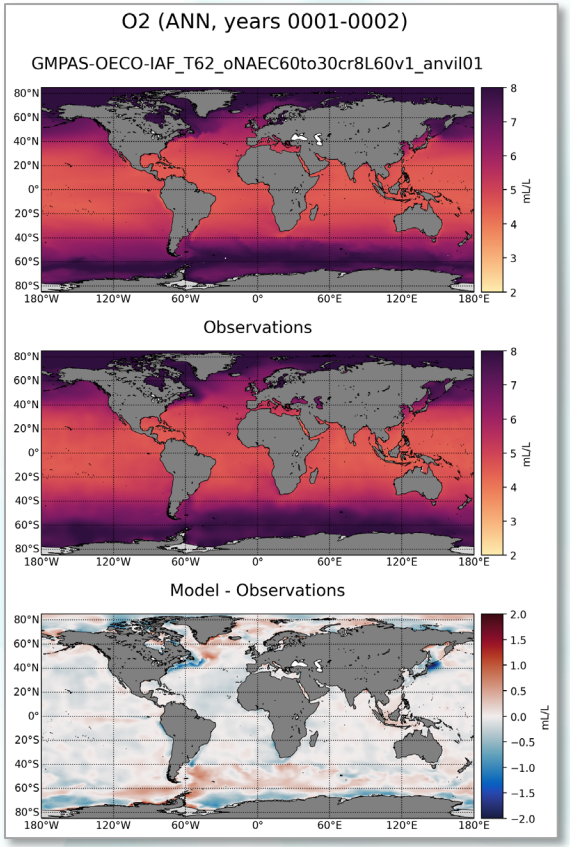


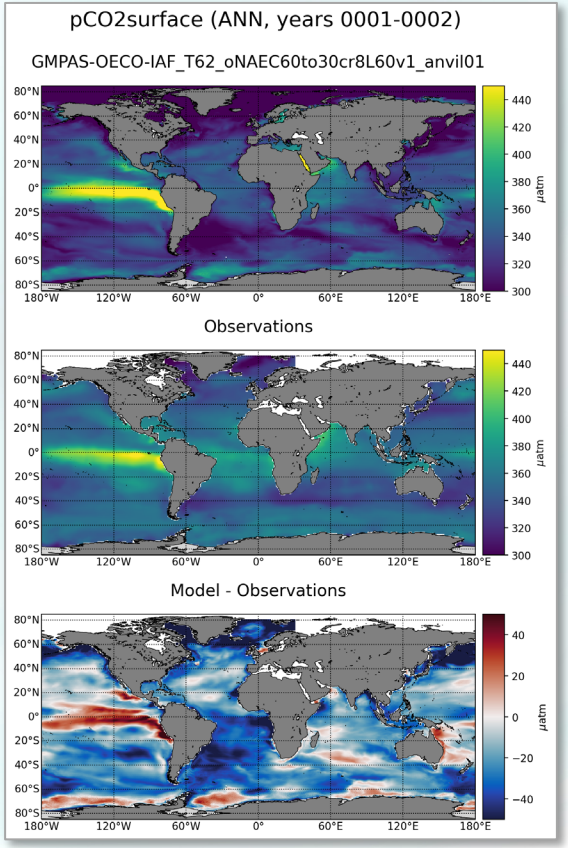
Arctic and Antarctic Maps

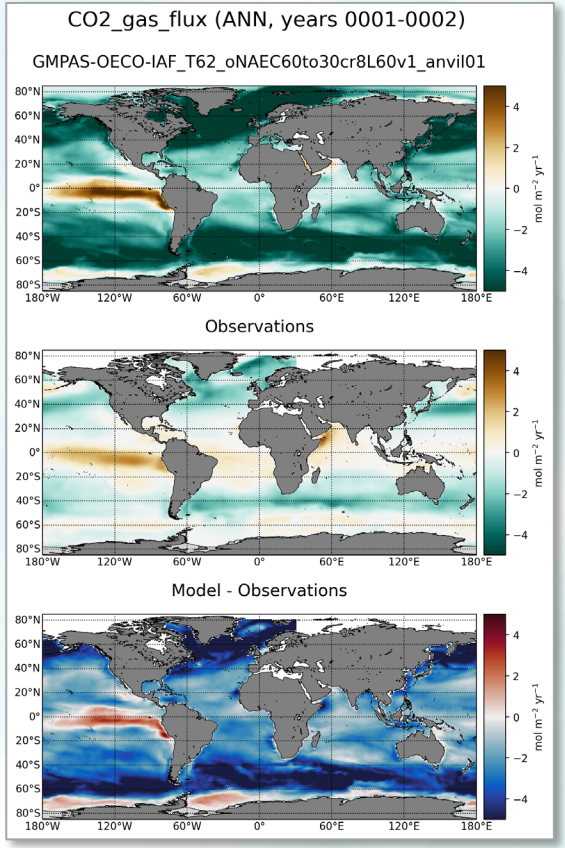




Biogeochemistry Surface Maps







Outline

- Setting Up and Running MPAS-Analysis
- Example Results
- Most Common Config Options
- Under the Hood
- Future Plans

Example Configuration Files

- config.example from the repo:

```
[runs]
# mainRunName is a name that identifies the simulation being analyzed.
mainRunName = runName

[execute]
# the number of parallel tasks (1 means tasks run in serial, the default)
parallelTaskCount = 1

# the parallelism mode in ncclimo ("serial" or "bck")
ncclimoParallelMode = serial

[diagnostics]
# The base path to the diagnostics directory
baseDirectory = /path/to/diagnostics

[input]
# directory containing model results
baseDirectory = /dir/for/model/output

# names of ocean and sea ice meshes (e.g. oEC60to30, oQU240, oRRS30to10, etc.)
mpasMeshName = mesh
…
```

Polar Regions Configuration File

- configs/polarRegions.conf for Cryosphere Campaign and High-Lat Project
- Defines “standard” config options for shared analysis

```
[climatologyMapWoa]
# comparison grid(s) ('latlon', 'antarctic', 'arctic') on which to plot analysis
comparisonGrids = ['antarctic', 'arctic']

# list of depths in meters (positive up) at which to analyze
depths = ['top', -50, -200, -400, -600, -800]

[climatologyMapWoaTemperature]
# A dictionary with keywords for the norm
normArgsResult = {'vmin': -2., 'vmax': 2.}

[climatologyMapWoaSalinity]
# A dictionary with keywords for the norm
normArgsResult = {'vmin': 33.8, 'vmax': 35.0}

[regionalTSDiagrams]
# the names of region groups to plot, each with its own section below
regionGroups = ['Antarctic Regions', 'Ocean Basins']
…
```

Example Job Scripts

- On most HPC machines, submit a batch job with a job script:

```
#!/bin/bash -l
#SBATCH --partition=regular
#SBATCH -C haswell
#SBATCH --nodes=1
#SBATCH --time=1:00:00
#SBATCH --account=e3sm
#SBATCH --job-name=mpas_analysis
#SBATCH --output=mpas_analysis.o%j
#SBATCH --error=mpas_analysis.e%j
#SBATCH -L cscratch1,SCRATCH,project

run_config_file="config.run_name_here"

export OMP_NUM_THREADS=1

source /global/cfs/cdirs/e3sm/software/anaconda_envs/load_latest_e3sm_unified.sh
export HDF5_USE_FILE_LOCKING=FALSE

srun -N 1 -n 1 mpas_analysis configs/polarRegions.conf $run_config_file
```

Most Common Configuration Options

```
[runs]
# mainRunName is a name that identifies the simulation being analyzed.
mainRunName = A_WCYCL1850-DIB-ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM
…
```

Most Common Configuration Options

```
[runs]
# mainRunName is a name that identifies the simulation being analyzed.
mainRunName = A_WCYCL1850-DIB-ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM

[execute]
# the number of parallel tasks (1 means tasks run in serial, the default)
parallelTaskCount = 1

# the parallelism mode in ncclimo ("serial" or "bck")
ncclimoParallelMode = serial
…
```

Most Common Configuration Options

```
[runs]
# mainRunName is a name that identifies the simulation being analyzed.
mainRunName = A_WCYCL1850-DIB-ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM

[execute]
# the number of parallel tasks (1 means tasks run in serial, the default)
parallelTaskCount = 1

# the parallelism mode in ncclimo ("serial" or "bck")
ncclimoParallelMode = serial

[diagnostics]
# The base path to the diagnostics directory
baseDirectory = /path/to/diagnostics
…
```

Most Common Configuration Options

```
…
[input]
# directory containing model results
baseDirectory = /global/cscratch1/sd/sprice/e3sm_scratch/cori-
knl/20200610.A_WCYCL1850-DIB-ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM
…
```

Most Common Configuration Options

```
…
[input]
# directory containing model results
baseDirectory = /global/cscratch1/sd/sprice/e3sm_scratch/cori-
knl/20200610.A_WCYCL1850-DIB-ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM

# subdirectory containing restart files
runSubdirectory = run
# subdirectory for ocean history files
oceanHistorySubdirectory = archive/ocn/hist
# subdirectory for sea ice history files
seaIceHistorySubdirectory = archive/ice/hist
…
```

Most Common Configuration Options

```
…
[input]
# directory containing model results
baseDirectory = /global/cscratch1/sd/sprice/e3sm_scratch/cori-
knl/20200610.A_WCYCL1850-DIB-ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM

# subdirectory containing restart files
runSubdirectory = run
# subdirectory for ocean history files
oceanHistorySubdirectory = archive/ocn/hist
# subdirectory for sea ice history files
seaIceHistorySubdirectory = archive/ice/hist

# names of namelist and streams files, either a path relative to baseDirectory
# or an absolute path.
oceanNamelistFileName = run/mpaso_in
oceanStreamsFileName = run/streams.ocean
seaIceNamelistFileName = run/mpassi_in
seaIceStreamsFileName = run/streams.seaice
…
```

Most Common Configuration Options

```
…
[input]
# directory containing model results
baseDirectory = /global/cscratch1/sd/sprice/e3sm_scratch/cori-
knl/20200610.A_WCYCL1850-DIB-ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM

# subdirectory containing restart files
runSubdirectory = run
# subdirectory for ocean history files
oceanHistorySubdirectory = archive/ocn/hist
# subdirectory for sea ice history files
seaIceHistorySubdirectory = archive/ice/hist

# names of namelist and streams files, either a path relative to baseDirectory
# or an absolute path.
oceanNamelistFileName = run/mpaso_in
oceanStreamsFileName = run/streams.ocean
seaIceNamelistFileName = run/mpassi_in
seaIceStreamsFileName = run/streams.seaice

# names of ocean and sea ice meshes (e.g. oEC60to30, oQU240, oRRS30to10, etc.)
mpasMeshName = ECwISC30to60E1r2
…
```

Most Common Configuration Options

```
…
[output]
# directory where analysis should be written
baseDirectory = /project/projectdirs/m3412/sprice/analysis/output/20200610.A_WCYCL1850-
DIB-ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM/yrs91-100
…
```

Most Common Configuration Options

```
…
[output]
# directory where analysis should be written
baseDirectory = /project/projectdirs/m3412/sprice/analysis/output/20200610.A_WCYCL1850-
DIB-ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM/yrs91-100

# provide an absolute path to put HTML in an alternative location (e.g. a web
# portal)
htmlSubdirectory = /project/projectdirs/m3412/www/xylar/20200610.A_WCYCL1850-DIB-
ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM/yrs91-100
…
```

Most Common Configuration Options

```
…
[output]
# directory where analysis should be written
baseDirectory = /project/projectdirs/m3412/sprice/analysis/output/20200610.A_WCYCL1850-
DIB-ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM/yrs91-100

# provide an absolute path to put HTML in an alternative location (e.g. a web
# portal)
htmlSubdirectory = /project/projectdirs/m3412/www/xylar/20200610.A_WCYCL1850-DIB-
ISMF_CMIP6.ne30_ECwISC30to60E1r2.cori-knl.maint1p2-3DGM/yrs91-100

# tags and analysis tasks to generate or explicitly skip
generate = ['all', 'no_BGC', 'no_icebergs', 'no_index', 'no_eke', 'no_min', 'no_max']
…
```

Most Common Configuration Options

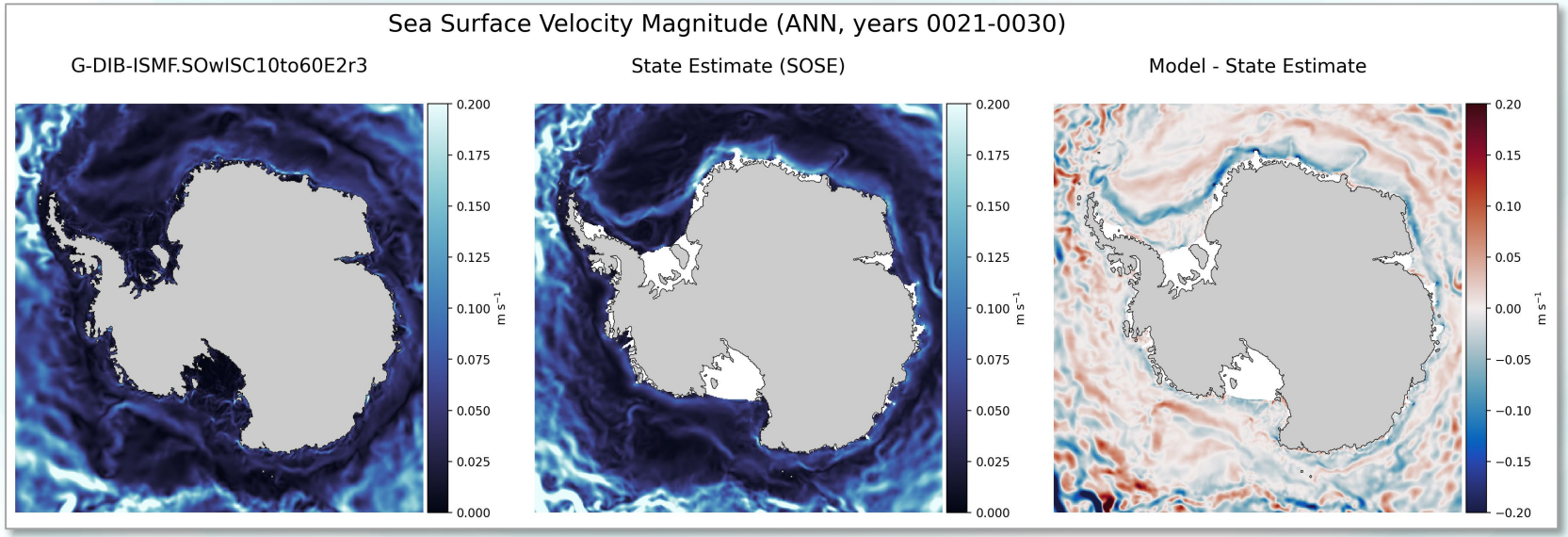
```
…
[climatology]
# the first and lat year over which to average climatologies
startYear = 91
endYear = 100

[timeSeries]
# start and end years for timeseries analysis.
startYear = 1
endYear = 100

[index]
# start and end years for the nino 3.4 analysis
startYear = 1
endYear = 100
…
```

Outline

- Setting Up and Running MPAS-Analysis
- Example Results
- Most Common Config Options
- Under the Hood
- Future Plans

Computing Climatologies

- Climatologies computed as needed automatically
- Using ncclimo or xarray/dask in python
- No need to precompute (different from E3SM_Diags)



Observations, Mapping Files and Region Masks

- Diagnostics data is here: https://web.lcrc.anl.gov/public/e3sm/diagnostics

Observations, Mapping Files and Region Masks

- Diagnostics data is here:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics
- Most observations are publicly available via Anvil:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/observations/Ocean
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/observations/SeaIce/

Observations, Mapping Files and Region Masks

- Diagnostics data is here:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics
- Most observations are publicly available via Anvil:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/observations/Ocean
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/observations/SeaIce/
- Mapping files from obs and oEC60to30v3 mesh to comparison grids (Global 0.5x0.5, Arctic and Antarctic) are here:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/mpas_analysis/maps/

Observations, Mapping Files and Region Masks

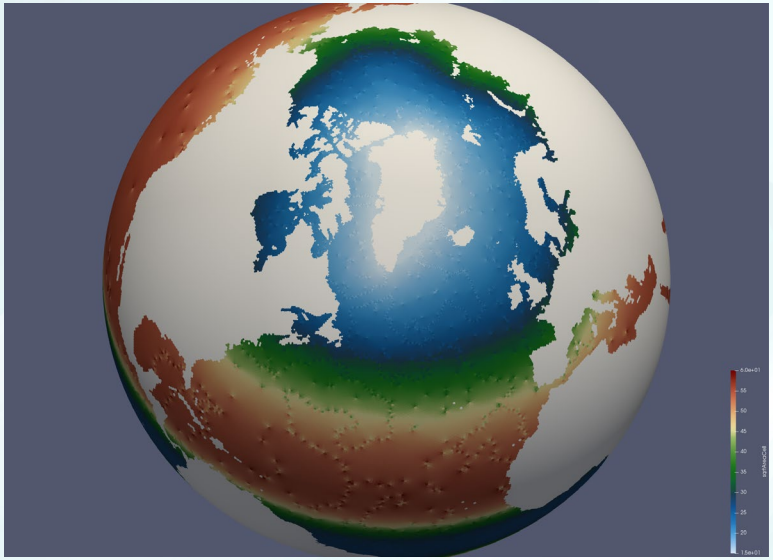
- Diagnostics data is here:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics
- Most observations are publicly available via Anvil:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/observations/Ocean
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/observations/SeaIce/
- Mapping files from obs and oEC60to30v3 mesh to comparison grids (Global 0.5x0.5, Arctic and Antarctic) are here:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/mpas_analysis/maps/
- Masks of MOC regions, global ocean basins and Antarctic basins are here:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/mpas_analysis/region_masks/

Observations, Mapping Files and Region Masks

- Diagnostics data is here:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics
- Most observations are publicly available via Anvil:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/observations/Ocean
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/observations/SeaIce/
- Mapping files from obs and oEC60to30v3 mesh to comparison grids (Global 0.5x0.5, Arctic and Antarctic) are here:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/mpas_analysis/maps/
- Masks of MOC regions, global ocean basins and Antarctic basins are here:
  https://web.lcrc.anl.gov/public/e3sm/diagnostics/mpas_analysis/region_masks/
- Some observations and meshes are not public, so data is on Anvil at:
  /lcrc/group/acme/diagnostics_private

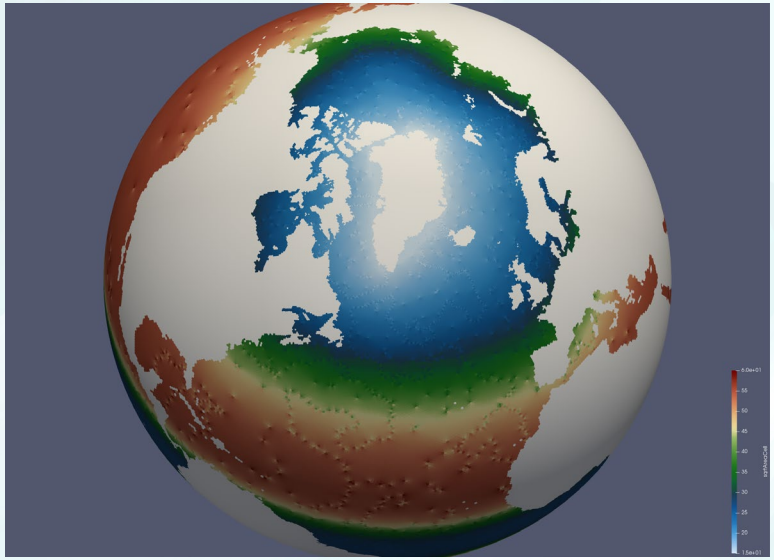
Supported Meshes

- We publicly support only:
  - oEC60to30v3



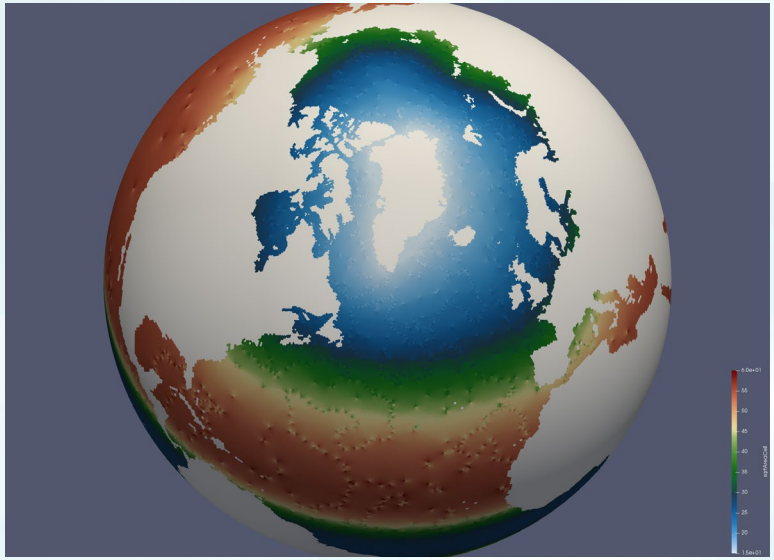
Supported Meshes

- We publicly support only:
  - oEC60to30v3
- We have mapping files and region masks on E3SM supported machines for:
  - oEC60to30v3wLI
  - oGNLD30to10 (incomplete)
  - oNAEC60to30cr8L60v1 (incomplete)
  - oQU240wLI
  - oRRS30to10v3wLI



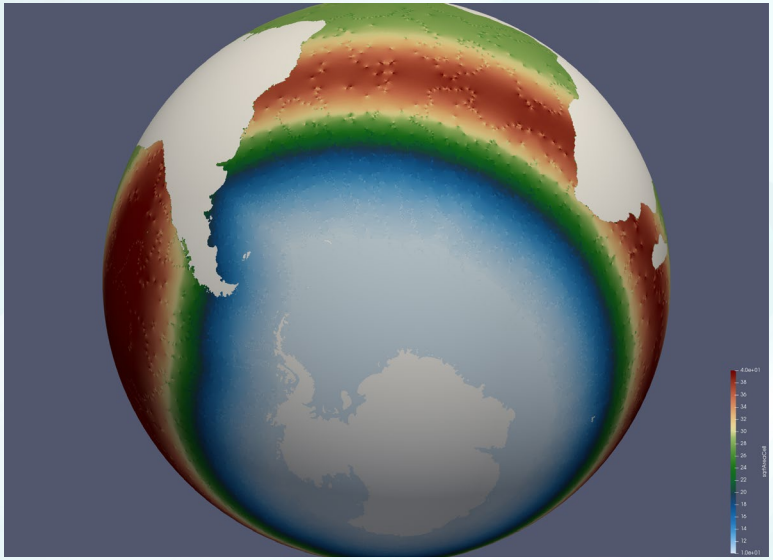
Supported Meshes

- We publicly support only:
  - oEC60to30v3
- We have mapping files and region masks on E3SM supported machines for:
  - oEC60to30v3wLI
  - oGNLD30to10 (incomplete)
  - oNAEC60to30cr8L60v1 (incomplete)
  - oQU240wLI
  - oRRS30to10v3wLI
- We will add v2 MPAS-Ocean/Seaice meshes as they are finalized



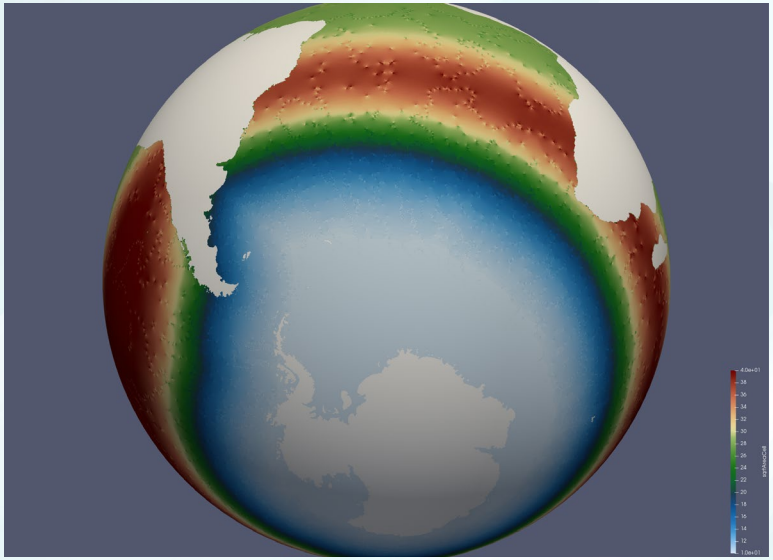
Support for New Meshes

- MPAS-Analysis can handle new meshes on the fly!
- …but generating mapping files and region masks can be **slow** and **memory intensive**



Support for New Meshes

- MPAS-Analysis can handle new meshes on the fly!
- …but generating mapping files and region masks can be **slow** and **memory intensive**
- We are happy to help with generating and caching them
- Tutorial coming soon

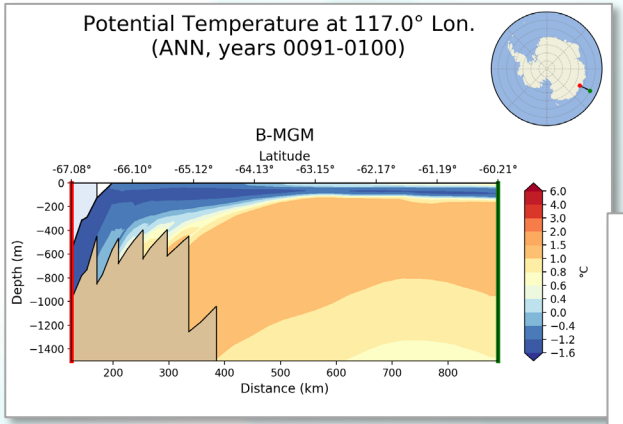


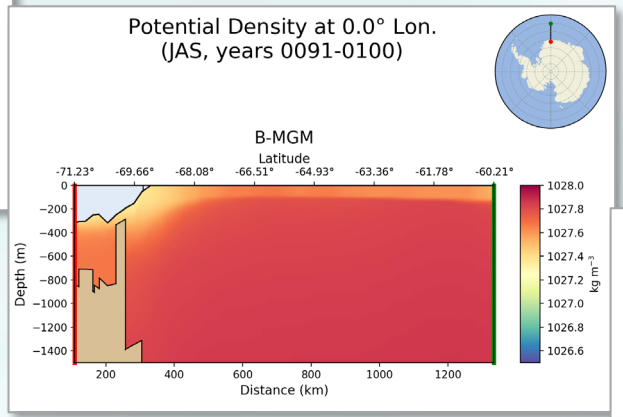
Outline

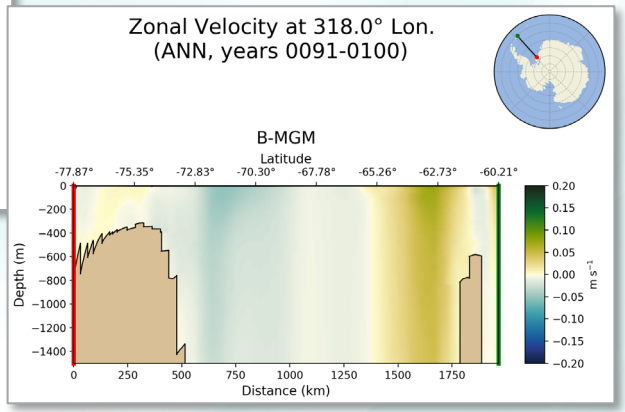
- Setting Up and Running MPAS-Analysis
- Example Results
- Most Common Config Options
- Under the Hood
- Future Plans

Future Plans: Transects on Native MPAS Meshes

- Complete and in final testing:

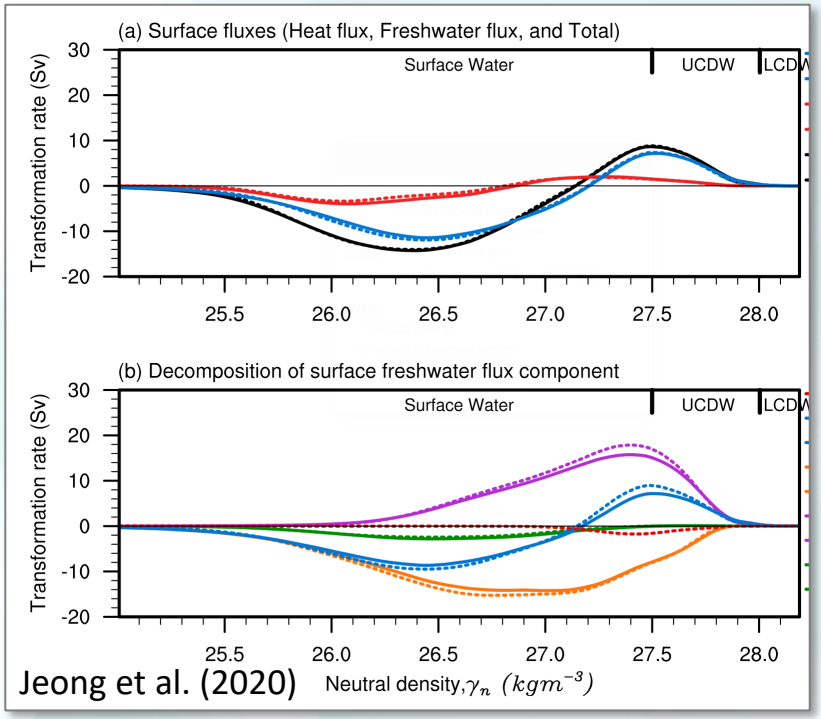




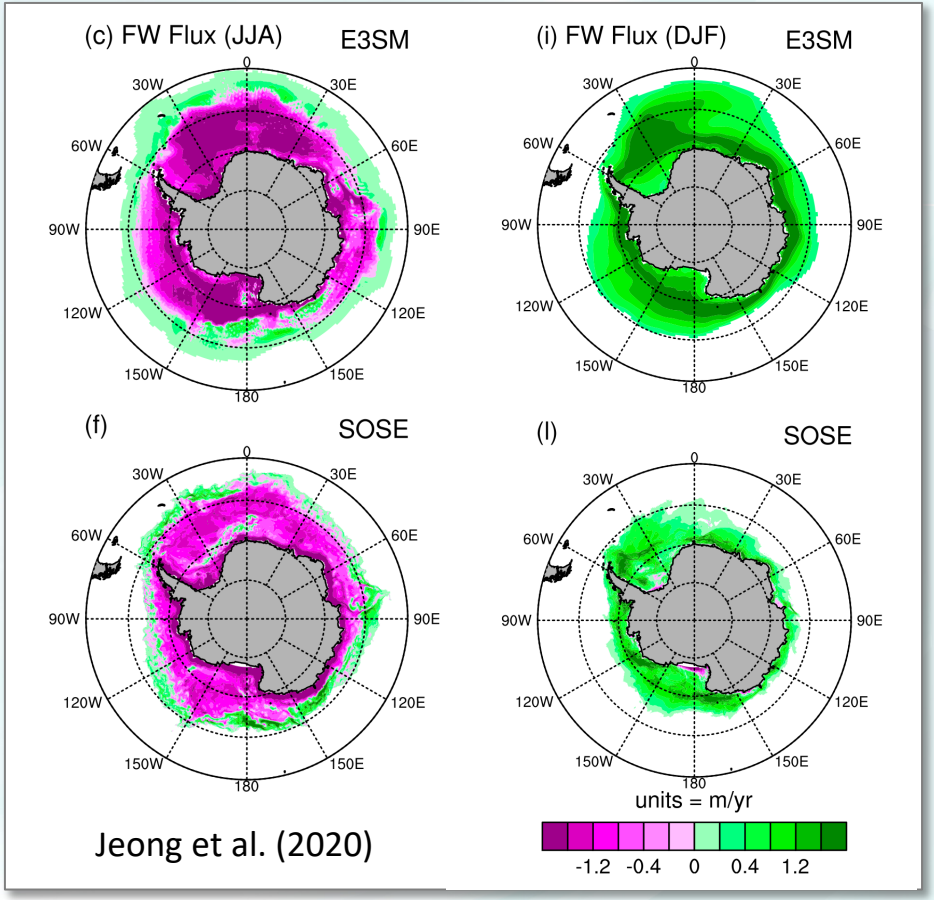


Future Plans: New Analysis

- Maps of sea ice formation and melting
- Water-mass transformation



Jeong et al. (2020)



Jeong et al. (2020)

Future Plans: Node Parallelism

- Currently, MPAS-Analysis only runs on one node
- Memory limitations (particularly at high res)
- Analysis takes several hours (particularly for long runs)
- Possible solution: parsl