The Arctic Subpolar gyre sTate Estimate
ASTE Release 1 User Guide

An T. Nguyen, Patrick Heimbach, Timothy A. Smith, Helen Pillar

atnguyen@oden.utexas.edu

Oden Institute for Computational Engineering and Sciences
Jackson School of Geosciences
Institute for Geophysics
The University of Texas at Austin, Austin, TX

March 04, 2021

1 Introduction[1]

This note describes the directory structure and content of the Arctic Subpolar sTate Estimate, ASTE Release 1 (ASTE_R1) data portal https://arcticdata.io/catalog/portals/ASTE/ (and mirrored at https://web.corral.tacc.utexas.edu/OceanProjects/ASTE/Release1/). Covering the period 01/Jan/2002 through 31/Dec/2017, ASTE_R1 synthesizes most of available satellite and in situ data into a coherent, dynamically and kinematically self-consistent framework provided by a general circulation model (the MITgcm). The coupled ocean-sea ice estimate so produced is free of artificial internal sources or sinks of mass, momentum, heat, or salt, thus enabling the calculation of accurate, time-evolving property budgets. The data used to constrain the model include satellite altimetry-derived sea surface height (SSH), GRACE ocean bottom pressure anomalies (OBP), AMSR-E and WinSat sea surface temperature (SST), in-situ hydrographic profiles from Argo, CTD, XBT, ITP, APB, Glider, moorings at important Arctic gateways, and sea-ice concentration measurements. The estimate uses the adjoint method to iteratively minimize the squared sum of weighted model-data misfits and control adjustments. A more detailed description of the estimate can be found in Nguyen et al. (2021).

This document is a modified version of the user guide of the Estimating the Circulation and Climate of the Ocean, ECCO Version 4, Release 4 (v4r4; Fukumori et al., 2019). We discuss the model configuration in Section 2 and describe the data portal in Section 3. Monthly and daily averaged model fields, as well as monthly snapshots, are discussed in Section 4. Finally, we describe how to calculate closed budgets and available analysis software in Sections 5 and 6.

 NSF sponsorship acknowledged.

2 Model

The model version that is used to produce ASTE_R1 is checkpoint65q of the MIT general circulation model (MITgcm). Appendix A1 in Nguyen et al. (2021) describes how to download the code, the data, and any needed auxiliary files to reproduce ASTE_R1 through re-running the forward model with corresponding input fields.

The grid used in ASTE_R1 is based on the global Lat-Lon-Cap 270 grid (Fig. 1a), akin to a cubed-sphere topology. It has five faces, with a simple latitude-longitude grid between 70°S and 57°N and an Arctic cap (Forget et al., 2015). ASTE covers the Atlantic Ocean (faces 1 and 5), the entire Arctic Ocean and surrounding marginal seas (faces 1, 4, 5), and the Bering Sea (face 4). The dimensions for the regional domain in the five faces are [270x450], [0x0], [270x270], [180x270], and [450x270], and are broken down further into tiles of dimensions either 270x270 or 90x90 (Figs. 1 & 2). The horizontal resolution varies spatially from 13 km in the high Arctic to 30 km in the tropical North Atlantic. The deepest ocean bottom is set to 6000 m below the surface, with the vertical grid spacing increasing from 10 m near the surface to 457 m near the ocean bottom.

3 Data Server

The data are hosted by the Arctic Data Center (arcticdata.io) at https://arcticdata.io/data/10.18739/A2CV4BS5K/ and are mirrored at https://web.corral.tacc.utexas.edu/OceanProjects/ASTE/Release1/ of the NSF-supported Texas Advanced Computing Center (TACC). Metadata can be viewed at https://doi.org/10.18739/A2CV4BS5K/. The Arctic Data Center portal offers an interface for users to browse and download data through their browser. It allows scripted data extraction via a command line interface through a range of web services, e.g. using wget to download the data.

A sample wget command to download ASTE_R1's monthly potential temperature fields on the native grid is as follows:

```
wget -r --no-parent
https://arcticdata.io/data/10.18739/A2CV4BS5K/nctiles_monthly/THETA/
```

4 Directory Structure

In this section, we describe the directory structure of ASTE_R1 in https://arcticdata.io/data/10.18739/A2CV4BS5K/ (and mirrored in https://web.corral.tacc.utexas.edu/OceanProjects/ASTE/Release1/). All links below will be listed as from arcticdata.io with the mirrored equivalents in parenthesis. The directory structure is similar to that of ECCOv4r4 (Fukumori et al., 2019).

4.1 Documentation

The directory **doc** (or **here**) contains useful documents that include

- a description of ASTE_R1's directory and file structures (ASTE_R1_user_guide.pdf, this document),

- a citation (or here) of ASTE_R1 (Nguyen et al., 2021),
- instructions on how to reproduce ASTE_R1 results (readme_rerun_ASTE_R1.txt, (or here) see also Appendix A1 of Nguyen et al., 2021),
- a note on analyzing budgets (evaluating_budgets_in_eccov4r3.pdf, Piecuch, 2017) with additional addendum as provided in the ASTE_R1_user_guide.pdf (this document).
- a summary file of all cost functions in costfunction0062 (or here),
- a "standard output file" STDOUT.0000 (or here) that the model creates during its integration with information about the model configuration and useful measures of the model state.
- a available_diagnostics.log (or here) file providing the model's variables info and unit.

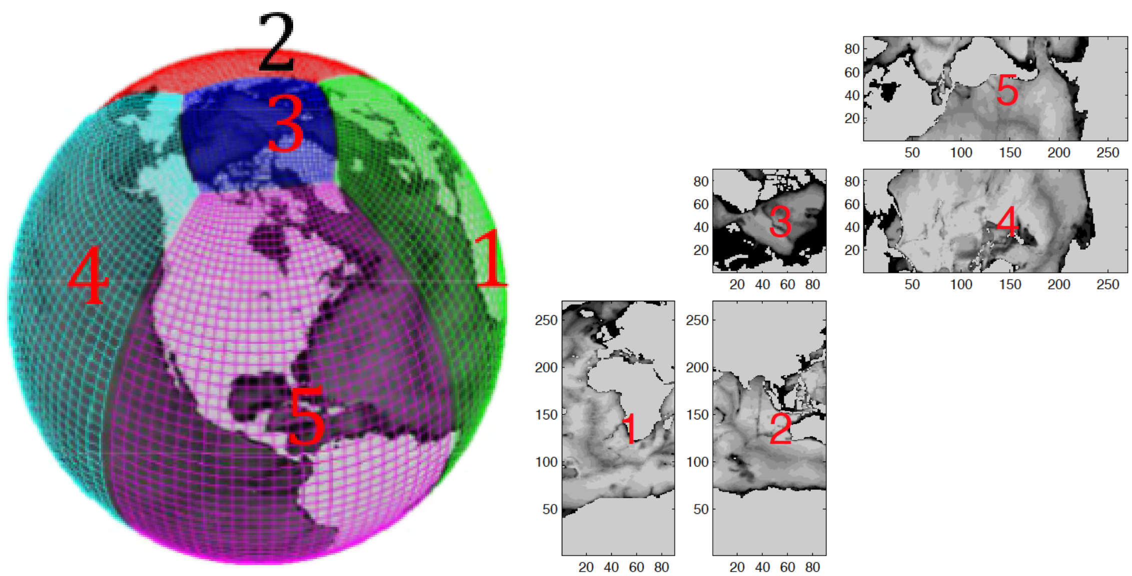

Figure 1. a) LLC270's lat-lon-cap (llc) grid with five faces; b) The dimensions of the five faces of the global grid. The figures are modified from Forget et al. (2015). For ASTE, the domain is limited as shown in Figure 2.

4.2 Model Grid

The model grid information can be found in the subdirectory nctiles_grid (or here). The ASTE domain is split into 29 regional tiles (Fig. 2), with each variable saved in 29 separate files in NetCDF format (GRID.0001.nc to GRID.0029.nc). These NetCDF files can be read by using various NetCDF tools from different programming languages and platforms, such as Python, MATLAB, FORTRAN. See more details in Section **Software**.

4.3 Introduction to Fields

ASTE_R1 provides the model state and forcing adjustments on the model native grid. Complete monthly fields are provided, including time averaged (Section 4.4) and instantaneous fields (Section 4.5) that allow one to close property budgets on a monthly basis at each grid cell. For depth-integrated budgets, two additional sets of monthly fields are provided: the depth-integrated time-averaged and instantaneous fields (Section 4.6). In addition, a set of daily 2D and 3D fields (Section 4.7) and twelve-month climatological means (Section 4.8) are also provided to make it possible to study higher frequency processes, e.g. mixed-layer dynamics, and look at climatological properties . The model time stamps are included in each NetCDF file header. For a quick reference, they are also listed in the time_list.txt ascii file within each of the directories described in Section 4.4-4.8.

4.4 Monthly Average Fields

The primary product consists of monthly-averaged model fields (nctiles_monthly, or here). Each subdirectory inside nctiles_monthly contains NetCDF files for each model state variable, as indicated by the name of the subdirectory. The files of each variable are organized by tiles. Each file contains all 192 records of monthly output, spanning 2002-2017. Thus, each file contains a single tile ranging from 0001-0029 (e.g., inside ETAN/ (or here) there are ETAN.0001.nc to ETAN.0029.nc, see Fig 2). Some of the most commonly used fields, such as velocity components, potential temperature, salinity, SSH, and OBP anomaly are UVELMASS, VVELMASS, THETA, SALT, ETAN, and PHIBOT. There are two 3D fields (WSLTMASS, WTHMASS) that, for the purpose of closing mass, heat, and salt budgets given the parameter choices in ASTE_R1, are only needed for the surface level and therefore provided as 2D fields.

Monthly averaged fields can be accessed with MATLAB or python tools, see Section **Software**.

4.4.1 Corrected Sea Level

ASTE_R1 variable ETAN (or here) is the height of the model's liquid ocean surface in unit meter. In the absence of sea ice, this is equivalent to sea surface height (SSH). In the presence of sea ice, a correction for the ice and snow loading is done as follows to recover the equivalent SSH:

$$ssh(i,j) = (ETAN(i,j) + rho_{ice}/rho_0 \cdot SIheff(i,j) + rho_{snow}/rho_0 \cdot SIhsnow(i,j)) \cdot hFacC(i,j,1)$$

where rho_{ice} = 910 kg/m^3, rho_{snow}=330 kg/m^3, and rho_0=1029 kg/m^3.

4.4.2 Native and Geographical Velocity Components

Users are advised to be aware of the directional convention of the vector fields of the model's LLC grid (Figure 1b). Within each face (tile), the x- and y-directions point left-to-right and bottom-to-top in the figure, respectively. As such, in faces 4 and 5, the x- and y-directions point to the south and to the east, respectively. In face 3, the x-direction points to the Pacific Ocean, away from the Atlantic, whereas y-direction points to North America, away from Asia. For users who wish to convert the model's velocity to true eastward and northward velocity components, we provide here the standard formula to

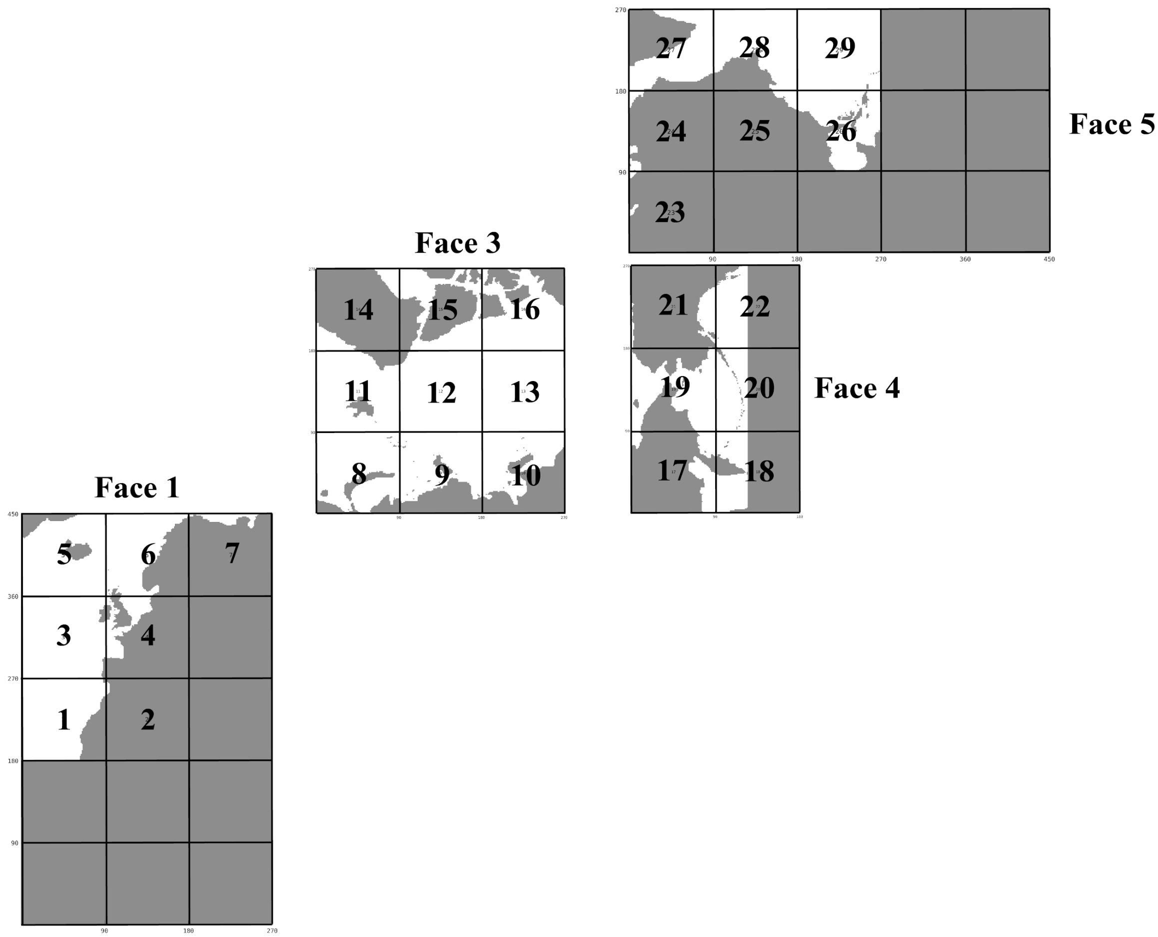

Figure 2 The partitioning of the ASTE domain into 29 regional tiles.

compute the geographical velocity components using the model's velocity (Table 1) and model projection geometry AngleCS and AngleSN provided in the GRID file:

```
u_center(i,j,k,itime) = (UVELMASS(i,j,k,itime)+UVELMASS(i+1,j,k,itime)) / 2
v_center(i,j,k,itime) = (VVELMASS(i,j,k,itime)+VVELMASS(i,j+1,k,itime)) / 2
u_East(i,j,k,itime) = u_center(i,j,k,itime)*AngleCS(i,j) - v_center(i,j,k,itime)*AngleSN(i,j)
v_North(i,j,k,itime) = u_center(i,j,k,itime)*AngleSN(i,j) + v_center(i,j,k,itime)*AngleCS(i,j)
```

Table 1 Native velocity components.

Filename	Description
UVELMASS (or here)	X-component of velocity, mass weighted (m/s).
VVELMASS (or here)	Y-component of velocity, mass weighted (m/s).

4.4.3 Advective and Diffusive Fluxes

The files with their names starting with "ADV" and "DF" indicate advective and diffusive fluxes, respectively. Similar to velocity, the horizontal components of the native fluxes also follow the model's directional convention. For instance, DFxE_TH means diffusive flux ("DF"), in the model's x-direction ("x"), evaluated explicitly ("E") for potential temperature ("TH"). Table 3 lists all the flux terms for potential temperature. See Piecuch (2017) for how to make use of the flux terms along with forcing terms to close budgets.

Table 2 Advective and diffusive flux terms for potential temperature.

Filename	Description
ADVx_TH	X-component ("x") of ADVective flux of potential temperature ("TH") (°C m^3/s) at a particular grid (i,j,k). Positive to increase temperature at (i,j,k).
ADVy_TH	Y-component ("y") of ADVective flux of potential temperature (°C m^3/s).
ADVr_TH	Z-component ("r") of ADVective flux of potential temperature (°C m^3/s).
DFxE_TH	X-component of DiFfusive flux of potential temperature (°C m^3/s). Explicit part ("E").
DFyE_TH	Y-component of DiFfusive flux of potential temperature (°C m^3/s). Explicit part.
DFrE_TH	Z-component of DiFfusive flux of potential temperature (°C m^3/s). Explicit part.
DFrI_TH	Z-component of DiFfusive flux of potential temperature (°C m^3/s). Implicit part ("I").

4.4.4 Non-local Diffusive Fluxes

When KPP is used, two KPP non-local terms, KPPg_TH for heat and KPPg_SLT for salt, need to be added to the vertical diffusive fluxes term in the corresponding budgets. Table 3 lists the information for two flux terms.

Table 3 KPP non-local terms.

Filename	Description
KPPg_TH (or here)	KPP non-local flux of potential temperature (°C m^3/s).
KPPg_SLT (or here)	KPP non-local flux of salinity (psu m^3/s).

4.5 Instantaneous Monthly Model Fields

Besides monthly averages, ASTE_R1 also provides monthly snapshots in the subdirectory nctiles_monthly_snapshots (or here) for THETADR (or here, THETA * drF, ºC m), SALTDR (or here, SALT * drF, psu m), and ETAN (or here). The main purpose of these snapshots is to support accurate budget calculations (see Section **Budget Calculation**); specifically, monthly mean fluxes that are provided equal changes between these snapshots (as opposed to changes between monthly average states of Section 4.4).

4.6 Depth-Integrated Monthly Average and Instantaneous Model Fields

To close the depth-integrated budgets, the model 2D depth-integrated monthly average nctiles_monthly_integrated (or here) and instantaneous nctiles_monthly_snapshots_integrated (or here) fields are also provided. See Section **Budget Calculation** and supported matlab code for closing these budgets.

4.7 Daily Average Fields

A subset of 2D and 3D daily averages are also available on the ASTE portal for select variables in directory nctiles_daily (or here, see Table 4).

Table 4 Daily averages on ASTE portal

Directory name	Description
ETAN (or **here**)	Model sea surface height anomaly (m).
PHIBOT (or **here**)	model ocean bottom pressure potential anomaly (P/rho), (m^2/s^2) .
SIarea (or **here**)	Fractional sea-ice covered area (m^2/m^2)
SIheff (or **here**)	Effective sea-ice thickness (m) that is defined as actual sea-ice thickness scaled by fractional sea-ice area (SIarea, or here).
SIhsnow (or **here**)	Effective snow thickness (m).
sIceLoad (or **here**)	Sea-ice and snow loading defined as mass of sea-ice & snow over area (kg/m^2).
oceTAUX (or **here**)	X-component of surface wind stress (N/m^2).
oceTAUY (or **here**)	Y-component of surface wind stress (N/m^2).
oceQnet (or **here**)	Net surface heat flux into ocean (W/m^2)
oceFWflx (or **here**)	Net surface FW flux into ocean ($kg/m^2/s$)
SIuice (or **here**)	X-component of sea ice velocity (m/s)
SIvice (or **here**)	Y-component of sea ice velocity (m/s)
MXLDEPTH (or **here**)	Mixed layer depth (m)
THETA (or **here**)	Ocean potential temperature (°C).
SALT (or **here**)	Ocean salinity (psu).
UVELMASS (or **here**)	X-component of velocity, mass weighted (m/s)
VVELMASS (or **here**)	Y-component of velocity, mass weighted (m/s)

4.8 Model Monthly climatology

The model 12-month climatology fields nctiles_climatology (or here), with each model field containing 12 records, each of the monthly means Jan, Feb, Mar, ... , Dec. The year range used to create the climatology is 2002-2017. Each subdirectory inside nctiles_climatology contains NetCDF files for each model state variable, as indicated by the name of the subdirectory. Note that, as described in Section 4.4, the two fields WSLTMASS (or here) and WTHMASS (or here) are stored only for the surface layer and accessible as 2D fields.

4.9 Model Equivalent of In-situ Data

The model equivalent of vertical profiles of in situ data in NetCDF format are in directory profiles (or here). The model fields are sampled on the fly during model integration at the time and location of the in situ data to generate the model equivalent. Hydrographic observations are stored in variables prof_T and prof_S, and the model's equivalents are in prof_Testim, prof_Sestim. The corresponding weights (squares of 1/uncertainty) can be found in prof_Tweight and prof_Sweight. See Section 6 for example matlab scripts to read and calculate model-obs misfits.

4.10 Atmospheric Forcing

The atmospheric forcing used in ASTE_R1 is JRA55 (Kobayashi et al., 2015). It is hosted at the JPL portal, which requires the users to create an account with NASA Earthdata. Adjustments to this forcing is included in the input files described in Section 4.11.

4.11 Input Files

The subdirectory rerun_ASTE_R1 (or here) includes all files in raw binary formats (except for the atmospheric forcing, see 4.10) that are needed to reproduce ASTE_R1 (Table 5).

Table 5 Input files are needed to reproduce ASTE_R1.

Directory or file name	Description
readme_rerun_ASTE_R1.txt (or here)	A readme file on how to download the code and link input files.
NAMELIST_ASTE_R1 (or here)	Namelist such as file "data", "data.ctrl", etc.
input_weight (or here)	Control weight.
bathy_fill9iU42Ef_noStLA_v1.bin (or here)	Bathymetry (m).
pickup*0000000007* files	Initial condition for ASTE_R1.
xx* files	Control adjustments, including for atmospheric forcing, on ASTE grid.
logdiffkr*, kap*	Mixing coefficients.
tile* files	Grid files needed to run the model.
smooth* files	Smoothing operator related files.
visc_v1up9_cap0jy460GoM600_Pac0.bin (or here)	Harmonic coefficients (m^2/s).
fekete_runoff_SMOOTH_llc270_v3.bin (or here)	Climatology river runoff (m/s). Positive to increase sea level.

5 Budget Calculation

Monthly mean advective and diffusive fluxes are provided in directory nctiles_monthly (or here). Piecuch (2017) describes how to perform an accurate volume, heat, and salt budget analysis using these fields, both in pseudo code and in MATLAB (see **Software** below). Although the note was written for ECCOv4r3, the calculation is similar for ASTE_R1, but with several amendments to take into account additional vertical diffusive heat and salt terms associated with KPP non-local fluxes as described section 4.4.4, and the exclusion of bottom geothermal flux which was not used in ASTE_R1. While the fields available in nctiles_monthly are appropriate for closing budgets at individual grid box within the full 3D domain, 2D depth-integrated budgets can also be computed using the provided nctiles_monthly_integrated (or here). Sample matlab scripts provided in Section 6 for closing both the full 3D and depth-integrated 2D budgets.

6 Software

Analysis tools in MATLAB and python have been developed specifically to analyze the ASTE_R1 output.

Python users can access the output with the xmitgcm python package (Abernathey et al, 2020), which allows users to interact with the ASTE_R1 output without the need to download it directly (although this package can be used to download the data as well). Additionally, the ECCOv4-py python package can be used to make standard plots and calculations (e.g. net volume transport, MOC, etc) easy even with the LLC grid topology. This package can be accessed at https://github.com/ECCO-GROUP/ECCOv4-py, with links to documentation therein. An interactive demonstration of these python packages being used to plot and perform calculations with the ASTE_R1 output can be found at https://github.com/crios-ut/aste (Smith, 2020). Finally it is worth noting that while the NetCDF files described in section 4 can be read with various python libraries, these files are designed for MATLAB users (see below). The output attained with the xmitgcm capabilities will have 270x270 grid tiles, rather than the 90x90 size in the NetCDF files.

Matlab users can refer to the matlab_tools (or here) for relevant functions inside ASTE_mlibrary (or here) and several examples on how to:

a. get started with read/plot ASTE_R1 NetCDF fields (example_read_plot_ASTE, or here);
b. perform tracer budgets (example_budget_tracers, or here);
c. look at model-data misfits in in situ profiles (or here) data (example_profiles, or here);
d. assemble the tiles to form a Nordic Seas region (example_NordicSeas, or here).

Some of the functions used in the above-mentioned examples are modifications of Gaël Forget original MATLAB/Octave toolboxes gcmfaces and MITprof (Forget 2017), which were originally created to handle ECCOv4 output. The modifications (gcmfaces_mod and MITprof_toolbox, or here and here) by An Nguyen were needed to reduce the memory footprint of the higher resolution ASTE configuration.

In each of the examples folders a-d, there are scripts to guide users, as well as sample figures. Specifically for budget (b), each of the matlab scripts for mass (do_ASTEr1_budget_mass_ts.m, or here), heat (do_ASTEr1_budget_heat_ts.m, or here), and salt (do_ASTEr1_budget_salt_ts.m, or here) is self-contained, with dependent functions already included in the scripts (e.g., nctiles2aste_v2.m, or here), and is set for only 1 time-step to demonstrating that budgets are closed at individual grid box for mass, heat, and salt budgets.

Questions:

atnguyen@oden.utexas.edu

Acknowledgement

The research was carried out at the University of Texas at Austin with funding from NSF grants NSF-OPP-1603903, NSF-OPP-1708289, and NSF-OCE-1924546. Additional funding was provided from the NASA-supported ECCO project through a JPL/Caltech subcontract. Computing resources were provided by NASA Advanced Supercomputing Division at the Ames Research Center and the Texas Advanced Computing Center (TACC).

References:

Ryan Abernathey, Raphael Dussin, Timothy Smith, Pascal Bourgault, Stickler Bot, Ed Doddridge, …, Martin Losch. (2020, November 24). MITgcm/xmitgcm: v0.5.0 (Version v0.5.0). Zenodo. http://doi.org/10.5281/zenodo.4288974

Forget, G., 2017: gcmfaces, A generic treatment of gridded earth variables in Matlab. (Available at ftp://mit.ecco-group.org/ecco_for_las/version_4/release2/doc/ECCOV4R2_gcmfaces.pdf.)

Forget, G., J.-M. Campin, P. Heimbach, C. N. Hill, R. M. Ponte, and C. Wunsch, 2015: ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation. Geoscientific Model Development, 8, 3071-3104, https://doi.org/10.5194/gmd-8-3071-2015

Fukumori, I., O. Wang, I. Fenty, G. Forget, P. Heimbach, R. M. Ponte, 2019: ECCO Version 4 Release 4. Available at https://ecco.jpl.nasa.gov/drive/files/Version4/Release4/doc/v4r4_synopsis.pdf.

Griffies, S. M., & Greatbatch, R. J. (2012). Physical processes that impact the evolution of global mean sea level in ocean climate models. Ocean Modelling, 51(C), 37–72. http://doi.org/10.1016/j.ocemod.2012.04.003

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., Takahashi, K. (2015). The JRA-55 Reanalysis: General Specifications and Basic Characteristics.

Journal of the Meteorological Society of Japan, 93 (1), 5–48, https://doi.org/10.2151/jmsj.2015-001 .

Nguyen, A. T., H. Pillar, V. Ocana, A. Bigdeli, T. A. Smith, and P. Heimbach, 2021: The Arctic Subpolar gyre sTate Estimate (ASTE): Description and assessment of a data-constrained, dynamically consistent ocean-sea ice estimate for 2002-2017. *J. Adv. Model. Earth Syst.*, in revision. https://doi.org/10.1002/essoar.10504669.4

Piecuch, C. G., 2017: A note on evaluating budgets in ECCO Version 4 Release 3. http://hdl.handle.net/1721.1/111094

Timothy Smith. (2020, October 26). crios-ut/aste: ASTE Release 1 (Version v1.0.0). Zenodo. http://doi.org/10.5281/zenodo.4132216

Wang, O., 2019: Instructions for reproducing ECCO Version 4 Release 4. Available at https://ecco.jpl.nasa.gov/drive/files/Version4/Release4/doc/v4r4_reproduction_howto.pdf .

Wang, O., I. Fukumori, and I. Fenty, 2017, An overview of ECCO Version 4 Release 3's ftp site. Available at https://ecco.jpl.nasa.gov/drive/files/Version4/Release3/doc/v4r3_overview.pdf.