UFS Land-DA Workflow

- Day 3: JEDI-bundle & Pre/post-processing -

Chan-Hoo Jeon, Jong Kim, Gillian Petro, Eddie Snyder, Kris Booker

(Key stakeholder groups: NOAA/EMC, PSL, GSL, NESDIS ; NCAR ; JCSDA)

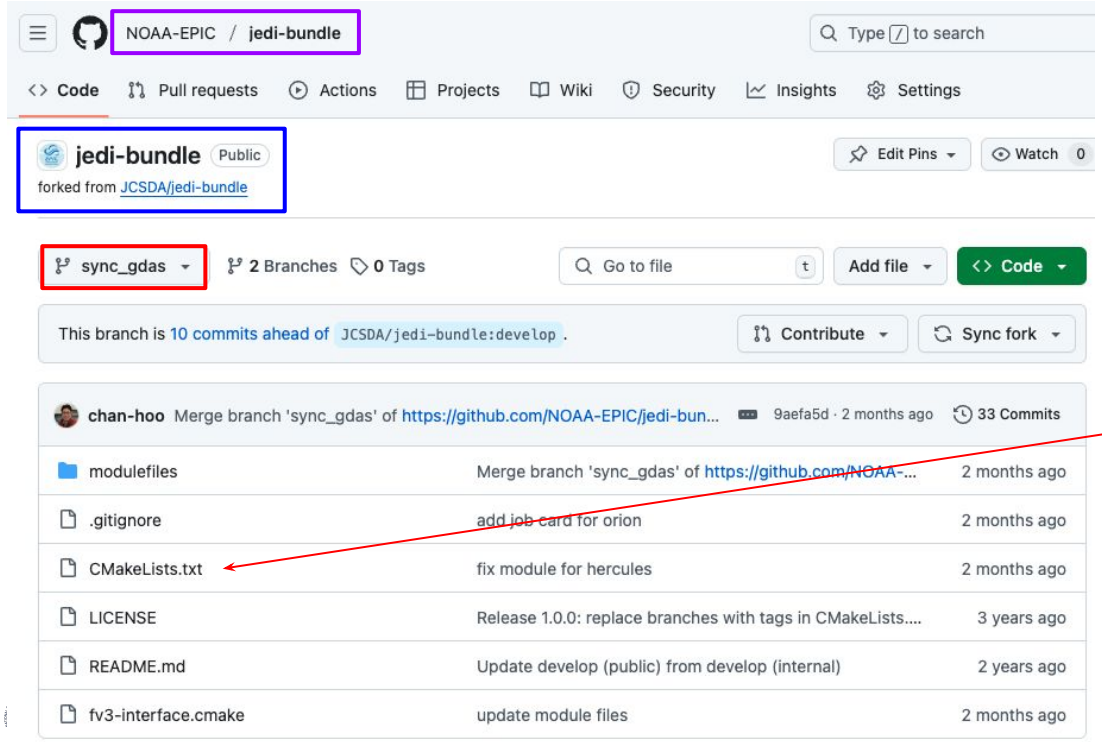
NOAA Earth Prediction Innovation Center (NOAA/EPIC)

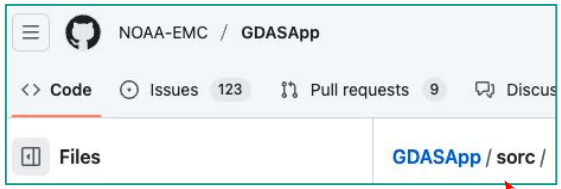
Outline

- JEDI-bundle synced with GDAS App
- Pre-processing: Snow observation data / ERA5 DATM forcing data
- Diagnostic plots in Analysis Task
- Built-in tile-to-tile converter
- Post-processing
- GitHub Repository: Issue / Pull Request (PR) / Projects
- Next steps (future plan)

JEDI-bundle Synced with GDAS App

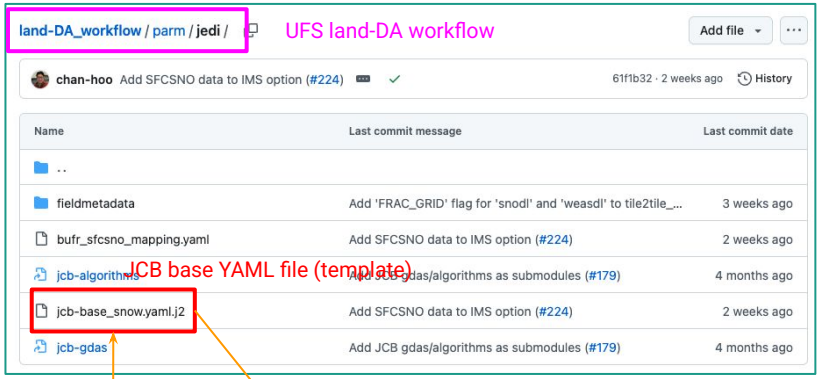
- The JEDI bundle synced with GDAS (NOAA/EMC) is provided in the ‘sync_gdas’ branch under the NOAA-EPIC fork of the JCSDA/jedi-bundle repository:

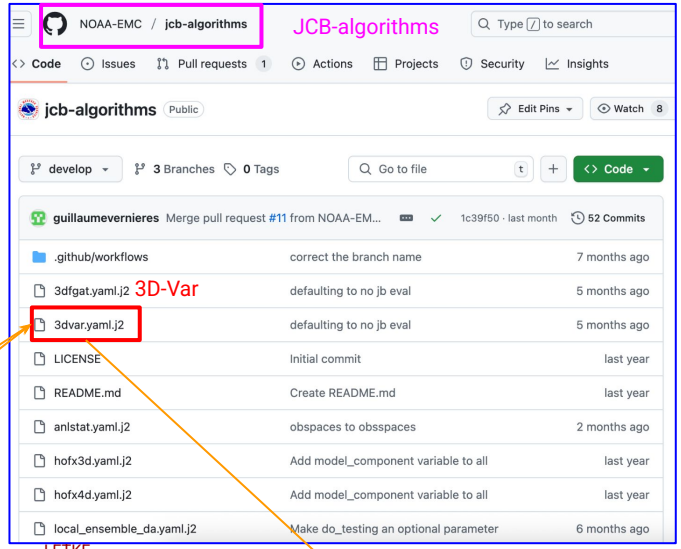


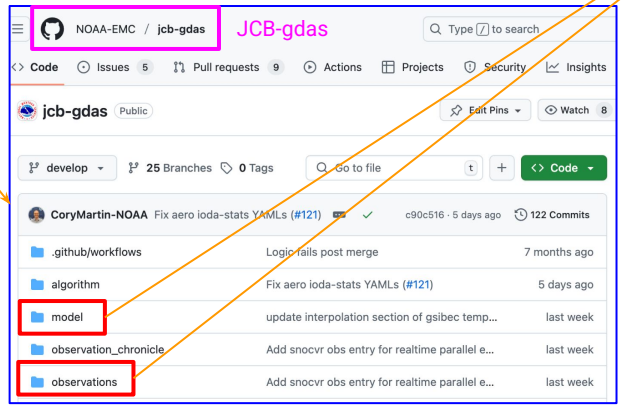


JEDI Configuration by JCB

- The JEDI input configuration YAML is created by JCB in the 'jcb' workflow task.







Task: jcb

land_analysis.xml

Rocoto XML file

setup_wflow_env.py

config.yaml

Workflow configuration file

jedi_3dvar_snow.yaml

JEDI input YAML file

If you want to tune the configuration in JCB-gdas, you can fork the 'NOAA-EPIC/jcb-gdas' repository and change it in a new branch.

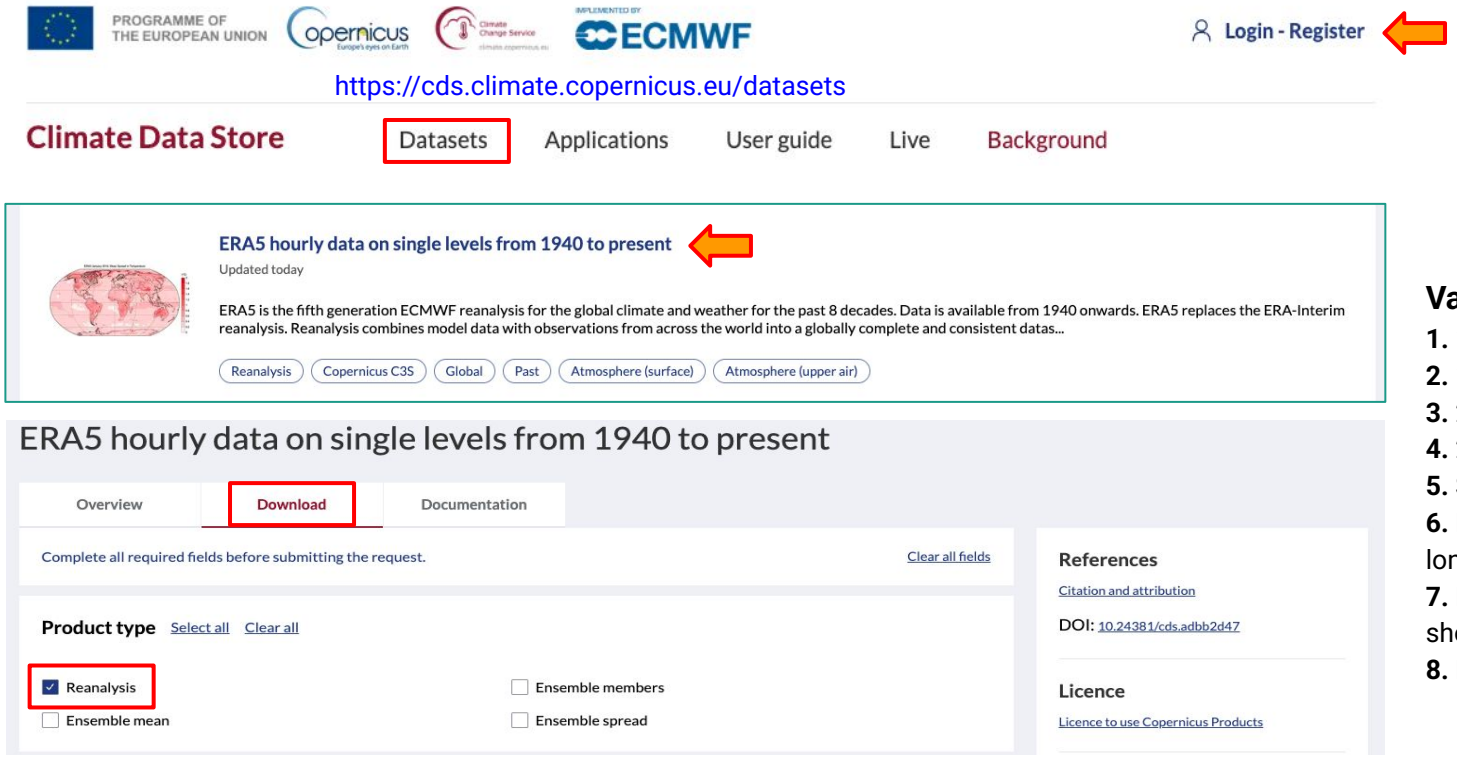
https://epic.noaa.gov/

Pre-processing: IODA Converters for Snow Observation Data

- IODA (Interface for Observation Data Access) converters: convert the snow observation raw data into the format of UFO and OOPS in JEDI.
  - UFO (**U**nified **F**orward **O**perator): Computes the simulated observations H(x) and performs the quality control (QC) and filtering of the observation data.
  - OOPS (**O**bject-**O**riented **P**rediction **S**ystem): provides data assimilation algorithms.
- IODA converter scripts used in the 'prep_data' task of the UFS land-DA workflow:
  - **GHCN** data: 'land-DA_workflow/ush/ghcn_snod2ioda.py'
    - Input: {year}.csv, ghcnd-stations.txt
    - Output: ghcn_snwd_ioda_{date}.nc (JEDI format)
  - **IMS** data: 'land-DA_workflow/ush/imsfv3_scf2ioda.py'
    - Input: ims_{date}.asc (ASCII file) obtained from the archive of GFS/GDAS.
    - Intermediate converter '**calfIMS.exe**': Convert the raw ASCII file to NetCDF file.
    - Output: obs_{date}.ims_snow.tm00.nc (JEDI format)

Pre-processing: ERA5 (DATM Forcing Data)

- The ERA5 forcing data can be downloaded from the Climate Data Store (CDS):



Variables:

1. 10m u-component of wind
2. 10m v-component of wind
3. 2m dewpoint temperature
4. 2m temperature
5. Surface pressure
6. Mean surface downward long-wave radiation flux
7. Mean surface downward short-wave radiation flux
8. Mean total precipitation rate

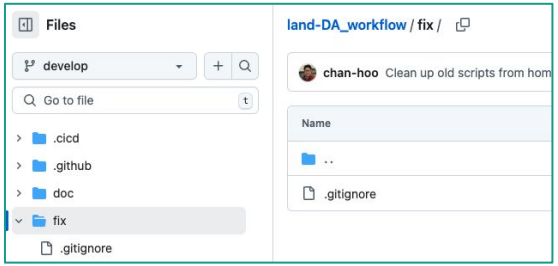
The downloaded NetCDF data files should be moved to the 'fix' directory

https://epic.noaa.gov/

Static Data (FIX) Directory

- The input static data (fix) files are soft-linked into the 'fix' directory by the build script.

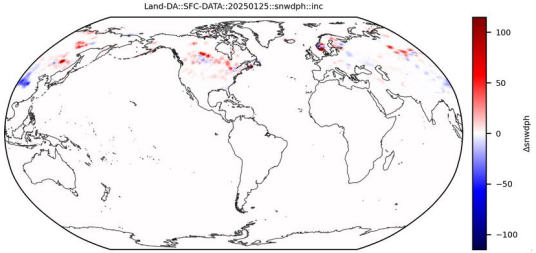
```
land-DA_workflow
└── fix
    ├── DA_obs                  ← Observation data files (pre-processed)
    │   ├── GHCN
    │   └── IMS
    ├── DATA_gdas               ← GDAS output data files
    ├── DATA_gfs                ← GFS output data files
    ├── DATA_ghcn               ← GHCN raw data files
    ├── DATA_RESTART            ← RESTART files for warm start
    ├── DATM_input_data         ← DATM forcing data files
    │   ├── era5
    │   │   └── raw_data        ← ERA5 raw data files (downloaded from CDS)
    │   └── gswp3
    ├── FV3_fix_global          ← Global static data for FV3ATM (UFS weather model)
    ├── FV3_fix_tiled           ← Tiled (6) static data for FV3ATM (UFS weather model)
    │   └── C{RES}
    ├── NaturalEarth            ← Natural Earth data for python cartopy background
    ├── NOAHMP_IC
    └── test_base               ← Base files for WE2E tests
```



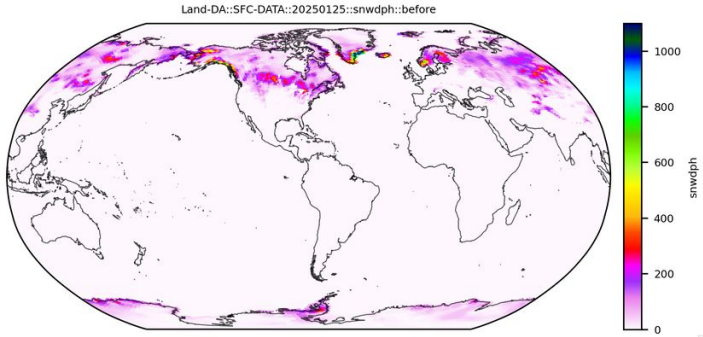
Initially empty in GitHub repository

https://epic.noaa.gov/

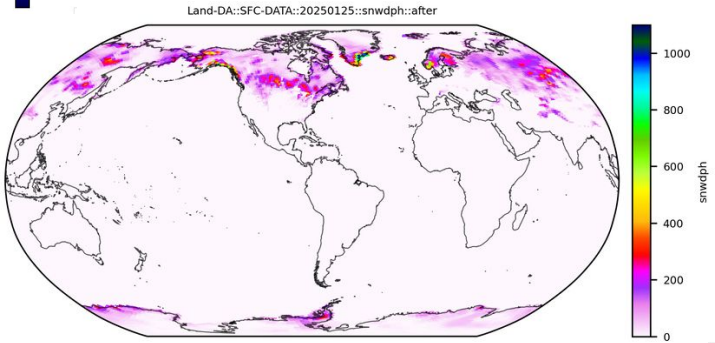
Diagnostic Plots in Analysis Task



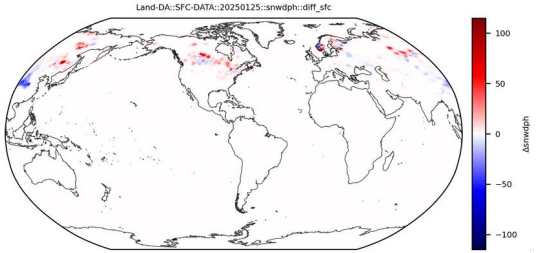
JEDI increment file



Before ‘analysis’ task



After ‘analysis’ task



Real increment by ‘**apply_incr**.exe’

https://epic.noaa.gov/

Tile-to-tile Converter

- Acts like a mediator between JEDI and the land model (Noah-MP) of the ufs weather model.
- The built-in tool 'tile2tile_converter' runs in two workflow tasks:
  - In 'pre_anal': Creates the 'sfc_data' files from the restart files for the 'analysis' task.
  - In 'post_anal': Creates the restart files for warm-start from the 'sfc_data' and restart files for the 'forecast' task.
- Two key variable names do not match between JEDI (sfc_data files) and the land model (restart files):

Variable name in 'tile2tile_converter'	Description	Noah-MP (restart)	JEDI (sfc_data)
swe	Snow water equivalent	weasd	sheleg / weasdl
snow_depth	Snow depth over land	snwdph	snwdph / snodl

- Source code: land-DA_workflow/sorc/tile2tile_converter.fd

Post-processing: ‘plot_stats’

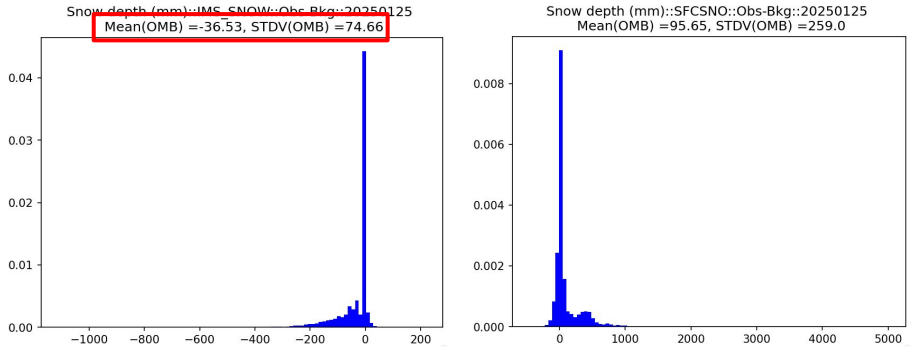
- Result file of JEDI: ‘diag.{obs_type}_{cycle_date}.nc’

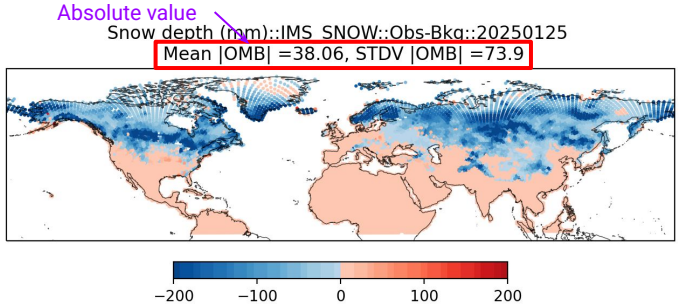
Group name	Description	LETKF	3D-Var
EffectiveError0	UFO’s computed effective ObsError value	○	○
EffectiveQC0	UFO’s final QC value given by ‘QCflags.h’	○	○
MetaData	Ancillary data such as stationElevation, longitude, and latitude	○	○
ObsBias0	Bias correction of observation data	○	○
ObsError	Observation errors from upstream data sources	○	○
ObsValue	Directly measured observation values	○	○
hofx0	End product of the forward operator; H(x)		○
hofx0_1	H(x) of Ensemble member 1	○	
hofx0_2	H(x) of Ensemble member 2	○	
hofx_y_mean_xb0	Mean of H(x)	○	
ombg	**Observation - H(Background)**	○	○

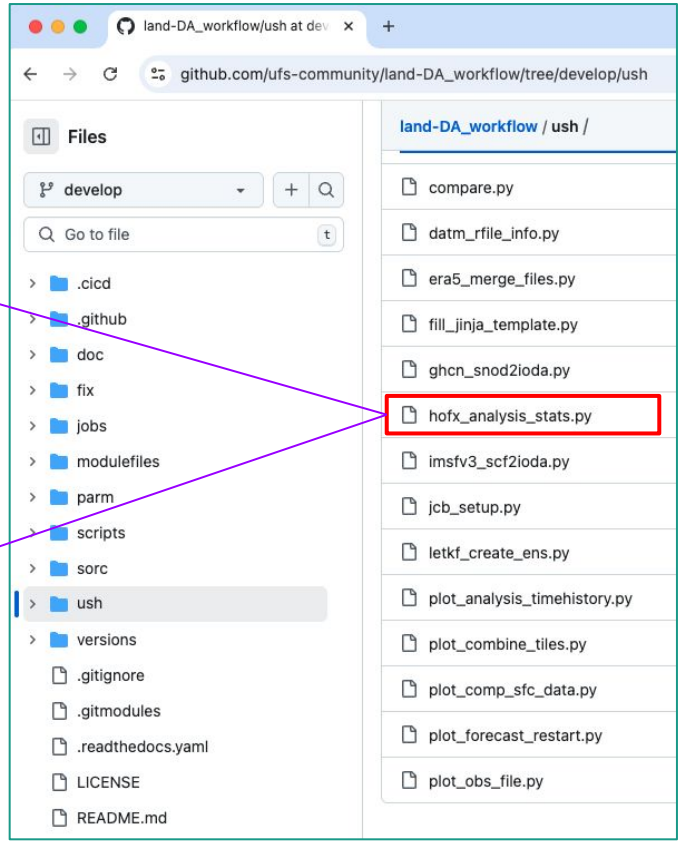
Post-processing: Histogram and Scatter Plots

- Task ‘plot_stats’: plots of ‘**ombg**’ from JEDI output (histogram / scatter plot)

CADRE 1: IMS + SFCSNO







Post-processing: Time-history Plot

- Task ‘plot_stats’: Time-history plots of ombg over cycles

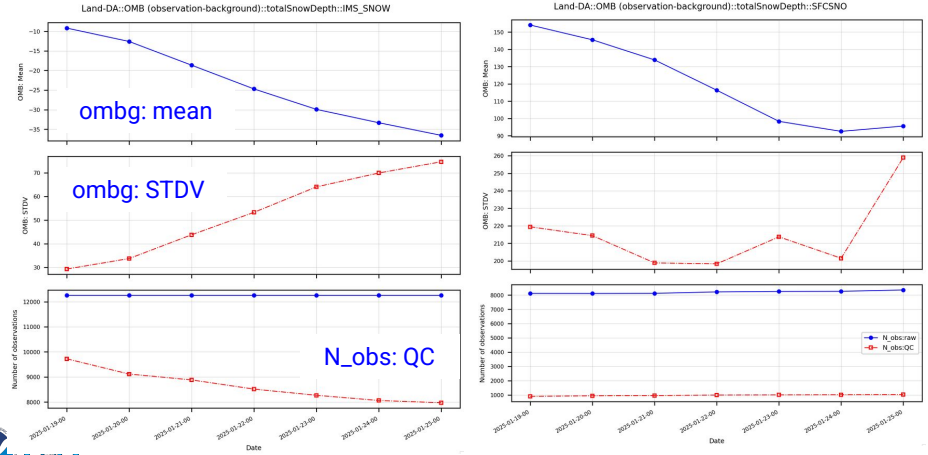
com_dir/landda.{date}/hofx/{time_history_data}.txt

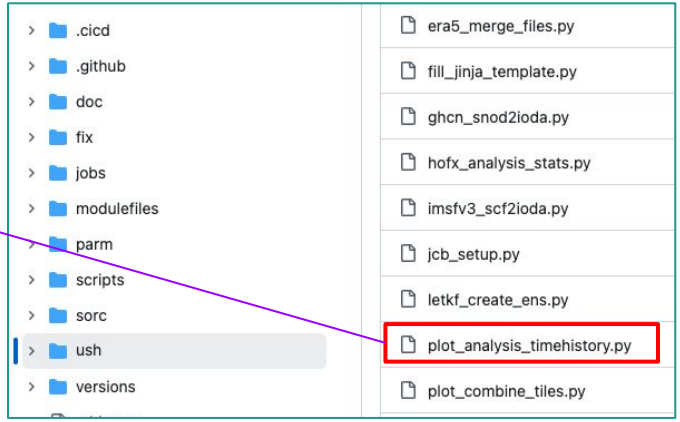
```
epic 1084787 May  1 20:57 diag.ghcn_snow_2025012500.nc
epic     277 May  1 21:04 hofx_omb_timehis_abs_ghcn_snow.txt
epic     301 May  1 21:04 hofx_omb_timehis_ghcn_snow.txt
```

log_dir/analysis_{date}.log

```
QC ims_snow snowCoverFraction: 12254 rejected as processed but not assimilated.
QC ims_snow snowCoverFraction: 0 passed out of 12254 observations.
QC ims_snow totalSnowDepth: 3 missing values.
QC ims_snow totalSnowDepth: 3272 out of bounds.
QC ims_snow totalSnowDepth: 319 black-listed.
QC ims_snow totalSnowDepth: 43 rejected by first-guess check.
QC ims_snow totalSnowDepth: 8617 passed out of 12254 observations.
```

CADRE 1: IMS + SFCSNO

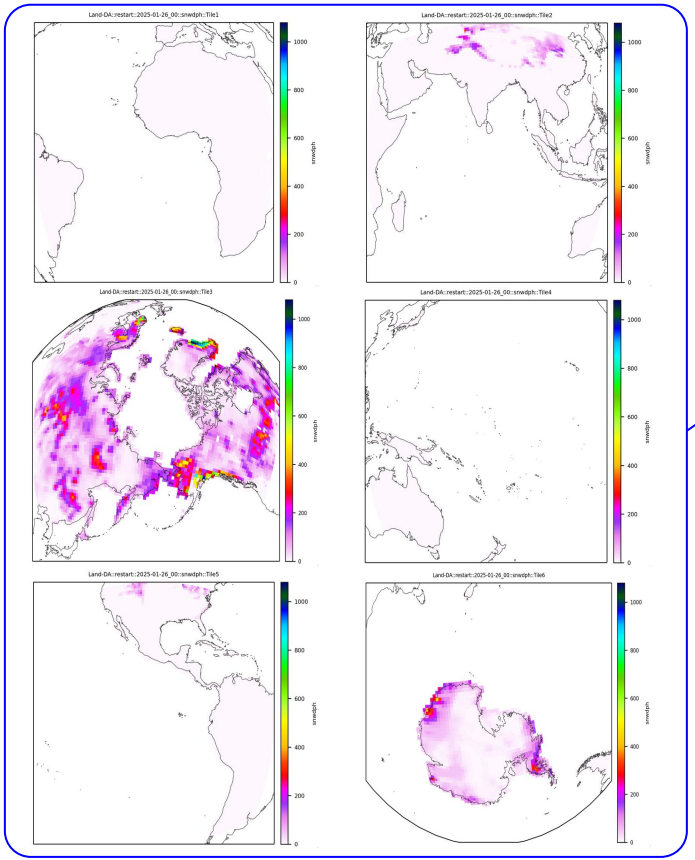


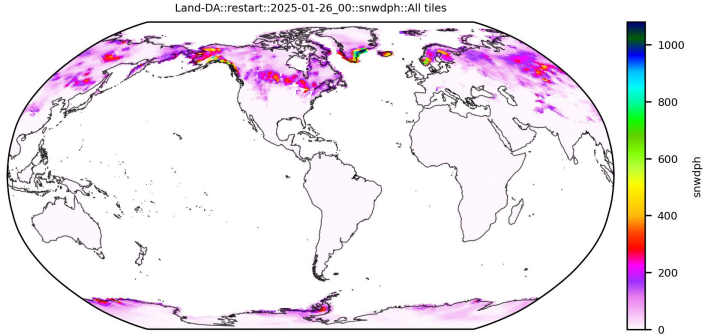


https://epic.noaa.gov/

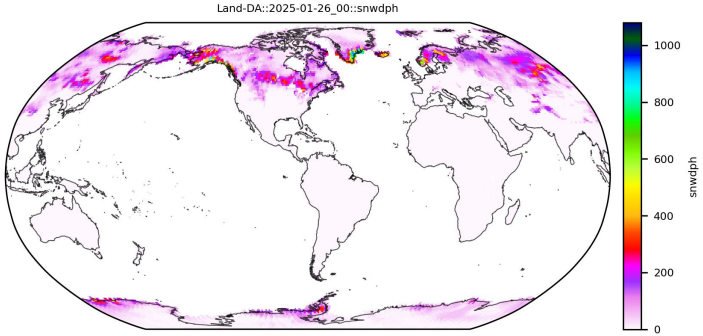
Post-processing: Plots of Restart Files

- Task ‘plot_stats’: three types of ‘**snow depth**’ plots for restart files (result of UFS weather model)





All tiles on the globe

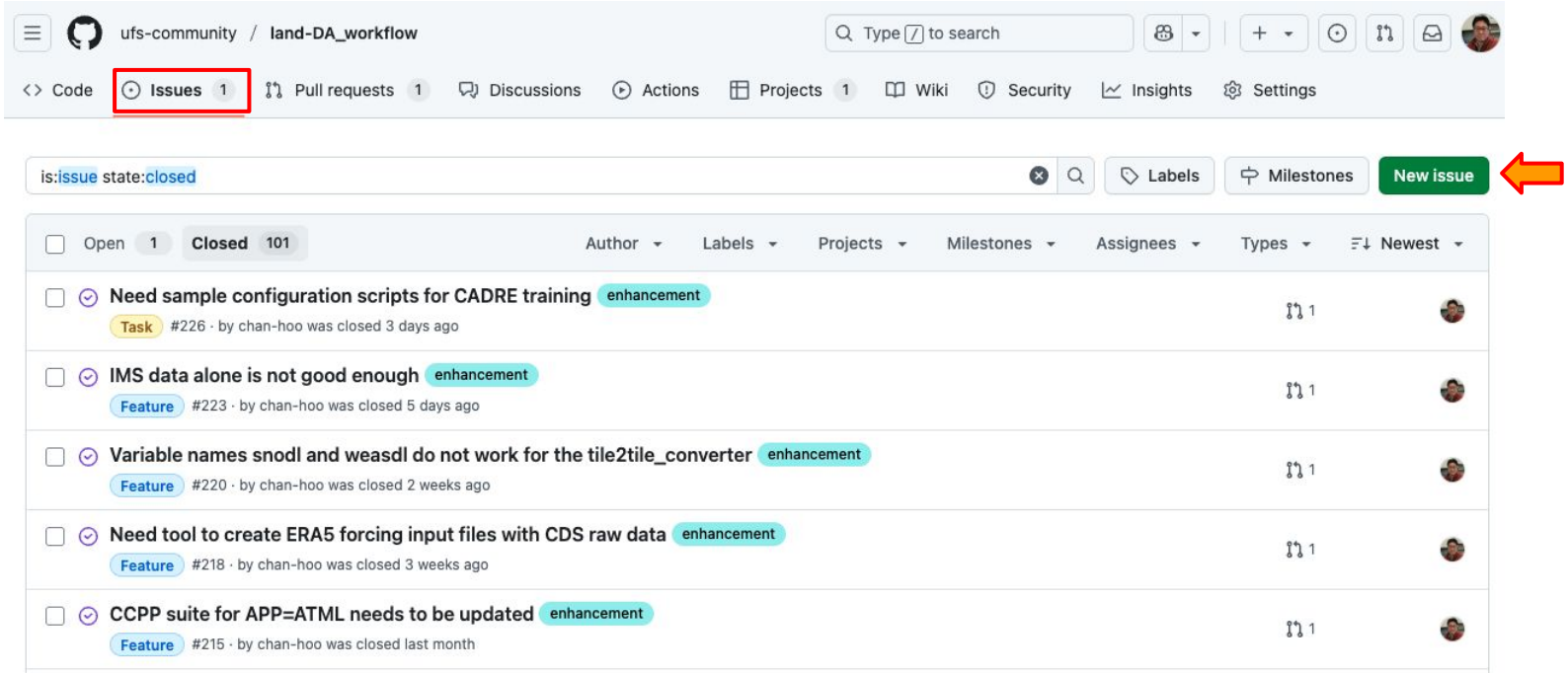


Interpolated onto the coarse grid (400x200)

https://epic.noaa.gov/

Git Issues to GitHub Repository

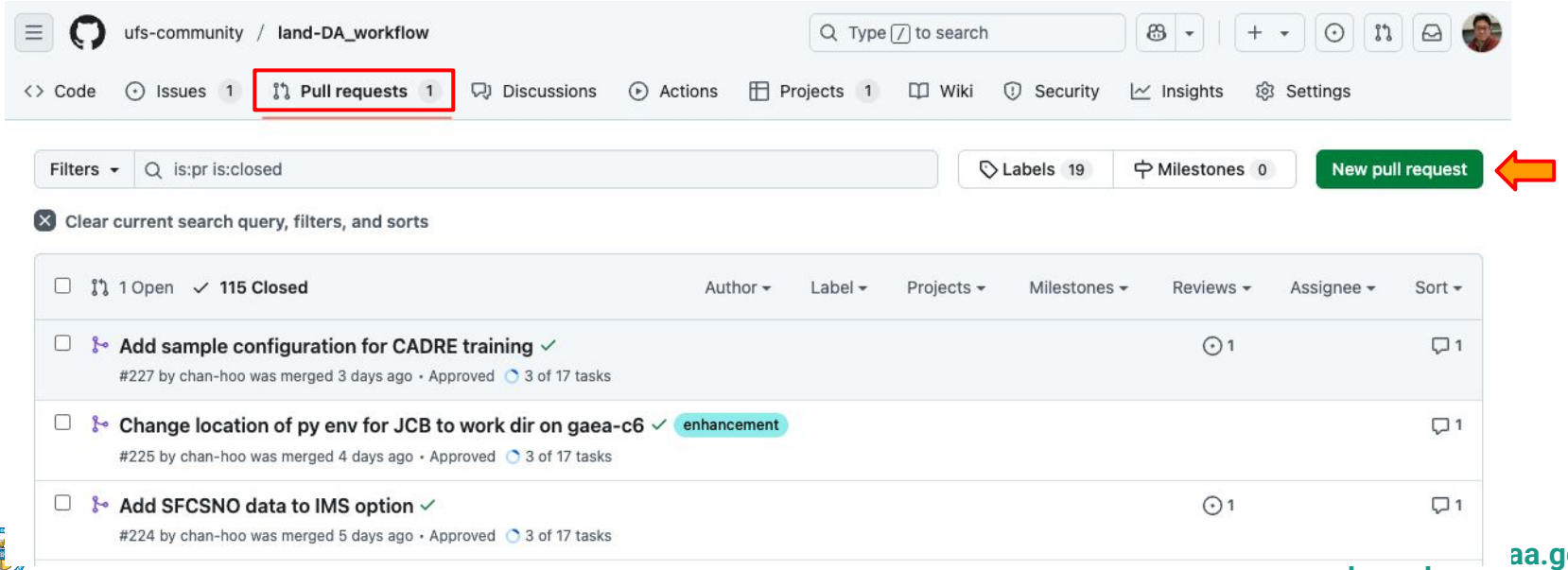
- You can open any issues on the land-DA workflow to the GitHub repository:



Github repository of UFS land-DA workflow: https://github.com/ufs-community/land-DA_workflow

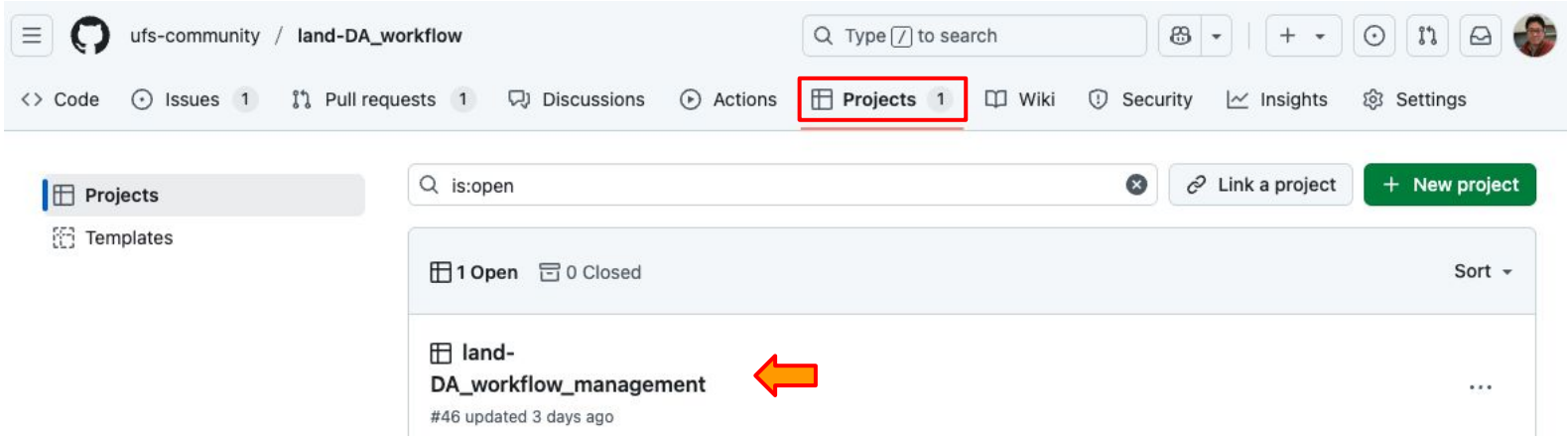
Pull Request (PR) to GitHub Repository

- Every update/development is integrated into the repo through a pull request (PR).
- PRs are reviewed by the code managers and should be approved by at least one.
- All PR commit history can be checked in the GitHub repository.



Project Management in GitHub Repository

- The GitHub issues and PRs are managed in the GitHub projects.



1	Executable 'apply_incr' does not work correctly in certain case #172	chan-hoo	On Hold		bug
2	Build-related files and directories should be located inside sorc to meet NCO standards #48	chan-hoo	Done	#49	NCO EE2 compliance
3	template files should be relocated #50	chan-hoo	Done	#49	NCO EE2 compliance
4	Module and configuration files need to be re-organized #51	chan-hoo	Done	#93	enhancement
5	Versions of modules should be controlled in centralized location #52	chan-hoo	Done	#63	NCO EE2 compliance

https://epic.noaa.gov/

Next Steps

- Continue to enhance the coupling options of Noah-MP (land) and FV3ATM (atmosphere) in the sync with UFS Weather Model.
- Continuous alignment of the GFSv17 JEDI land-DA operational features though the JCB analysis configuration: analysis control variables and observation database management.
- Extend analysis experiment capability including high-resolution model configuration and re-analysis capability: a complete set of pre- and post-processing utilities.
- Continue to enhance DA workflow structure and framework reproducible for additional applications.

Q / A

Community feedbacks: https://github.com/ufs-community/land-DA_workflow

or

Emails: jong.kim@noaa.gov / chan-hoo.jeon@noaa.gov