Land DA System Training

Running the Containerized Land DA System

By: Gillian Petro, Chan-Hoo Jeon, Edward Snyder
(Key stakeholder groups: NOAA/EMC, PSL, GSL, NESDIS ; NCAR ; JCSDA)

https://epic.noaa.gov/

How to Get Help!

- Raise your hand! :)
- Slack Channel: #cadre-epic-data-assimilation-training
  - For real-time assistance during the CADRE-EPIC DA Training
- Land DA GitHub Discussions forum
  - For more in-depth inquiries about the science behind the Land DA workflow
  - For post-training assistance using the Land DA System

Preliminary Steps

- SSH to your instance:
  - `ssh student(#)@jump.epic.noaa.gov`

    OR
  - `ssh student(#)@137.75.93.46`
- Run `ls` to see what is there. You should see:

```
[ubuntu@ip-10-29-93-178:~$ ls
Land-DA_v2.1_inputs.tar.gz  land-DA_workflow    ubuntu22.04-intel-landda-cadre25.img
inputs                      rocoto
jedi-bundle                 setup_container.sh
```

Raise your hand if:

1. You have NOT ssh'd to the system or
2. You do NOT see the expected files

Overview

- Six (6) available cases
  - Will cover 2-3 cases:

Filename prefix	App	DATM forcing	JEDI algorithm	Observation	Cold/Warm start	Number of cycles	Note
cadre1	**LND**	**ERA5**	**3D-Var**	**IMS + SFCSNO**	**Warm start**	**7**	**Benchmark case**
cadre2	**LND**	**GSWP3**	**LETKF**	**GHCN**	**Warm start**	**2**	**Different forcing and observation options**
cadre3	**ATML**	**N/A**	**3D-Var**	**GHCN**	**Cold start**	**2**	**Option 2: Noah-MP + FV3ATM**

Documentation Resources

- Follow along in the documentation
  - Copy-paste from commands in repo: cadre-epic-commands.md
- Prerequisites: preinstalled
- Data: prestaged
- Container: predownloaded
- Container setup: Performed yesterday in Edward Snyder's container presentation

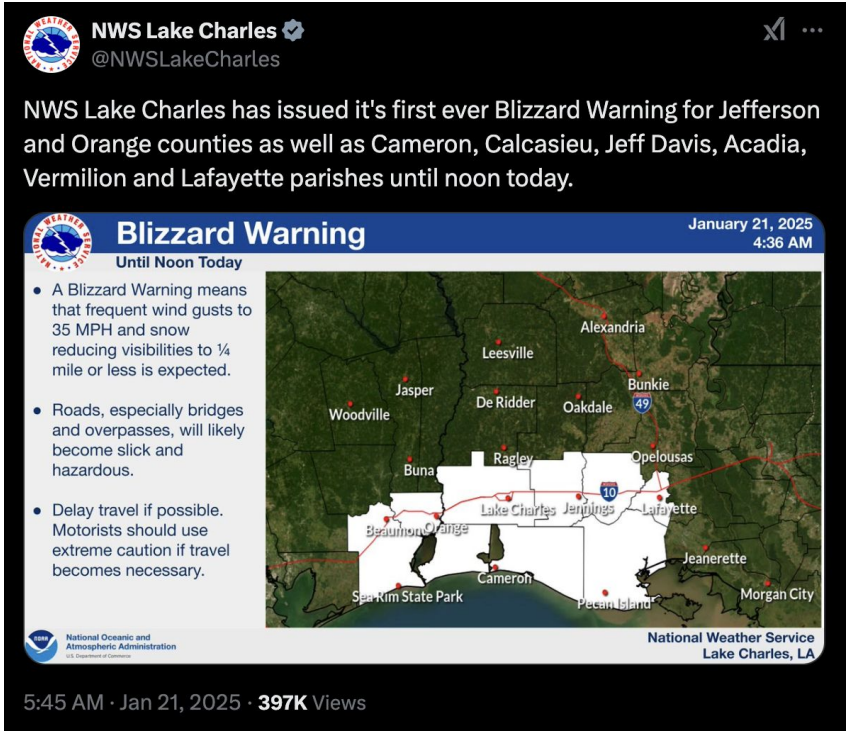
Repositories:

- CADRE-DA-training: https://github.com/NOAA-EPIC/CADRE-DA-training
- land-DA_workflow: https://github.com/ufs-community/land-DA_workflow

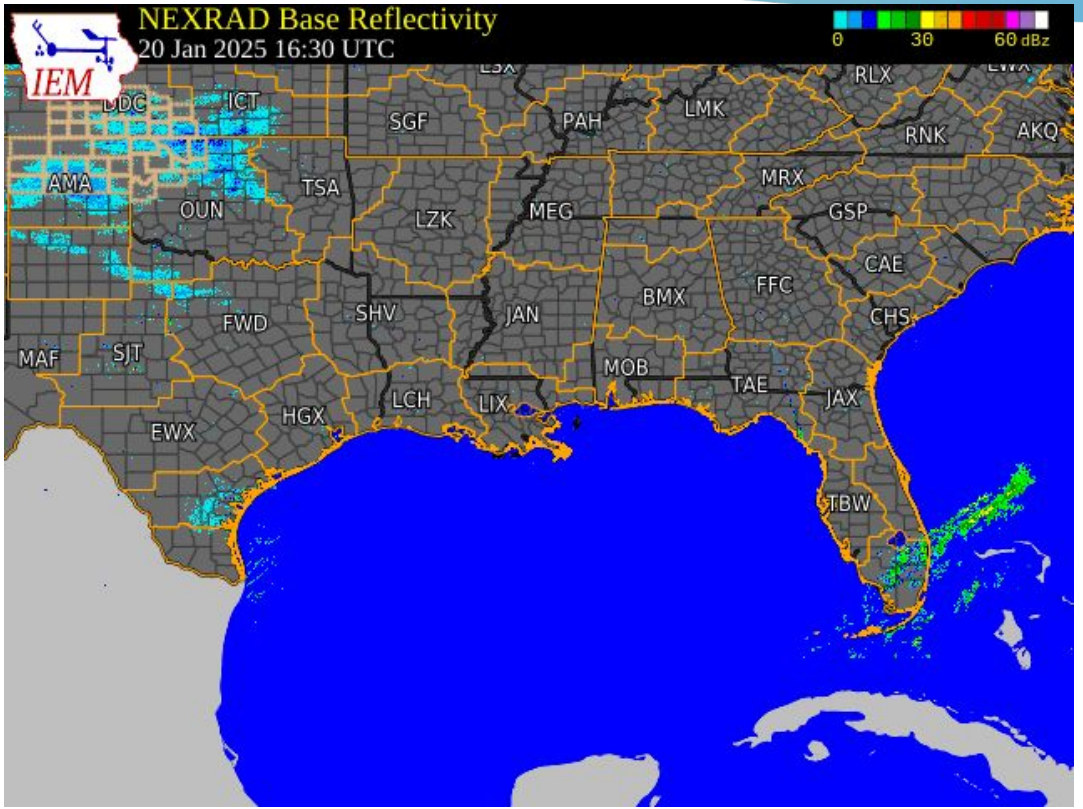
cadre1: Gulf Coast Blizzard

cadre1: Gulf Coast Blizzard

- Polar vortex stretched far south
- Unusually warm gulf waters
- Result: Unprecedented snowfall across Gulf Coast
- National Weather Service (NWS):
  - Winter Storm Warning
  - Extreme Cold Warning
  - First ever blizzard warning in Lake Charles, LA area



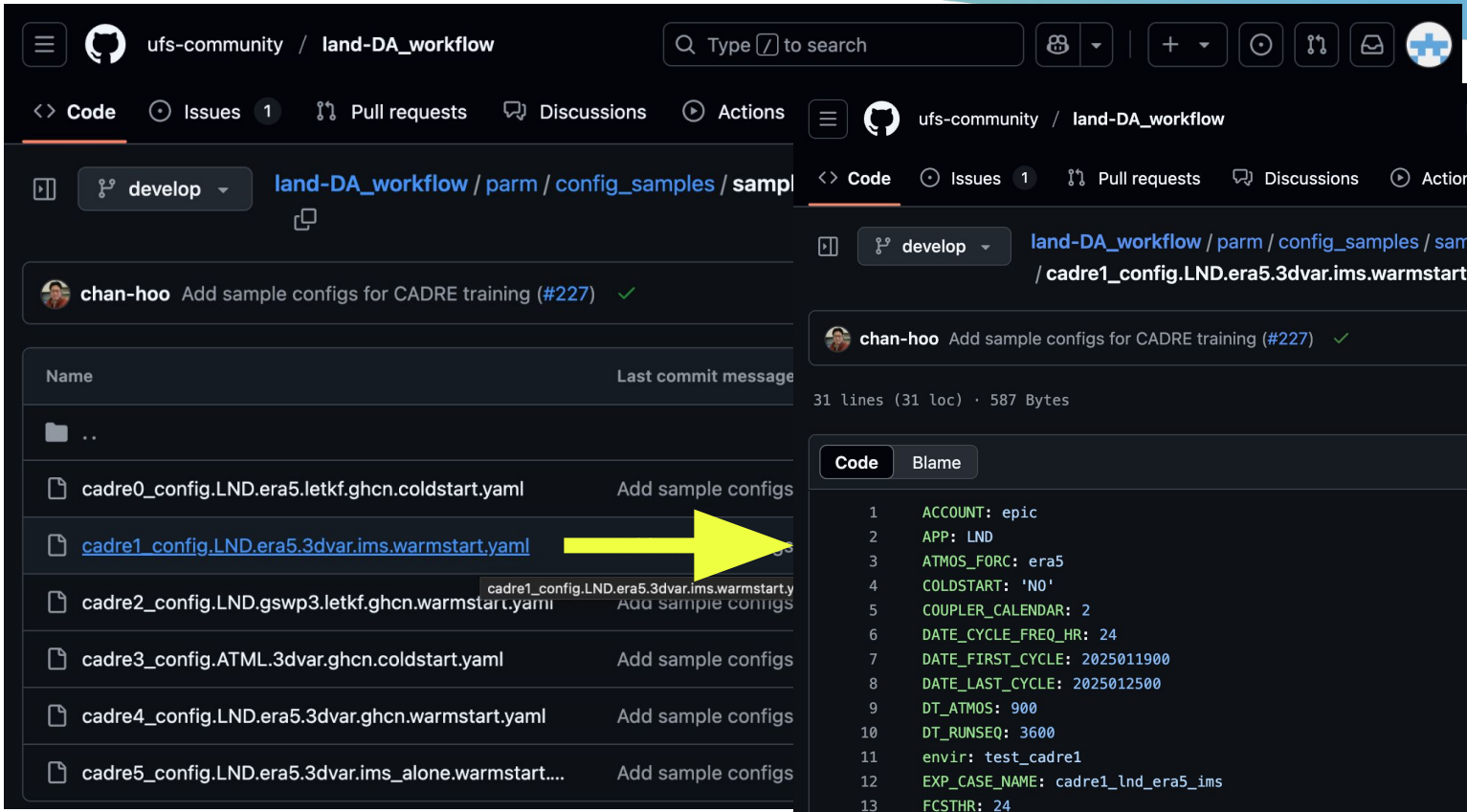
https://epic.noaa.gov/



Iowa State University Iowa Environmental Mesonet

cadre1 configuration

- 2025-01-19 0z to 2025-01-26 0z
  - 24-hour forecasts starting Jan. 19, 2025
  - Data assimilation every 24 hours
  - Relevant variables:
    - DATE_CYCLE_FREQ_HR: 24
    - DATE_FIRST_CYCLE: 2025011900
    - DATE_LAST_CYCLE: 2025012500
    - FCSTHR: 24
- NOTE: Storm hit the Gulf Coast late Jan. 20 and left land by Jan. 23


ufs-community / land-DA_workflow
Type / to search
Code
Issues 1
Pull requests
Discussions
Actions
develop
land-DA_workflow / parm / config_samples / sampl
chan-hoo Add sample configs for CADRE training (#227)
Name
Last commit message
..
cadre0_config.LND.era5.letkf.ghcn.coldstart.yaml
Add sample configs
cadre1_config.LND.era5.3dvar.ims.warmstart.yaml
cadre1_config.LND.era5.3dvar.ims.warmstart.y
cadre2_config.LND.gswp3.letkf.ghcn.warmstart.yaml
cadre3_config.ATML.3dvar.ghcn.coldstart.yaml
Add sample configs
cadre4_config.LND.era5.3dvar.ghcn.warmstart.yaml
Add sample configs
cadre5_config.LND.era5.3dvar.ims_alone.warmstart....
Add sample configs

ufs-community / land-DA_workflow
Code
Issues 1
Pull requests
Discussions
Actions
Projects 1
Wiki
develop
land-DA_workflow / parm / config_samples / samples_cadre
/ cadre1_config.LND.era5.3dvar.ims.warmstart.yaml
Go to file
chan-hoo Add sample configs for CADRE training (#227)
83645e4 · 2
31 lines (31 loc) · 587 Bytes
Code
Blame
Raw
1 ACCOUNT: epic
2 APP: LND
3 ATMOS_FORC: era5
4 COLDSTART: 'NO'
5 COUPLER_CALENDAR: 2
6 DATE_CYCLE_FREQ_HR: 24
7 DATE_FIRST_CYCLE: 2025011900
8 DATE_LAST_CYCLE: 2025012500
9 DT_ATMOS: 900
10 DT_RUNSEQ: 3600
11 envir: test_cadre1
12 EXP_CASE_NAME: cadre1_lnd_era5_ims
13 FCSTHR: 24
14 JEDI_ALGORITHM: 3dvar
15 LND_CALC_SNET: .true.
16 LND_IC_TYPE: custom

Configure the experiment

- NOTE: Container setup performed yesterday in Edward Snyder's container presentation
- Load the workflow environment:
  - `cd land-DA_workflow`
  - `module use modulefiles`
  - `module load wflow_singularity`
- Copy the sample configuration file into `config.yaml`:
  - `cd parm`
  - `cp config_samples/samples_cadre/cadre1_config.LND.era5.3dvar.ims.warmstart.yaml config.yaml`

Configure the experiment cont'd

- Open config.yaml
  - `vim config.yaml`
- Edit (use `i` key to enter insert mode):
  - **DATE_LAST_CYCLE**: 2025012200
  - Change DATE_LAST_CYCLE
- Generate the experiment directory by running:
  - `./setup_wflow_env.py -p=singularity`

```
ACCOUNT: epic
APP: LND
ATMOS_FORC: era5
COLDSTART: 'NO'
COUPLER_CALENDAR: 2
DATE_CYCLE_FREQ_HR: 24
DATE_FIRST_CYCLE: 2025011900
DATE_LAST_CYCLE: 2025012500
DT_ATMOS: 900
DT_RUNSEQ: 3600
envir: test_cadre1
EXP_CASE_NAME: cadre1_lnd_era5_ims
FCSTHR: 24
JEDI_ALGORITHM: 3dvar
LND_CALC_SNET: .true.
LND_IC_TYPE: custom
LND_INITIAL_ALBEDO: 0.25
LND_LAYOUT_X: 1
LND_LAYOUT_Y: 2
LND_OUTPUT_FREQ_SEC: 21600
MED_COUPLING_MODE: ufs.nfrac.aoflux
model_ver: v2.1.0
net: landda
NPROCS_ANALYSIS: 6
NPZ: 127
OBS_GHCN_SNOW: 'NO'
OBS_IMS_SNOW: 'YES'
OBS_SFCSNO: 'YES'
RES: 96
RUN: landda
WE2E_TEST: 'NO'
```

setup_wflow_env.py Output (excerpt)

```
ubuntu@ip-10-29-93-226:~/land-DA_workflow/parm$ ./setup_wflow_env.py -p=singularity
 Python Log Level= str: INFO, attr: 20
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L34:: Current directory (PARMdir): /home/ubuntu/land-DA_workflow/parm
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L36:: Home directory (HOMEdir): /home/ubuntu/land-DA_workflow
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L38:: Experimental base directory (exp_basedir): /home/ubuntu
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L168:: Experimental case directory /home/ubuntu/exp_case/cadre1_lnd_era5_ims has been cre
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L175:: Rocoto YAML template: /home/ubuntu/land-DA_workflow/parm/templates/template.land_a
**************************************************
Overriding              ACCOUNT = epic
Overriding                  APP = LND
Overriding           ATMOS_FORC = era5
...
Overriding        queue_default = batch
Overriding              res_p1 = 97
**************************************************
               KEEPDATA: YES
                    RUN: landda
     nprocs_forecast_lnd: 36
       MED_COUPLING_MODE: ufs.nfrac.aoflux
           EXP_CASE_NAME: cadre1_lnd_era5_ims
                     NPZ: 127
...
             exp_basedir: /home/ubuntu
                     RES: 96
            ATM_LAYOUT_X: 3
          native_default: None
            ATM_LAYOUT_Y: 8
DATM_STREAM_FN_LAST_DATE:
            LND_LAYOUT_Y: 3
     LND_OUTPUT_FREQ_SEC: 21600
INFO::/home/ubuntu/land-DA_workflow/sorc/conda/envs/land_da/lib/python3.12/site-packages/uwtools/config/validator.py::L76::0 schema-validation errors fo
INFO::/home/ubuntu/land-DA_workflow/sorc/conda/envs/land_da/lib/python3.12/site-packages/uwtools/rocoto.py::L66::0 Rocoto XML validation errors found
ubuntu@ip-10-29-93-226:~/land-DA_workflow/parm$
```

Questions?

The Experiment Directory

- Navigate to the experiment directory:
  - `cd ../../exp_case/cadre1_lnd_era5_ims/`
- Run the workflow launch command:
  - `rocotorun -w land_analysis.xml -d land_analysis.db`
- To see jobs that are in progress:
  - `squeue -u $USER`
- Check progress with `rocotostat`:
  - `rocotostat -w land_analysis.xml -d land_analysis.db`
  - STATE could be QUEUED, SUBMITTING, RUNNING, SUCCEEDED, DEAD, UNAVAILABLE
    - Let us know if a task goes DEAD or UNAVAILABLE!!!

CYCLE	TASK	JOBID	STATE	EXIT STATUS	TRIES	DURATION
202501200000	jcb	4	QUEUED	-	0	0.0
202501200000	prep_data	14	QUEUED	-	0	0.0
202501200000	pre_anal	15	QUEUED	-	0	0.0
202501200000	analysis	-	-	-	-	-
202501200000	post_anal	-	-	-	-	-
202501200000	forecast	-	-	-	-	-
202501200000	plot_stats	-	-	-	-	-
202501210000	jcb	5	QUEUED	-	0	0.0
202501210000	prep_data	-	-	-	-	-
202501210000	pre_anal	-	-	-	-	-
202501210000	analysis	-	-	-	-	-
202501210000	post_anal	-	-	-	-	-
202501210000	forecast	-	-	-	-	-
202501210000	plot_stats	-	-	-	-	-
202501220000	jcb	6	QUEUED	-	0	0.0
202501220000	prep_data	-	-	-	-	-
202501220000	pre_anal	-	-	-	-	-
202501220000	analysis	-	-	-	-	-
202501220000	post_anal	-	-	-	-	-
202501220000	forecast	-	-	-	-	-
202501220000	plot_stats	-	-	-	-	-

Running the Experiment

- Normally, users need to issue the **`rocotorun`** command multiple times.
  - **`rocotorun`** initiates the next task once its dependencies have completed successfully.
  - **`rocotostat`** will only show updates after a **`rocotorun`** command is issued.
- Automate the workflow using the **`run_expt.sh`** script:
  - **`wget https://raw.githubusercontent.com/NOAA-EPIC/CADRE-DA-training/refs/heads/main/Day2/run_expt.sh /home/ubuntu/exp_case/cadre1_lnd_era5_ims`**
  - **`chmod 755 run_expt.sh`**
  - **`./run_expt.sh`**

Directory Structure

```
$LANDDAROOT (<exp_basedir>): Base directory
 ├── land-DA_workflow (<HOMElandda> or <CYCLEDIR>): Home directory of the land DA workflow
 │      ├── jobs
 │      ├── modulefiles
 │      ├── parm
 │      ├── scripts
 │      ├── sorc
 │      └── ush
 ├── exp_case
 │      └── EXP_CASE_NAME
 │            ├── com_dir --> symlinked to ptmp/test_*/com/landda/v2.1.0
 │            ├── land_analysis.yaml
 │            ├── land_analysis.xml
 │            ├── launch_rocoto_wflow.sh
 │            ├── log_dir --> symlinked to ptmp/test_*/com/output/logs
 │            └── tmp_dir --> symlinked to ptmp/test_*/com/tmp
 └── ptmp (<PTMP>)
       └── test_* (<envir> or <OPSROOT>)
             └── com (<COMROOT>)
             │     ├── landda (<NET>)
             │     │     └── vX.Y.Z (<model_ver>)
             │     │           └── landda.YYYYMMDD (<RUN>.<PDY>): Directory containing the output files
             │     │                 ├── hofx
             │     │                 └── plot
             │     └── output
             │           └── logs (<LOGDIR>): Directory containing the log files for the Rocoto workflow
             └── tmp (<DATAROOT>)
                   ├── <jobid> (<DATA>): Working directory
                   └── DATA_SHARE
                         ├── YYYYMMDD (<PDY>): Directory containing the intermediate or temporary files
                         ├── hofx: Directory containing the soft links to the results of the analysis task for plotting
                         └── DATA_RESTART: Directory containing the soft links to the restart files for the next cycles
```

Troubleshooting

- Navigate to the log directory:
  - `cd log_dir`

```
ubuntu@ip-10-29-93-209:~/exp_case/cadre1_lnd_era5_ims/log_dir$ ls
analysis_2025011900.log  forecast_2025012100.log  plot_stats_2025011900.log  post_anal_2025012100.log  prep_data_2025011900.log
analysis_2025012000.log  forecast_2025012200.log  plot_stats_2025012000.log  post_anal_2025012200.log  prep_data_2025012000.log
analysis_2025012100.log  jcb_2025011900.log       plot_stats_2025012100.log  pre_anal_2025011900.log   prep_data_2025012100.log
analysis_2025012200.log  jcb_2025012000.log       plot_stats_2025012200.log  pre_anal_2025012000.log   prep_data_2025012200.log
forecast_2025011900.log  jcb_2025012100.log       post_anal_2025011900.log   pre_anal_2025012100.log   workflow.log
forecast_2025012000.log  jcb_2025012200.log       post_anal_2025012000.log   pre_anal_2025012200.log
```

- View the log file for the tasks that went DEAD
  - `vim prep_data_2025012000.log`
  - Scroll to the bottom for the last output
  - Search for error or failure messages
- Rewind a task using Rocoto (see reference guide)

Is anyone seeing DEAD tasks?

Compare Plots

Plots

- Navigate to the plots directory:
  - `cd ../com_dir/landda.202501##/plot`
  - One directory per cycle
- Each directory will contain several plots:

```
ubuntu@ip-10-29-93-226:~/exp_case/cadre1_lnd_era5_ims/com_dir/landda.20250120/plot$ ls
hofx_omb_ims_snow_20250120_histogram.png      landda_out_restart_2025-01-21_00_snwdph_alltiles.png
hofx_omb_ims_snow_20250120_scatter.png        landda_out_restart_2025-01-21_00_snwdph_tile1.png
hofx_omb_sfcsno_20250120_histogram.png        landda_out_restart_2025-01-21_00_snwdph_tile2.png
hofx_omb_sfcsno_20250120_scatter.png          landda_out_restart_2025-01-21_00_snwdph_tile3.png
landda_comp_sfc_20250120_snwdph_after.png     landda_out_restart_2025-01-21_00_snwdph_tile4.png
landda_comp_sfc_20250120_snwdph_before.png    landda_out_restart_2025-01-21_00_snwdph_tile5.png
landda_comp_sfc_20250120_snwdph_diff_sfc.png  landda_out_restart_2025-01-21_00_snwdph_tile6.png
landda_comp_sfc_20250120_snwdph_inc.png       landda_timehistory_omb_totalSnowDepth_ims_snow.png
landda_obs_ims_20250120_SnowDepth.png         landda_timehistory_omb_totalSnowDepth_sfcsno.png
landda_out_combined_2025-01-21_00_snwdph.png
```

Downloading Plots

- Open a new terminal window.
- Type **`bash`** to ensure a bash shell.
- Add your private key (e.g., **`ssh-add ~/.ssh/id_ed25519_student1`**).
- For each directory of plots, run:
  - **`rsync -v --rsh "ssh student#@jump.epic.noaa.gov ssh" ubuntu@controller:/home/ubuntu/exp_case/cadre1_lnd_era5_ims/com_dir/landda.202501##/plot/* ./plots/cadre1/202501##`**
- In the command, replace:
  - **`student#`** with your actual student number,
  - **`landda.202501##`** with the cycle date, and
  - **`./plots/202501##/`** with the cycle date.

Questions?

cadre2

cadre2

- 2000-02-02 0z - 2000-02-03 0z
  - 24-hour forecasts starting Feb. 2, 2000
  - Data assimilation every 24 hours
  - Relevant variables:
    - DATE_CYCLE_FREQ_HR: 24
    - DATE_FIRST_CYCLE: 2000020200
    - DATE_LAST_CYCLE: 2000020300
    - FCSTHR: 24
  - **GSWP3 forcing data:** Daily-resolution observed climate data on a global (land and ocean) 0.5°x0.5° lat-lon grid from the Global Soil Wetness Project Phase 3 (GSWP3). The data set covers the period 1901-2010.
- https://github.com/ufs-community/ufs-weather-model/blob/develop/tests/tests/datm_cdeps_lnd_gswp3

cadre2

Before continuing previously started work:

- Open a new terminal window and ssh to system:
  - `ssh student(#)@jump.epic.noaa.gov / ssh student(#)@137.75.93.46`
- Load the workflow environment:
  - `cd land-DA_workflow`
  - `module use modulefiles`
  - `module load wflow_singularity`

Configure cadre2

- Navigate to the **`parm`** directory to configure the new experiment:
  - **`cd parm`**
- Copy the new experiment configuration into **`config.yaml`**:
  - **`cp config_samples/samples_cadre/cadre2_config.LND.gswp3.letkf.ghcn.warmstart.yaml config.yaml`**
- Generate the experiment directory by running:
  - **`./setup_wflow_env.py -p=singularity`**

```
ubuntu@ip-10-29-93-178:~/land-DA_workflow/parm$ ./setup_wflow_env.py -p=singularity
 Python Log Level= str: INFO, attr: 20
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L34:: Current directory (PARMdir): /home/ubuntu/land-DA_workflow/parm
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L36:: Home directory (HOMEdir): /home/ubuntu/land-DA_workflow
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L38:: Experimental base directory (exp_basedir): /home/ubuntu
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L168:: Experimental case directory /home/ubuntu/exp_case/cadre2_lnd_gswp3_ghcn has been creat
ed.
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L175:: Rocoto YAML template: /home/ubuntu/land-DA_workflow/parm/templates/template.land_analy
sis.yaml
**************************************************
Overriding              ACCOUNT = epic
Overriding                  APP = LND
Overriding           ATMOS_FORC = gswp3
Overriding        queue_default = batch
Overriding               res_p1 = 97
**************************************************
               WE2E_TEST: NO
               JEDI_PATH: /home/ubuntu
           ATM_LAYOUT_Y: 8
                 FCSTHR: 24
       MED_COUPLING_MODE: ufs.nfrac.aoflux
                     APP: LND
             CCPP_SUITE: FV3_GFS_v17_p8_ugwpv1
                  envir: test_cadre2
              KEEPDATA: YES
            JEDI_PY_VER: python3.10
              DT_RUNSEQ: 3600
     nprocs_forecast_lnd: 36
          JEDI_ALGORITHM: letkf
            WRITE_GROUPS: 1
    WRITE_TASKS_PER_GROUP: 6
INFO::/home/ubuntu/land-DA_workflow/sorc/conda/envs/land_da/lib/python3.12/site-packages/uwtools/config/validator.py::L76::0 schema-validation errors f
in Rocoto config
INFO::/home/ubuntu/land-DA_workflow/sorc/conda/envs/land_da/lib/python3.12/site-packages/uwtools/rocoto.py::L66::0 Rocoto XML validation errors found
ubuntu@ip-10-29-93-178:~/land-DA_workflow/parm$
```

cadre2: *setup_wflow_env.py* output (excerpts)

The Experiment Directory

- Navigate to the experiment directory:
  - `cd ../../exp_case/cadre2_lnd_gswp3_ghcn/`
- Run the workflow launch command:
  - `rocotorun -w land_analysis.xml -d land_analysis.db`
- To see jobs that are in progress:
  - `squeue -u $USER`
- Check progress with `rocotostat`:
  - `rocotostat -w land_analysis.xml -d land_analysis.db`
  - STATE could be QUEUED, SUBMITTING, RUNNING, SUCCEEDED, DEAD, or UNAVAILABLE
    - Let us know if a task goes DEAD or UNAVAILABLE!!!

Running the Experiment

- Normally, users need to issue the **`rocotorun`** command multiple times.
  - **`rocotorun`** initiates the next task once its dependencies have completed successfully.
  - **`rocotostat`** will only show updates after a **`rocotorun`** command is issued.
- Automate the workflow using the **`run_expt.sh`** script:
  - **`wget https://raw.githubusercontent.com/NOAA-EPIC/CADRE-DA-training/refs/heads/main/Day2/run_expt.sh .`**
  - **`chmod 755 run_expt.sh`**
  - **`./run_expt.sh`**

Questions?

Compare Plots

Plots

- Navigate to the plots directory:
  - `cd com_dir/landda.2000020#/plot`
  - One directory per cycle
- Each directory will contain several plots:

```
[ubuntu@ip-10-29-93-209:~/exp_case/cadre2_lnd_gswp3_ghcn/com_dir/landda.20000202/plot$ ls
hofx_omb_ghcn_snow_20000202_histogram.png     landda_out_restart_2000-02-03_00_snwdph_alltiles.png
hofx_omb_ghcn_snow_20000202_scatter.png       landda_out_restart_2000-02-03_00_snwdph_tile1.png
landda_comp_sfc_20000202_snwdph_after.png     landda_out_restart_2000-02-03_00_snwdph_tile2.png
landda_comp_sfc_20000202_snwdph_before.png    landda_out_restart_2000-02-03_00_snwdph_tile3.png
landda_comp_sfc_20000202_snwdph_diff_sfc.png  landda_out_restart_2000-02-03_00_snwdph_tile4.png
landda_comp_sfc_20000202_snwdph_inc.png       landda_out_restart_2000-02-03_00_snwdph_tile5.png
landda_obs_ghcn_20000202_SnowDepth.png        landda_out_restart_2000-02-03_00_snwdph_tile6.png
landda_out_combined_2000-02-03_00_snwdph.png  landda_timehistory_omb_totalSnowDepth_ghcn_snow.png
```

Downloading Plots

- Open a new terminal window.
- Type **`bash`** to ensure a bash shell.
- Add your private key (e.g., **`ssh-add ~/.ssh/id_ed25519_student1`**).
- For each directory of plots, run:
  - **`rsync -v --rsh "ssh student#@jump.epic.noaa.gov ssh" ubuntu@controller:/home/ubuntu/exp_case/cadre2_lnd_gswp3_ghcn/com_dir/landda.2000020#/plot/* ./plots/cadre2/2000020#`**
- In the command, replace:
  - **`student#`** with your actual student number
  - **`landda.2000020#`** with the cycle date
  - **`./plots/2000020#`** with the cycle date

Questions?

cadre3: Building ATML executables

cadre3

- 2022-12-21 0z - 2022-12-22 0z
  - 24-hour forecasts starting Dec. 21, 2022
  - Data assimilation every 24 hours
  - Relevant variables:
    - DATE_CYCLE_FREQ_HR: 24
    - DATE_FIRST_CYCLE: 2022122100
    - DATE_LAST_CYCLE: 2022122200
    - FCSTHR: 24
- Same date as the sample configuration of the global workflow.

cadre3

- Open a new terminal window and ssh to system:
  - **`ssh student(#)@jump.epic.noaa.gov / ssh student(#)@137.75.93.46`**
  - (If continuing in same terminal window, navigate to **`/home/ubuntu`** instead.)

Build ATML executables for cadre3

- Shell into the container:
  - `singularity shell -B /home:/home /home/ubuntu/ubuntu22.04-intel-landda-cadre25.img`
- Go to the `land-DA_workflow/sorc` directory that was copied out of the container.
  - `cd /home/ubuntu/land-DA_workflow/sorc`
- Set up the environment
  - `source /opt/spack-stack/spack-stack-1.6.0/envs/fms-2024.01/.bashenv-fms`
  - `module use ../modulefiles`
  - `module load build_singularity_intel`

Build ATML executables for cadre3

- Build the model using **`app_build.sh`**.
  - Options:
    - ATML Configuration:
      - **`./app_build.sh -p=singularity -a=ATML --conda=off --build`**
    - LND (default) Configuration:
      - `./app_build.sh -p=singularity --conda=off --build`
- NOTES:
  - **`-p=singularity`** indicates that the platform uses a container build
  - **`--conda=off`** prevents rebuilding conda, which was pre-built in previous steps.
  - **`--build`** keeps the executables in the **`build`** directory under **`bin`**.

Configure cadre3

- Type **`exit`** to exit the container.
- Load the workflow environment:
  - **`cd land-DA_workflow`**
  - **`module use modulefiles`**
  - **`module load wflow_singularity`**
- Navigate to the **`parm`** directory to configure the new experiment:
  - **`cd /home/ubuntu/land-DA_workflow/parm`**
- Copy the new experiment configuration into **`config.yaml`**:
  - **`cp config_samples/samples_cadre/cadre3_config.ATML.3dvar.ghcn.coldstart.yaml config.yaml`**
- Generate the experiment directory by running:
  - **`./setup_wflow_env.py -p=singularity`**

Point to ATML Executables

- For ATML configurations, users must modify `run_container_executable.sh`. For example:
  - `vim run_container_executable.sh`
- Uncomment the second-to-last line of the script:
  - `#export SINGULARITYENV_PREPEND_PATH=/home/ubuntu/land-DA_workflow/src/build/bin:$SINGULARITYENV_PREPEND_PATH`
- This line adds the ATML executables to the container by exporting the `SINGULARITYENV_PREPEND_PATH` variable.

setup_wflow_env.py Output (excerpt)

```
 Python Log Level= str: INFO, attr: 20
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L34:: Current directory (PARMdir): /home/ubuntu/land-DA_workflow/parm
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L36:: Home directory (HOMEdir): /home/ubuntu/land-DA_workflow
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L38:: Experimental base directory (exp_basedir): /home/ubuntu
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L168:: Experimental case directory /home/ubuntu/exp_case/cadre3_atml has been
created.
INFO::/home/ubuntu/land-DA_workflow/parm/./setup_wflow_env.py::L175:: Rocoto YAML template:
/home/ubuntu/land-DA_workflow/parm/templates/template.land_analysis.yaml
************************************************
Overriding                  ACCOUNT = epic
Overriding                      APP = ATML
Overriding               ATMOS_FORC = gswp3
...
Overriding            queue_default = batch
Overriding                  res_p1 = 97
************************************************
                DATE_CYCLE_FREQ_HR: 24
          DATM_STREAM_FN_LAST_DATE:
               nprocs_forecast_atm: 54
                           APP: ATML
                    LND_CALC_SNET: .false.
...
                     ATM_LAYOUT_Y: 4
                          MACHINE: singularity
INFO::/home/ubuntu/land-DA_workflow/sorc/conda/envs/land_da/lib/python3.12/site-packages/uwtools/config/validator.py::L76::0 schema-validation
errors found in Rocoto config
INFO::/home/ubuntu/land-DA_workflow/sorc/conda/envs/land_da/lib/python3.12/site-packages/uwtools/rocoto.py::L66::0 Rocoto XML validation errors
found
```

The Experiment Directory

- Navigate to the experiment directory:
  - `cd ../../exp_case/cadre3_atm1/`
- Run the workflow launch command:
  - `rocotorun -w land_analysis.xml -d land_analysis.db`
- To see jobs that are in progress:
  - `squeue -u $USER`
- Check progress with `rocotostat`:
  - `rocotostat -w land_analysis.xml -d land_analysis.db`
  - STATE could be QUEUED, SUBMITTING, RUNNING, SUCCEEDED, DEAD, or UNAVAILABLE
    - Let us know if a task goes DEAD or UNAVAILABLE!!!

Running the Experiment

- Normally, users need to issue the **`rocotorun`** command multiple times.
  - **`rocotorun`** initiates the next task once its dependencies have completed successfully.
  - **`rocotostat`** will only show updates after a **`rocotorun`** command is issued.
- Automate the workflow using the **`run_expt.sh`** script:
  - **`wget https://raw.githubusercontent.com/NOAA-EPIC/CADRE-DA-training/refs/heads/main/Day2/run_expt.sh .`**
  - **`chmod 755 run_expt.sh`**
  - **`./run_expt.sh`**

Questions?

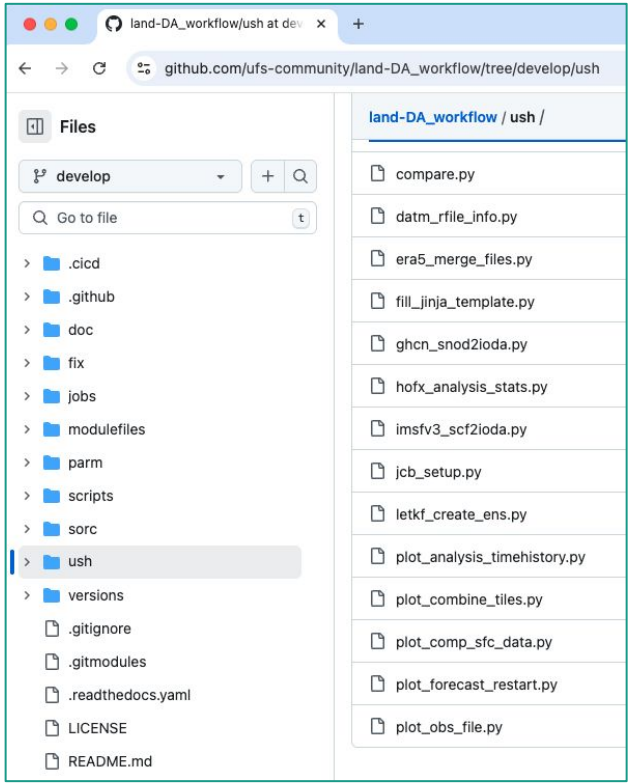
Downloading cadre3 Plots

- Open a new terminal window.
- Type **`bash`** to ensure a bash shell.
- Add your private key (e.g., **`ssh-add ~/.ssh/id_ed25519_student1`**).
- For each directory of plots, run:
  - **`rsync -v --rsh "ssh student#@137.75.93.46 ssh" ubuntu@controller:/home/ubuntu/exp_case/cadre3_atml/com_dir/landda.20221222/plot/* plots/cadre3/20221222`**
- In the command, replace:
  - **`student#`** with your actual student number
- Only one cycle has plots

Questions?

Post-Processing: cadre1 case

- The **`plot_stats`** task plots output from the experiment, including:
  - Scatter plots and histograms (hofx_analysis_stats.py)
  - Restart output files (plot_forecast_restart.py)
  - Time history of analysis output (plot_analysis_timehistory.py)
  - Comparison of **`sfc_data`** files by analysis task (plot_comp_sfc_data.py)
  - Plot observation data files (plot_obs_file.py)

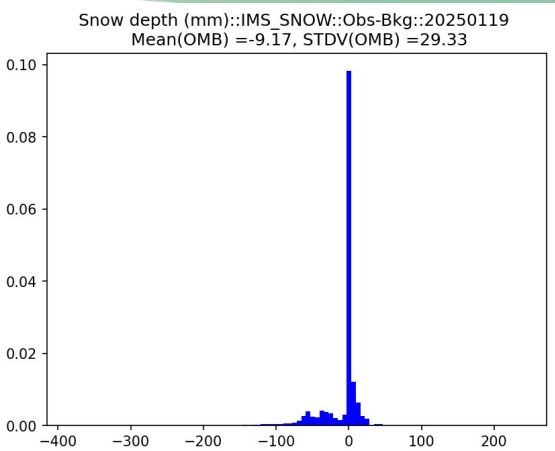


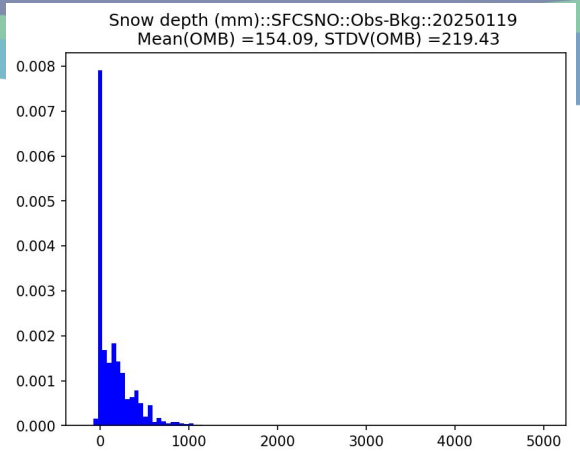
https://epic.noaa.gov/

Histograms

- Histograms show Observation Minus Background (OMB) from JEDI output
  - OMB values on the x-axis
  - Frequency density values on the y-axis
  - Title lists real value of OMB mean and standard deviation
- OMB values are skewed:
  - Left for IMS data
  - Right for SFCSNO data
- Indicates a bias/issue in the data

1/19

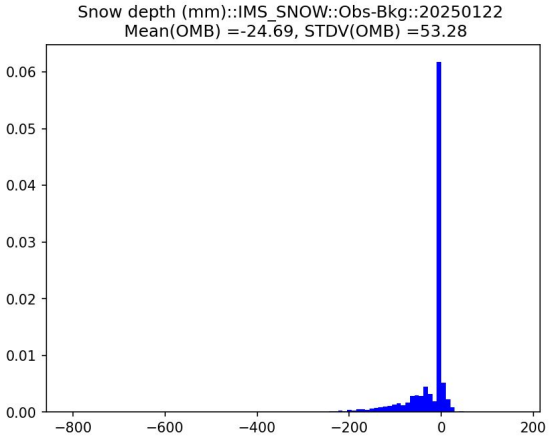


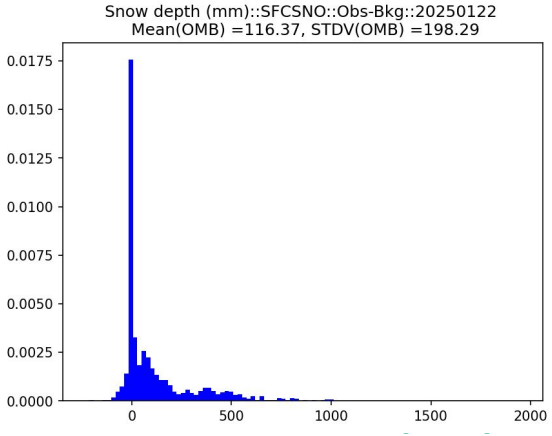


IMS

SFCSNO

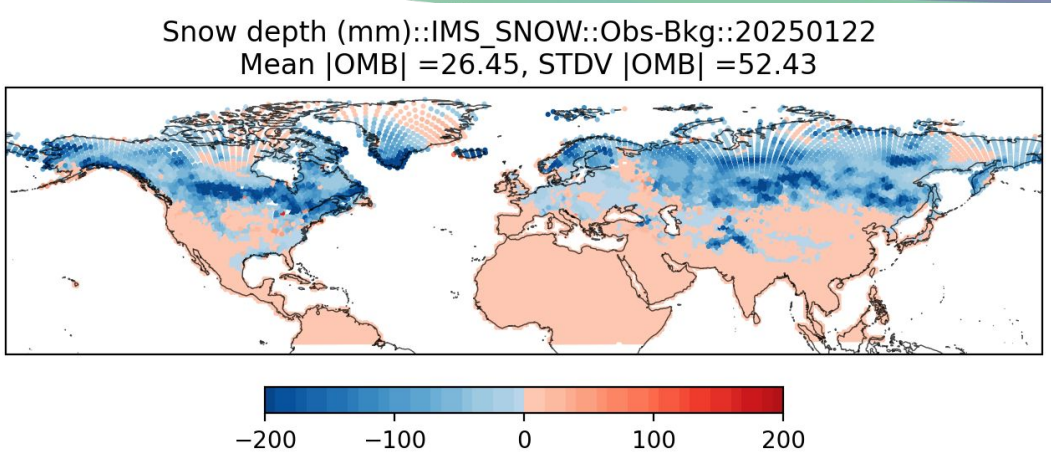
1/22



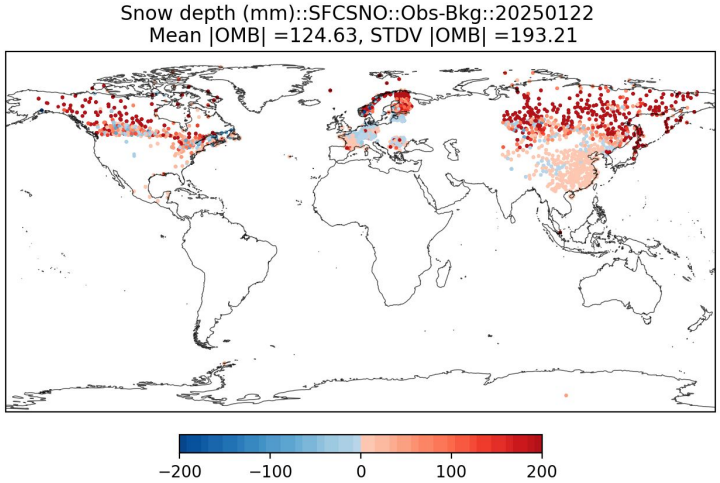


Scatterplots

- Scatter plots show Observation Minus Background (OMB) from JEDI output
  - Blue points: Observed values are less than the background values
  - Red points: Observed values are greater than the background values.
  - Title indicates mean and standard deviation of absolute value of OMB.
- Bias:
  - **IMS:** Preponderance of blue
  - **SFCSNO:** Preponderance of red
- Both shifting more blue over time



IMS observed snow depth is less than the background; the model therefore seems to be overforecasting snow depth.

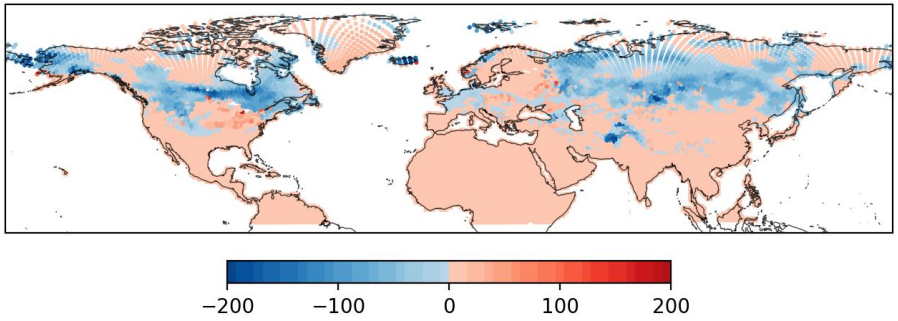


SFCSNO observed snow depth is greater than the background; the model seems to be underforecasting snow depth.

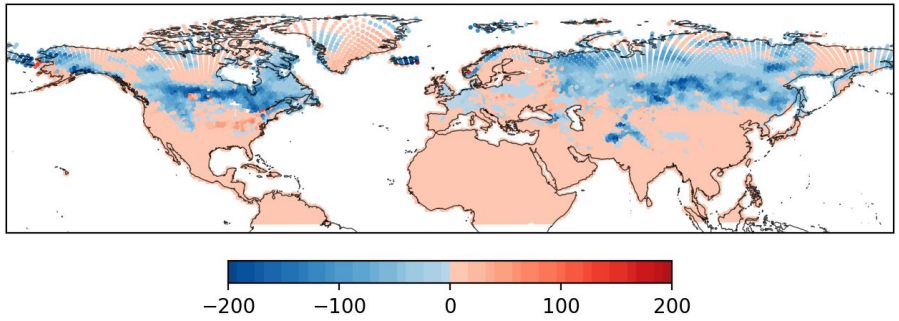
https://epic.noaa.gov/

Analysis Increment Over Time: IMS

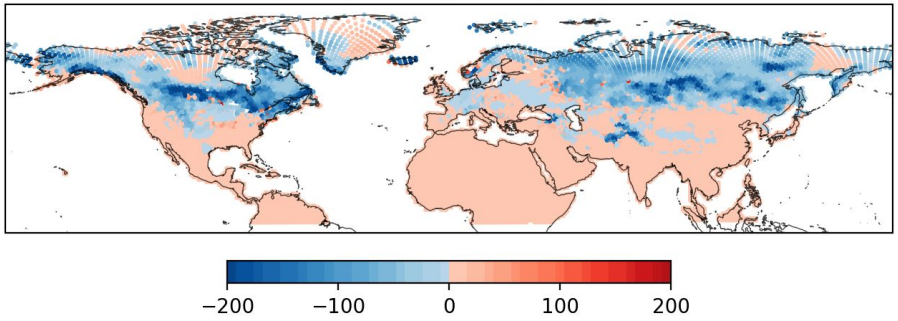
Snow depth (mm)::IMS_SNOW::Obs-Bkg::20250119
Mean |OMB| =13.62, STDV |OMB| =27.55



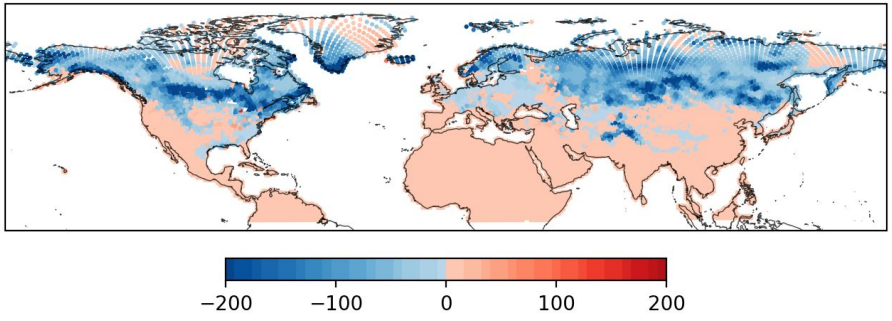
Snow depth (mm)::IMS_SNOW::Obs-Bkg::20250120
Mean |OMB| =16.22, STDV |OMB| =32.17



Snow depth (mm)::IMS_SNOW::Obs-Bkg::20250121
Mean |OMB| =21.29, STDV |OMB| =42.57

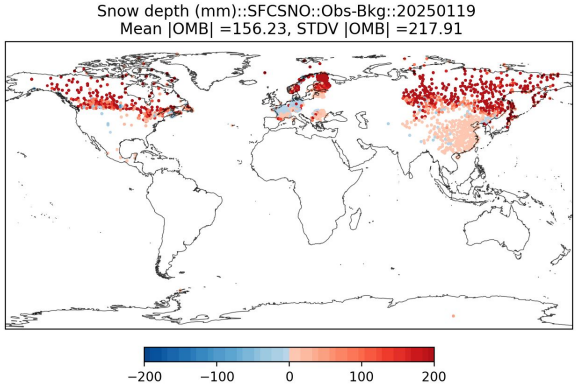


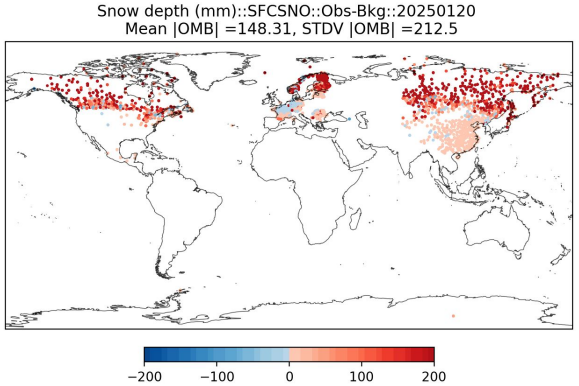
Snow depth (mm)::IMS_SNOW::Obs-Bkg::20250122
Mean |OMB| =26.45, STDV |OMB| =52.43

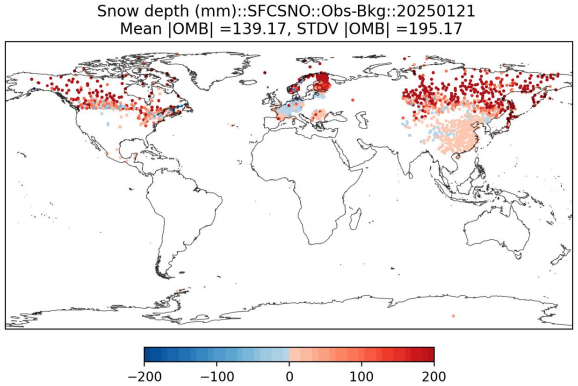


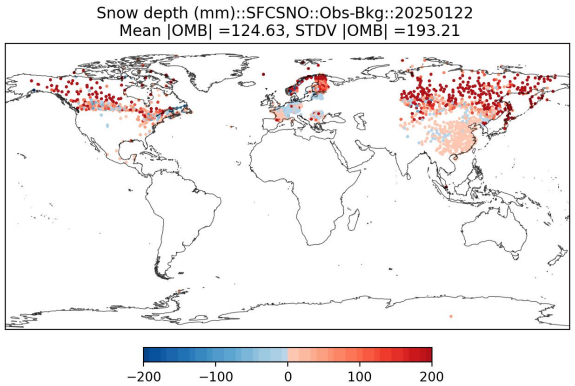
https://epic.noaa.gov/

Analysis Increment Over Time: SFCSNO









https://epic.noaa.gov/

Time-History Plots

- Time history data: `com_dir/landda.{date}/hofx/{time_history_data}.txt`

```
epic 1652462 Jun  2 16:36 diag.ims_snow_2025012200.nc
epic  988066 Jun  2 16:36 diag.sfcsno_2025012200.nc
epic     148 Jun  2 16:44 hofx_omb_timehis_abs_ims_snow.txt
epic     159 Jun  2 16:44 hofx_omb_timehis_abs_sfcsno.txt
epic     165 Jun  2 16:44 hofx_omb_timehis_ims_snow.txt
epic     174 Jun  2 16:44 hofx_omb_timehis_sfcsno.txt
```

  - Date – mean – STD

```
2025-01-19-00 -9.17 29.33 241.9 -384.58
2025-01-20-00 -12.59 33.76 199.67 -392.38
2025-01-21-00 -18.63 43.8 235.18 -586.18
2025-01-22-00 -24.69 53.28 166.54 -809.05
```

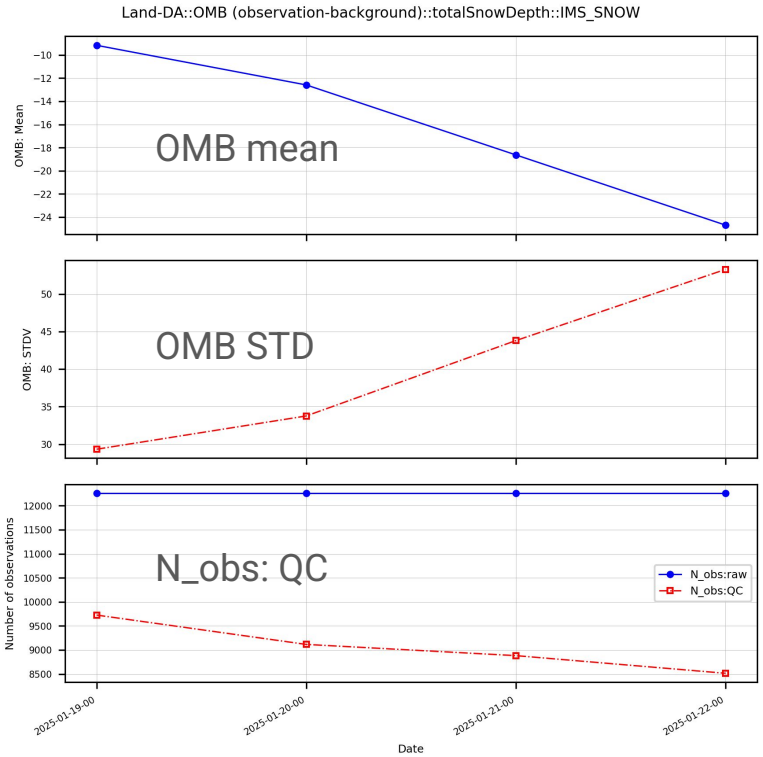
- Analysis task log file: `log_dir/analysis_{date}.log`
  - Contains info on quality control (QC) for observations

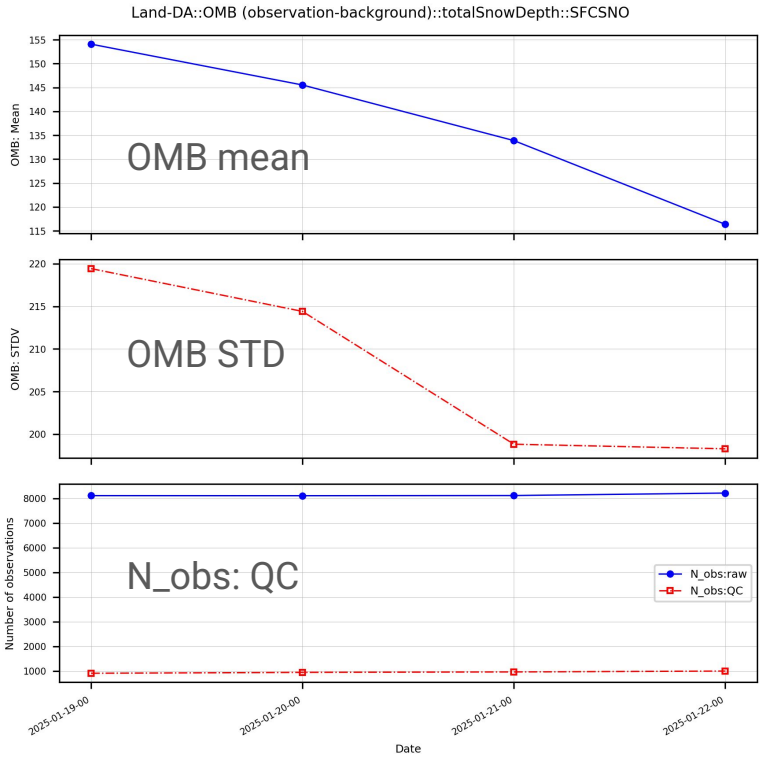
Time-History Plots

- Example: `log_dir/analysis_20250122.log`

```
OOPS_STATS Variational end                          - Runtime:    80.86 sec,  Local Memory:     2.37 Gb
ims_snow: save database to diags/diag.ims_snow_2025012200.nc (io pool size: 4)
sfcsno: save database to diags/diag.sfcsno_2025012200.nc (io pool size: 4)
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: 5 missing values.
QC ims_snow totalSnowDepth: 3347 out of bounds.
QC ims_snow totalSnowDepth: 319 black-listed.
QC ims_snow totalSnowDepth: 66 rejected by first-guess check.
QC ims_snow totalSnowDepth: 8517 passed out of 12254 observations.
QC sfcsno totalSnowDepth: 17 out of bounds.
QC sfcsno totalSnowDepth: 71 out of domain of use.
QC sfcsno totalSnowDepth: 16 black-listed.
QC sfcsno totalSnowDepth: 6329 removed by thinning.
QC sfcsno totalSnowDepth: 214 rejected by first-guess check.
QC sfcsno totalSnowDepth: 575 rejected by difference check.
QC sfcsno totalSnowDepth: 998 passed out of 8220 observations.
QC sfcsno totalSnowDepth: 214 assigned the background_check diagnostic flag.
QC sfcsno totalSnowDepth: 0 assigned the buddy_check diagnostic flag.
QC sfcsno totalSnowDepth: 575 assigned the elevation_bkgdiff diagnostic flag.
QC sfcsno totalSnowDepth: 1 assigned the invalid_elevation diagnostic flag.
QC sfcsno totalSnowDepth: 17 assigned the invalid_snowdepth diagnostic flag.
QC sfcsno totalSnowDepth: 27 assigned the land_check diagnostic flag.
QC sfcsno totalSnowDepth: 10 assigned the landice_check diagnostic flag.
QC sfcsno totalSnowDepth: 43 assigned the missing_elevation diagnostic flag.
QC sfcsno totalSnowDepth: 0 assigned the missing_snowdepth diagnostic flag.
QC sfcsno totalSnowDepth: 6 assigned the rejectlist diagnostic flag.
QC sfcsno totalSnowDepth: 6329 assigned the temporal_thinning diagnostic flag.
```

cadre1: IMS + SFCSNO





https://epic.noaa.gov/

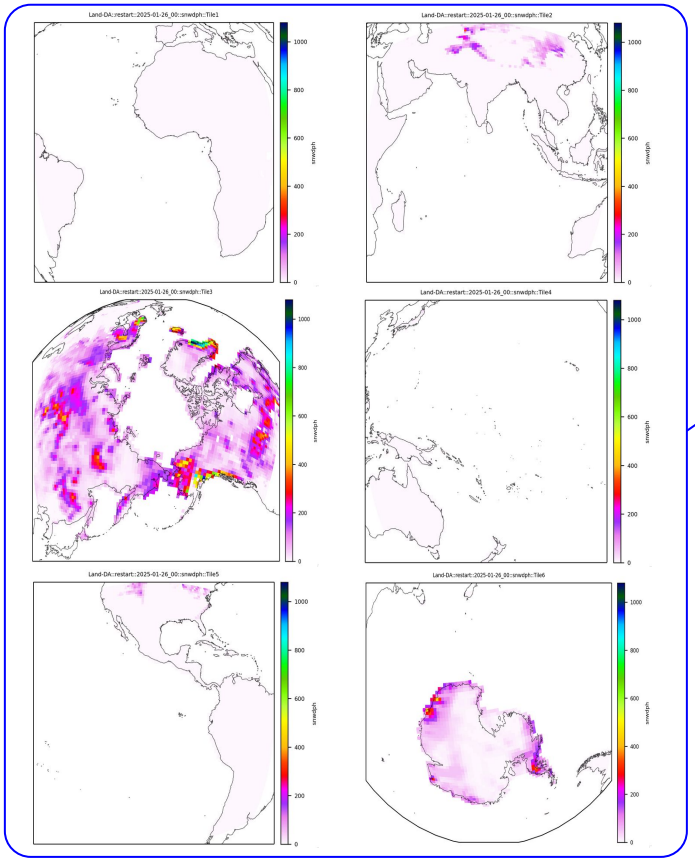
Post-processing: ‘plot_stats’

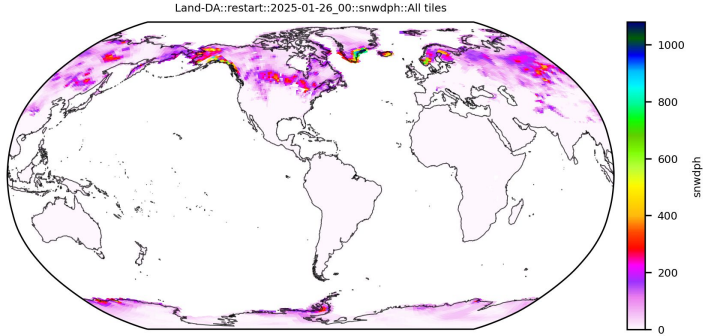
- Result file of JEDI: ‘diag.{obs_type}_{cycle_date}.nc’

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)**	○	○

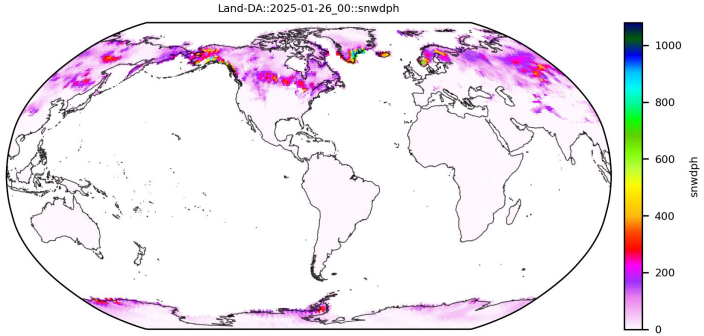
Post-processing: Plots of Restart Files

- 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/