

Script for GraphCast demonstration (full anaconda install version)

ATM419/563 Fall 2024

This demonstration involves the Google GraphCast machine learning weather prediction model, operated through ECMWF's ai-models front-end.

THIS VERSION IS AVAILABLE IN CASE THE STREAMLINED VERSION FAILS.

* ----- preliminaries (only done once) ----- *

- move to your atm419 lab space

- Install Anaconda Python into your lab scape

```
$ wget https://repo.anaconda.com/archive/Anaconda3-2023.09-0-Linux-x86_64.sh
```

```
$ bash Anaconda3-2023.09-0-Linux-x86_64.sh
```

→ hit ENTER, then space bar repeatedly to scroll through license, finally type **yes**

→ when asked, specify install directory as

/network/rit/lab/atm419lab/fall2024/**yournetid**/anaconda3

→ don't hit ^C when you see "done" but it appears to hang. Just hang on...

→ Do you wish to update your shell profile to automatically initialize conda?

no

- activate your new conda, create and populate a new environment "ai"

```
$ source /network/rit/lab/atm419lab/fall2024/yournetid/anaconda3/bin/activate
```

```
$ conda create -n GCTEST
```

```
$ conda activate GCTEST
```

```
$ conda install python=3.10
```

```
$ pip install ai-models
```

```
$ pip install ai-models-graphcast
```

```
$ pip install flax==0.8.2
```

```
$ pip install jax==0.4.23
```

```
$ pip install jaxlib==0.4.23
```

- make a new directory and move into it

```
$ mkdir GRAPHCAST
```

```
$ cd GRAPHCAST
```

- get the graphcast model. Adding ".git" avoid github logon request some encountered

```
$ git clone https://github.com/ecmwf-lab/ai-models-graphcast.git
```

- This creates the new directory **ai-models-graphcast**. Move into it

```
$ cd ai-models-graphcast
```

- we are installing the non-gpu version

```
$ pip install -r requirements.txt
```

```

47
48 • reiterate the versions of jax and jaxlib we need
49 $ pip install jax==0.4.23
50 $ pip install jaxlib==0.4.23
51
52 • copy model data, including training data
53 cp -rf /network/rit/lab/atm419lab/GRAPHCAST/params .
54 cp -rf /network/rit/lab/atm419lab/GRAPHCAST/stats .
55
56 • copy the submit script
57 cp /network/rit/lab/atm419lab/GRAPHCAST/submit_graphcast .
58
59 • edit this line in the submit_graphcast script
60 source /network/rit/lab/atm419lab/fall2024/yournetid/anaconda3/bin/activate
61
62 * ----- running the model ----- *
63 • you are in your $LAB/fall2024/yournetid/GRAPHCAST/ai-models-graphcast
64 directory
65
66 • submit the batch job
67 sbatch -p burst-daes submit_graphcast
68
69 • outputs with names like “out-XX.grib” will be created. The out-0.grib file will
70 appear first, and the rest of the files all at once when the model finishes
71
72 • Inside the submit_graphcast script, this line runs the model
73 ai-models --input file --file
74 /network/rit/lab/atm419lab/GRAPHCAST/ERA5_2021123012.grib --date
75 20211230 --time 12 --path 'out-{step}.grib' --lead-time 18 graphcast
76
77 → The file ERA5_2021123012.grib contains a subset of ERA5 reanalysis fields from
78 06 and 12 UTC 12/30/2021. GraphCast requires two initialization times 6 h apart
79 → We specify the start time as 20211230 at 12 UTC
80 → We specify the run time as 18 hours
81
82 • Model run takes about 10 min (absent resource competition)
83 • tail -f sbatch.err.out ... will see some complaints like this. Ignore them
84
85 2024-04-24 10:49:30,062 INFO Unable to initialize backend 'cuda':
86 2024-04-24 10:49:30,062 INFO Unable to initialize backend 'rocm': module
87 'jaxlib.xla_extension' has no attribute 'GpuAllocatorConfig'
88 2024-04-24 10:49:30,063 INFO Unable to initialize backend 'tpu': INTERNAL: Failed
89 to open libtpu.so: libtpu.so: cannot open shared object file: No such file or directory
90
91
92

```

93

94 • When you see something like the below, the model is wrapping up

95

96 2024-04-24 10:56:35,856 INFO Doing full rollout prediction in JAX: 7 minutes 5

97 seconds.

98 2024-04-24 10:56:35,856 INFO Converting output xarray to GRIB and saving

99

100 • At the end of the model run, the output files are created, all at once

101 → outputs will be out-0.grib, out-12.grib, out-18.grib, out-6.grib

102 → they are GRIB2 files

103

104 * ----- **analyzing the output** ----- *

105 • we will combine the model outputs into a single GRIB2 file “combined.grib2”

106

107 \$ grib_copy out-0.grib out-6.grib out-12.grib out-18.grib combined.grib2

108

109 • copy the notebook

110 \$ cp \$LAB/GRAPHCAST/GRIB_plot_GC.ipynb .

111

112 • launch ARCC Jupyterlab. Batch or snow. Minimal resources suffice.

113 • execute notebook

114

115 • If you cannot run the model for some reason, you can edit the notebook to point to

116 the pre-made output file:

117 /network/rit/lab/atm419lab/GRAPHCAST/combined.grib2

118

119 * ----- **running the model in the future** ----- *

120

121 • **To run the model in the future, this is all the setup you need**

122 source /network/rit/lab/atm419lab/fall2024/**yournetid**/anaconda3/bin/activate

123 conda activate ai

124 → go to your GRAPHCAST/ai-models-graphcast folder and proceed

125

126

127 * ----- **next steps (on your own)** ----- *

128 (1) you need to establish an account on ECMWF CDS

129 <https://cds.climate.copernicus.eu/how-to-api>

130 (2) then you can run ai-models-graphcast fetching ERA5 data from CDS

131 → this can be **very, very slow** to retrieve data from CDS

132 → Edit submit_graphcast to alter the “ai-models” line

133 ai-models --input cds --date 20211230 --time 12 --path 'out-{step}.grib' --

134 lead-time 18 graphcast &

135 (3) get it working on GPU instead so it runs very fast!!!

136

137 {Coureurs des GraphCast}