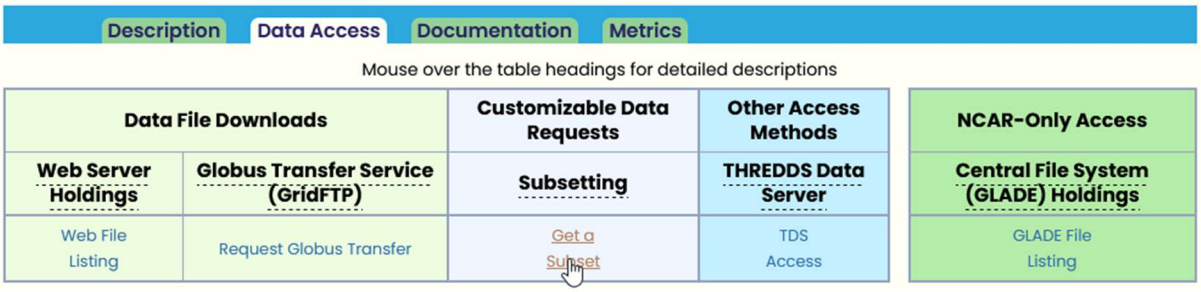
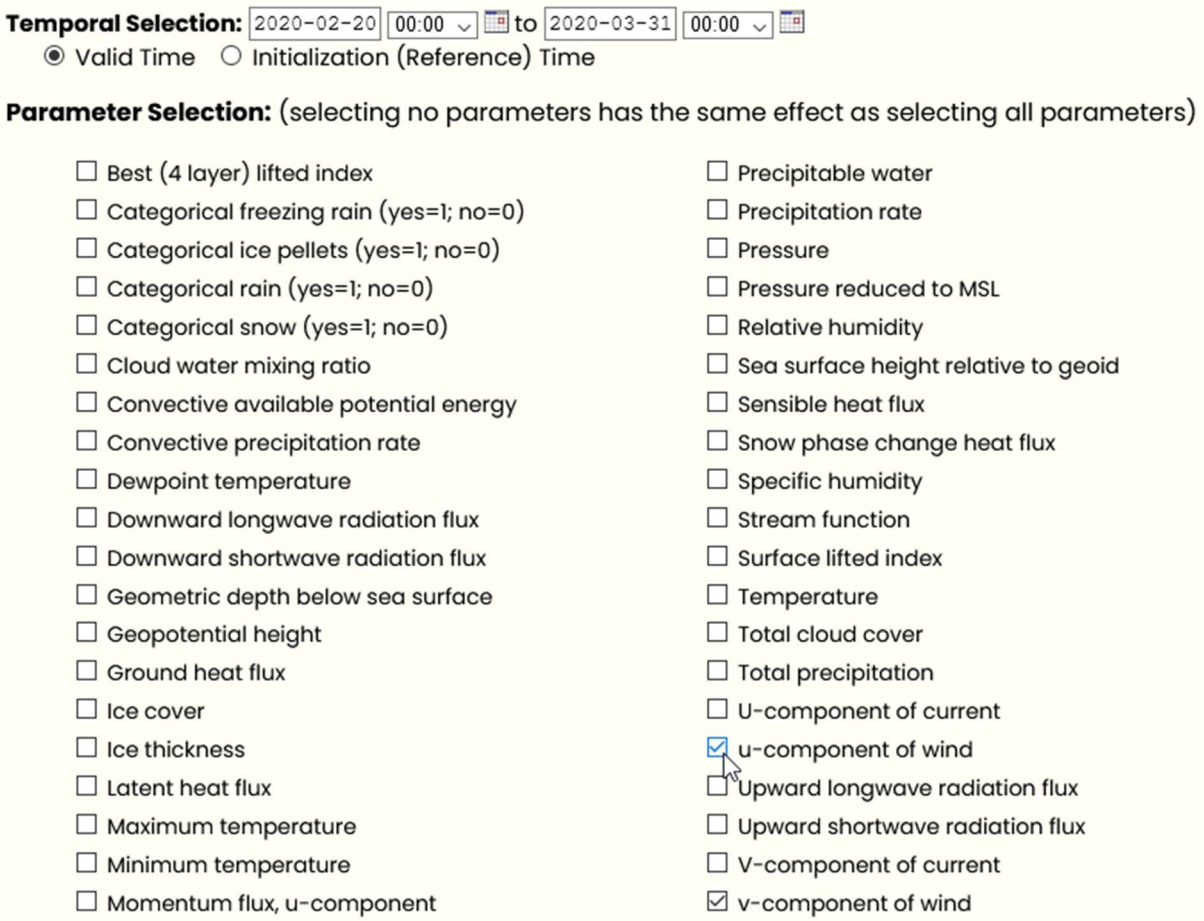
Download and convert wind data archive for running WW3 model

1. CFSR products

- go to https://rda.ucar.edu/datasets/ds094.1/index.html#!access
- select “Get a Subset”



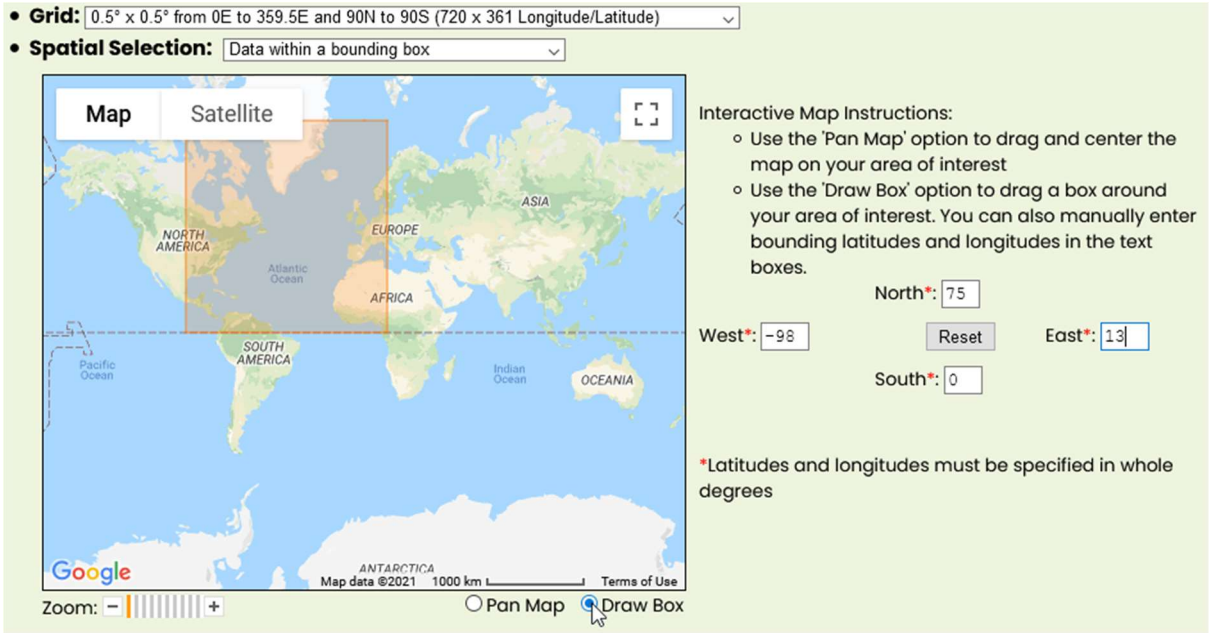
- select the desired time interval and the northward (v) and eastward (u) components of the windfield:



Then you may want to:

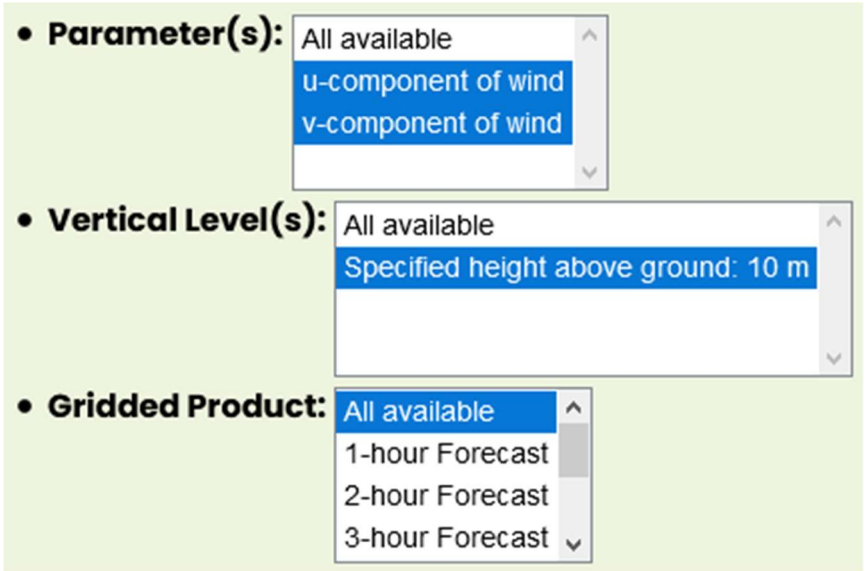
- download the data directly in netcdf (instead of .grib2):

- and select the data within a box:



The required data are:

- wind components 10 m above ground
- for 1h forecast to 6h forecast



Then you may want to:

- compress your monthly netcdf files
- and group them into a single archive file

Submit Your Request:

According to the selections that you have made, your subset request matches **123 RDA data files**. You will receive this number of files, except that these files will only contain data records that match your selections. The total uncompressed volume is estimated at **261.86 GBytes**. Please make sure that you can handle this volume of data before you submit your request or choose a compression option below .

Data Compression Options:

○none ◉gzip (.gz) ○bzip2 (.bz2) ○Unix (.Z) ○Zip (.zip)

File Combination Options:

If you choose an option, many smaller files will be combined into fewer larger files using the method you choose. This will reduce the number of files that you will need to download from our server, but you will need to be able to separate the files on your end.

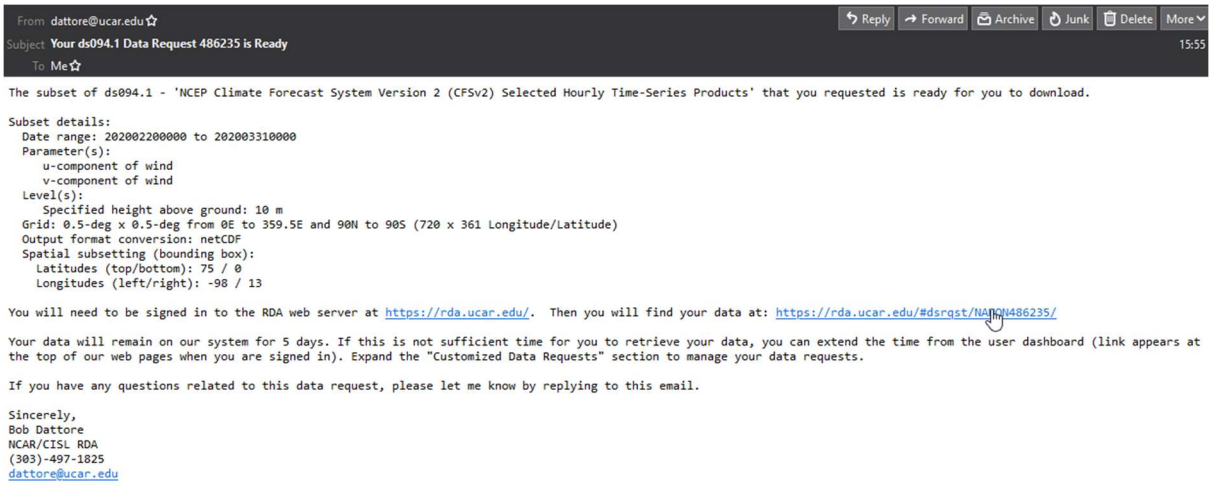
○none ◉Unix tar

Download Method:

◉Web download
○Globus transfer: You will be prompted to select a destination endpoint after submitting your request. The Globus transfer will be submitted automatically when your data are ready.

Submit Request | Start Over

After your data was processed you will receive an email with a link for downloading the data:



From dattore@ucar.edu ☆ | Reply | Forward | Archive | Junk | Delete | More

Subject **Your ds094.1 Data Request 486235 is Ready** 15:55

To Me ☆

The subset of ds094.1 - 'NCEP Climate Forecast System Version 2 (CFSv2) Selected Hourly Time-Series Products' that you requested is ready for you to download.

Subset details:
Date range: 202002200000 to 202003310000
Parameter(s):
u-component of wind
v-component of wind
Level(s):
Specified height above ground: 10 m
Grid: 0.5-deg x 0.5-deg from 0E to 359.5E and 90N to 90S (720 x 361 Longitude/Latitude)
Output format conversion: netCDF
Spatial subsetting (bounding box):
Latitudes (top/bottom): 75 / 0
Longitudes (left/right): -98 / 13

You will need to be signed in to the RDA web server at https://rda.ucar.edu/. Then you will find your data at: https://rda.ucar.edu/#dsrqst/NAMN486235/

Your data will remain on our system for 5 days. If this is not sufficient time for you to retrieve your data, you can extend the time from the user dashboard (link appears at the top of our web pages when you are signed in). Expand the "Customized Data Requests" section to manage your data requests.

If you have any questions related to this data request, please let me know by replying to this email.

Sincerely,
Bob Dattore
NCAR/CISL RDA
(303)-497-1825
dattore@ucar.edu

The the archive file will contain monthly netcdf files :
[anahon@centaurus 2019_JFM]$ **tar xvf** wnd10mx0.5.cdas1.201812-201903.grb2.nc.tar
[anahon@centaurus 2019_JFM]$ **gzip -d** wnd10mx0.5.cdas1.201*
> wnd10mx0.5.cdas1.201812.grb2.nc
> wnd10mx0.5.cdas1.201901.grb2.nc
> wnd10mx0.5.cdas1.201902.grb2.nc
> wnd10mx0.5.cdas1.201903.grb2.nc

then the following python code should help to read the time and required variables for ww3

```
for n, file in enumerate(files):
        print(file)
        nc = netcdf4.Dataset(file)
        t_ref =  np.datetime64('1900-01-01T00:00:00')
        valid_date_str = ["".join(nc.variables['valid_date_time'][i].astype(str).tolist()) for
                          i in range(len(nc.variables['valid_date_time']))]
```

```
        valid_date_time = np.array([datetime.strptime(d, '%Y%m%d%H') for d in
                                    valid_date_str], dtype='datetime64[s]')
        valid_date_dt = valid_date_time - t_ref
        time = valid_date_dt.astype(float)/24/60/60
        print(time[0])
        if n == 0:
                idate = np.array([1900, 1, 1, 0])
                lat_vec = nc.variables["lat"][:]
                lon_vec = nc.variables["lon"][:]
        u = nc.variables["U_GRD_L103"][:]
        v = nc.variables["V_GRD_L103"][:]
        lon = lon_vec
        lat = lat_vec
        for i in range(len(lat_vec)-1):
                lon = np.append(lon, lon_vec)
        lon = np.reshape(lon, (len(lat_vec), len(lon_vec)))

        for i in range(len(lon_vec)-1):
                lat = np.append(lat, lat_vec)
        lat = np.reshape(lat, (len(lon_vec), len(lat_vec)))
        lat = np.transpose(lat)
```

2. GFS products

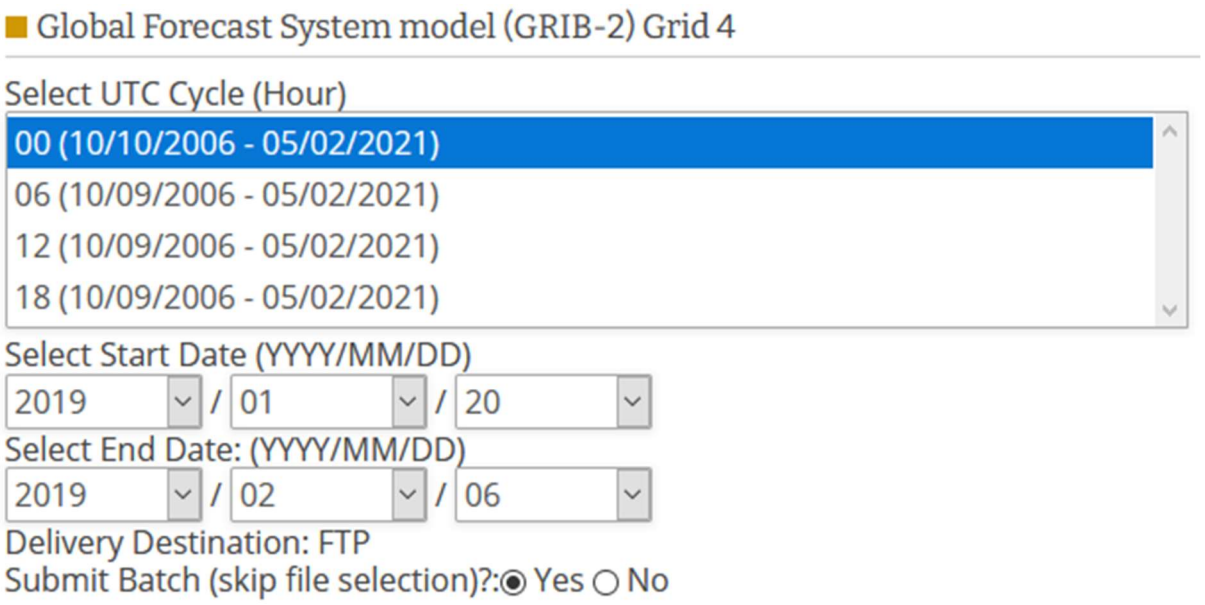
The GFS forecast and analyse archive can be access here:

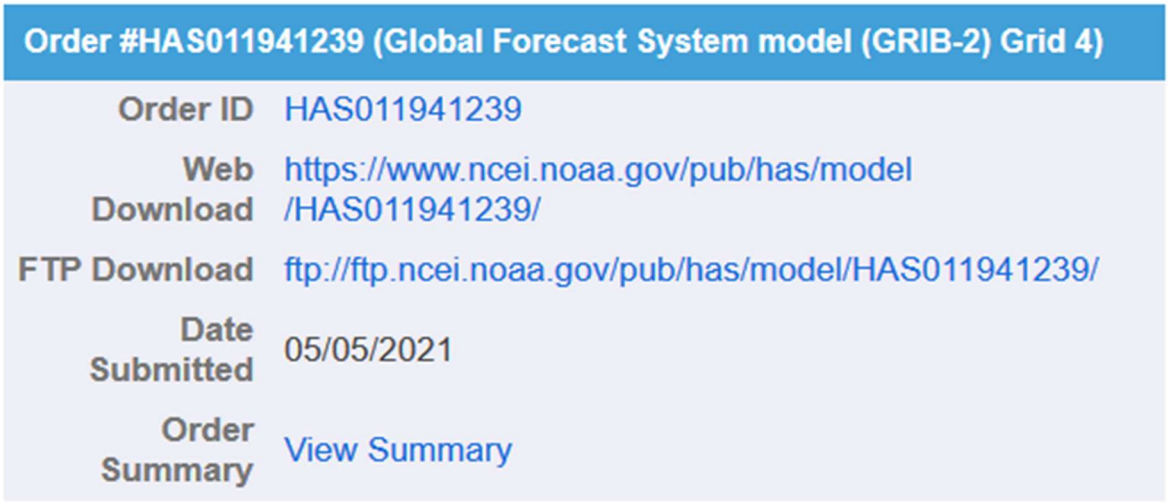
https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs

GFS Forecasts

Model	Grid/Scale	Period of Record	Model Cycle	Output Timestep	Data Access Links
GFS	004 (0.5°) - Domain	18May2020–Present online, since 10Oct2006 in archive	4/day: 00, 06, 12, 18UTC	3-hourly, +000 to +192 hours	HTTPS TDS AIRS

For the ~last 12 months, the data is accessible through an ftp server (HTTPS option). The data is splitted into model output timestep (~90MB). Older data is only accessible through the AIRS archive. It is necessary to create a request to access the archive.





Files in the archive are available in grib2 format with a single file for each of the four daily runs of the model. Each file contains ALL atmospheric variables and for the ENTIRE globe, which represents ~6GB per file.

*grib2 format can be process efficiently with the NOAA wgrid2 software, accessible here: https://www.cpc.ncep.noaa.gov/products/wesley/wgrib2/index.html

wgrib2.exe allows the user to efficiently split the downloaded archive, select the desired model timesteps, variables, and geographical box. Then it is possible to concatenate various files into a single grib2 archive and to convert it into netcdf file.

So to download the wind data for running the equivalent of a 24h OPENCoastS forecast, the following script was written which accept in argument the ftp directory accessible after the user has created a request to AIRS. As in the above figures, the script assumes the request only consists of the '00' cycle, i.e., model run at 00UTC.
curl_wgrib2.sh:

```
rm temp*file *.nc
rm *.grb2 *.tar
rm file_list
curl --output file_list --list-only $1

counter=0
while read f
do
   (( counter++ ))
   echo $counter
   echo $1$f
   file_to_download="curl -O $1$f"
   eval ${file_to_download}
   tar_archive="tar -xvf $f"
   eval ${tar_archive}
   rm *.tar
   rm temp_1_file
   for g in *.grb2
   do
            echo $g
            ../wgrid2/v3.0.2/wgrib2.exe $g -match '(:(UGRD|VGRD):10 m above ground:|:PRMSL:mean sea
level:|:TMP:surface:)' -match ':(anl|(3|6|9|12|15|18|21) hour fcst):' -append -new_grid_winds earth -new_grid latlon -
98:223:0.5 0:151:0.5 temp_1_file
   done
   ../wgrid2/v3.0.2/wgrib2.exe temp_1_file -append -new_grid_winds earth -new_grid latlon -98:223:0.5 0:151:0.5
temp_2_file
   rm *.grb2
done <./file_list
../wgrid2/v3.0.2/wgrib2.exe temp_2_file -netcdf file.nc
```

For this particular request the script was run as follow:
./curl_wgrib2.sh ftp://ftp.ncei.noaa.gov/pub/has/model/HAS011911356/

It was tested on windows, in a Git_Bash shell, after having previously downloaded the version 3.0.2 of wgrid2 (**../wgrid2/v3.0.2/wgrib2.exe**).

The created file.nc should then be slightly modified to be used for ww3, as the time is given in seconds after 01-01-1970, when ww3 expects days after 01-01-1900. Also, the lat and lon vectors were transformed into matrix as done in the python script for the CFSR data.