

NMME Forecast Model Updates in the IRI Data Library

Michael Bell

May 19, 2020

The North American Multi-model Ensemble (NMME) seasonal forecasts, which serve as inputs to the IRI seasonal forecasts, are updated on a monthly basis in the IRI Data Library. Currently the NMME includes updated forecast output from the following suite of eight models from North American modeling centers:

CMC-Canada: GEM-NEMO, CanCM4i (10 ensemble members each from CanSIPSv2; ensemble members 1-10 from CanSIPSv2 are GEM-NEMO, 11-20 from CanSIPSv2 are CanCM4i)

COLA-RSMAS: COLA-RSMAS-CCSM4

GFDL: GFDL-CM2.1-aer04, GFDL-CM2.5-FLOR-A06, GFDL-CM2.5-FLOR-B01

(According to Tom Delworth at GFDL, tom.delworth@noaa.gov, the GFDL CM2.1 and FLOR NMME models will be replaced by a new model named SPEAR by January 2021)

NASA: NASA-GEOSS2S

NOAA/NCEP: NCEP-CFSv2

More information about the NMME project and its component models can be found at the following NOAA/NCEP/CPC web page:

https://www.cpc.ncep.noaa.gov/products/NMME/users_guide.html

The modeling centers that produce the NMME forecasts attempt to distribute new forecasts by the 8th day of each month. However, the modeling centers occasionally encounter difficulties that may delay the release of new forecast data past the 8th. The scripts to update these data in the Data Library are run by the datag cron on at least a daily basis near the beginning of each calendar month.

NMME Data Entries in the IRI Data Library

In the IRI Data Library datasets for these models are found at the following URL:

<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/>

Although many more models or previous versions of NMME models are accessible at this URL, only the eight models outlined above are current and are being updated monthly.

Depending upon how each model's output is structured and served from each model source's distribution site, the hindcast and forecast for each model may be separate sub-datasets, or may be in a single subdataset entry continuous along a single start time grid (S). The hindcast period across the models is generally 1982 (or 1981) to 2010. The hindcast period for the CFSv2 model extends slightly into 2011 (through the 01 March 2011 start; with forecasts starting from the 01 April 2011 start). In all cases the hindcasts are fixed, not on-the-fly (i.e., none of the NMME models has a new hindcast produced for every new monthly forecast start). In the case of all currently-updated NMME models, except for the NASA-GEOSS2S, whose earliest forecast start is 01 November 2017, the hindcast and forecast datasets are actually continuous across the start time grid, even if the model has separate hindcast and forecast sub-datasets.

Updating NMME Forecasts in the IRI Data Library

A set of perl scripts update the NMME forecast data in the Data Library. The following cron entries owned by datag on gfs2mon1 run the perl scripts that update the NMME forecasts:

```
50      9      5-15    *      *      /home/datag/cronjobs/getCOLA_CCSM4p0.pl
30     10      1-6    *      *      /home/datag/cronjobs/getCanSIPSV2.pl
30     17      5-15    *      *      /home/datag/cronjobs/getGFDL_CM2p1_aer04.pl
35     17      5-15    *      *      /home/datag/cronjobs/getGFDL_A06_FLOR.pl
40     17      5-15    *      *      /home/datag/cronjobs/getGFDL_B01_FLOR.pl
45     17      5-15    *      *      /Data/data23/NMME/NASA_GEOSS2S.D/Forecast.D/getNASA_GEOSS2S.pl
```

The getCanSIPSV2.pl script updates the CanSIPSV2 dataset entry from which the GEM-NEMO and CanCM4i entries sample ensemble members 1-10 and 11-20, respectively, to form their own entries.

The CFSv2 NMME forecast is a bit different in that the NMME entry samples data from the full CFSv2 dataset in the Data Library.

The full CFSv2 6-hourly forecast dataset (six-hourly starts every day of monthly mean forecast values) is updated using the following script run by datag's cron on gfs2mon1:

```
32 6,14,23 * * * /home/datag/cronjobs/getdata_cfsv2_gfs2.pl
```

Additional Information for Each NMME Model

CanSIPSV2 (GEM-NEMO, CanCM4i)

These forecast models are updated by the script
`/home/datag/cronjobs/getCanSIPSV2.pl`

The primary contact person at CMC for this model is Bill Merryfield, bill.merryfield@canada.ca . Other contact persons at CMC for this model include Benoit Archambault, benoit.archambault2@canada.ca , Manuel Ferreira, manuel.ferreira@canada.ca , and Nicole Bois, nicole.bois@canada.ca .

This model's forecasts are initialized and released monthly, typically on the first day of each month, or the last day of the previous month.

The data files are downloaded from the following remote http location at CMC:

<http://collaboration.cmc.ec.gc.ca/cmc/cmoi/GRIB/NMME/1p0deg/>

As of August 2019, CanSIPSV2, GEM-NEMO, and CanCM4i have taken the place of CMC's previous NMME models, Cansips, CMC1-CanCM3, and CMC2-CanCM4, which were discontinued after the July 2019 start.

Every month there is a separate file for each variable, each containing all 20 ensemble members. The data files are downloaded locally on gfs2mon1 to
`/Data/data23/NMME/CanSIPSV2.D/Forecast.D/`

The CanSIPSV2 forecast dataset in the Data Library is here:

<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/.CanSIPSV2/.FORECAST/.MONTHLY/>

However, the datasets most relevant to users are the GEM-NEMO and CanCM4i datasets, which each draw their 10 ensemble members from the 20-member CanSIPSV2 forecast entry (CanSIPSV2 members 1-10 for GEM-NEMO and CanSIPSV2 members 11-20 for CanCM4i).

The CanSIPSV2 files we download include both an initialization date at the very beginning of the file name, which can actually be a day in the previous month, and the NMME start month and year later in the file name.

The GEM-NEMO dataset entry is placed here:

<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/.GEM-NEMO/.FORECAST/.MONTHLY/Y/>

The CanCM4i dataset entry is placed here:

<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/.CanCM4i/.FORECAST/.MONTHLY/>

COLA-RSMAS CCSM4

This forecast model is updated by the script
`/home/datag/cronjobs/getCOLA_CCSM4p0.pl`

The primary contact person for this model at RSMAS is Dughong Min, dmin@rsmas.miami.edu .
The secondary contact person at RSMAS is Ben Kirtman, bkirtman@rsmas.miami.edu .

This model's forecasts are initialized and released monthly.

The data files are downloaded from the following remote ftp location at RSMAS:

<ftp://decadal.rsmas.miami.edu/NMME/>

with month-of-year subdirectories containing data files for each variable and each of 10 ensemble members.

The data files are downloaded locally on gfs2mon1 to
`/Data/data23/NMME/CCSM4.0.D/`
with month-of-year subdirectories holding the files for each start month but for all years.

The COLA-RSMAS-CCSM4 forecast dataset in the Data Library is here:

<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/.COLA-RSMAS-CCSM4/.MONTHLY/>

GFDL SPEAR

The GFDL **S**eamless **P**rediction **E**arth **R**esearch System is slated to replace the GFDL CM2.1 and FLOR NMME models by January 2021, though forecasts are expected to

become available by the summer or fall of 2020 even as the CM2.1 and FLOR forecasts continue to be released, according to Tom Delworth at GFDL.

A journal article documenting this model has been published in the AGU's Journal of Advances in Modeling Earth Systems (JAMES), and can be found here:

<https://doi.org/10.1029/2019MS001895>

The primary contact person for this model at GFDL is Tom Delworth, tom.delworth@noaa.gov.

GFDL CM2.1-aer04

This forecast model is updated by the script
`/home/datag/cronjobs/getGFDL_CM2p1_aer04.pl`

The primary contact person for this model at GFDL is Tom Delworth, tom.delworth@noaa.gov.

This model's forecasts are initialized and released monthly.

The data files are downloaded from the following remote ftp locations at GFDL:

[ftp://ftp.gfdl.noaa.gov/pub/Oar.Gfdl.Nmme/Vandendool/\\$month_retrospective_v3.1_aer04/](ftp://ftp.gfdl.noaa.gov/pub/Oar.Gfdl.Nmme/Vandendool/$month_retrospective_v3.1_aer04/)

and

[ftp://ftp.gfdl.noaa.gov/pub/Oar.Gfdl.Nmme/Vandendool/\\$month_retrospective_v3.1_aer04/additional_variables/](ftp://ftp.gfdl.noaa.gov/pub/Oar.Gfdl.Nmme/Vandendool/$month_retrospective_v3.1_aer04/additional_variables/)

where `$month` is the lowercase full name of the start month, such as

`august_retrospective_v3.1_aer04/`

These month-of-year subdirectories contain data files for each variable and each of 10 ensemble members. Please note that GFDL removes data files from subdirectories of previous start months very quickly.

The data files are downloaded locally on gfs2mon1 to
`/Data/data23/NMME/GFDL_CM2.1_v3.1_aer04.D/`

with month-of-year subdirectories holding the files for each start month but for all years.

The GFDL CM2.1-aer04 forecast dataset in the Data Library is here:

<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/.GFDL-CM2p1-aer04/.MONTHLY/>

GFDL CM2.5-FLOR-A06

This forecast model is updated by the script

/home/datag/cronjobs/getGFDL_A06_FLOR.pl

The primary contact person for this model at GFDL is Tom Delworth, tom.delworth@noaa.gov.

This model's forecasts are initialized and released monthly.

The data files are downloaded from the following remote ftp locations at GFDL:

[ftp://ftp.gfdl.noaa.gov/pub/Oar.Gfdl.Nmme/Vandendool/\\$month_retrospective_A06_FLOR/](ftp://ftp.gfdl.noaa.gov/pub/Oar.Gfdl.Nmme/Vandendool/$month_retrospective_A06_FLOR/)

and

[ftp://ftp.gfdl.noaa.gov/pub/Oar.Gfdl.Nmme/Vandendool/\\$month_retrospective_A06_FLOR/additional_variables/](ftp://ftp.gfdl.noaa.gov/pub/Oar.Gfdl.Nmme/Vandendool/$month_retrospective_A06_FLOR/additional_variables/)

where *\$month* is the lowercase full name of the start month, such as

august_retrospective_A06_FLOR/

These month-of-year subdirectories contain data files for each variable and each of 12 ensemble members. Please note that GFDL removes data files from subdirectories of previous start months very quickly.

The data files are downloaded locally on gfs2mon1 to
/Data/data23/NMME/GFDL_CM2.5_FLOR_A06.D/

with month-of-year subdirectories holding the files for each start month but for all years.

The GFDL CM2.5-FLOR-A06 forecast dataset in the Data Library is here:

<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/.GFDL-CM2p5-FLOR-A06/.MONTHLY/>

GFDL CM2.5-FLOR-B01

This forecast model is updated by the script
`/home/datag/cronjobs/getGFDL_B01_FLOR.pl`

The primary contact person for this model at GFDL is Tom Delworth, tom.delworth@noaa.gov .

This model's forecasts are initialized and released monthly.

The data files are downloaded from the following remote ftp locations at GFDL:

`ftp://ftp.gfdl.noaa.gov/pub/Oar.Gfdl.Nmme/Vandendool/$month_retrospective_B01_FLOR/`

and

`ftp://ftp.gfdl.noaa.gov/pub/Oar.Gfdl.Nmme/Vandendool/$month_retrospective_B01_FLOR/additional_variables/`

where `$month` is the lowercase full name of the start month, such as

`august_retrospective_B01_FLOR/`

These month-of-year subdirectories contain data files for each variable and each of 12 ensemble members. Please note that GFDL removes data files from subdirectories of previous start months very quickly.

The data files are downloaded locally on gfs2mon1 to
`/Data/data23/NMME/GFDL_CM2.5_FLOR_B01.D/`

with month-of-year subdirectories holding the files for each start month but for all years.

The GFDL CM2.5-FLOR-B01 forecast dataset in the Data Library is here:
<http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/.GFDL-CM2p5-FLOR-B01/.MONTHLY/>

NASA-GEOSS2S

This forecast model is updated by the script

`/Data/data23/NMME/NASA_GEOSS2S.D/Forecast.D/getNASA_GEOSS2S.pl`

The primary contact person for this model at NASA is Kazumi Nakada,
kazumi.nakada@nasa.gov.

More information about this model can be found in the following documentation file:

`/Data/data23/NMME/NASA_GEOSS2S.D/Forecast.D/dataset_documentation.txt`

This model's forecasts are initialized and released monthly.

The data files are downloaded from the following remote https location at NASA:

`https://gmao.gsfc.nasa.gov/gmaoftp/gmaofcst/seasonal/GEOSS2S-2_1/NMME/forecast/monthly/YYYY/mmm/`

Where `YYYY` is the 4-digit year of the NMME forecast start month, and `mmm` is the lowercase, 3-letter English abbreviation of the NMME forecast start month, for example:

`https://gmao.gsfc.nasa.gov/gmaoftp/gmaofcst/seasonal/GEOSS2S-2_1/NMME/forecast/monthly/2018/apr/`

These subdirectories contain data files for each variable and each of 10 ensemble members. The files representing different ensemble members are labelled according to the forecast initialization date and the ensemble member. Generally, there are 7 ensemble members (4-10) associated with initializations on the last day of the month before the NMME forecast start date (00Z on the first day of the month), and 3 ensemble members (1-3) associated with initializations every 5 previous days (e.g., the 15th, 20th, and 25th days of the month previous to the NMME forecast start date).

The update script uses the last initialization dates in the forecast file names that are downloaded (dated with the last day of the month previous to the NMME forecast start date) to determine the month and year of the NMME forecast start, and writes those to the file
`/Data/data23/NMME/NASA_GEOSS2S.D/Forecast.D/endmonth.txt`

This “endmonth.txt” file is read by the index.tex file for this dataset to specify the end date of the start time grid “S” for this model’s forecast.

Please note that the earliest available NMME forecasts for this model are from the November 2017 start date. Unlike the other NMME models which have hindcasts ending in 2010, this model includes hindcasts from 1981 to January 2017. However, this also means that there is a gap of several months between the end of the hindcast period and the start of the forecast period.

The data files are downloaded locally on gfs2mon1 to
`/Data/data23/NMME/NASA_GEOSS2S.D/Forecast.D/`

with month-of-year subdirectories holding the files for each start month, but for all years.

The NASA-GEOSS2S forecast dataset in the Data Library is here:

<http://iridl.ideo.columbia.edu/SOURCES/.Models/.NMME/.NASA-GEOSS2S/.FORECAST/.MONTHLY/>

NOAA/NCEP CFSv2

As mentioned above, the CFSv2 NMME forecast is a bit different from the other models in that the NMME entry samples data from the full CFSv2 dataset in the Data Library.

The full CFSv2 6-hourly forecast dataset (six-hourly starts every day of monthly mean forecast values) is updated using the following script run by datag's cron on gfs2mon1:

`/home/datag/cronjobs/getdata_cfsv2_gfs2.pl`

The contact for the CFSv2 model at NOAA is cfs@noaa.gov.

More information about the CFSv2 model can be found at the following page on the NCEP website:

<https://cfs.ncep.noaa.gov/>

The update script downloads the CFSv2 files from the following 7-day rotating archive:

<https://nomads.ncep.noaa.gov/pub/data/nccf/com/cfs/prod/cfs/>

If there is a problem that prevents the forecast files from being downloaded from this short-lived archive in a timely manner, the following location at NCEI is a full archive of the forecast dataset, and is kept up to date in near real time:

ftp://nomads.ncdc.noaa.gov/modeldata/cfsv2_forecast_mm_9mon/

The data files are downloaded locally on gfs2mon1 to

/Data/data22/noaa/ncep/cfsv2/fcst/

The IRI Data Library entry for the full CFSv2 forecast dataset is here:

http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.EMC/.CFSv2/.MONTHLY_REALTIME/

There are two CFSv2 NMME forecast entries included within the NMME dataset:

Early Month Samples:

http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/.NCEP-CFSv2/.FORECAST/.EARLY_MONTH_SAMPLES/

Pentad Samples:

http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/.NCEP-CFSv2/.FORECAST/.EARLY_MONTH_SAMPLES/

The short dataset descriptions for each of these sub-datasets explain a little bit about how they differ from the full CFSv2 entry, but not everything. The NMME CFSv2 forecast entries differ from the full CFSv2 entry in the Data Library in three primary ways:

- Only the geopotential height at 200 hPa, 2-meter temperature, SST, and precipitation variables (and forecast validation time T) are included in the NMME entries
- The NMME entries sample subsets of starts from the full set of CFSv2 files.
 - The “Early Month Samples” NMME forecast entry includes 4-times-daily forecast starts from the last day of the month previous to the month of the NMME monthly start and the first seven days of the same month of the NMME monthly start. In the NMME entry, these 32 forecasts show up as 32 ensemble members in the NMME monthly start. For example, for the 01 June 2019 NMME start, the 32 ensemble members come from the 00Z, 06Z, 12Z, and 18Z starts of the 31 May 2019 full CFSv2 forecast and the 00Z, 06Z, 12Z, and 18Z starts of the 01 June 2019 to 07 June 2019 full CFSv2 forecast dataset. The Sampling Table in the Early Month Samples NMME forecast entry shows how the sampling is done:
http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.EMC/.CFSv2/.NMME_R

[EALTIME_ENSEMBLE/.FLXF/mark/sampleS/M/4/STEP/S/exch\[M\]table:/mark:/table/.html](http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.EMC/.CFSv2/.REALTIME_ENSEMBLE/.FLXF/mark/sampleS/M/4/STEP/S/exch[M]table:/mark:/table/.html)

- The “Pentad Samples” NMME forecast entry includes 4-times-daily forecast starts sampled every five days from the full CFSv2 forecast dataset files. These 4-times-daily forecasts become individual ensemble members of the NMME monthly start. Pentad starts between the 8th day of the month previous to the NMME monthly start and the 7th day of the same month as the NMME monthly start are assigned to the NMME monthly start (00Z on the first day of the month). This results in 24 ensemble members for all NMME start months except November, which has 28 ensemble members. The Sampling Table in the Pentad Samples NMME forecast entry shows how the sampling is done:
[http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.EMC/.CFSv2/.REALTIME_ENSEMBLE/.FLXF/mark/sampleS/M/4/STEP/S/exch\[M\]table:/mark:/table/.html](http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.EMC/.CFSv2/.REALTIME_ENSEMBLE/.FLXF/mark/sampleS/M/4/STEP/S/exch[M]table:/mark:/table/.html)
- The variables in the NMME entry are spatially regridded to the 1 deg. lat/lon gridding used by all the NMME models (if necessary, i.e., the ocean and pressure-level variables are already on the correctly-centered 1 deg. grid, though the Y (latitude) coordinate orientation is reversed), and the Y (latitude) coordinate orientation is reversed to run from 90 deg. S to 90 deg. N. Units conversions may also be applied to the variables.

As an intermediate step between the full CFSv2 forecast dataset and the NMME CFSv2 forecast datasets, the full CFSv2 forecasts (all variables) are resampled in the manner explained above as seen in the “Early Month Samples” and “Pentad Samples” NMME forecasts. These intermediate forecast datasets appear as their own sub-datasets as part of the full CFSv2 dataset entry:

Full CFSv2 NMME_REALTIME_ENSEMBLE (Early Month Samples):

http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.EMC/.CFSv2/.NMME_REALTIME_ENSEMBLE/

Full CFSv2 REALTIME_ENSEMBLE (NMME Pentad Samples):

http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.EMC/.CFSv2/.REALTIME_ENSEMBLE/

The CFSv2 NMME forecast entries for the precipitation, 2-meter temperature, SST, and 200 hPa geopotential height make reference to these two sub-sampled versions of the full CFSv2 forecast datasets and apply any required spatial resampling to 1 deg. lat/lon, reversal of Y axis orientation, and units conversions. The CFSv2 NMME forecast precipitation, SST, and 2-meter temperature variables are also cached using “defasvarcachesilent”.

