FV3-GEFS/Sub-seasonal

- Configurations

Yuejian Zhu

Environmental Modeling Center
NCEP/NWS/NOAA

Status Update: July 19th 2018

Major Milestones

- **Q2FY18 -** Prepare FV3-GFS for reanalysis project: Develop and test low-resolution version of FV3-GFS and FV3-GDAS, and configure the model for reanalysis project.

- **Q4FY18 -** Determine ensemble configuration for FV3-GEFS: Configure for optimum ensemble size (# members), resolution, physics, and coupling to Ocean, Ice, Land and Wave models using NEMS/NUOPC mediator; conduct testing for quality assurance and computational efficiency.

- **Q3FY19 -** Produce ~20-year reanalysis datasets: Mainly ESRL/PSD activity. Determine configuration of the reanalysis system; develop observational database for reanalysis; prepare observational inputs; and produce reanalysis suitable for reforecasts and calibration.

- **Q4FY19 -** Produce ~30-year reforecast datasets for FV3-GEFS: Finalize ensemble configuration and produce reforecasts consistent with the reanalysis data; extend the reforecast length to 35 days.

- **Q4FY19** – Produce 2-3 year retrospective forecast for FV3-GEFS: Use the same configuration as real-time, and retrospective FV3GFS/EnKF analysis.

- **Q1FY20 -** Transition FV3-GEFS into operations: Conduct pre-implementation T&E; transition the system for operational implementation. Replace GEFSv11 and stop GEFSv10 (legacy run to support NWC) ???

FV3GEFS Implementation Plan

June 2018

Implementation Plan for FV3-GEFS (FY2017-2020)

FV3GEFS	FY17				FY18				FY19				FY20				
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	% complete
FV3GEFS Reanalysis Development			Develop and test low resolution FV3GFS with FV3GDAS, configure it for reanalysis (ESRL)														100%
FV3GEFS Ensemble Configuration		Configure FV3GFS ensemble resolution, members, physics, coupling to ocean and sea-ice, and extend forecasts to weeks 3&4 (EMC)															95%
FV3GEFS Reanalysis Production							Produce ~20-year reanalysis datasets using FV3GFS/GDAS (ESRL)										15%
FV3GEFS Reforecasts								Finalize FV3GEFS V12 configuration & produce 30-year reforecasts (extended to 35 days) for calibration and validation of HEFS/OWP									0%
FV3GEFS V12 Evaluation										Evaluate FV3GEFS V12 forecast performance out to weeks 3&4			◆				0%
FV3GEFS V12 Implementation													FV3GEFS V12 in operation				0%
Advancement of FV3GEFS							Today						Further advancements of FV3GEFS (GFS/GEFS unification, ensemble based coupled modeling)				

* Proposed changes for GEFS V12: 1) Produce FV3 based reanalysis in FY18 using the same configuration as Q2FY19 FV3GFS (ESRL); 2) Reforecasts will be based on FV3GEFS configured with 2-Tier SST approach; and 3) FV3GEFS Reforecasts extended to 35 days to include weeks 3&4 guidance.

FV3-GEFS V12 Operational configuration – 1

- Resolutions
  - C384 (25km) horizontal resolution and 64 hybrid vertical levels
- Ensemble size
  - **30+1** members per cycle (preferred, depends on resources availability)
- Frequency
  - 0000; 0600; 1200 and 1800UTC
- Forecast lead-time
  - 16 days for every cycle, except for
  - 35 days once per day at 00UTC
    - CPC request to have extended forecasts once per day at 00UTC.
    - Finish runs before coming morning (e.g. 8am), there are 5-hours time window (+5:45 -- +10:45)
- Output
  - 3 hourly at 0.25d out to 10 days (**25% additional CPU resource for hourly output**)
    - **0.25d output will be in-house only, no public access. 0.5d output will be for public access.**
  - 6 hourly at 0.5d from 10 to 35 days.
- Computation cost (mainly 16-day forecast)
  - For un-coupled GEFS only
  - Task for 1 member: ~38 (I/O + Integration)+10 (post)=48 nodes for 16-day forecast within 1-hour (include I/O + forecast + NCEP post + ENS post)
    - **Reference - 3 CPU hour node/per day/per member or 72 CPU hour/per day/per member**
  - Final: 30+1 members: 31x48 =1488 nodes for integration and post within 1-hour real-time operation (WCOSS) – The testing is based on Cray (totally available nodes -2048)
  - Expect to enhance computation efficacy by 20% --- **1190 nodes**
  - **Update: targeting <= 40 nodes (integration+I/O+UPP+SPP+ENSSTAT) for one member**
  - **Note: extra costs for wave ensemble (???), and NGAC (control only: 20 nodes?)**

FV3-GEFS V12 Operational configuration – 2

- Model
  - FV3GEFS (C384L64) with GFDL MP
- Initial perturbations
  - FV3EnKF analysis (early run) or FV3EnKF f06
  - No TS relocation for FV3EnKF
- Stochastic perturbations for atmosphere
  - SKEB, SPPT (5-Scales) and SHUM (0.004)
- Stochastic perturbations for land
  - No stochastic perturbations
- Boundary forcing for weeks 3&4 forecast
  - 2-tiered SST approach
- **Timeline for finalizing configuration – Q4FY18**

FV3-GEFS reforecast configuration

- Model configurations
  - The same as real-time GEFSv12 (C384L64)

- Period of retrospective
  - 30 years (1989 – 2018)
    - 1989 – 1999 (11 years) CFS analysis
    - 2000 – 2018 (19 years) Hybrid FV3 GFS/EnKF reanalysis (ESRL/PSD)
    - **Initial analyses are in-consistent**

- Frequency and ensemble size
  - Configuration: 30 years, initialized at 00UTC for every day; runs 5 members out to 16 days, except for 11 members out to 35 days every 7 days.
    - **Cost - ~ 715 nodes (Cray) for 9 months (7/24) === project to 1 year to finish (consider 25% extra time)**

- Output data
  - Format – GRIB2
  - Frequency and resolution
    - 3 hourly out to 10 days at 0.25 degree resolution
    - 6 hourly beyond 10 days at 0.5 degree resolution
    - **Hourly precipitation – has discussed with NCEP post team, but needs extra effort, no decision yet**
  - Save all variables at above resolution on HPSS for 5-year
  - Save selected variables on disk for CPC, MDL and NWC (depends on HRAC approving)
    - Currently, combined all three centers --- about 75 variables without considering BMOS variables
    - **Will produce 6-hourly and 3-hourly precipitation data alternately as current operation**
    - **Will provide relative humidity instead of specific humidity**
  - ESRL/PSD will convert GRIB format data to NetCDF for public access
  - Note: size of C384 master file for one forecast lead-time at 0.25 degree = **345mb**

FV3-GEFS retrospective runs

- Model configurations
  - The same as real-time GEFSv12 (means dynamics/physics and horizontal/vertical resolution)
- Period of retrospective runs
  - 2 years (plus one summer) to support WPC, SPC, AWC, NHC, CPC and MDL's evaluation
- Initial analysis and perturbations
  - Latest hybrid 4DEnsVar analysis and EnKF analysis (full quality of forecast)
- Frequency and ensemble size (cost)
  - Reference of cost: 72 CPU (3 nodes) hours /per day/per member
  - Configuration: Once per day for 2 years and 1 summer (2x365+183 days) with **31/21** members, out to 16 days*
    - **~310 nodes (Cray) for 6 months (7/24) === project to 8 months to finish**
    - ***35-day forecasts are included in the reforecast experiments**
- Output files
  - **Frequency and resolution**
    - 3 hourly out to 10 days at 0.25 degree resolution
    - 6 hourly beyond 10 days at 0.5 degree resolution
  - Save all variables at above resolution on HPSS
  - Save immediately useful variables on disk for WPC, SPC,AWC, CPC and MDL (pending on disk space we have)
    - Plan to save total 73 variables on disk (52 upper air + 21 surface – **see slides 11&12**)
  - Note: size of C384 master file for one forecast lead-time at 0.25 degree = **345mb**

HPC Resource Request for Reforecast
(original proposal)

- **715** Cray nodes for one year (slide #15)
  - Q4FY18 – Q4FY19 (one year)

- **2270** TB tape storage (HPPS) (slide #16-18)
  - Grib2 format
  - Data will be stored in HPPS for 5 years

- **389** TB disk storage (slide #16-18)
  - Grib2 format
  - To save selected variables on disk for CPC, MDL and NWC to access directly

HPC Resource Request for Reforecast (new/modified plan)

- **320** "Dell" nodes for one year (slide #15 - notes) to guarantee 30 years reforecasts from CPC's requirement
  - Q4FY18 – Q4FY19 (one year)

- **2270** TB tape storage (HPPS) (slide #16-18)
  - Grib2 format
  - Data will be stored in HPPS for 5 years

- **389** TB disk storage (slide #16-18)
  - Grib2 format
  - To save selected variables on disk for CPC, MDL and NWC to access directly

HPC Resource Request for Retrospective Runs (new/modified Plan)

- **310** "Dell" nodes for 9 months (slide #19 - notes) to guarantee two and half years retrospective runs for user evaluation
  - Q1FY19 – Q4FY19 (9 months)

- **670** TB tape storage (HPPS) (slide #19)
  - Grib2 format
  - Data will be stored in HPPS for 1 year

- **55** TB disk storage (slide #19)
  - Grib2 format
  - To save selected variables (0.5 degree) on disk for EMC's performance evaluation

HPCRAC Approval

Recommendation…..

- Current FV3 and GEFS v12 footprints will not fit together on WCOSS
  - Begin GEFS v12 retrospectives as soon as FV3 restrospectives are complete in September
  - GEFS v12 can use the resources on the Dell system that will be vacated by the FV3
  - The Dell no longer needs to be held from the rest of the development community, but GEFS v12 should be given a fence, reservation or very high priority queue in order to guarantee node availability

Support slides (12-19) for GEFS reforecast/retrospective forecast configuration, and resource

Summary from HPCRAC Meeting 6/18/18

- The discussion point was that after all the reforecast data is provided to OWP they will still need to spend some months in building their calibration. They will not be ready by the time FV3GEFS v12 is ready to go in and they will continue using GEFS v10 till there calibrations are ready.
- So the question that came back to us was that we need two solutions.
  - One is having the full reforecast resources timeline (we already have that) and a
  - Second solution which identifies the minimum reforecast resources necessary for the implementation of fv3gefs v12 (with the remaining reforecasts being delivered on a delayed time line).

Proposal 1: Allocate requested resources for completion of 20-year reanalysis and 30-year reforecast experiments by Q4FY19

- Pros:
  - Meet NWS AOP Milestone
  - Complete the project w/CPO-SLA funding (ends in FY19)
  - Prepare for implementation of GEFS V12 in Q2FY20
  - Ramp-up efforts for coupled model development for S2S predictions beginning with GEFSV13.0
  - Work with OWP and CPC and MDL for calibration of products (OWP needs all 30-year reforecasts, and 6-9 months for development)
  - Remove GEFS V10 from operations soon after OWP moves to GEFS V12
- Cons:
  - **1/3rd of Cray (715 nodes) not available for any development***

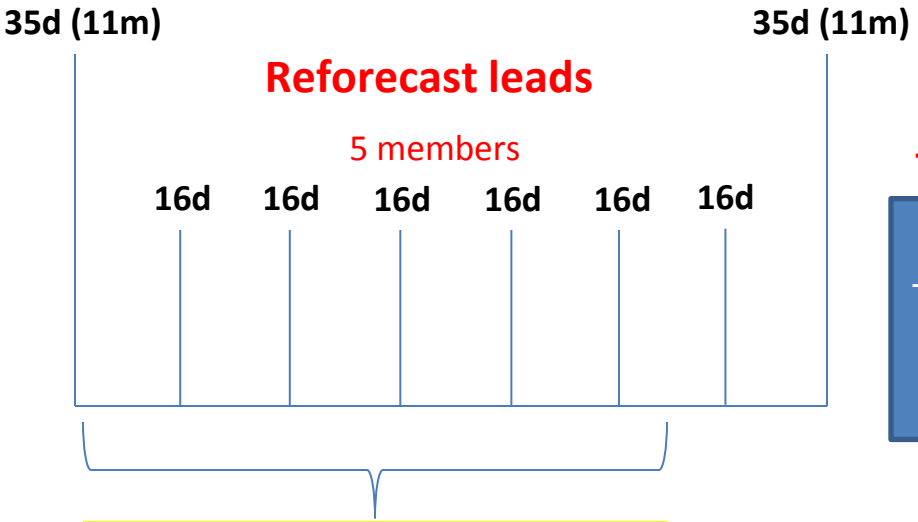
***We could find alternate resources on RDHPC to reduce the use of WCOSS resources**

Proposal 2: Implement GEFS V12.0 with minimum necessary retrospectives; Extend the time for generating 30-year reforecast experiments

- Pros:
  - Complete retrospective experiments for 3 summers and 2 winters (requirement for 16-day forecast evaluation by the field)
  - **Requires only 310 Cray nodes reserved for 8 months**
  - Accelerate implementation of GEFSV12.0 in operations
  - Stagger reforecast generation based on resource availability
- Cons:
  - Funding for reforecast project need to be extended beyond FY19
  - OWP, CPC and MDL wait longer to develop calibrations
  - Coupled model development for S2S predictions will be delayed
  - No scientific changes can be made for longer time (forecast model configuration already frozen)
  - GEFS V10 will need to run for a longer period in operations

11/5 members, every day at 00UTC

30 years reforecast



Every 7 days (period) as a unit
= 96 days (24 hrs fcst) and 5 m
+ 35 days and 11 m
= 96x5x72 + 35x11x72 (CPU hrs)
= 62280 CPU hrs

Project 650 Cray nodes for 1 year

Note: 7/19/2018
To satisfy CPC's requirement of 30 years reforecast – once per week, 11 members, out to 35 days
Cost: 45% of full reforecast configuration

Reference: FV3GEFS (C384 – 25km)
72 numbers of process (CPUs)
Per 1 member / per 1 day fcst

Summary

30 years = 1565 units

Need to run 6 units
Per day, last 9 months

HPCARC request:

Estimated - based on current option
Max. N. of Pros = 15570 CPUs (~650 nodes)
Estimated length of run = 4 hrs
Job submission = 6/day

Upper Air Variables (selected #1)

	U	V	T	RH	Height	VV	O3MR
10hPa	C,E	C,E	C,E		C,E		C
50hPa	C,E	C,E	C,E		E		C
100hPa	E	E	E		E		C
200hPa	C,M,E	C,M,E	C,M,E	C,M	C,M,E		
250hPa	M,E	M,E	M,E	M	M,E		
500hPa	C,M,E	C,M,E	C,M,E	C,M	C,M,E		
700hPa	C,M,E	C,M,E	C,M,E	C,M	C,M,E		
850hPa	C,M,E	C,M,E	C,M,E	C,M	M,E	E	
925hPa	M,E	M,E	M,E	M	M,E		
1000hPa	M,E	M,E	M,E	M	M,E		
0.996 (hybrid)	C	C	C	C			

Total: 55 variables to support CPC, MDL and EMC (NAEFS), but not for MDL's BMOS

C – CPC; M – MDL; N – NWC; E - EMC (the same for next slide)

Surface and other variables (Selected #2)

Variables	Requested	total	Notes
PMSL, Surface Pressure	C,M,N,E	2	
T2m, Tmax, Tmin	C,M,N,E	3	Tmax and Tmin for 6-hr
2m RH	M,N,E	1	Could convert to Td or q
U10m, V10m	C,N,E	2	
QPF	C,M,N,E	1	3-hr accumulation
Precipitation Types	C,M,E	4	Rain, Freezing rain, Ice Pellets, Snow
PWAT	M	1	
CAPE	C,M,E	1	
CIN	C,M,E	1	
Total sky cover (TCDC)	M,E	1	
Snow water equivalent	C	1	
OLR	C,E	1	
SDLR	N	1	
SDSR	N	1	

Total 21 variables, the BMOS variables are not counted in this list

Estimation of Tape and Disk Storage

- Reforecast

- HPSS (estimate – 2270TB)
  - 3hr to 10 days (0.25degree)
    - 81x365x5x30x350mb=1552TB
    - 81x73x6x30x350mb=372TB
  - 6hr from 10 days to 35 days (0.5degree)
    - 24x365x5x30x100mb=131TB
    - 24x73x6x30x100mb=32TB
    - 76x73x11x30x100mb=183TB
- Disk storage (estimate – 380TB)
  - 75 variables (3hr to 10 days) – 320TB
  - 75 variables (6hr from 10 to 35 days) - 58TB

Estimation of Tape and Disk Storage

- Retrospective forecast

- HPSS (estimate – 670TB) for short term
  - 3hr to 10 days (0.25degree)
    - 81x31x365x2.5x350mb=802TB
  - 6hr from 10 days to 16 days (0.5degree)
    - 24x31x365x2.5x100mb=68TB

- Disk storage (estimate – 55TB)
  - ~90 variables (3hr to 10 days – 0.5 degree) – 40TB
  - ~90 variables (6hr from 10 to 16 days – 0.5 degree) - 15TB

Example of one output file with selected variables

One member, f27, 3hr accu. period – f24-27

```
1:0:d=2017060100:GUST:surface:27 hour fcst:ENS=+1
2:513979:d=2017060100:HGT:10 mb:27 hour fcst:ENS=+1
3:1100800:d=2017060100:TMP:10 mb:27 hour fcst:ENS=+1
4:1403043:d=2017060100:UGRD:10 mb:27 hour fcst:ENS=+1
5:2131454:d=2017060100:VGRD:10 mb:27 hour fcst:ENS=+1
6:2702181:d=2017060100:HGT:50 mb:27 hour fcst:ENS=+1
7:3525243:d=2017060100:TMP:50 mb:27 hour fcst:ENS=+1
8:3829030:d=2017060100:UGRD:50 mb:27 hour fcst:ENS=+1
9:4466612:d=2017060100:VGRD:50 mb:27 hour fcst:ENS=+1
10:5076251:d=2017060100:HGT:100 mb:27 hour fcst:ENS=+1
11:5917644:d=2017060100:TMP:100 mb:27 hour fcst:ENS=+1
12:6226233:d=2017060100:UGRD:100 mb:27 hour fcst:ENS=+1
13:6891039:d=2017060100:VGRD:100 mb:27 hour fcst:ENS=+1
14:7522280:d=2017060100:HGT:200 mb:27 hour fcst:ENS=+1
15:8344493:d=2017060100:TMP:200 mb:27 hour fcst:ENS=+1
16:8656132:d=2017060100:RH:200 mb:27 hour fcst:ENS=+1
17:9253767:d=2017060100:UGRD:200 mb:27 hour fcst:ENS=+1
18:10063206:d=2017060100:VGRD:200 mb:27 hour fcst:ENS=+1
19:10780125:d=2017060100:HGT:250 mb:27 hour fcst:ENS=+1
20:11591062:d=2017060100:TMP:250 mb:27 hour fcst:ENS=+1
21:11902423:d=2017060100:RH:250 mb:27 hour fcst:ENS=+1
22:12577422:d=2017060100:UGRD:250 mb:27 hour fcst:ENS=+1
23:13005646:d=2017060100:VGRD:250 mb:27 hour fcst:ENS=+1
24:13437590:d=2017060100:HGT:500 mb:27 hour fcst:ENS=+1
25:14377305:d=2017060100:TMP:500 mb:27 hour fcst:ENS=+1
26:14712696:d=2017060100:RH:500 mb:27 hour fcst:ENS=+1
27:15431090:d=2017060100:UGRD:500 mb:27 hour fcst:ENS=+1
28:16213726:d=2017060100:VGRD:500 mb:27 hour fcst:ENS=+1
29:16996838:d=2017060100:HGT:700 mb:27 hour fcst:ENS=+1
30:17396265:d=2017060100:TMP:700 mb:27 hour fcst:ENS=+1
31:17408604:d=2017060100:RH:700 mb:27 hour fcst:ENS=+1
32:18155024:d=2017060100:UGRD:700 mb:27 hour fcst:ENS=+1
33:18164957:d=2017060100:VGRD:700 mb:27 hour fcst:ENS=+1
34:18174873:d=2017060100:HGT:850 mb:27 hour fcst:ENS=+1
35:18425852:d=2017060100:TMP:850 mb:27 hour fcst:ENS=+1
36:18441715:d=2017060100:RH:850 mb:27 hour fcst:ENS=+1
37:19226662:d=2017060100:VVEL:850 mb:27 hour fcst:ENS=+1
38:19249951:d=2017060100:UGRD:850 mb:27 hour fcst:ENS=+1
39:19272126:d=2017060100:VGRD:850 mb:27 hour fcst:ENS=+1
40:19294790:d=2017060100:HGT:925 mb:27 hour fcst:ENS=+1
41:19526152:d=2017060100:TMP:925 mb:27 hour fcst:ENS=+1
42:19575607:d=2017060100:RH:925 mb:27 hour fcst:ENS=+1
43:20349823:d=2017060100:UGRD:925 mb:27 hour fcst:ENS=+1
44:20400060:d=2017060100:VGRD:925 mb:27 hour fcst:ENS=+1
45:20449228:d=2017060100:TMP:1000 mb:27 hour fcst:ENS=+1
46:20599946:d=2017060100:RH:1000 mb:27 hour fcst:ENS=+1
47:21281923:d=2017060100:VVEL:1000 mb:27 hour fcst:ENS=+1
48:21524859:d=2017060100:UGRD:1000 mb:27 hour fcst:ENS=+1
49:21686118:d=2017060100:VGRD:1000 mb:27 hour fcst:ENS=+1
50:21844415:d=2017060100:HGT:1000 mb:27 hour fcst:ENS=+1
51:22281750:d=2017060100:PRES:surface:27 hour fcst:ENS=+1
52:23010038:d=2017060100:SOILW:0-0.1 m below ground:27 hour fcst:ENS=+1
53:23293271:d=2017060100:SOILW:0.1-0.4 m below ground:27 hour fcst:ENS=+1
54:23579656:d=2017060100:SOILW:0.4-1 m below ground:27 hour fcst:ENS=+1
55:23869506:d=2017060100:SOILW:1-2 m below ground:27 hour fcst:ENS=+1
56:24157866:d=2017060100:WEASD:surface:27 hour fcst:ENS=+1
57:24288073:d=2017060100:SNOD:surface:27 hour fcst:ENS=+1
58:24396748:d=2017060100:TMP:2 m above ground:27 hour fcst:ENS=+1
59:25139389:d=2017060100:RH:2 m above ground:27 hour fcst:ENS=+1
60:25796452:d=2017060100:TMAX:2 m above ground:24-27 hour max fcst:ENS=+1
61:26530478:d=2017060100:TMIN:2 m above ground:24-27 hour min fcst:ENS=+1
62:27265800:d=2017060100:UGRD:10 m above ground:27 hour fcst:ENS=+1
63:28075711:d=2017060100:VGRD:10 m above ground:27 hour fcst:ENS=+1
64:28859977:d=2017060100:APCP:surface:24-27 hour acc fcst:ENS=+1
65:29291583:d=2017060100:APCP:surface:0-27 hour acc fcst:ENS=+1
66:29717788:d=2017060100:ACPCP:surface:24-27 hour acc fcst:ENS=+1
67:30081025:d=2017060100:ACPCP:surface:0-27 hour acc fcst:ENS=+1
68:30440045:d=2017060100:CSNOW:surface:24-27 hour ave fcst:ENS=+1
69:30459840:d=2017060100:CICEP:surface:24-27 hour ave fcst:ENS=+1
70:30460170:d=2017060100:CFRZR:surface:24-27 hour ave fcst:ENS=+1
71:30461530:d=2017060100:CRAIN:surface:24-27 hour ave fcst:ENS=+1
72:30548054:d=2017060100:LFTX:surface:27 hour fcst:ENS=+1
73:30981145:d=2017060100:PWAT:entire atmosphere (considered as a single
layer):27 hour fcst:ENS=+1
74:31431089:d=2017060100:TCDC:entire atmosphere:24-27 hour ave fcst:ENS=+1
75:31978461:d=2017060100:DSWRF:surface:24-27 hour ave fcst:ENS=+1
76:32350542:d=2017060100:DLWRF:surface:24-27 hour ave fcst:ENS=+1
77:32855824:d=2017060100:ULWRF:top of atmosphere:24-27 hour ave fcst:ENS=+1
78:33286469:d=2017060100:HLCY:3000-0 m above ground:27 hour fcst:ENS=+1
79:33301527:d=2017060100:4LFTX:surface:27 hour fcst:ENS=+1
80:33343930:d=2017060100:CAPE:180-0 mb above ground:27 hour fcst:ENS=+1
81:34128735:d=2017060100:CIN:180-0 mb above ground:27 hour fcst:ENS=+1
82:34279457:d=2017060100:HPBL:surface:27 hour fcst:ENS=+1
83:35583253:d=2017060100:TMP:0.995 sigma level:27 hour fcst:ENS=+1
84:35663998:d=2017060100:RH:0.995 sigma level:27 hour fcst:ENS=+1
85:36050417:d=2017060100:UGRD:0.995 sigma level:27 hour fcst:ENS=+1
86:37592788:d=2017060100:VGRD:0.995 sigma level:27 hour fcst:ENS=+1
87:39061744:d=2017060100:ALBDO:surface:24-27 hour ave fcst:ENS=+1
88:39656014:d=2017060100:PRMSL:mean sea level:27 hour fcst:ENS=+1
```


Will take highlight variables (15) away; totally = 73 variables (52 upper air + 21 surface)

Sample data for GEFSv12 reforecast – contributed by Hong Guan

All (CPC, MDL and NWC/OWP);

As we promised before, we will send out a sample data for selected variables to allow all our stakeholders to test/valid. Dr. Hong Guan is our contact (cced), please let us know if there is any question. We'd like to have your confirmation before next reanalysis/reforecast meeting (current schedule - July 17 2018)

We have saved 74 variables (see attached slides - sample for you to verify):
1. Five ensemble members include ensemble control
2. 0.25 degree for 0-10 days every 3 hours
3. 0.5 degree for 10-35 days every 6 hours.
4. We have 2 QPF records in this sample, but will delete duplicate one later.

Notes for CPC: we will add on O3MR for 10hPa, 50hPa and 100hPa later
Notes for MDL: sample has excluded your BMOS request
Notes for NWC/OWP: you need to have WCOSS access soon, ftp sample here for validation/demonstration only. Currently, EMC does not have ftp disk storage for public access, except for future coordination/discussion with ESRL/PSD

To access sample data through website:
0.25 degree data: ftp://ftp.emc.ncep.noaa.gov/gc_wmb/wd20hg/FV3GEFS_rfcst/2017060100/pgrb2ap25
0.5 degree data: ftp://ftp.emc.ncep.noaa.gov/gc_wmb/wd20hg/FV3GEFS_rfcst/2017060100/pgrb2ap50

or anonymous ftp:
ftp ftp.emc.ncep.noaa.gov ID: anonymous PW: your email
cd gc_wmb/wd20hg/FV3GEFS_rfcst/2017060100 (you will see two subsets)

To access sample data from WCOSS directly (luna machine):
0.25 degree: /gpfs/hps3/emc/ensemble/noscrub/emc.enspara/FV3GEFS_rfcst/2017060100/pgrb2ap25
0.5 degree: /gpfs/hps3/emc/ensemble/noscrub/emc.enspara/FV3GEFS_rfcst/2017060100/pgrb2ap50

See an inventory of one forecast (lead), and one member:
http://www.emc.ncep.noaa.gov/gmb/wd20hg/FV3_anl/rfcst_output_0p25
http://www.emc.ncep.noaa.gov/gmb/wd20hg/FV3_anl/rfcst_output_0p50

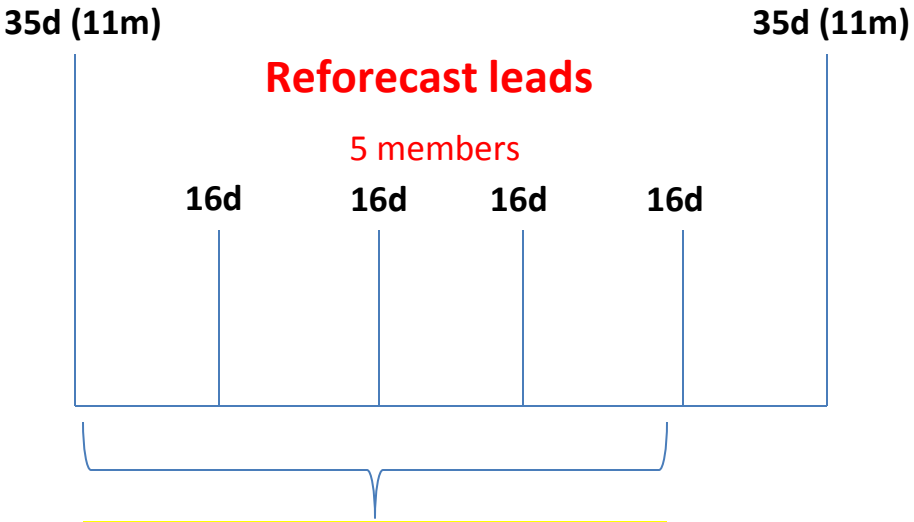
Receives confirmation of sample output data

- MDL – John Wagner for EKDMOS
  - Hi Yuejian, I believe the sample data will be good for EKDMOS. I have not been able to test everything as the control member is encoded as a low-res control (even though its 0.25 degrees) and my code is expecting the high-res control member. I will need to make some changes to get this data into TDLPACK, which I haven't had time to because of the WCOSS outages. I was able to convert the other members to TDLPACK without error. I see no reason not to proceed with these settings. Thanks. - John
- CPC – no confirmation yet for posted sample data
- OWC – Mark Fresch (future POC: Dr. Kaksu Lee)
  - Yuejian, The sample GEFSv12 reforecast is acceptable to OWP. Thanks, especially for Hong's help. -MarkF

Background!!!

30 years reforecast

11/5 members, every day at 00UTC



Every 5 days (period) as a unit
= 64 days (24 hrs fcst) and 5 m
+ 35 days and 11 m
= 64x5x72 + 35x11x72 (CPU hrs)
= 50760 CPU hrs

Project 715 Cray nodes for 1 year

Reference: FV3GEFS (C384 – 25km)
72 numbers of process (CPUs)
Per 1 member / per 1 day fcst

Summary

30 years = 2190 units

Need to run 8 units
Per day, last 9 months

HPCARC request:

Estimated - based on option one
Max. N. of Pros = 17160 CPUs (715 nodes)
Estimated length of run = 3 hrs
Job submission = 8/day

Upper Air Variables (selected #1)

	U	V	T	RH	Height
50hPa	C	C	C	C	
200hPa	C,M	C,M	C,M	C,M	C,M
250hPa	M	M	M	M	M
500hPa	C,M	C,M	C,M	C,M	C,M
700hPa	C,M	C,M	C,M	C,M	C,M
850hPa	C,M	C,M	C,M	C,M	M
925hPa	M	M	M	M	M
1000hPa	M	M	M	M	M
0.996 (hybrid)	C	C	C	C	

Total: 43 variables to support CPC and MDL, but not for MDL's BMOS

C – CPC; M – MDL; N – NWC (the same for next slide)

Surface and other variables (Selected #2)

Variables	Requested	total	Notes
PMSL, Surface Pressure	C,M,N	2	
T2m, Tmax, Tmin	C,M,N	3	Tmax and Tmin for 6-hr
2m RH	M,N	1	Either RH or q
U10m, V10m	C,N	2	
QPF	C,M,N	1	3-hr accumulation
Precipitation Types	C,M	4	Rain, Freezing rain, Ice Pellets, Snow
PWAT	M	1	
Convective precipitation (*BMOS)	M	1	3-hr
CAPE	C,M	1	
CIN	C,M	1	
Storm Relative Helicity (*BMOS)	M	1	
Total sky cover	M	1	
SFC albedo (*BMOS)	M	1	3-hr
Surface wind gust (*BMOS)	M	1	
PBL height (*BMOS)	M	1	
Lifted index (*BMOS)	M	1	
Water Equiv. Accum. Snow Depth	C,M	1	CPC request snow water equivalent
995hPa VV (*BMOS)	M	1	
Soil moisture (*BMOS)	M	4	
OLR	C	1	
SDLR	N	1	
SDSR	N	1	

Total 32 variables include BMOS variables, 21 variables without BMOS variables

Request Summary from Stakeholders

	CPC	NWC	MDL	Others
Years	30y (1989-2018)	30y (1989-2018)	3-5 years (EKDMOS) 15-25 years (BMOS)???	
Forecast frequency	Every 5 days at 00UTC	Every day at 00UTC	Every day for EKDMOS Every 5 days (BMOS)??	
Membership	5-11 members, more is better	5 members		
Forecast length	35 days			
Output frequency	6 hourly – week 1 6 hourly – week 2 Daily - weeks 3&4	6 hourly	1-hourly QPF out to 264h 3-hourly out to 240h 6-hourly out to 384h	
Output resolution	Match real-time operation	Match real-time operation	EKDMOS – 0p50 BMOS - match GFS FV3 (0p25, 0p50, 1p00)	
Output variables	Upper air ~24 Surface ~11	10 surface variables	Upper air ~45 (EKDMOS) Surface ~ 22 (EKDMOS) 100+ for BMOS???	

Note: EMC will work with stakeholders to make sure for saving all the variables for them

Length of reforecast -

1. Needless to say, having 11 members twice per week to 35 days will be better than having 11 members once per week to 35 days, but in the end, it will all depend on resources, and having something (11 members once per week to 35 days) will be better then nothing (no forecasts to 35 days).
2. That said, finding optimal solution between the trade space of number of hindcast years, ensemble size each day, and hindcast frequency is a tough question. An alternative to having 11 members once per week to 35 days would be having 5 members to 35 days twice per week. This should be same resource but might provide better sampling across ICs and a better fit for climatology.

Temporal output

Week-1 time frame - 6 hourly
Week-2 time frame - 6 hourly
Week-3/4 time frame - daily.

Variables

U,V,T,q
At vertical levels - 50, 250, 500, 700, 850, 0.996 sigma
PMSL
Snow water equivalent
2-m temperature
Max temp in 6 hour
Min temp in 6 hour
10-m u component
10-m v component
QPF in 6 hours
CAPE
CIN
Upward longwave radiation (OLR)

CPC's request

Horizontal resolution (0.25degree, 0.5degree and 1 degree outputs)

Whatever matches the corresponding model.
0.5 degree would be plenty.

1) Length of reforecast / # of members -
5 members, spanning 30 years, issued once per day at 0Z

2) Temporal output
Week-1 time frame - 6 hourly
Week-2 time frame - 6 hourly
Week-3/4 time frame - 6 hourly

Note: 6 hourly will allow full utility of the reforecasts through the extended GEFSv12 forecast horizon, as we certainly want to leverage GEFSv12 forecasts through full forecast horizon to the extent there is useful skill. But as useful skill may not extend beyond week 2 for the prospective GEFSv12 (we can only guess without analysis), we could forego reforecast data altogether beyond day 16 if resource constraints dictated. This would effectively preclude use of the GEFSv12 in HEFS or any potential NWM application beyond day 16.
We could also consider reducing the reforecasts temporal resolution to 24 hourly beyond day 16. This would allow us to potentially use the GEFSv12 temp forcing in weeks 3-4 for HEFS, but would still preclude use of the precip forcing in HEFS and all GEFS forcings for NWM application beyond day 16.

3) Variable output
* 2-m temperature
* Max temp in 6 hour
* Min temp in 6 hour
* QPF in 6 hours
** long-wave radiation (surface downward)
** short-wave radiation (surface downward)
** 2m specific humidity
** surface pressure
** (10 m) U wind
** (10 m) V wind

* required for HEFS
** additional forcings needed for any NWM application

NWC's request

4) Horizontal resolution (0.25degree, 0.5degree and 1 degree outputs)
Whatever matches the corresponding operational model

5) A consistent reanalysis methodology
Differences in observations accepted, the methodology should be as consistent as possible over the reanalysis period and with the operational analysis methodology.

2nd level priority - Desirable additional attributes to support potential NWM applications (if resources allow):

1) Higher temporal resolution (3 hourly or 1 hourly output if possible)

2) Higher spatial resolution (max model resolution, 0.25 degrees?)

EKDMOS/BMOS Reforecast Needs

1. Length of reforecast
   a. 3-5 years
   b. Preferably, data would be every 5th day and span 15-25 years
2. Temporal output
   a. 3-hourly from 6-hr to 240-hr
   b. 6-hourly from 240-hr to 384-hr
   c. Exception - 1-hourly QPF from 1 to 264 hours (NWM)

1. Variables
   a. Temperature, U-Wind, V-Wind, RH, Height
      i. Surface (2-m or 10-m) for T, U, V, and RH
      ii. Vertical levels for EKDMOS - 1000, 925, 850, 700, 500, 250, 200
      iii. Vertical levels for BMOS - prefer if this matches GFS FV3 (every 50 mb from 100-1000 mb, plus 925, 975, and 30 mb)
   b. Surface Pressure
   c. PMSL
   d. 6-hr Max T
   e. 6-hr Min T
   f. Cat Rain
   g. Cat Freezing Rain
   h. Cat Ice Pellets
   i. Cat Snow
   j. 3-hr and 6-hr Total Precip
   k. PWAT
   l. Total Sky Cover
   m. CAPE
   n. 3-hr and 6-hr Conv Precip - BMOS
   o. 6-hr SFC Albedo - BMOS
   p. Surface Wind Gust - BMOS
   q. PBL Height - BMOS
   r. Lifted Index - BMOS
   s. CIN - BMOS
   t. Water Equiv Accum Snow Depth - BMOS
   u. Storm Relative Helicity - BMOS
   v. 995 mb VV - BMOS
   w. Soil Moisture (0-.1M, .1-.4M, .4-1M) - BMOS
2. Horizontal Resolution
   a. EKDMOS - 0p50
   b. BMOS - match GFS FV3 (0p25, 0p50, 1p00)

MDL's request

Still need to know how we want to proceed with BMOS. Do we want to continue to use GEFS as we do with EKDMOS? Or do we want to try to get them to match the output from the GFS? We can not match what GFSMOS has with what currently comes out of the GEFS.

NAEFS bias corrected variables for 0.5d

Update: June 15 2017

Variables	pgrba_bc file	Total 53 (1)
GHT	10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa	10
TMP	2m, 2mMax, 2mMin, 10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa	13
UGRD	10m, 10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa	11
VGRD	10m, 10, 50, 100, 200, 250, 500, 700, 850, 925, 1000hPa	11
VVEL	850hPa	1
PRES	Surface, PRMSL	2
FLUX (top)	ULWRF (toa - OLR)	1
Td and RH	2m (April 8 2014)	2
TCDC	Total cloud cover (March 29 2016)	1
WIND	10 meter Wind speed (this upgrade)	0(1)
Notes	CMC do not apply for last 4 variables FNMOC data is in process now	

Implementation Plan for FV3-GEFS (FY2017-2020)

FV3GEFS	FY17				FY18				FY19				FY20			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
FV3GEFS Reanalysis Development			Develop and test low resolution FV3GFS with FV3GDAS, configure it for reanalysis (ESRL)													
FV3GEFS Ensemble Configuration		Configure FV3GFS ensemble resolution, members, physics, coupling to ocean and sea-ice, and extend forecasts to weeks 3&4 (EMC)														
FV3GEFS Reanalysis Production							Produce ~20-year reanalysis datasets using FV3GFS/GDAS (ESRL)									
FV3GEFS Reforecasts								Finalize FV3GEFS V12 configuration* & produce ~20-year reforecasts (extended to 35 days)								
FV3GEFS V12 Evaluation										Evaluate FV3GEFS V12 forecast performance out to weeks 3&4						
FV3GEFS V12 Implementation											Transition FV3GEFS V12 into operations					
Advancement of FV3GEFS													Further advancements of FV3GEFS (GFS/GEFS unification, ensemble based coupled modeling for 35-day weather outlook guidance)			

Delayed about 2-3 months

* Proposed changes for GEFS V12: 1) Produce FV3 based reanalysis in FY18 using the same configuration as Q2FY18 FV3GFS (ESRL); 2) Reforecasts will be based on FV3GEFS configured with either coupled to Ocean and Sea-Ice models or use 2-Tier SST approach; and 3) FV3GEFS Reforecasts extended to 35 days to include weeks 3&4 guidance.

WCOSS Statistics

WCOSS Component	Compute nodes	Peak TFs	Processor Type	Storage	Cores/ node	Memory/node
IBM - P1	640 / 504	149	Sandy Bridge	1.2 PB	16	64G
IBM - P2	1080	563	Ivy Bridge	2.0 PB	24	64G
Cray	2048	2,037	Haswell	3.5 PB	24	64G
DELL	**1212**	**1,400**	**Broadwell**	**5.4 PB***	**28**	**128G**

Tide/Gyre; Surge/Luna; Mars/Venus;

HPCRAC Request will be submitted in May after finalizing the requirements and model configuration for reforecasts and retrospectives.

* Disk will be divided into 3 filesystems on Day 1

Disk storage estimations (old for lower resolution)

Based on 1*1 degree NAEFS exchange file (89 variables)
1 file size = 4 mb

Total files = 365(days)*30(years)*65(lead times for 16 days)*5(members)*4(mb)=14235000mb ~15TB

Add 0.5*0.5 degree data for 0-10 days (every 3 hours)
Around factor 4 = 15*4 = 60TB

Add extended forecast at 1*1 degree ~5TB

Request totally 80TB disk storage

Request totally 80*10=800TB tape storage

Master files: 300-320 MB (542 variables) – Dingchen (6/15/2018)

0p25 grid,	37 MB (pgrb2a, 87 variables) ,	221 MB	(pgrb2b, 455 variable)
0p50 grid,	12 MB	68 MB	
1p00 grid	4 MB	21 MB	
2p50 grid	1 MB	4 MB	