FV3GFS Workflow and Forecast-Only Experiments

Fanglin Yang
IMSG - Environmental Modeling Center
National Centers for Environmental Prediction

FV3GFS Weekly Meeting
February 13, 2017

Acknowledgments: *Contributions and /or insightful comments made by Rusty Benson, Jun Wang, George Gayno, Vijay Tallapragada, Shian-Jiann Lin, Lucas Harris, Jeff Whitaker, Philip Pegion, shrinivas Moorthi, Hui-ya Chuang, Rahul Mahajan, Dusan Jovic etc are gratefully acknowledged.*

Outline

- Workflow and utilities
- Forecast-only experiment
- Hydrostatic vs non-Hydrostatic

Q3FY17 GFS Implementation Superstructure SVN Repository

https://svnemc.ncep.noaa.gov/projects/gfs/branches/

Revision 80517
/gfs/branches/gfs_q3fy17
[Parent Directory]
gdas.v14.1.0/
gfs.v14.1.0/
gfs_workflow.v14.1.0/
global_shared.v14.1.0/

/gfs/branches/
gfs_q3fy17/
global_shared.v14.1.0
[Parent Directory]
docs/
exec/
fix/
modulefiles/
parm/
scripts/
sorc/
ush/

Revision 80517
/gfs/branches/gfs_q3fy17/
global_shared.v14.1.0/sorc
[Parent Directory]
emcsfc_ice_blend.fd/
emcsfc_snow2mdl.fd/
gettrk.fd/
global_chgres.fd/
global_cycle.fd/
global_fcst.fd/
global_sfchdr.fd/
global_sighdr.fd/
gsi.fd/
ncep_post.fd/
………… ~17 programs

/gfs/branches/gfs_q3fy17/
gfs_workflow.v14.1.0/para
[Parent Directory]
bin/
exp/
jobs/
scripts/
ush/
util

Revision 80517
/gfs/branches/gfs_q3fy17/
gfs_workflow.v14.1.0/para/jobs
[Parent Directory]
anal.sh
angu.sh
arch.sh
avrg.sh
copy.sh
d3dp.sh
dcop.sh
dump.sh
earc.sh
ecen.sh
echk.sh
efcs.sh
efmn.sh ……………………….

~30 jobs

The branches are set up to run on all NCEP production and development machines to support both operation and research development !

FV3GFS Superstructure SVN Repository

https://svnemc.ncep.noaa.gov/projects/fv3gfs/

/fv3gfs/trunk

gfs_workflow.v15.0.0/

global_shared.v15.0.0/

gdas.v15.0.0/

gfs.v15.0.0/

lib/

gfs_workflow.v15.0.0/:

- /bin/
- exp/
- exp_fv3gfs/
- jobs/
- scripts/
- ush/
- util/

global_shared.v15.0.0/:

- docs/
- exec/
- fix/
- modulefiles/
- parm/
- scripts/
- sorc/
- ush/

https://svnemc.ncep.noaa.gov/projects/fv3/

NEMS FV3
only contains forecast model source code

Currently managed by Jun Wang

https://**svnemc.ncep.noaa.gov/projects/fv3gfs/trunk/global_shared.v15.0.0/**

New & Updated Source Code

(fv3gfs.fd moved to nems fv3)

fv3nc2nemsio.fd/
fre-nctools.fd
global_chgres.fd
orog.fd
regrid_nemsio.fd/

New USH Scripts

fv3gfs_driver_chgres.sh
fv3gfs_chgres.sh

fv3gfs_remap.sh
fv3gfs_remap_weights.sh

fv3gfs_nc2nemsio.sh
fv3gfs_regrid_nemsio.sh

fv3gfs_driver_grid.sh
fv3gfs_filter_topo.sh
fv3gfs_make_grid.sh*
fv3gfs_make_orog.sh*

New fcst script

exglobal_fcst_fv3gfs.sh.ecf

New parm/ parm_fv3diag

diag_table
diag_table_history
variable_table.txt

New fix fields

C1152/
C192/
C3072/
C384/
C48/
C768/
C96/

C768_grid.tile[1-6].nc	remap_weights_C768_0p125deg.nc
C768_grid_spec.tile[1-6].nc	remap_weights_C768_0p25deg.nc
C768_oro_data.tile[1-6].nc	remap_weights_C768_0p5deg.nc
C768_mosaic.nc	remap_weights_C768_1deg.nc

fv3_SCRIP_C768_GRIDSPEC_lon3072_lat1536.gaussian.bilinear.nc
fv3_SCRIP_C768_GRIDSPEC_lon3072_lat1536.gaussian.neareststod.nc

https://svnemc.ncep.noaa.gov/projects/fv3gfs/trunk/gfs_workflow.v15.0.0/

Restructured and simplified "jobs" scripts

Fcst.sh
Post.sh
Vrfy.sh
Arch.sh

New para_config structure ./exp_fv3gfs

para_config – master
(reduced from 2252 lines to 377 lines)
config.fcst
config.post
config.vrfy
config.arch

config.gsi config.prep
config.dump config.nsst

Does not use reconcile.sh

submit_fv3gfs.sh for running forecast-only experiment

Users can use this script to run forecast-only experiments.
It reads from /com dir or extracts from HPSS current operational GFS ICs, converts to FV3GFS cold start ICs, submits any given number of forecast-only runs at a time, runs post, vrfy and arch steps, and submits a batch job to wait for a few hours to submit next group for forecast-only runs. Currently the workflow is controlled by **psub and pend**.

https://svnemc.ncep.noaa.gov/projects/fv3gfs/trunk/gfs_workflow.v15.0.0/

para_config --master config

```
……………………
Set up environment and running directories
………………………..
# -------------------------------------------------------------
#  settings that are used by more than one steps
# -------------------------------------------------------------
gfs_cyc=1                    # GFS cycles (00, 06, 12 and 18Z), defaults to 1 (00Z) cycle
gdas_cyc=4                   # number of GDAS cycles
fseg=1                       # number of AM forecast segments for gfs
FHCYC=24                     # Surface cycle calling interval

LEVS=64                      # number of AM levels
CASE1=C192                   # 1st segment resolution (0-240 hr)
CASE2=C192                   # 2nd segment resolution (240-384 hr)
CASE_ENKF=C382               # ENKF resolution
CASE=$(eval echo \${CASE$nknd})
if [ $CSTEP = efmn -o $CSTEP = epos ]; then CASE=$CASE_ENKF; fi

case $CASE in
  C48)  DELTIM=3600; layout_x=4  ; layout_y=8  ;;
  C96)  DELTIM=1800; layout_x=4  ; layout_y=8  ;;
  C192)  DELTIM=900 ; layout_x=4  ; layout_y=8  ;;
  C384)  DELTIM=450 ; layout_x=4  ; layout_y=8  ;;
  C768)  DELTIM=225 ; layout_x=8  ; layout_y=16 ;;
  C1152) DELTIM=150 ; layout_x=8  ; layout_y=16 ;;
  C3072) DELTIM=90  ; layout_x=16 ; layout_y=32 ;;
  *)    echo "grid $CASE not supported, exit"
        exit ;;
esac
…………………
```

```
fmax1=240; ; fmax2=384
for cyc in 00 06 12 18; do
eval FHMAXFCST${cyc}GFS1=$fmax1
eval FHMAXFCST${cyc}GFS2=$fmax2    # maximum hour 2st segment
 eval FHMAXFCST${cyc}GDAS=9   # maximum forecast hour for GDAS
 eval FHOUTFCST${cyc}GFS1=6
 eval FHOUTFCST${cyc}GFS2=12
 eval FHOUTFCST${cyc}GDAS=1
 eval FHZERFCST${cyc}GFS1=6
 eval FHZERFCST${cyc}GFS2=12
 eval FHZERFCST${cyc}GDAs=6
 eval MFCST${cyc}GFS=$fseg              #number of GFS forecast segments
 eval MFCST${cyc}GDAS=1                 #number of GDAS forecast segments
done
cdump=$(echo $CDUMP|tr '[a-z]' '[A-Z]')
FHMAX=$(eval echo \${FHMAXFCST$cycn$cdump$nknd})
FHOUT=$(eval echo \${FHOUTFCST$cycn$cdump$nknd})
FHZER=$(eval echo \${FHZERFCST$cycn$cdump$nknd})


#---if fdiag is given, it overwrites FHOUT
fh00=$(echo $DELTIM 3600|awk '{printf "%f", $1/$2}')
fdiag="$fh00,6.,12.,18.,24.,30.,36.,…………………

REMAP_GRID=latlon      #gaussian or latlon for using fregrid or regrid_nemsio for remapping
……………
```

https://svnemc.ncep.noaa.gov/projects/fv3gfs/trunk/gfs_workflow.v15.0.0/

Config.fcst

```
……………..
IC_DIR=/gpfs/hps/ptmp/$LOGNAME/FV3IC/ICs
if [ $REMAP_GRID = latlon ]; then
 DIAGTABLE=$BASE_GSM/parm/parm_fv3diag/diag_table
else
 DIAGTABLE=$BASE_GSM/parm/parm_fv3diag/diag_table_history
fi

npes=$(( ${layout_x} * ${layout_y} * 6 ))
tasks=$npes                      # number of PEs for 1st segment
nth_f=2                   # number of threads for AM forecast
npe_node_f=$pe_node              # number of pes per node for AM forecast
task_per_node=$((npe_node_f/nth_f))

MODE=64bit          # choices:  32bit, 64bit
TYPE=nh             # choices:  nh, hydro
HYPT=off            # choices:  on, off  (controls hyperthreading)
COMP="prod"         # choices:  debug, repro, prod
if [ ${HYPT} = on ]; then
  export hyperthread=".true."
  export j_opt="-j 2"
else
  export hyperthread=".false."
  export j_opt="-j 1"
fi
FCSTEXEC=fv3_gfs_${TYPE}.${COMP}.${MODE}.x
APRUN="aprun -n $tasks -N $task_per_node -d $nth_f $j_opt -cc depth“
…………………………………………..
```

https://svnemc.ncep.noaa.gov/projects/fv3gfs/trunk/gfs_workflow.v15.0.0/

Fcst.sh

```
….
set -a;. $CONFIG;set +a
echo "-----end of $CONFIG ------------"
echo

export CKSH=${CKSH:-$(echo $CSTEP|cut -c-4)}
export CKND=${CKND:-$(echo $CSTEP|cut -c5-)}
export machine=${machine:-WCOSS_C}
machine=$(echo $machine|tr '[a-z]' '[A-Z]')
eval export DATA=$DATATMP
rm -rf $DATA||exit 1; mkdir -p $DATA||exit 1; cd $DATA||exit 1
chmod ${permission:-755} $DATA

#
export PBEG=${PBEG:-$SHDIR/pbeg}
export PEND=${PEND:-$SHDIR/pend}
export PERR=${PERR:-$SHDIR/perr}
export VERBOSE=YES


$FORECASTSH
rc=$?
if [[ $rc -ne 0 ]];then $PERR;exit 1;fi

if [ ${KEEPDATA:-NO} != YES ] ; then rm -rf $DATA ; fi
$PEND
```

Fcst.sh has been reduced from more than 1000 lines to 39 lines

It directly calls exglobal_fcst_fv3gfs.sh.ecf, which can be used for operation.

https://svnemc.ncep.noaa.gov/projects/fv3gfs/trunk/gfs_workflow.v15.0.0/

Config.post

```
POSTSH=$BASEDIR/jobs/post.sh                  #post workflow
NODES=4                               #number of nodes for all post jobs
if [ $CKND -eq 2 ]; then NODES=3 ; fi
POST_MEMORY=3072

#-------------------------------------------------------------------------------
```

#--use ESRL ESMF regrid tool to remap forecast 6-tile netcdf files to global Gaussian grids and to write output in nemsio format.

```
……
REGRIDNEMSIOSH=$BASE_GSM/ush/fv3gfs_regrid_nemsio.sh
………."
#---------------------------------------------------------------------------
```

#--**use GFDL fregrid tool to remap forecast 6-tile netcdf files to global lat-lon grids but still in netcdf format, then use NC2NEMSIO to convert to nemsio.**

```
…
REMAPSH=$BASE_GSM/ush/fv3gfs_remap.sh
#remap 6-tile output to global array in netcdf
REMAPEXE=$BASE_GSM/exec/fregrid_parallel
master_grid=0p25deg                  #1deg 0p5deg 0p25deg 0p125deg etc
……………………………..
```

```
POSTJJOB=$BASEDIR/jobs/JGFS_POST.sh
global_nceppost.sh and down-stream jobs
POSTGPSH=$BASE_POST/ush/global_nceppost.sh
POSTGPEXEC=$BASE_POST/exec/ncep_post
npe_node_po=6                        #number of tasks per node for UPP
npe_po=$((NODES*npe_node_po))            #total number of tasks for UPP
NTHRPOST=1
APRUN_NP="aprun -n $npe_po -N $npe_node_po -j 1 -d $NTHRPOST -cc depth"


GFS_DOWNSTREAM=YES                      #run downstream jobs
GFSDOWNSH=../gfs_downstream_nems.sh
GFSDWNSH=$USHDIR/gfs_dwn_nems.sh
downset=1
npe_node_dwn=8                      #number of tasks per node
npe_dwn=$((NODES*npe_node_dwn))
nthread_dwn=$((pe_node/npe_node_dwn))
APRUN_DWN="aprun -n $npe_dwn -N $npe_node_dwn -j 1 -d $nthread_dwn cfp"


………………………
```

https://svnemc.ncep.noaa.gov/projects/fv3gfs/trunk/gfs_workflow.v15.0.0/

post.sh

```
set -a;. $CONFIG;set +a

export CKSH=$(echo $CSTEP|cut -c-4)
export CKND=$(echo $CSTEP|cut -c5-)
export machine=${machine:-WCOSS}
export machine=$(echo $machine|tr '[a-z]' '[A-Z]')
eval export DATA=$DATATMP
rm -rf $DATA||exit 1;mkdir -p $DATA||exit 1;cd $DATA||exit 1
chmod ${permission:-755} $DATA
#
export PBEG=${PBEG:-$SHDIR/pbeg}
export PEND=${PEND:-$SHDIR/pend}
export PERR=${PERR:-$SHDIR/perr}
$PBEG

if [ ${REMAP_GRID:-latlon} = latlon ]; then
 echo
 $REMAPSH                         #remap 6-tile output to global array in netcdf
 if [[ $? -ne 0 ]];then $PERR;exit 1;fi

 echo
 $NC2NEMSIOSH                      #convert netcdf to nemsio
 if [[ $? -ne 0 ]];then $PERR;exit 1;fi
else
 echo
 $REGRIDNEMSIOSH
 if [[ $? -ne 0 ]];then $PERR;exit 1;fi   ##use ESMF regrid to convert directly to gaussian grid
fi

echo
$POSTJJOB                         #converts nemsio to grib2 and run down-stream jobs
if [[ $? -ne 0 ]];then $PERR;exit 1;fi

……..
```

Post.sh has been reduced from 1089 lines to ~50 lines

It uses operational-alike JGFS_POST.sh and exgfs_nceppost.sh.ecf.

Source code ncep_post.fd was updated to handle nemiso input on lat-lon grid, and for any model vertical layers.

NEMS GSM Input and Output

Initial Conditions:

gfnanl.gfs(gdas).$cdate

sfnanl.gfs(gdas).$cdate

Where
gfnanl – atmosphere states
sfnanl – land surface states

In binary nemsio format with meta data header, on model native grid

Forecast direct output:

gfnf$fh.gfs(gdas).$cdate

sfnanl$fh.gfs(gdas).$cdate

flnf$fh.gfs(gdas).$cdate

Where
fh is forecast hour; fln is flux variables such as radiation and precipitation etc

Also all in binary nemsio format, and on model native grid

NCEP Unified Post reads in model direct output, then 1) performs vertical interpolation to place variables on standard isobaric layers, 2) derives other output products, and 3) writes out grib2 products.

FV3GFS Input and Output

Cold RESTART
gfs_ctrl.nc sfc_ctrl.nc
gfs_data.tile$n.nc
sfc_data.tile$n.nc
Where n=1,2,…6

Created by CHGRES using operational GFS IC as input

Warm RESTART
coupler.res fv_core.res.nc
fv_core.res.tile$.nc
fv_srf_wnd.res.tile$n.nc
fv_tracer.res.tile$n.nc
sfc_data.tile$n.nc

Written out at the end of forecast

Forecast history: latlon=grid
$cdate.atmos_4xdaily.tile$n.nc
$cdate.nggps2d.tile$n.nc
$cdate.nggps3d_4xdaily.tile$n.nc
atmos_static   grid_spec

Written out at $fdiag interval

Forecast history: latlon=gaussian
$cdate.fv3_historyd.tile$n.nc
$cdate.fv3_history2d.tile$n.nc

Written out at $fdiag interval

```
#---if fdiag is given, it overwrites FHOUT
fh00=$(echo $DELTIM 3600|awk '{printf "%f", $1/$2}')
fdiag="$fh00,6.,12.,18.,24.,30.,36.,42.,48.,54.,60.,66.,72.,78.,84.,90.,96.,102.,108.,114.,120.,126.,132.,138.,144.,150.,
156.,162.,168.,174.,180.,186.,192.,198.,204.,210.,216.,222.,228.,234.,240."
NFCST=$(echo $fdiag |awk -F '[\t,]' '{print NF}')   ;#number of forecast output
```

Resolution, Physics Grid, and Run-time on Cray
10-d forecast, 6-hourly output, 3.75-minute time step
C768, 13km, 3,538,944 points

Hydro/ non-hydro	precision	threads	nodes	CPU (min/ 10day)
Non-hydro	32-bit	2	64	89
Non-hydro	64-bit	2	64	137
Non-hydro	64-bit	2	144	69
Non-hydro	64-bit	4 Hyper-Thread	64	135
hydro	64-bit	2	64	95
hydro	64-bit	2	144	51

T1534 NEMS GFS (~13 km, 3072x1536), 61 nodes, 73 minutes

FV3GFS is about 1.5 ~2 times slower than NEMS GFS

FV3GFS Forecast-only benchmark test

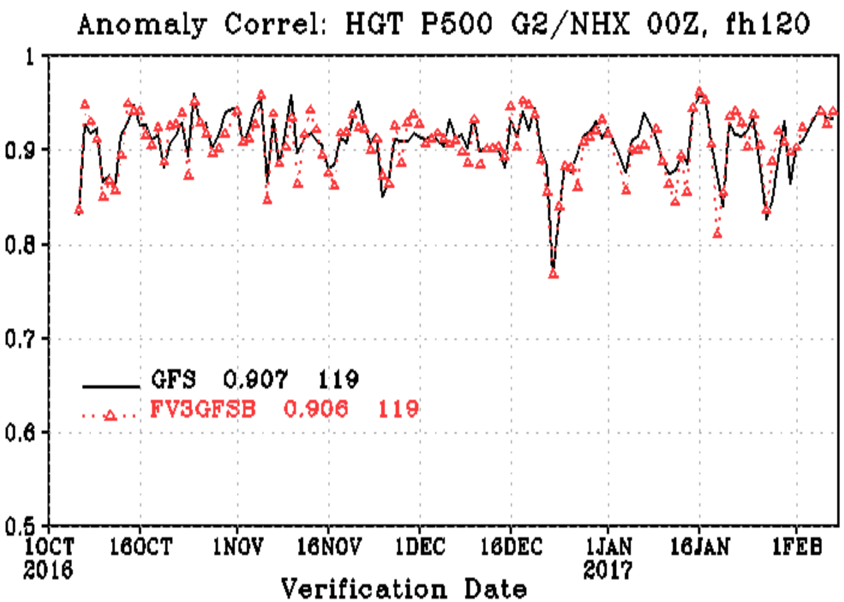
The version used for this test is the same as the one used for NGGPS dycore phase-II comparison, except for the following changes

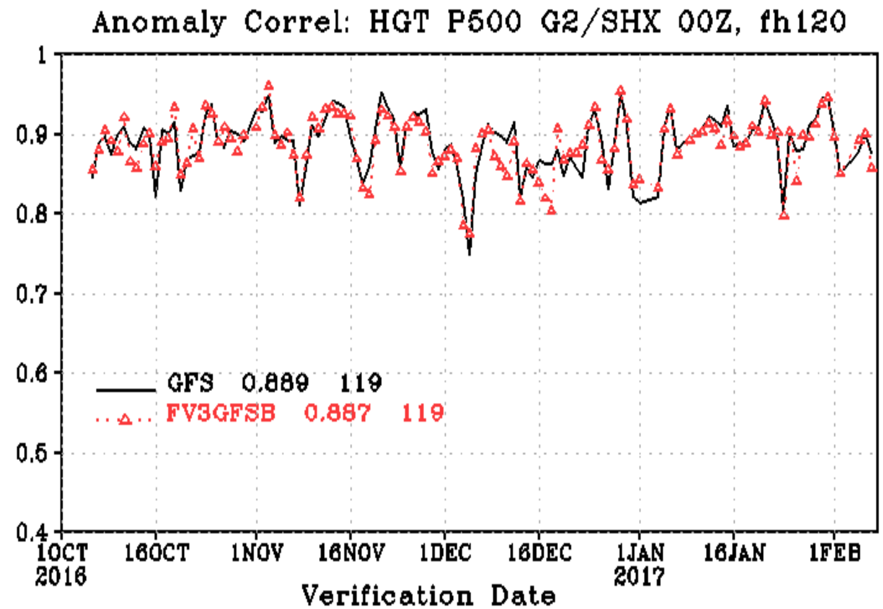
- **TKE dissipative heating:** F in fv3gfs, changed to T as in ops GFS
- **shallow convection:** F in fv3gfs, changed to T as in ops GFS
- **PBL scheme:** Moninq in fv3gfs, changed to hybrid EDMF as in ops GFS
- **sfc emmisivity:** black body (IEMS=0) in fv3gfs, changed to grey as in ops GFS**.**
- **mountain block and GWD:** 3.5,0.25 in fv3gfs, (2.0,0.25 in ops GFS)

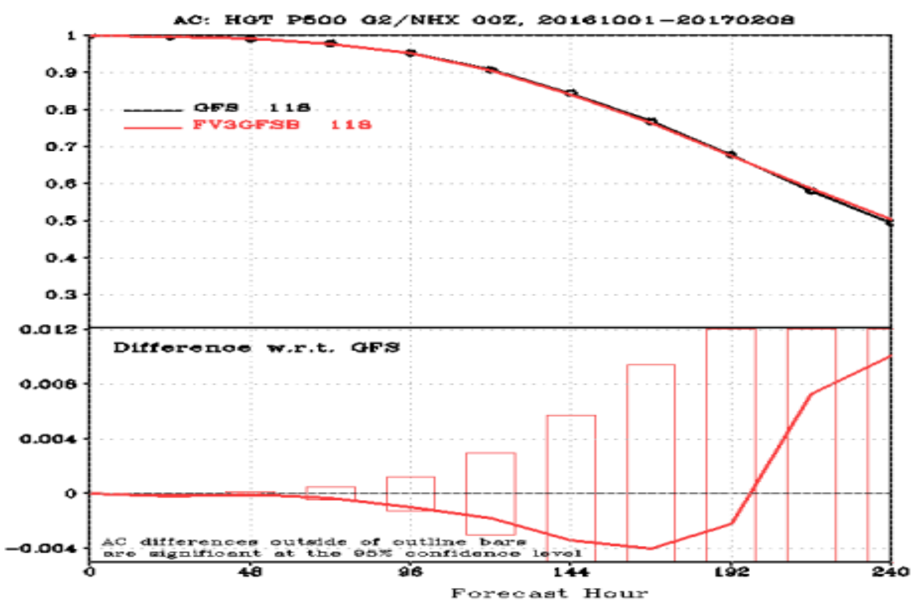
Period: 20161001 ~ 20170208, initialized with current ops GFS ICs
Resolution: C768 (~13km), 64 Layer
Configuration: non-hydrostatic, 64-bit precision
CPU: on CRAY with 64 nodes takes ~13.7 minutes per day

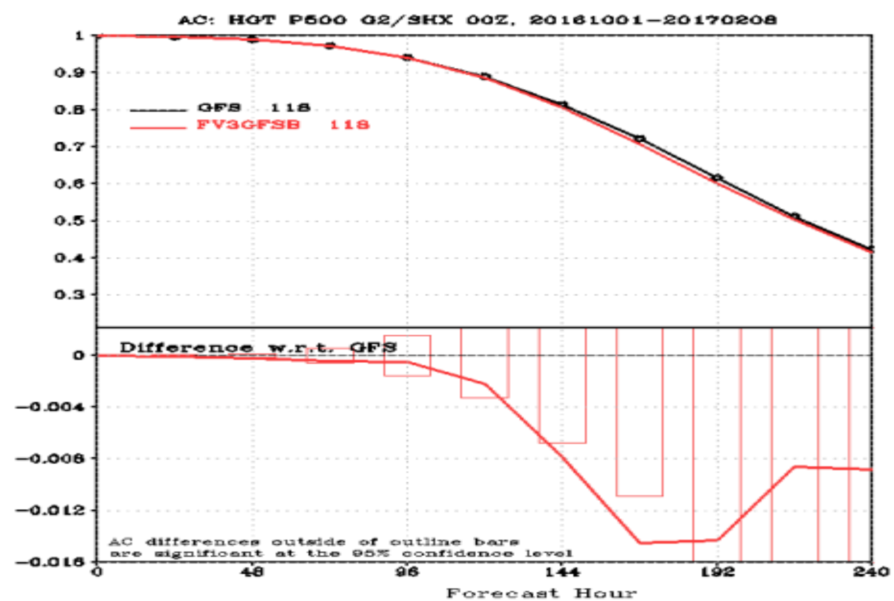
http://www.emc.ncep.noaa.gov/gmb/wx24fy/NGGPS/fv3gfsb/

500hPa HGT AC

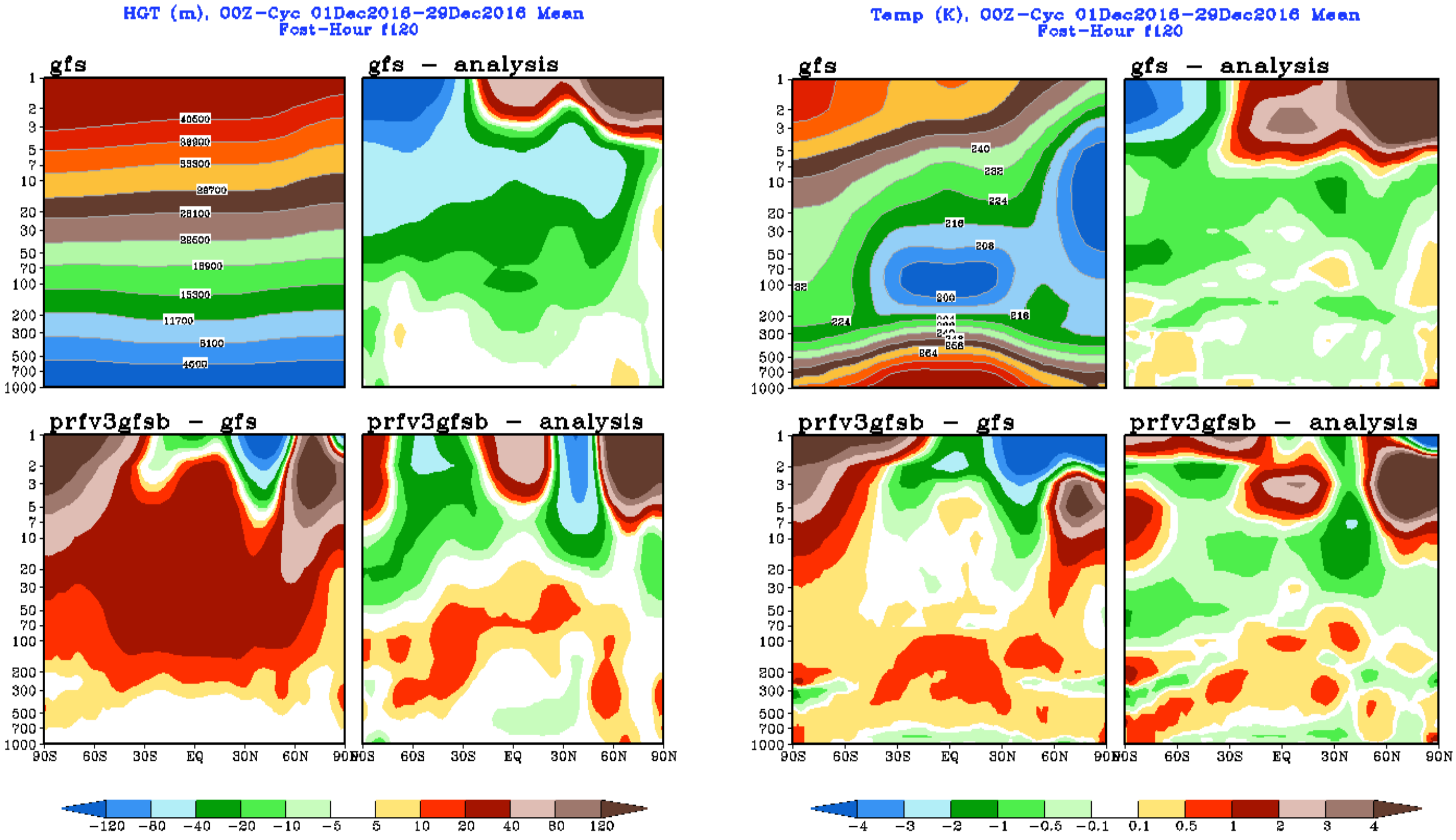






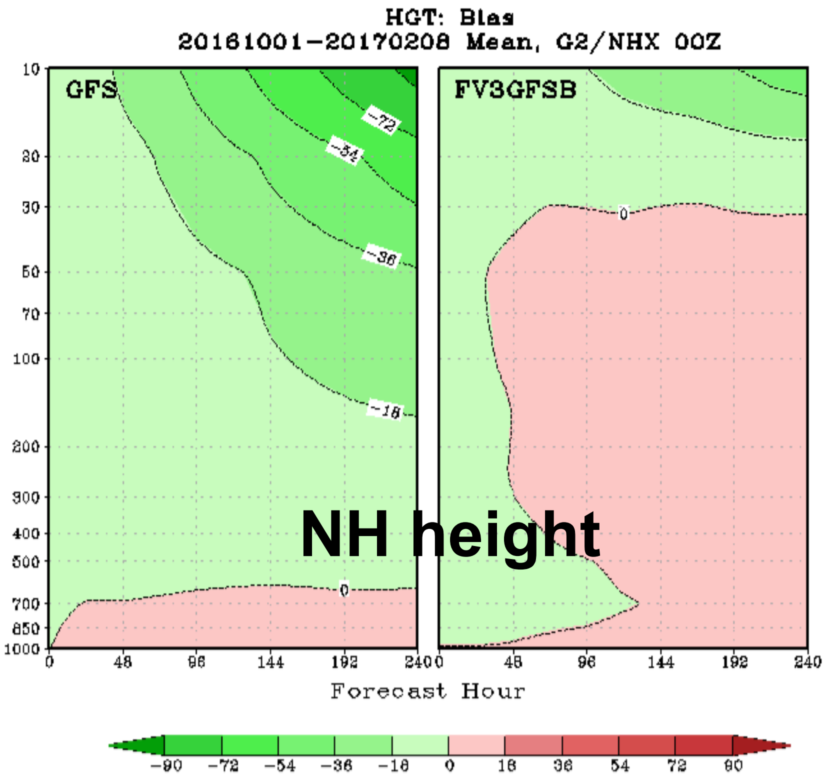


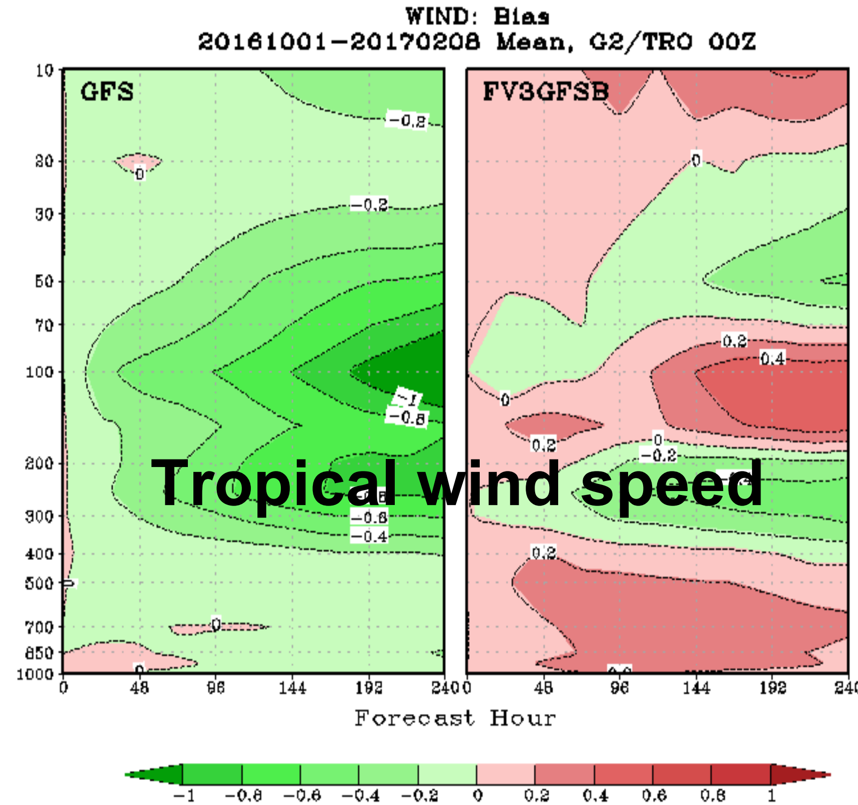
Zonal Mean HGT and T, 120-hr forecast



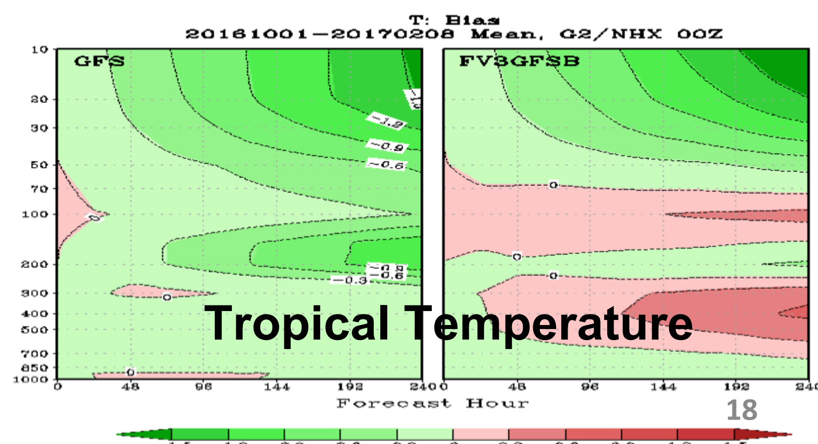
FV3GFS reduced stratospheric cold bias
Slightly too warm in the troposphere

BIASES, verified against GFS analysis



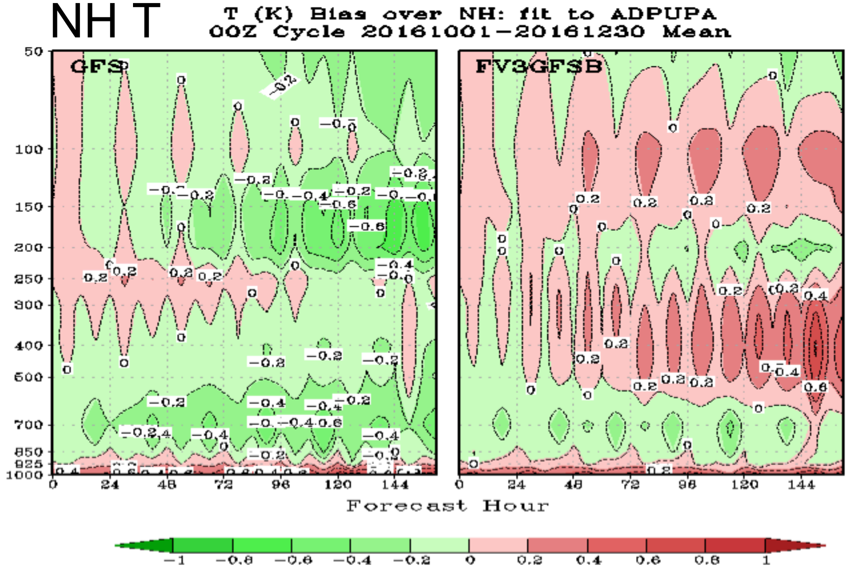


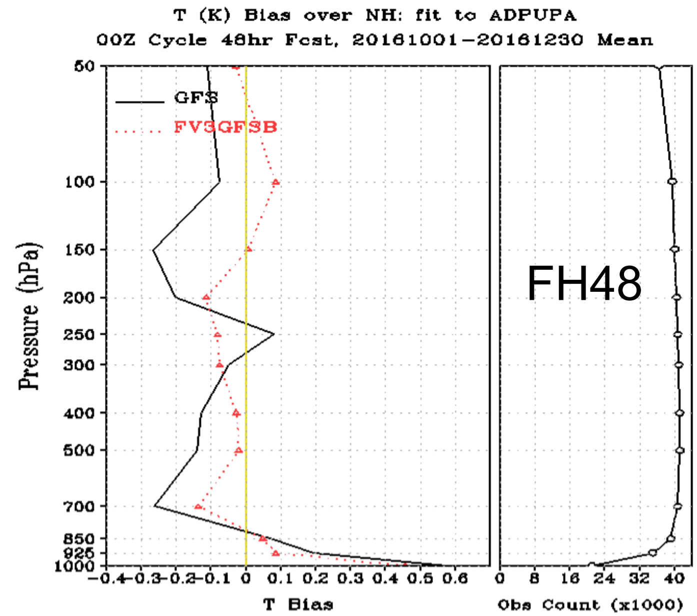
FV3GFS reduced cold bias in the stratosphere, and maintains better tropospheric jets.



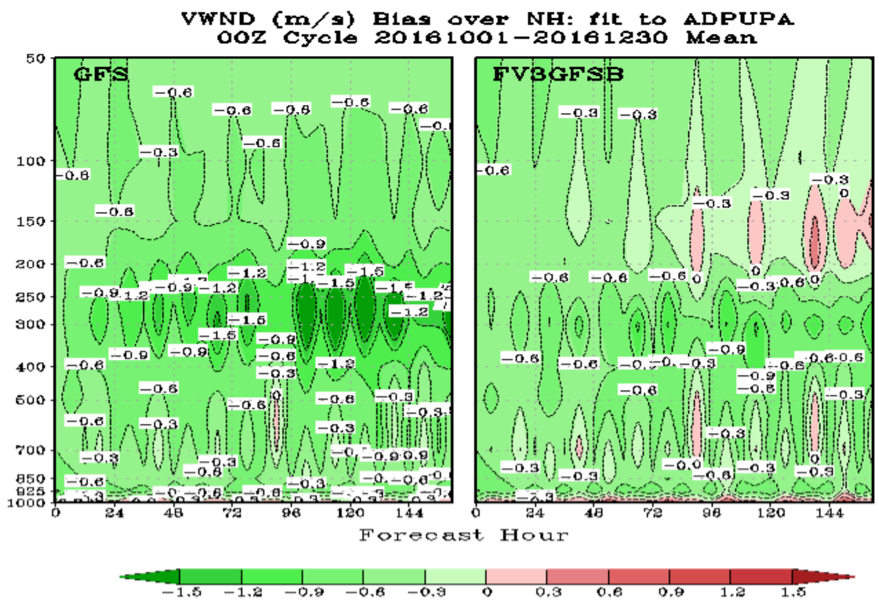
Verified against ROBS, BIAS

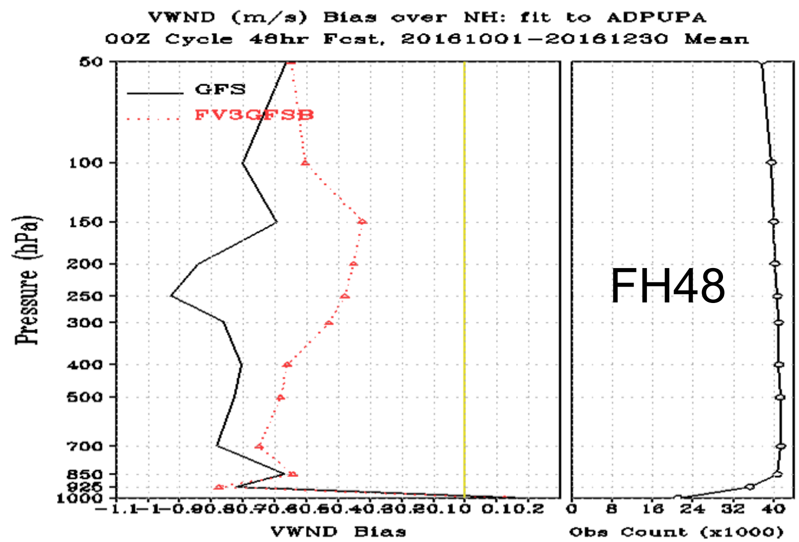
NH T



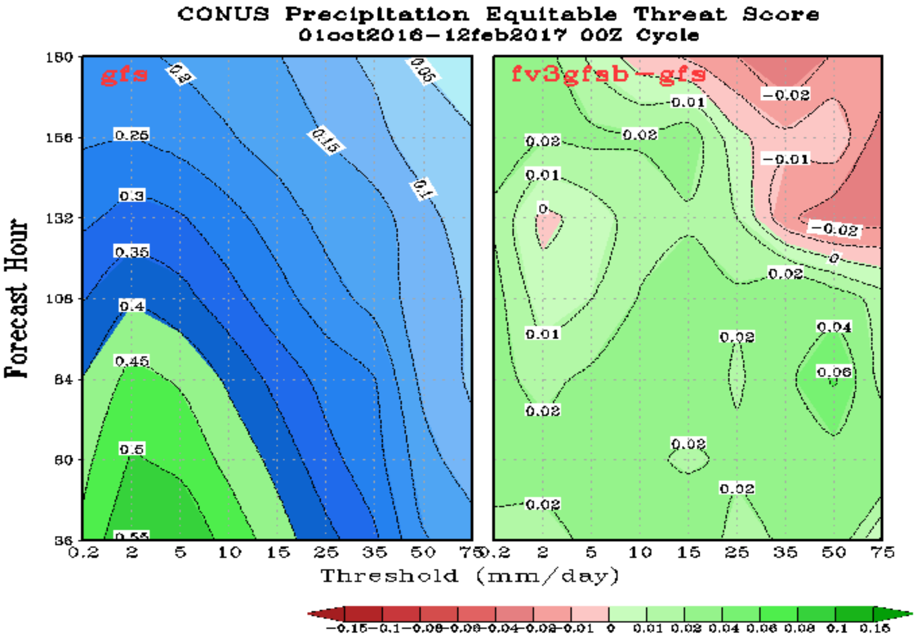


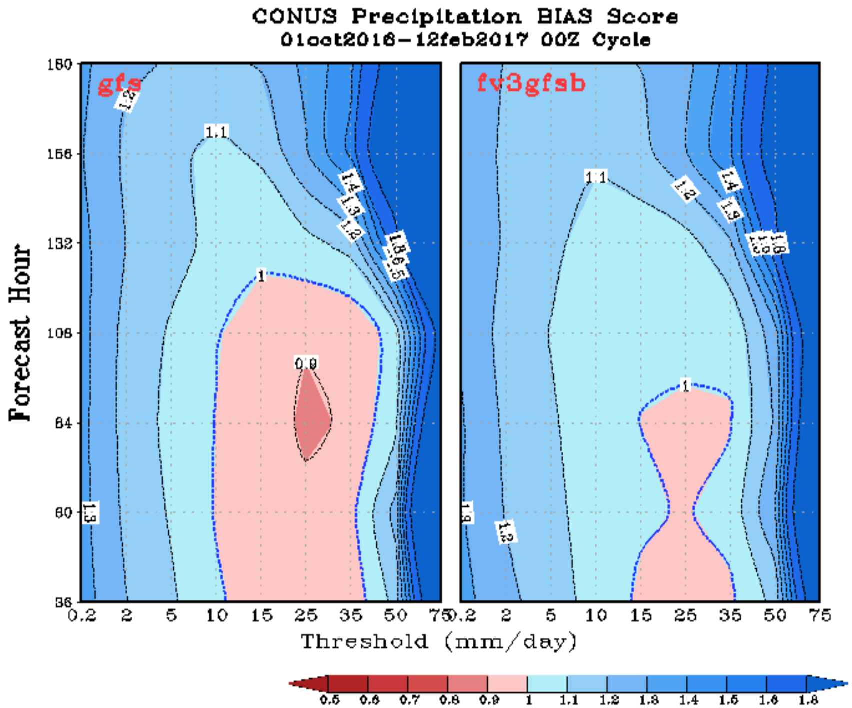
NH Wind



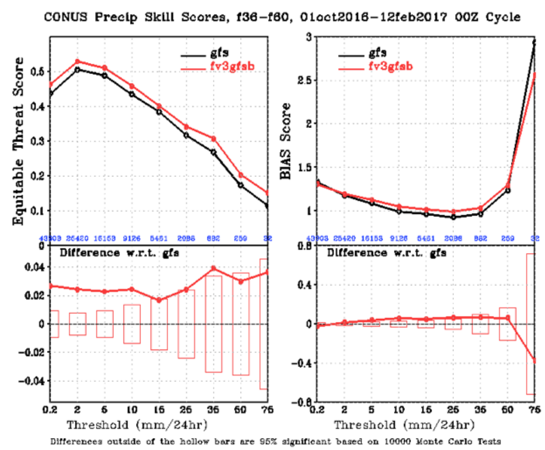


precipitation





FV3GFS improved rainfall forecast skills

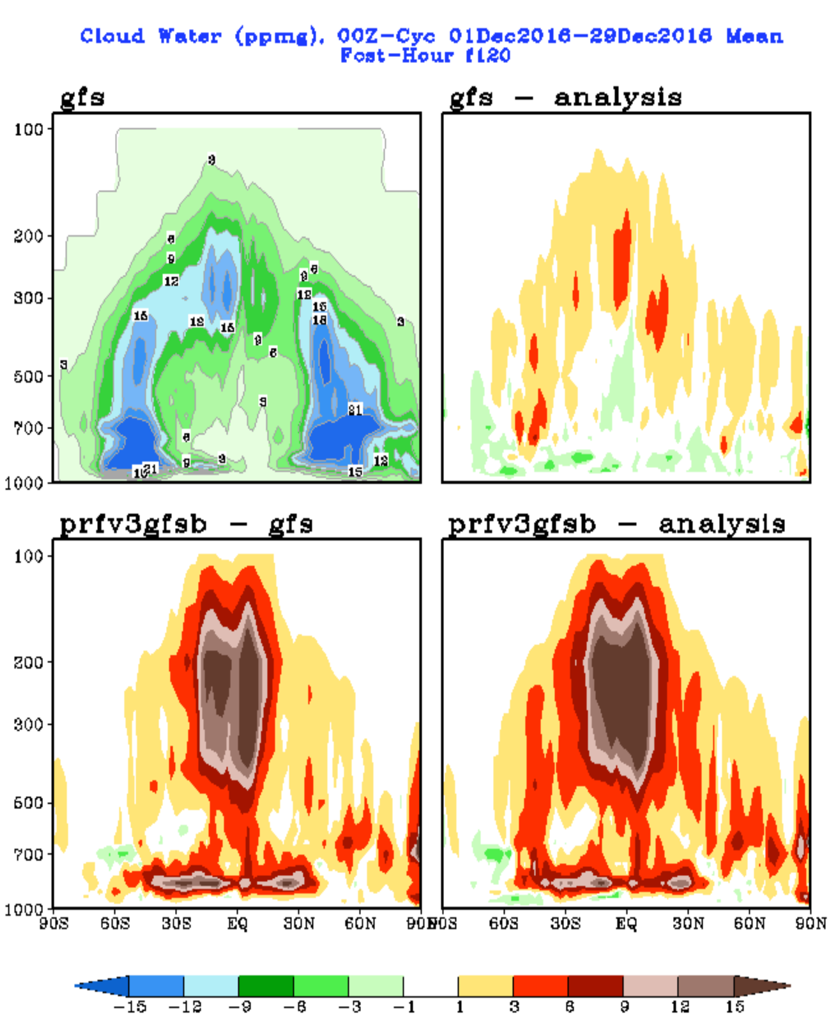


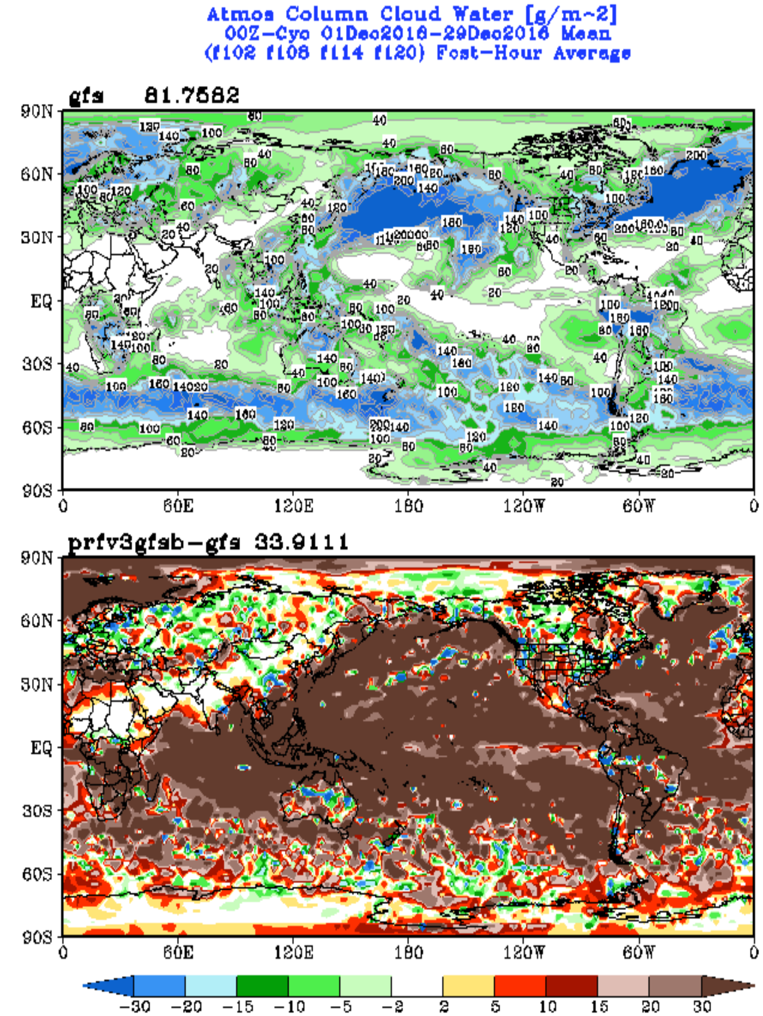
However: EMC physics developers recently reported that the deep convection scheme used in FV3GFS for phase-II dycore comparison is different from the one in the current operational GFS. The improvement in fv3gfs precipitation likely came from both the different convection scheme and new model dynamics.

Need to use IPD v4.0 to establish a new benchmark

Cloud Water

More cloud water

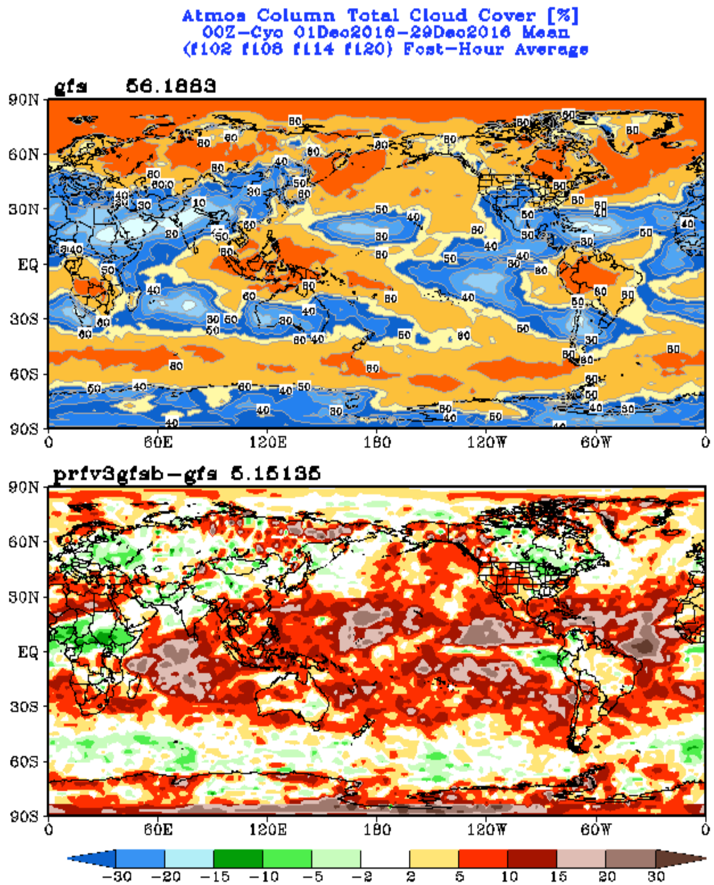


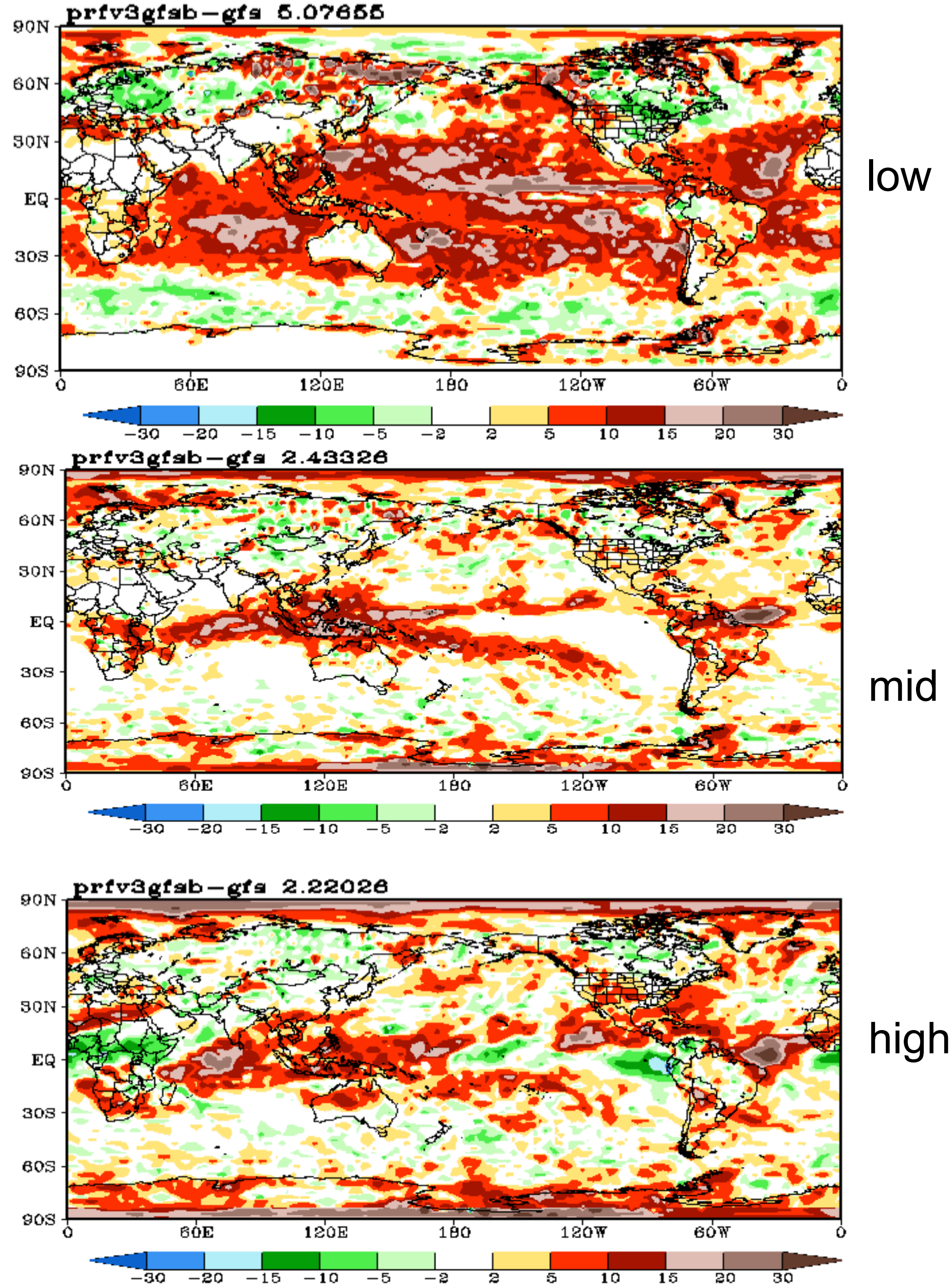


prautco(2) = (/1.0e-4, 1.0e-4/)
psautco(2) = (/6.0e-4,3.0e-4/)
Same as ops GFS

Clouds

total

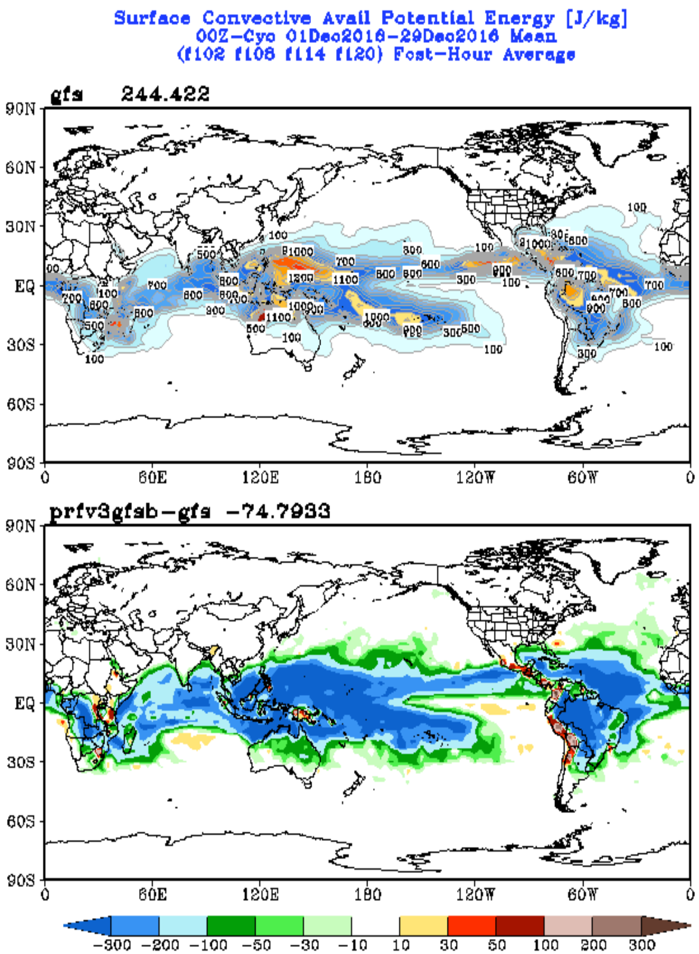


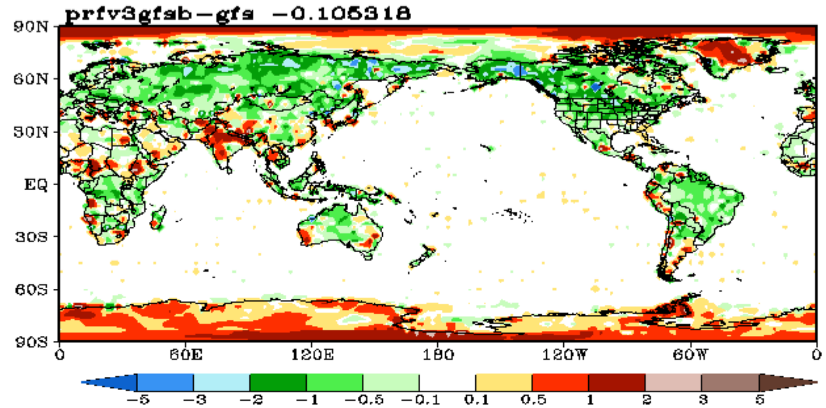


Xu-Randall Cloud cover scheme has been tuned. Even though cloud water increases, high cloud fraction remains almost unchanged

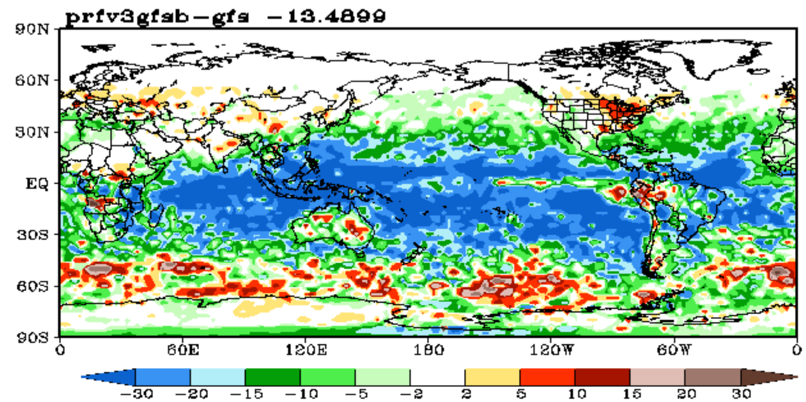
Much Smaller CAPE

Larger LH
Smaller downward SW

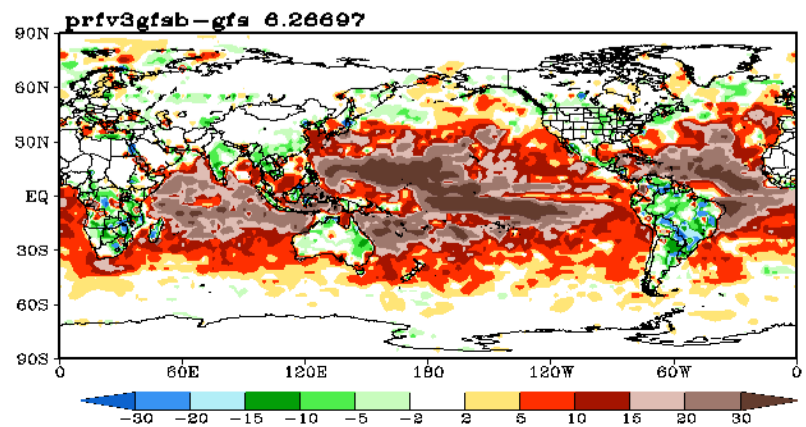




Tskin



Sfc down SW



Sfc Latent heat flux

Hydrostatic vs Non-hydrostatic

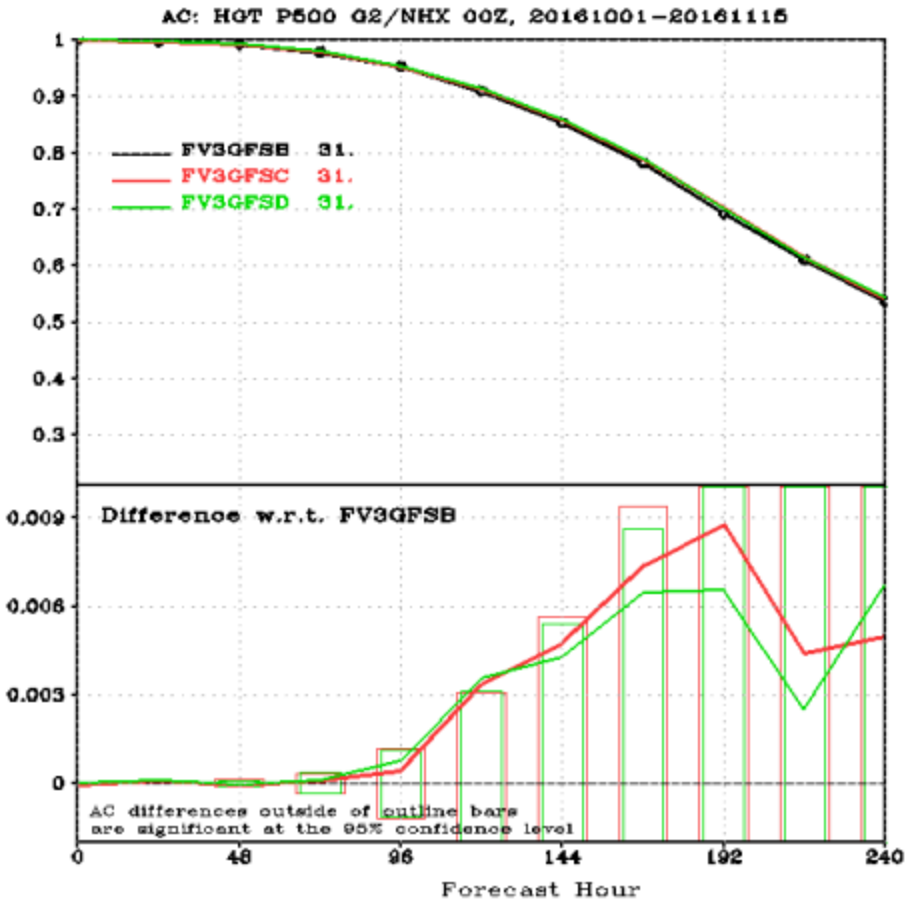
- Prfv3gfsb: non-hydrostatic, non-mono
- Prf3gfsc: hydrostatic, non-mono
- Prfv3gfsd: hydrostatic, mono

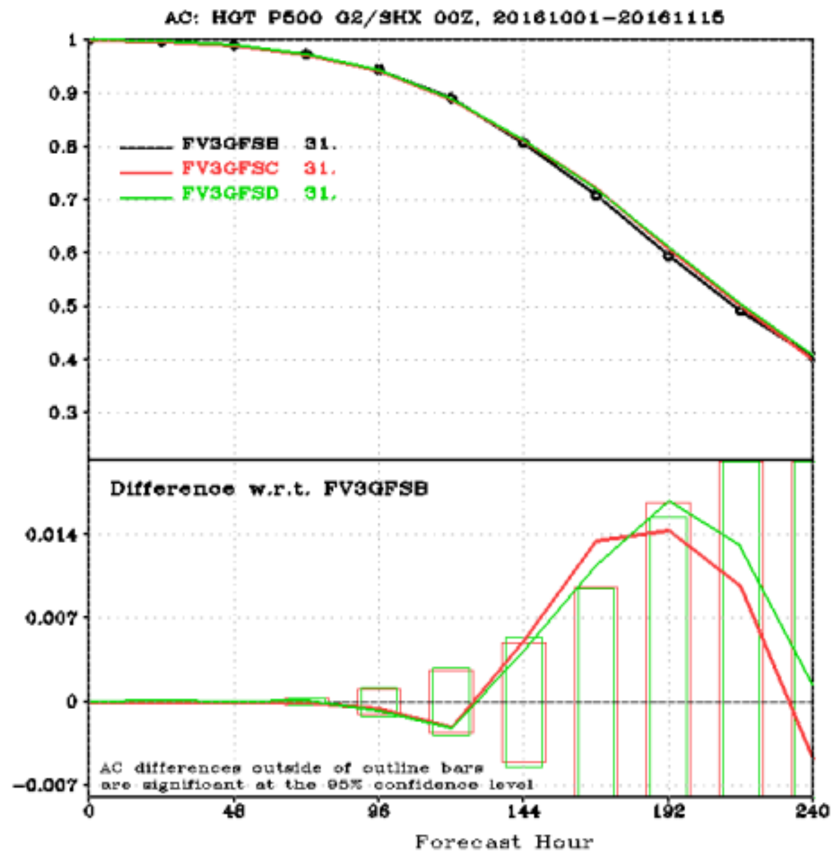
All ran at 64-bit, C768 resolution, with operational GFS ICs for the 01oct2016-06nov2016 period.

SJ-Lin and Lucas Harris:

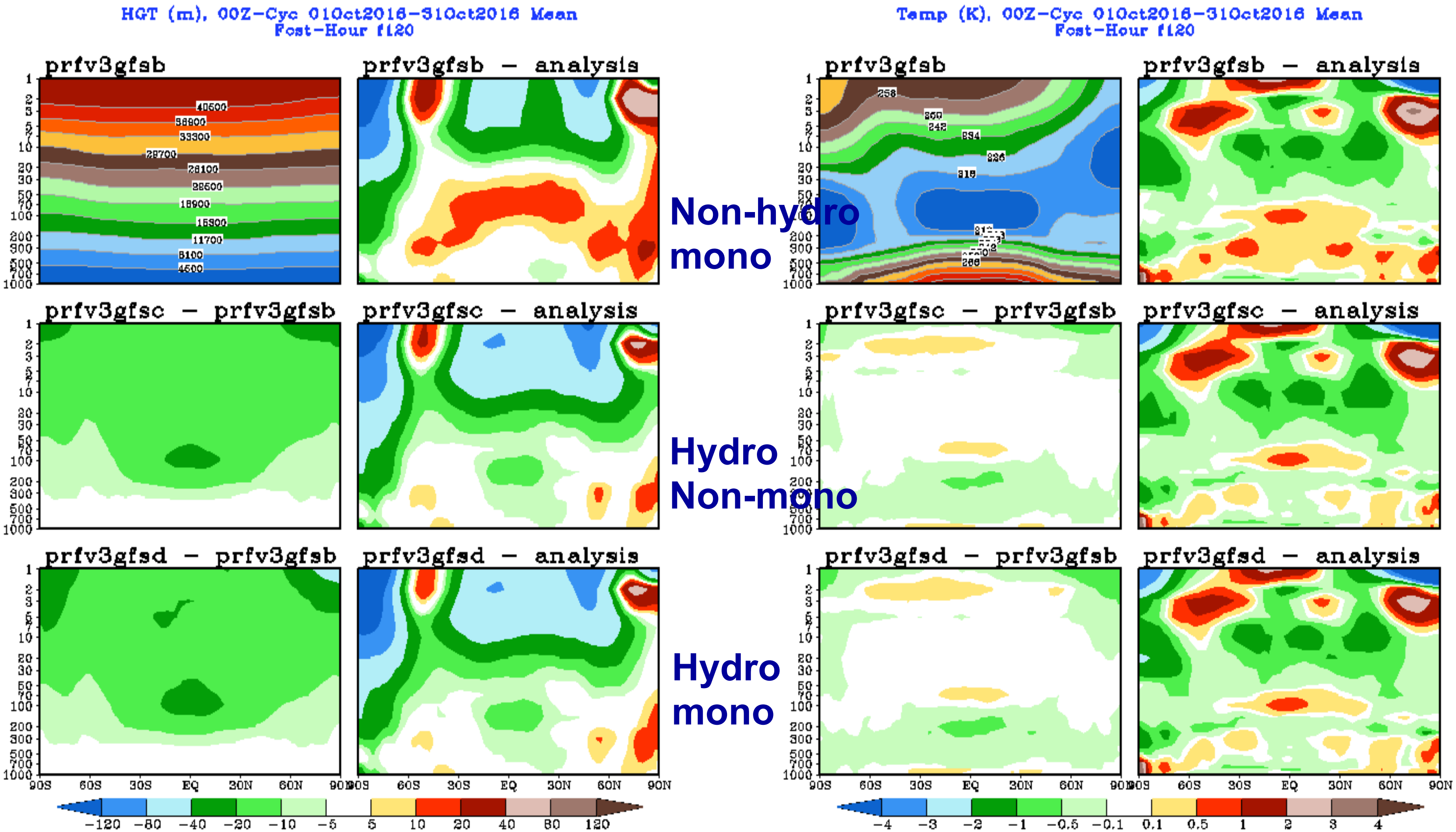
- **Monotonic** Scheme -- is intrinsically diffusive to 2-delta waves. Needs explicit horizontal 4th-order divergence damping. No vorticity damping.
- **Non-monotonic** Scheme – applies no monotonicity constraint. Needs damping to vorticity and momentum fluxes.

500hPa HGT AC

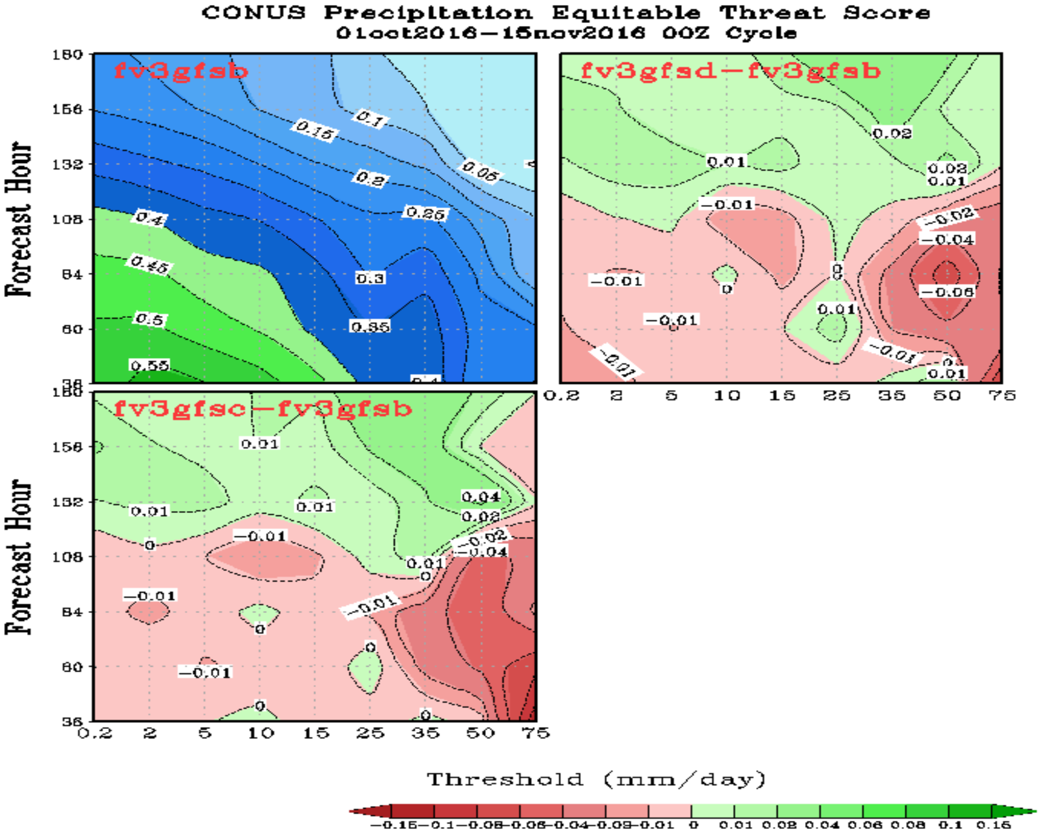


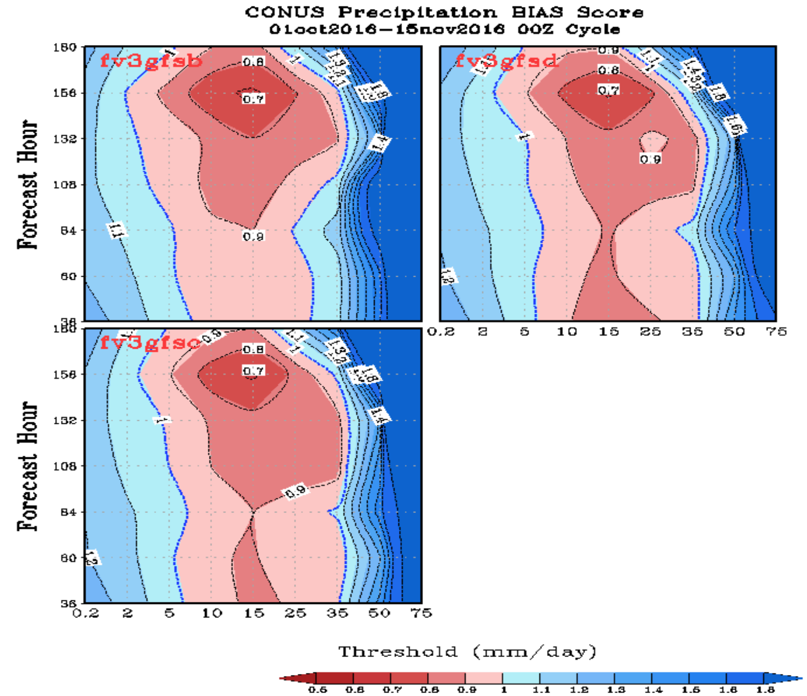


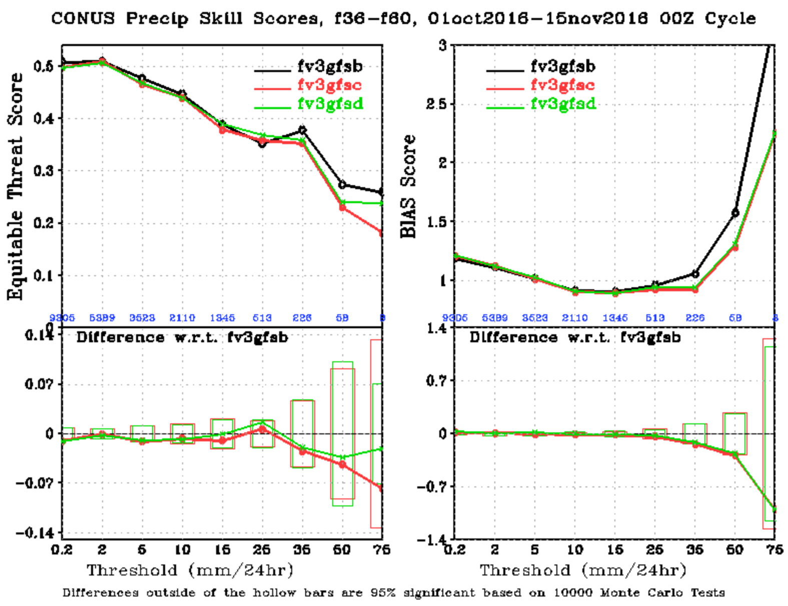
Hydro slightly better than non-hydro !



- **Hydro is colder than non-hydro in the troposphere**
- **Difference in the stratospehre between hydro and non-hydro is small.**
- **Hydro is too cold near the tropopause, while non-hydro is too warm**



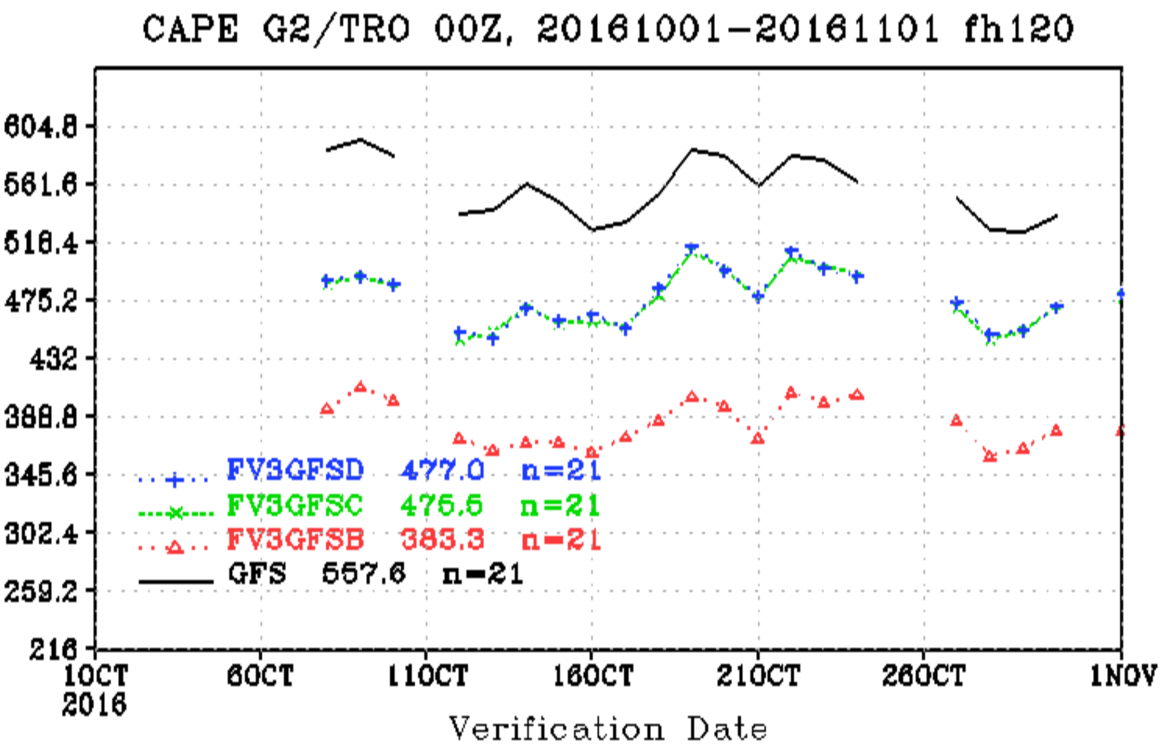


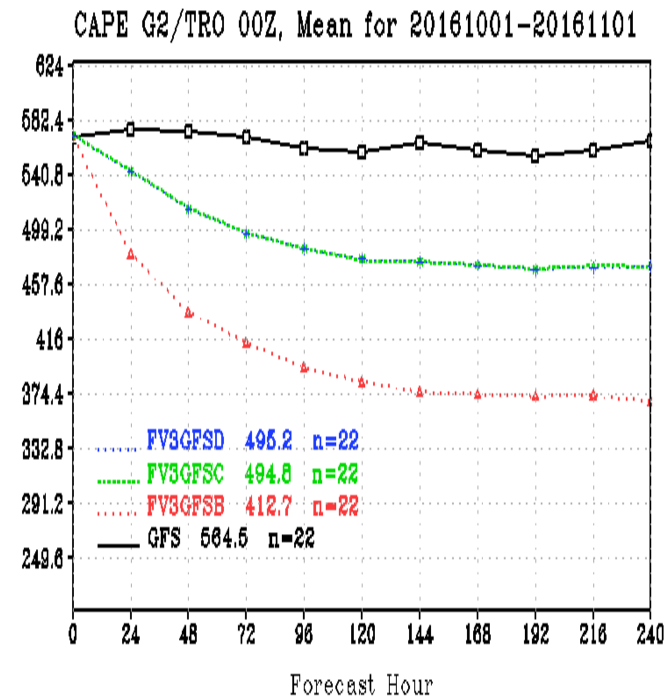


CONUS Precipitation Score:

Hydro is worse than non-hydro for short-range forecasts, but difference is insignificant. Need to test for summer, and with more samples

Tropical CAPE





- **Non-hydro CAPE is ~35% less than that ops GFS CAPE**
- **Hydro CAPE is much large than non-hydro CAPE**

Concluding remarks

- Development of workflow scripts and utilities for **data assimilation** are still ongoing.
- A **new benchmark** needs to be established using IPD v4.0 and compared to Q3FY17 NEMS GFS.
- More sensitivity experiments need to be carried out to better understand the model behavior with different parameter options. Different applications (e.g. deterministic vs ensemble forecast) might require the use of different settings (e.g., 32-bit vs 64-bit; non-hydro vs hydro) to achieve **best computing efficiency and forecast performance**.
- Given that the model behaves quite differently between hydro and non-hydro, **ENKF and high-res deterministic** may have to be run with the same option.

THANK YOU

Resolution, Physics Grid, and Run-time on Cray

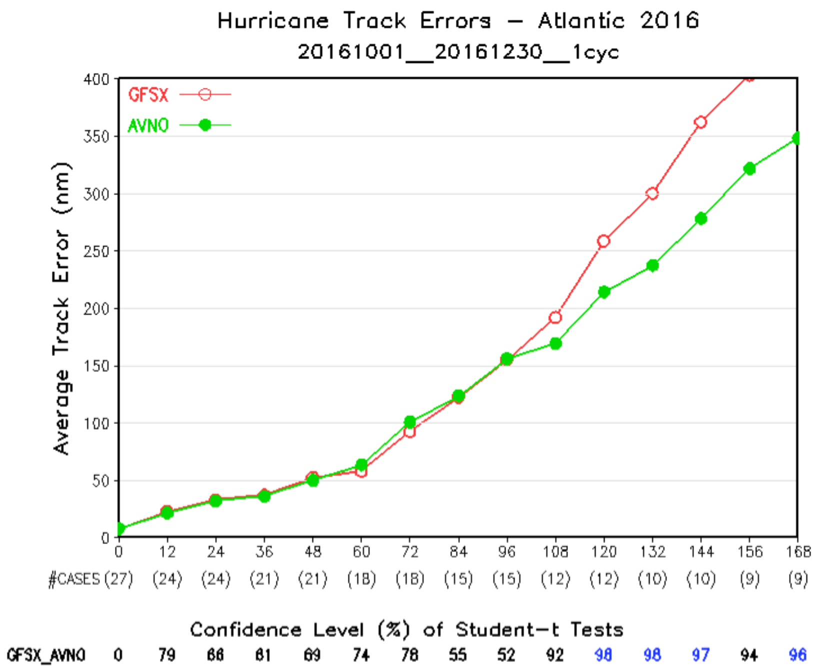
10-d forecast, 6-hourly output, non-hydrostatic, 3.75-minute time step,

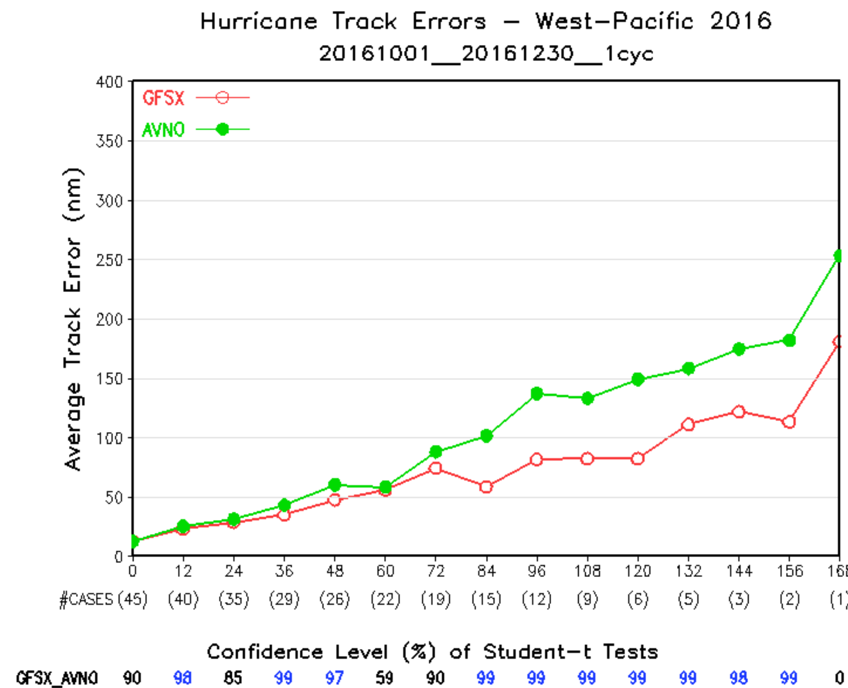
C#L63	Res	#Physics Grid (6xCxC)	Runtime (minutes) 2threads
C48	200km	13,824	**32-bit, 16 nodes: 3.6**
C96	100km	55,296	**32-bit, 16 nodes: 7.2**
C192	50km	221,184	**32-bit, 16 nodes: 21.0**
C384	25km	884,736	**32-bit, 16 nodes: 75.7**
C768	**13km**	**3,538,944**	**32-bit 16 nodes: 323**
C1152	8km	6,291,456	
C3072	3.5 km	9,830,400	

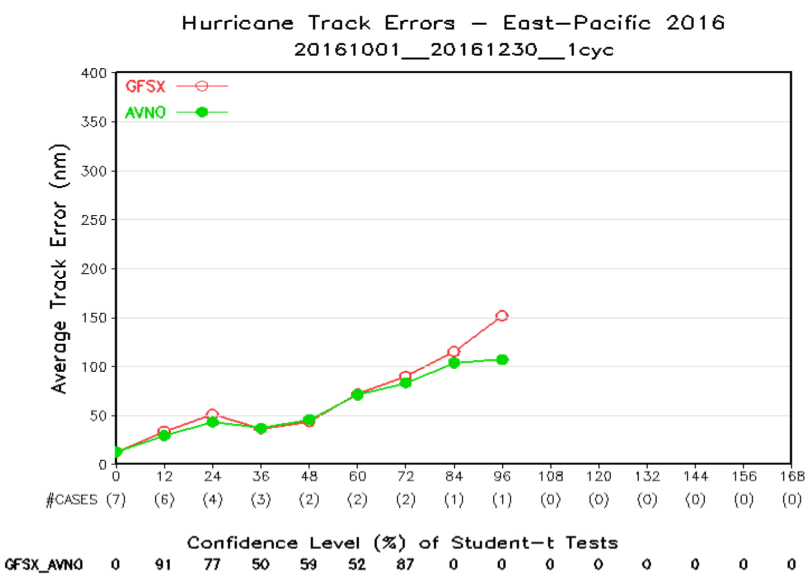
T1534 NEMS GFS (~13 km, 3072x1536), 61 nodes, 73 minutes

Thumb of rule : 10,000 / C# is grid size

Hurricane track

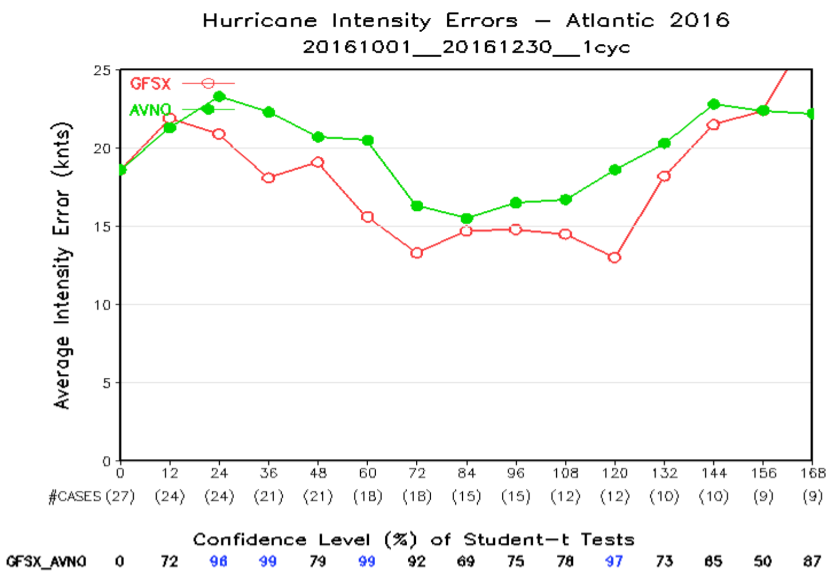


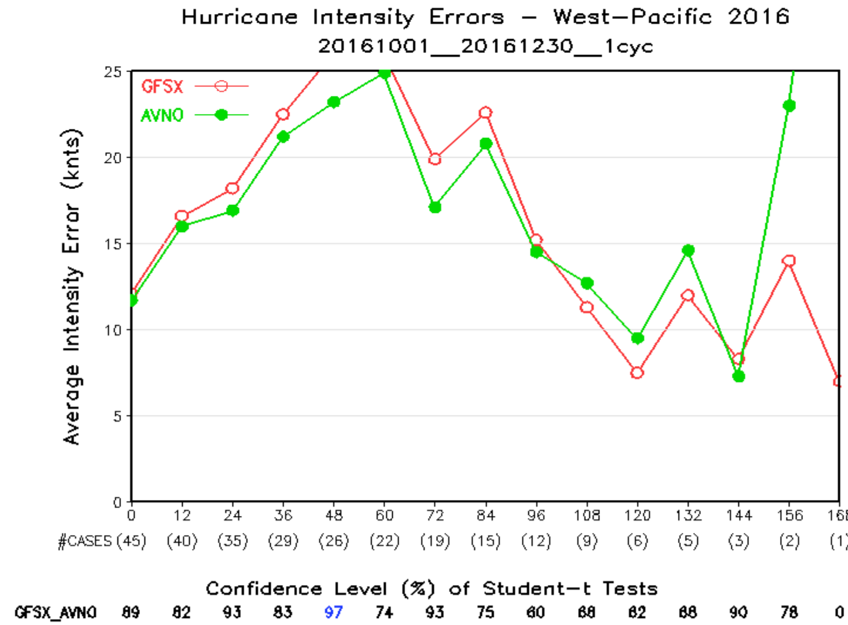


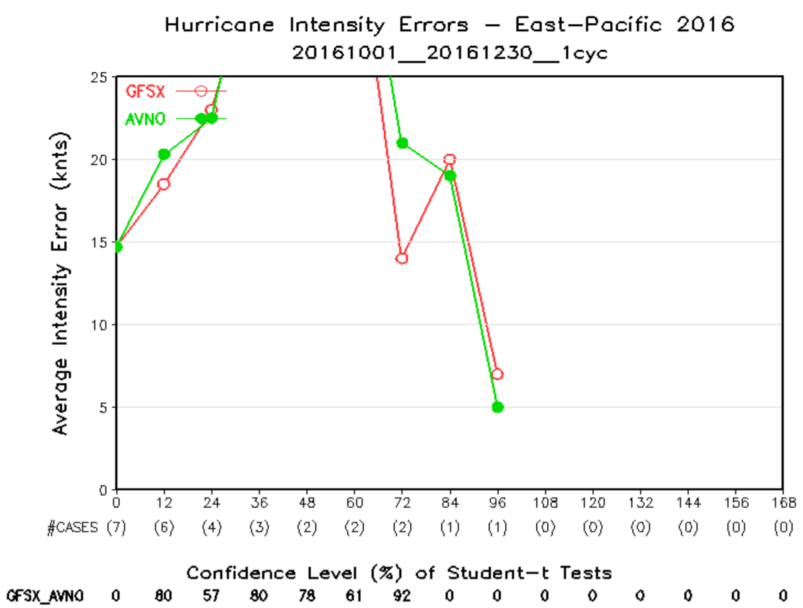


Limited sample size and only 00Z cycle !

Hurricane Intensity



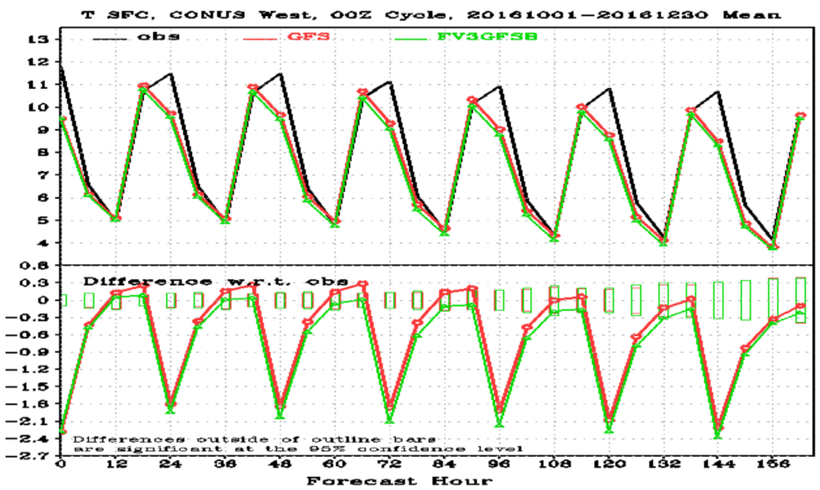




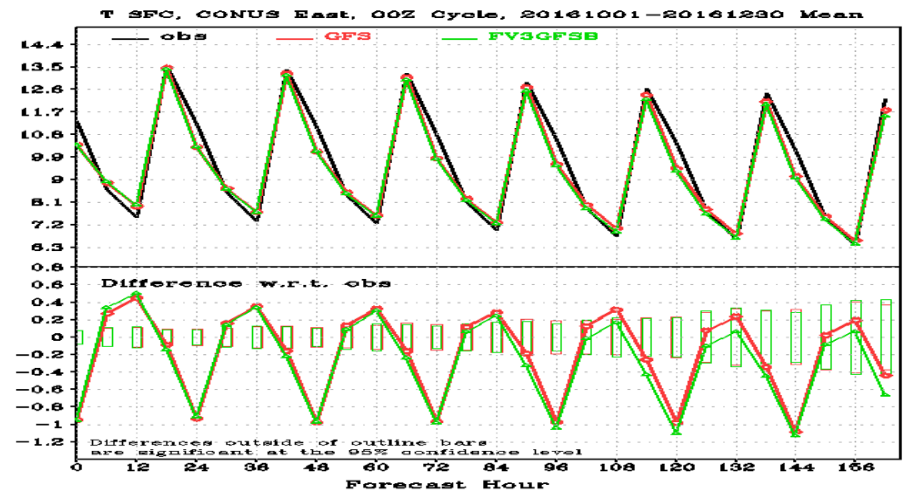
Limited sample size and only 00Z cycle !

Verified against Station Obs

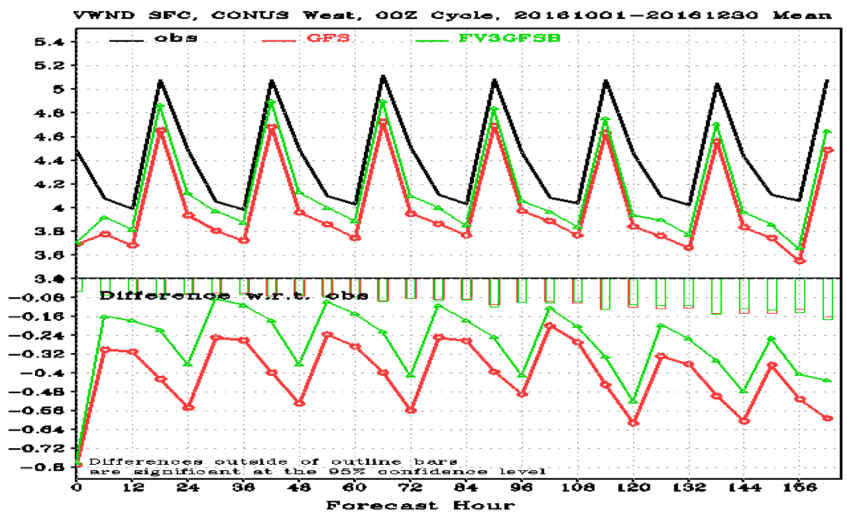
T2m, CONUS WEST



T2m, CONUS EAST



U10m, CONUS WEST



U10m, CONUS EAST