Data Assimilation for FV3GFS: Preliminary Results and Plans for JEDI
...and coupled modeling update

Daryl Kleist
NOAA/NWS/NCEP/EMC

The 2nd Taiwan West Pacific Global Forecast System Development Workshop
Central Weather Bureau
Taipei, June 6-9, 2017

Many contributions from EMC DA team and ESRL collaborators. I need to especially thank Rahul Mahajan, Cathy Thomas, Jeff Whitaker, and Phil Pegion.
UKMET NWP Index Time Series
Courtesy Gilbert Brunet & Dale Barker

* Parameters: Surface pressure, 500hPa geopotential height, 250hPa/850hPa Winds; Forecast ranges from T+24h to T+120h
DA Components of 2017 GFS Upgrade

• Upgrading to the NOAA Environmental Modeling System (NEMS) infrastructure
  – IO change

• Near Sea Surface Temperature (NSST) Analysis

• Turn on VIIRS AMVs
• Introduce Log-Normal wind QC for AMVs
• Use GOES clear-air water vapor winds

• Inclusion of extra GNSS-RO observations
## FV3-GFS Data Assimilation (DA) Plan (FY2017-2020)

<table>
<thead>
<tr>
<th>FY17 (Q1)</th>
<th>FY17 (Q2)</th>
<th>FY17 (Q3)</th>
<th>FY17 (Q4)</th>
<th>FY18 (Q1)</th>
<th>FY18 (Q2)</th>
<th>FY18 (Q3)</th>
<th>FY18 (Q4)</th>
<th>FY19 (Q1)</th>
<th>FY19 (Q2)</th>
<th>FY19 (Q3)</th>
<th>FY19 (Q4)</th>
<th>FY20 (Q1)</th>
<th>FY20 (Q2)</th>
<th>FY20 (Q3)</th>
<th>FY20 (Q4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NOW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NOW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NOW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **FY17**
  - Adopt GDAS (4D Hybrid En-VAR) DA for FV3GFS
  - Testing, Evaluation and Operational Implementation of new satellite datasets (GOES-16, JPSS, COSMIC-2 etc.)

- **FY18**
  - Increase vertical resolution to 127 levels and increase GDAS resolution to 35 km

- **FY19**
  - Incorporate JEDI Unified Forward Operator and Modular GSI infrastructure
  - Develop and implement DA on native cubed sphere grid

- **FY20**
  - Further advancements of FV3GDAS Global-Meso-Marine unification (Unified DA Development)

Experimental (beta) implementation of FV3GFS
Rocoto-based workflow 90% done. We can now fully cycle NEMS-FV3 similar to NEMS-GSM

- GSI requires rectilinear grid (Gaussian or regular lat-lon). Utilizing intermediate NEMSIO-Gaussian grid files for DA. No substantive changes to GSI.
  - FV3 receives Gaussian grid *increments* and interpolates to native grid.

Using GSM-based climatological background error, analysis is effectively hydrostatic

No fully functioning stochastic physics or initialization

- Collaborating with NOAA/ESRL on SPPT, SHUM, SKEBS (nearly ready) and IAU (shortly thereafter)
  - TLNMC still applied within assimilation itself

Cycling of surface fields slightly different for now (no global_cycle)
Preliminary Experiment:
Different than NGGPS Phase 2 Testing

- GSM (Control)
  - T670L64 (~0.3 hPa top mid layer), no DFI
  - T254L64, 80 member EnKF
- NEMS-FV3
  - C384L63 (~1 hPa top mid-layer) * See Fanglin Yang’s talk
  - C192L63, 80 member EnKF

- Other details about both experiments:
  - Late/Early cycles, Long forecasts at 00 UTC
  - Hybrid 3DEnVar (87.5% ensemble), Ensemble recentering
  - No stochastic physics, extra inflation for stability
  - No additional initialization (DFI)
  - Both experiments start from operational state on Nov 30, 2015
Model Equivalents (06hr FGAT)
Our first result at EMC*

RMSE O-F (2015120600-2015121318)
Next Steps

• Start testing of stochastic physics (this week!)
• Integration and testing of IAU
• 4DEnVar benchmark testing including real-time at full resolution
• Work toward higher vertical resolution, advanced physics configuration

• In Parallel
  – NGGPS-era offers both a challenge and an opportunity for broader coordination and significant generalization
  – A new effort in the US on coordinated data assimilation development is gaining traction: JEDI
  – All slides and discussion that follow are courtesy of Tom Auligne (JCSDA Director) and Yannick Tremolet (JEDI Master)
Joint Effort for Data assimilation Integration (JEDI)

STRATEGY

1. Collective path toward National Unified Next-Generation Data Assimilation
2. Modular, Object-Oriented code for flexibility, robustness and optimization
3. Mutualize **model-agnostic** components across
   - Applications (atmosphere, ocean, land, aerosols, etc.)
   - Models & Grids (regional/global, FV3)
   - Observations (past, current and future)

OBJECTIVES

1. Facilitate **innovation** to address next scientific grand challenges
2. Increase **R2O** transition rate
3. Increase **science productivity** and code **performance**
Abstract interfaces are the most important aspect of the design.
Example from NOAA Operations

Point of view

Ecosystem = collaborative environment
- Code repo & Reviews (Bitbucket)
- Issue tracking (JIRA)
- Testing (Bamboo)
- Documentation (Confluence)
- Support (JIRA Helpdesk)
- Governance
  - Identify code utility
  - Define interfaces
  - Specify roles + authorities
  - Allocate resources
Roadmap connecting two simultaneous approaches

• **Top-down approach:** *start with clean slate...*
  - Gather users’ needs and produce requirements document.
  - Evaluate ECMWF OOPS beta version.

• **Bottom-up approach:** *start from existing codes...*
  - Incremental refactoring of GSI code
  - High-level modularization and polymorphism (OO code)
1. Split Data Assimilation into Pre-processor, UFO, Solver, (+Utils) independent libraries
2. Develop **flexible** UFO infrastructure
   - Read/write observations and model equivalent
   - Match observation types with model variables
   - Encapsulate+generalize+optimize interpolation of model fields to observation locations
3. Polymorphic version including CRTM interface, bias correction, QC, error estimation, etc.
4. Expand UFO to sea-ice, ocean (NCODA?), ...
Coupled Modeling Effort at EMC

• CFSv2 was implemented in 2011. Work ongoing toward coupled system for S2S applications.

• GOAL: NEMS-based system with FV3, MOM6, CICE5
  – NOAH-land inline for now with transition to NOAH-MP

• CURRENT: NEMS-based benchmark system with GSM, MOM5.1, and CICE5
  – Benchmarked skill with this system to see where we are, demonstrate NEMS capability, develop workflow, and build validation tools
Experiment

- UGCSBench: T574L64 SL GSM, MOM5.1 on CFSv2 tripolar grid (0.25 degree tropics, 0.5 degree global), CICE5

- Cold Start from CFSR initial conditions.
  - April 2011-March 2017, twice per month, 35 day forecasts
    - 144 cases run
  - Uncoupled runs with same system using bias corrected SST for lower boundary as additional sanity check
  - Compared directly to CFSv2 forecasts (lower resolution)

- From 144 cases, significant effort put into creating calibration climatologies

- Credit: Chris Melhauser, Malaquias Pena, Suru Saha, and Huug van den Dool
### CONUS 2-meter temperature

<table>
<thead>
<tr>
<th></th>
<th>UGCSbench</th>
<th>UGCSbench</th>
<th>CFSv2ops</th>
<th>CFSv2ops</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw</td>
<td>Sec</td>
<td>Raw</td>
<td>Sec</td>
</tr>
<tr>
<td>week1</td>
<td>88.3</td>
<td>93.2</td>
<td>77.9</td>
<td>92.0</td>
</tr>
<tr>
<td>week2</td>
<td>50.6</td>
<td>57.1</td>
<td>45.9</td>
<td>55.9</td>
</tr>
<tr>
<td>week3</td>
<td>24.6</td>
<td>28.3</td>
<td>19.3</td>
<td>24.0</td>
</tr>
<tr>
<td>week4</td>
<td>14.8</td>
<td>17.1</td>
<td>2.3</td>
<td>3.4</td>
</tr>
<tr>
<td>Week3&amp;4</td>
<td>26.9</td>
<td>32.4</td>
<td>14.2</td>
<td>19.3</td>
</tr>
</tbody>
</table>

UGCSbench is better than the CFSv2ops for all lead times.

SEC: Systematic Error Correction
Week 3 & 4
T2m AC - mask

UGCSbench

CFSv2ops

UGCS ( NEMS GSM+MOM5.1+CICE5 )
Saha, Melhauser, Pena et al.
Week 2 & 3: UGCSbench has higher skill than Uncoupled and CFSv2
The current “UGCS” (NEMS GSM+MOM5.1+CICE5) is working coupled model

Additional (planned) enhancements will only serve to make it even more competitive:

1. Replacing the spectral model with the GFDL FV3 dynamic core for the atmospheric model component (work underway)

2. Replacing the MOM5.1 with the more advanced GFDL MOM6 for the ocean model component (work underway)

3. Working towards improving the coupling physics with the new FV3 dynamic core (work underway)

4. Working towards an FV3 based weakly coupled data assimilation system, based on the hybrid EnKF approach to all component systems (work underway).

5. Working towards a full ensemble of coupled model members with consistent initial perturbations to all components.

6. Reanalysis and Retrospective forecasts for consistent and appropriate systematic error correction, as well as skill estimation.

7. Working towards a full end-to-end workflow infrastructure that includes full validation metrics (work underway).
Final Thoughts

• Significant progress made on DA component for FV3-GFS
  – Using Gaussian grid as interim solution
  – Stochastic physics and IAU are critical, high priority developments

• Longer term
  – Higher vertical resolution, advanced physics (new variables such as hydrometeors), non-hydrostatic analysis

• JEDI
  – Significant change, Huge opportunity
    • Co-development of GSPRO component of UFO would be mutually beneficial to CWB and NOAA

• NEMS-based coupled modeling progress
  – Extending our hybrid-ensemble capabilities to coupled model (weakly first)
Thank you!
NCEP Coupled Hybrid Data Assimilation and Forecast System

Data Assimilation

Coupled Model Ensemble Forecast

Coupled Ensemble Forecast (N members)

Ensemble Analysis (N members)

INPUT

OUTPUT

Adapted from Saha
Coding Aspects

• **Flexibility**
  - It should be easy to modify the system (new science, new functionality, better scalability...)
  - Code modification should not imply changes all over the place: Locality in the source code

• **Reliability**
  - The code must run without crashing.
  - The code must do what the user thinks it does
  - Lots of testing:
    - Internal consistency and correctness of results (this is not meteorological evaluation),
    - Mechanism to run all the tests easily,
    - Tests run automatically on push to source repository.

• **Modularity**
  - The forecasting problem can be broken into manageable pieces:
    - DA can be described without knowing the specifics of a model or observations.
    - Solvers can be written without knowing the details of the matrices and vectors involved.
    - Development of a new model dycore/grid should not require knowledge of DA algorithm.

• **Separation of concerns**:
  - All aspects exist but scientists focus on one aspect at a time.
  - Different concepts should be treated in different parts of the code.
DATA ASSIMILATION COMPONENTS for Atmosphere, Ocean, Waves, Sea-ice, Land, Aerosols, Chemistry, Hydrology, Ionosphere

**Observations**
- Obs. Pre-processor
  - Reading
  - Data selection
  - Basic QC

**Model**
- Unified Forward Operator (UFO)
- (Observations)
- (Obs + model equivalent)

**Background & Obs Error**
- Solver
  - Variational/EnKF
  - Hybrid
- Analysis Increments
  - Model Initial Conditions
  - Observation Impact (OSE, OSSE)
  - Situational awareness
  - Reanalysis

**PRIORITY**
- Verification
- Model post-proc.
- Cal/Val, Monitoring
- Retrievals
- Simulated Obs.
JEDI: Project Schedule

- **Phase 0: FY16**  Spin-up with JCSDA funds
- **Phase 1: Q2 FY18**  Technical specs + Prototyping
- **Phase 2: Q4 FY19**  Mature system demonstration
- **Phase 3: Q2 FY21**  Transition to operations
Phase II (NGGPS) Testing

- MPAS and FV3 interfaced to operational 4D ensemble-variational DA system
- Due to time and HPC constraints, tests run at reduced resolution (~50 km)
- 80 member ensemble, cycle started at 2015090100
- Differences with operational configuration:
  - No high-resolution control analysis
  - No static background error component (full ensemble used to maximize feedback between dycore and DA)
  - No digital filter or tangent-linear balance constraint
  - No stochastic physics in ensemble (multiplicative inflation increased to compensate)
- Baseline GFS experiment at T382 resolution for reference
- Assessing:
  - Work required to replace spectral dycore in GDAS
  - Whether issues arise that may not be evident when models initialized from ‘foreign’ analysis
Observation Fits Comparison

Vector Wind (left) and Temp (right) O-F (2015090500-2015092900)

https://www.weather.gov/sti/stimodeling_nggps_implementation_atmdynamics