Overview of Data Assimilation and Operational Requirements at EMC

Jacob R. Carley$^{ab}$, Rahul Mahajan$^{ab}$, and John Derber$^b$

$^a$ IM Systems Group
$^b$ NOAA/NCEP/EnvironMental Modeling Center

jacob.carley@noaa.gov
A (Very) Broad Overview

- What kinds of DA applications are in operations at EMC?
  - Ocean, land, and atmosphere
  - Variety of spatial and temporal scales
  - Variety of applications

- What are some of EMC’s current DA requirements?
  - Infrastructure + Software, e.g.:
    - Coding standards, reliability, portability, etc.
  - Science + Capabilities, e.g.:
    - Ensemble, variational methods, bias correction, observations

- Continue unifying DA efforts in EMC
  - This workshop could be an excellent facilitator
Data Assimilation at EMC: **Ocean** + Waves

- Global Real Time Ocean Forecasting System (Hybrid Coordinate Ocean Model)
  - Initialization comes from partnership with Navy
    - Navy Coupled Ocean Data Assimilation
    - Currently 3DVar
      - Temp., salinity, geopotential, u, v
    - Evolve toward use for 7 overlapping regions
  - In development:
    - A Hybrid 3DVar/EnKF Ocean Assimilation System
    - Joint effort between University of Maryland and EMC
    - LETKF data assimilation system has been implemented with GFDL’s newly released MOM6, the SIS ice model, and the Navy’s HYCOM ocean model

Thanks to EMC Marine Modeling Branch colleagues for this material.
Data Assimilation at EMC: Ocean + Waves

Objective: Provide wave data assimilation products for operational wave models and guidance at the National Weather Service.

Approach: Apply DA methodologies with minimum changes to forward model, to assimilation mechanisms and to measurements.

Variational

Objective: Development of Variational DA systems for significant wave height (Hs)

Based on Behringer et al 1998

Previous Experience: 2D-Var DA for Hs (Chen et al 2004), 5D-Var DA for SWAN (Orzech et al 2014)

Deliverable: Operational Analysis of the Multi_1 and GLW (deterministic) wave field. T2O: FY17.

Kalman Filter

Objective: Development of LETKF – WW3 systems for Hs

Based on LETKF (Penny et. al 2016)

Previous experience: Development of experimental prototype by Etala (2015), (sponsored by JCSDA).


Gridpoint Statistical Interpolation (GSI) in collaboration with Mesoscale Group

2DVar, hourly 2.5 km significant wave height analysis (RTMA/URMA Pondeca et al 2011). T2O: FY16

Thanks to EMC Marine Modeling Branch colleagues for this material.
Data Assimilation at EMC: Land

- Land Data Assimilation at NCEP: Noah land surface model
- Global
  - Run in semi-coupled mode with Climate Data Assimilation System v2 (CDASv2)
  - Provides daily initial land states to Climate Forecast System version 2
  - **Forcing**: CDASv2 atmospheric output, “blended” precipitation, snow.
    - Precip: satellite, gauge, CDASv2
    - Snow: Interactive Multisensor Snow snow cover and AFWA snow depth
- North America
  - Run in “uncoupled” mode.
  - **Forcing**:
    - Precip: gauge-based, radar/satellite disaggregated
    - Atmosphere: North American Regional Climate Data Assimilation System.
  - Climatology NLDAS for 30+ years provides anomalies used in drought monitoring
- Plan to move toward NASA Land Information System for Global applications (EnkF based)

Thanks to Jiarui Dong for this material
• Variational Gridpoint Statistical Interpolation (GSI) + EnKF
• GSI Underpins the vast majority of the production suite
  ○ 2DVar - Hourly analyses
    ■ RTMA/URMA
  ○ Global - Weather, Climate, Reanalysis
    ■ GFS, CFS, CFSR (Global Spectral Model)
  ○ Regional - Weather, Aviation
    ■ NAM (Nonhydrostatic Multiscale Model on the B-grid)
  ○ Short term - Hourly, Aviation, Near-term hazards
    ■ RAP, HRRR (WRF-ARW)
  ○ Tropical Storms
    ■ Hurricane WRF (WRF-NMM)
• Nearly every system that does not use GSI has an upstream dependency on a system that does use GSI
The GSI + EnKF system has the following broad capabilities:

- 3DVar, 3DEnVar, and 4DEnVar (with hooks for 4DVar)
- Variational QC
- Multiple minimization options
- Support for many models
- Variational bias correction
  - Aircraft and satellite radiances
- Balance
  - Moving towards appropriately balanced model start-up
  - Tangent Linear Normal Mode Constraint (primarily global application)
  - Clouds/precip./moisture/etc.
- Assimilates wide variety of observations (radar, conventional, satellite, etc.)
- GSI’s observation operators used by EnKF system
- EnKF system features two algorithms:
  - EnSRF and LETKF

First Requirement: Must be able to perform all current analysis capabilities
Requirements: QC and Observations

- Quality control is necessary
  - May be a part of upstream processing
  - Variational QC
  - Gross error handling within DA system
- Ability to adjust observation errors by instrument, platform, type, level
- Switches for turning on/off observations
- Uses all observations currently assimilated in operations, e.g.:
  - Satellite radiances, mesonet, METAR, upper air, Doppler radar radial winds, profiler, buoy, aircraft, GPS RO, etc.
Requirements: Reliability, Resiliency, Consistency, Reproducibility

- Must be reliable, resilient, and consistent
  - Code must run **on time** with zero tolerance for failure.
  - System must have methods to handle complications in production *without failing*
    - e.g., an error in observation processing/transmission leads to an observation file >20 times its normal size (yes this happens)
  - Inconsistent timings are **very** difficult to plan around.
- Reproducible
  - Results should not change when re-run with identical inputs
- Ability to run on operational machine within available resources and time window.

For Feb. 2016 the **on time** number was **99.957%**.
● Produce diagnostics at observation locations
  ○ Monitoring for all observations
  ○ Bias correction terms
  ○ etc.
● Produce diagnostics for monitoring the solution algorithm (e.g. minimization)
● Diagnostics specific for so-called ‘unconventional’ observations, e.g.
  ○ Satellite Radiances
    ■ Bias correction terms, usage, counts, weighting function, etc.
  ○ GPS RO
  ○ Ozone
Requirements: I/O and Configuration

- Multi-resolution capabilities, e.g.,
  - Analysis, background, and components of background error may be at different resolutions
- Inclusion of bias correction in control vector
- Easily add state and control variables
- I/O must be compatible with upstream and downstream applications.
  - WMO standard BUFR
  - Various model data formats
  - Standards can be changed, but requires considerable work
The EMC operational DA environment is complex
  ● Many upstream and downstream dependencies

Requirements in a nutshell:
  ● Must meet/exceed everything currently done (within reason)
    ■ e.g., observations assimilated, quality of the analysis, etc.
  ● Benefit from more collaborative engagement with the wider community
    ○ New/Improved algorithms
    ○ New observation operators
      ■ For current or future observation platforms and data types
    ○ Code generalization
    ○ Simplification
    ○ etc.

Thank you! Questions?