Met Office Data Assimilation: Status And Plans

Dale Barker

Blueprints for Next Generation DA Workshop, NCAR, 10/3/16
Outline of talk

• DA At The Met Office: Current Status

• Global NWP: Hybrid 4DVar/4DEnVar, ETKF/En4DEnVar comparison

• What else for the next 4-5 years

• After 2020: ‘Exascale’ Programme
Unified Model For Weather And Climate

NWP - Atmos, Ocean, Waves
Deterministic MOGREPS

Component Models
GA6, GL6, GO5, GSI6

Coupled AOIL Model
GC2.0

GloSea
Decadal
Climate Change &
UK-ESM1

DA Systems:
OPS/VAR/ETKF
SURF->LIS?
NEMOVAR

Emerging DA: Wave, Sea-Ice, AQ, Space Weather, Hydrology
Copernicus Atmosphere Service

Air Quality

- Global air quality – ECMWF
- Regional air quality
  - Multi model ensemble

Main Developments
- Expanded domain
- European verification
- Pollen (Birch) forecast
- Dynamic biomass burning
- Composition DA (3DVar initially)
UK Environmental Prediction Prototype (UKC1):
High-Resolution (1.5km) Coupled O-A-L-H…

Coupled – Uncoupled SSH Differences:

River routing (RFM model):

Huw Lewis
Data Assimilation For NWP and Climate Reanalysis

Observations

Weather/Climate Model

Model Evaluation

Observation Network Design

NWP (predicting the future)

Climate Reanalysis (reviewing the past)

Analysis

OPS

VAR

ETKF

UM
SINGV DA: Observation Feed

Adam Maycock

[Diagram showing various data feeds and their subtypes]
Main NWP Configs At The Met Office

Global
- 17km resolution (33km ensemble)
- 70 vertical levels (80km top)
- 48 hour forecast twice/day
- 6 day forecast twice/day
- Hybrid 4DVar DA/ETKF

UKV
- 1.5km UK model (2.2km ensemble)
- 70 vertical levels (40km top)
- 36 hour forecast eight times/day
- 3DVar DA
Main NWP Configs At The Met Office

Global
- 17km resolution (33km ensemble)
- 70 vertical levels (80km top)
- 48 hour forecast twice/day
- 6 day forecast twice/day
- Hybrid 4DVar DA, ETKF->En4DEnVar in 2017

UKV
- 1.5km UK model (2.2km ensemble)
- 70 vertical levels (40km top)
- 36 hour forecast eight times/day
- 3DVar DA -> Hourly 4DVar in 2017 -> Var/Ens DA?
PS37 (March 2016) SA Package: Variational Bias Correction + New Obs

Verification based on 79 days (10 April – 26 June 2015)

Verification Against Own Analysis

Verification Against Observations

NWP Index Change = +3.31

NWP Index Change = +1.43

Saunders 2015: MOSAC paper 20.20
Data assimilation: Hybrid-4DVar → Hybrid-4DEnVar

Hybrid-4DVar

- Background error covariance at beginning of window:

\[ B = \beta_c^2 \mathbb{B}_c + \beta_e^2 \mathbb{B}_e \]

Climatological covariance  Ensemble covariance

\[ \mathbb{B}_e = (C) \circ \mathbb{P}_e^f \]

Spatial localisation covariance  Raw ensemble covariance
Data assimilation: Hybrid-4DVar → Hybrid-4DEnVar

Hybrid-4DEnVar

- No PF model, but much more IO required to read ensemble data.

- Analysis consists of two parts:
  - A 3DVar-like analysis based on the climatological covariance $B_c$
  - A 4D analysis consisting of a linear combination of the ensemble perturbations.

- Localisation is currently in space only: same linear combination of ensemble perturbations at all times.

- Can use as basis for EDA: Ensembles of 4DEnVar - En4DEnVar
Data assimilation: Hybrid-4DVar $\rightarrow$ Hybrid-4DEnVar

Trials using the current ETKF-based ensemble
(Adam Clayton, on secondment to KMA)

- Can get small improvements over the KMA operational configuration if we switch from 23 to 176 members in hybrid-4DVar.
- But hybrid-4DVar performance gets worse when we give the ensemble a high weight.
- However, hybrid-4DEnVar benefits.
- Hybrid-4DEnVar does not beat non-hybrid 4DVar.
Hybrid DA trials using the En-4DEnVar ensemble

Summary of results
(Trials run on ECMWF and KMA computers)

• Hybrid-4DVar performance is now best when we give the ensemble a high weight.
• Hybrid-4DEnVar can now beat non-hybrid 4D-Var.
• The gap between hybrid-4DVar and hybrid-4DEnVar is slightly larger.
Global Ensemble

- The initial version of the new En-4DEnVar ensemble
  - is a little worse than the current ETKF-based ensemble according to standard ensemble metrics,
  - but produces better error covariances for data assimilation.
- Further scientific and computational enhancements will be made to replace the ETKF with an En-4DEnVar ensemble in 2017-2018.

Global Data Assimilation

- Hybrid-4DEnVar is consistently ~2% worse than Hybrid-4DVar with same ensemble size.
- With large enough (N=200) ensemble, Hybrid-4DEnVar can beat non-hybrid 4DVar.
- Hybrid-4DVar will remain the operational global DA scheme until at least 2020.
- 4DEnVar development will focus on next-generation post-2020 ‘Exascale DA’ system.
Post 2020: Computational Challenges

• 1-4 Million cores
• 100-200 MW for full ECMWF suite on XC30 technology
• ‘Only 5 MW affordable’.

Two challenges:
• Alternative hardware
• Codes that can effectively use them
LFRic
Lewis Fry Richardson
Weather Prediction by Numerical Process, 1922

- Met Office programme of work to deliver replacement to Unified Model
- Born out of discussions at INI
- Builds on GungHo (new dynamical core) and in particular its Computational Science workpackage recommendations

工合
**Algorithm** - high level abstraction of the model express the algorithm in terms of global fields and operations on those fields (kernels)

**PSy** - auto generated parallel layer, unpacks data, inserts communication directives, optimises kernel calls and performs horizontal looping

**Kernel** - scientific code that operates column-wise on unpacked data. Many degrees of freedom per cell (column). Multiple contributions for DoFs on shared entities.
Met Office Requirements

• A suitable replacement for the UM, so:
  
  • Good software development practices:
    
    o Config management, review, testing, documentation, modular design, good style
    
    o Provide well-supported software environment
      
      ▪ Support on main platforms such as Linux desktop and Cray should be available early 2016
  
  • Support physics, DA requirements
  
  • Aligned with existing UM system working practices
    
    o Trac and Subversion for code management
    o Cylc and Rose for running
    o Rose Stem for integration tests
LFRic scope and relationship to wider Exascale work

LFRic
- Physical model
- LAMs
- Data structure
- I/O in the sense of interfacing to other models etc
- Support for diagnostics output and processing
- Idealized tests

LFRic System(?)
- CreateBC
- CAP
- RCF
- User experience
- Licences/IPR
- Fault tolerance?

Visualization and Diagnostics etc
- Links with AVD, Informatics Lab, etc

I/O
- Design and implementation of an exascale system.

Data Assimilation
- Ensemble processing
- Analysis
- Data processing
- ...??

Coupling issues
- Oversight of the coupling between the various models
- Maybe handling of external organisations (couping in a broader sense)?
Thanks for listening.

Any Questions?
Global Rapid Update Cycling (RUC) (Example 00UTC analysis shown)

Rapidly updating global provides:
• Much greater flexibility to decide when to run global forecasts
• Smoother transition between subsequent analyses.
• DA (affordable via preconditioning) benefits e.g. smaller increments
• Basis for unified global NWP/cloud analysis system