Joint Effort for Data Assimilation Integration

Status Update

Tom Auligné, Director, Joint Center for Satellite Data Assimilation with inputs from Y. Trémolet and JEDI team

NGGPS Bi-weekly meeting:: January 29, 2020
JEDI is a Joint Effort

**JEDI core-team:** Yannick Trémolet, Anna Shlyaeva, Benjamin Ménétrier, Clémentine Gas, Dan Holdaway, Mark Miesch, Mark Olah, Maryam Abdi-Oskouei, Ryan Honeyager, Steve Herbener, Xin Zhang

**JEDI contributors:** Andrew Collard, Ben Johnson, BJ Jung, Chris Harrop, Clara Draper, Cory Martin, David Davies, Emily Liu, François Vandenberghe, Guillaume Vernières, Hailing Zhang, Hui Shao, Jeff Whitaker, Jonathan Guerreette, Junmei Ban, Lou Wicker, Marek Wlasak, Mariusz Pagowski, Michael Cooke, Ming Hu, Rahul Mahajan, Ricardo Todling, Sarah King, Sergey Frolov, Steve Sandbach, Steve Vahl, Travis Sluka, Wojciech Śmigaj, Yali Wu, Yanqiu Zhu, Yunheng Wang...

**JEDI collaborators:** Chris Snyder, Dale Barker, Daryl Kleist, Nancy Baker, Ron Gelaro

**Representing:** JCSDA, NOAA/EMC, NOAA/ESRL, NASA/GMAO, NRL, USAF, NCAR, UKMO

And about 120 padawans who attended three JEDI Academies
Analyzed ice fraction aggregate

200 hPa T increment propagated 24h by GFSv15 on AWS (1,728 cores) in 7min20s

Adjoint Sensitivity to initial conditions @500 hPa
What we accomplished
How we got here
Assembling a Center of Excellence
OOPS (Object Oriented Prediction System)
Full data assimilation generic algorithms

UFO (Unified Forward Operator)
The ‘app-store’ of model-agnostic observation operators

CRTM (Community Radiative Transfer Model)
Accurately and efficiently simulate satellite radiances

SABER (System-Agnostic Background Error Representation)
Generic background error covariance modeling (incl. BUMP)

IODA (Interface for Observation Data Access)
Performs all the I/O of the observations

Abstract interfaces are the most important aspect of the design
The end of the monolithic gigantic jumble of code
Community Engagement and Support

3rd JEDI Academy

17th JCSDA Science Workshop

Marine IODA/UFO Code Sprint
<table>
<thead>
<tr>
<th>Step</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access <em>latest</em> code and build</td>
<td>10 min</td>
</tr>
<tr>
<td>Run test experiment (on laptop, Cloud, HPC)</td>
<td>10 min</td>
</tr>
<tr>
<td>Submit issue ticket and new development</td>
<td>5 min</td>
</tr>
<tr>
<td>Automatic testing</td>
<td>10 min</td>
</tr>
<tr>
<td>Peer-review</td>
<td>Same day</td>
</tr>
<tr>
<td>Merge code back to community code</td>
<td>5 sec</td>
</tr>
</tbody>
</table>

Automated testing tools: Travis-Cl, AWS CodeBuild, AWS ECR, Docker Hub.

Containers: AWS, ECR.

VCS: GitHub.

Code development process:
1. GitHub
2. Pull Request
3. Merge Code
Community Engagement and Support

- 7th JCSDA Symposium @Annual AMS Meeting. January 6-10, 2019. Phoenix, AZ
- Joint ECMWF/JCSDA Workshop. February 3-5, 2020. Reading, UK

Quarterly Newsletter, Seminars, Visiting Scientist Program, www.jcsda.org
Aug. 2017  First line of code, univariate B matrix, decision on OOPS
Nov. 2017  Introduction of Unified Forward Operator (UFO)
May 2018   Marine UFO
May 2018   4D-Var with FV3 (dry)
Aug. 2018   Multivariate B matrix
Dec. 2018  One month cycling 3D-Var (MPAS)
Aug. 2019   Introduction of generic QC filters
Oct. 2019  Marine DA transferred to EMC
Nov. 2019. 4D-Var with outer loops (Sept. for GEOS)
Dec. 2019   Generic QC filters for IR radiances
Jan. 2020   Cycling 4D-Var (FV3-GEOS)
Where we are today
# Models being Interfaced to JEDI

<table>
<thead>
<tr>
<th>MODEL</th>
<th>TYPE</th>
<th>INTERFACE</th>
<th>CENTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>FV3GFS</td>
<td>Atmosphere</td>
<td>fv3-jedi</td>
<td>NOAA-EMC</td>
</tr>
<tr>
<td>GEOS</td>
<td>Atmosphere</td>
<td>fv3-jedi</td>
<td>NASA-GMAO</td>
</tr>
<tr>
<td>FV3GFS GSDChem</td>
<td>Atmospheric chemistry</td>
<td>fv3-jedi</td>
<td>NOAA-ESRL</td>
</tr>
<tr>
<td>GEOS-AERO</td>
<td>Atmospheric aerosols</td>
<td>fv3-jedi</td>
<td>NASA-GMAO</td>
</tr>
<tr>
<td>MPAS</td>
<td>Atmosphere</td>
<td>mpas</td>
<td>NCAR</td>
</tr>
<tr>
<td>WRF</td>
<td>Atmosphere</td>
<td>wrf-jedi</td>
<td>NCAR</td>
</tr>
<tr>
<td>LFRic</td>
<td>Atmosphere</td>
<td>lfric</td>
<td>Met Office (UK)</td>
</tr>
<tr>
<td>MOM6</td>
<td>Ocean</td>
<td>soca</td>
<td>NOAA-EMC</td>
</tr>
<tr>
<td>SIS2</td>
<td>Sea ice</td>
<td>soca</td>
<td>NOAA-EMC</td>
</tr>
<tr>
<td>CICE6</td>
<td>Sea ice</td>
<td>soca-cice6</td>
<td>NOAA-EMC</td>
</tr>
<tr>
<td>NEPTUNE</td>
<td>Atmosphere</td>
<td>neptune</td>
<td>NRL</td>
</tr>
<tr>
<td>QG</td>
<td>Toy model</td>
<td>oops</td>
<td>ECMWF</td>
</tr>
<tr>
<td>Lorenz 95</td>
<td>Toy model</td>
<td>oops</td>
<td>ECMWF</td>
</tr>
<tr>
<td>ShallowWater</td>
<td>Toy model</td>
<td>shallow-water</td>
<td>NOAA-ESRL</td>
</tr>
</tbody>
</table>
## FV3 Model Interfacing Status

<table>
<thead>
<tr>
<th>Milestone</th>
<th>GFS</th>
<th>GEOS</th>
<th>FV3 Solo</th>
</tr>
</thead>
<tbody>
<tr>
<td>3DEnVar</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>4DEnsVar</td>
<td>✓</td>
<td>✓</td>
<td>NA</td>
</tr>
<tr>
<td>4DVar</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>4DVar with linear physics</td>
<td>X</td>
<td>✓</td>
<td>NA</td>
</tr>
<tr>
<td>Ensemble H(X)</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>4D H(x) in-core</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Multiple outer loops (IO)</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Multiple outer loops in-core</td>
<td>✓</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>Multiple resolutions</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>EDA</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Multiple resolution outer loops</td>
<td>X</td>
<td>X</td>
<td>✓ (simple B)</td>
</tr>
</tbody>
</table>
Static B and Cube-Sphere Poisson Solver

Initial D-Grid winds (correlation length scales ~200km)

\[ B = K_h K_v D C D^\top K_v^\top K_h^\top \]

- \( D \): Standard deviation
- \( C \): Correlation (BUMP)
- \( K_h \): Horizontal Balance (Poisson solver)
- \( K_v \): Vertical balance (BUMP)

Final D-Grid winds (correlation length scales ~4000km)

Stream function (and velocity potential)

Work done with John Thuburn (University of Exeter, UK) and Benjamin Menetrier (JCSDA)
<table>
<thead>
<tr>
<th>Observation Type (Instrument)</th>
<th>IODA obs file</th>
<th>$H(x)$</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aircraft</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Radiosonde</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Satwinds</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Additional conventional</td>
<td>✓</td>
<td>✓</td>
<td>Sfc obs, ship obs, wind profiler, etc.</td>
</tr>
<tr>
<td>AMSU-A</td>
<td>✓</td>
<td>✓</td>
<td>n15, n18, n19, metop-a, metop-b, aqua</td>
</tr>
<tr>
<td>AIRS</td>
<td>✓</td>
<td>✓</td>
<td>aqua</td>
</tr>
<tr>
<td>CRIS</td>
<td>✓</td>
<td>✓</td>
<td>npp</td>
</tr>
<tr>
<td>HIRS-4</td>
<td>✓</td>
<td>✓</td>
<td>metop-a, metop-b</td>
</tr>
<tr>
<td>IASI</td>
<td>✓</td>
<td>✓</td>
<td>metop-a, metop-b</td>
</tr>
<tr>
<td>MHS</td>
<td>✓</td>
<td>✓</td>
<td>n18, n19, metop-a, metop-b</td>
</tr>
<tr>
<td>VIIRS AOD</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>GNSSRO</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Marine (retrievals)</td>
<td>✓</td>
<td>✓</td>
<td>SST, SSS, SSH, Insitu Temp, Seaice (frac, thick)</td>
</tr>
<tr>
<td>Marine (radiiances)</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
UFO Status

- Specific observation operators being implemented
- Generic QC Filters
  - Entirely controlled from yaml configuration files (no coding, no compilation)
  - Filters are written once and used with many observation types
  - More filters will be developed as needed
- Generic observation bias correction
  - Use same generic concepts as QC (ObsDiag, ObsFunction)
  - Generic collection of predictors (controlled from yaml files)
- Current Status and Validation (vs. GSI):
  - Generic filters: 26 differences in QC flags out of 6M IASI obs (rounding errors)
  - VarBC predictors identical (machine precision) to GSI (IASI and AMSU-A)
- The system is getting mature
Generic QC Filters

About 18 pages (fairly small font!) of GNSSRO QC code replaced by generic filters and yaml configuration files

No factory reconfiguration (i.e. no recompilation)

**Future-proof:** can accommodate unplanned configurations

ObsTypes:
- **ObsOperator**:
  - name: GnssroBndGSI

ObsFilters:
- **Filter**: Domain Check
  - variable: impact_height
    - minvalue: 0
    - maxvalue: 50000

- **Filter**: Background Check
  - variables: [bending_angle]
    - threshold: 3.0

- **Filter**: ROobserror
  - variable: bending_angle
  - errmodel: ROPP

- **Filter**: Domain Check
  - variable: occulting_sat_id
    - is_in: 825

Rejecting KOMPSAT5 rising profiles

Using ROPP error estimates
Operational QC Procedures for IR

Operational (GSI) QC Flowchart for Infrared Sounders

- Wavenumber Check
  - Satellite Zenith Angle Check
    - GoES Sounder (true)
    - \( \theta_{sat} > 60^\circ \)
- Observational BT Check
  - Topography Check
    - Model Top Transmittance Check
  - Large sensitivity to Tskin
- Sensitivity Check
  - Cloud Check
    - Large sensitivity to Tskin
  - \( \tau (I_{clb} \rightarrow TOA) > 0.02 \)
- NSST Check
  - CrIS over Land (true)
  - CrIS over Land (false)
  - Tzr_qc = 1
- General Error Inflation
  - Based on Jacobians from sfc emissivity, skin temp and sfc type
  - Tzr_qc = 0
- Latitude Check
  - Tzr_qc = 1
  - Tzr_qc = 0

Minimization

- GoES Sounder (false)
- \( T_p > 550 \) or \( T_p < 50 \)
- \( \nu > 2400 \text{ cm}^{-1} \)

Each box should have its own check. Checks between GSI and UFO.

- varinv values are modified for all processes
- errf values are modified at processes marked with *
- Additional QC for Ozone:
  - Set Ozone Jacobian to zero for QC_NoIRJocO3_Pole is true and obs latitude > 60 degree

A process of observation **Error Inflation** or called **Inverse of Error** (varinv) **Reduction** and **Bound** (errf) **Tightening** from their original values
Section of YAML for Cloud Detection QC

```yaml
# Cloud Detection Check
- Filter: Bounds Check
  filter variables:
  - name: brightness_temperature
    channels: *all_channels
  test variables:
  - name: CloudDetect@ObsFunction
    options:
      channels: *all_channels
      use_flag: [ 1, -1, -1, -1, 1, -1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, -1, -1, 1, -1, -1, 1, -1, -1, 1, -1, -1, 1, -1, -1, 1, -1, -1, 1 ]
      use_flag_clddet: [ 1, -1, -1, -1, 1, -1, -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, -1, -1, 1, -1, -1, 1, -1, -1, 1, -1, -1, 1, -1, -1, 1, -1, -1, 1 ]
    obserr_demisf: [ 0.01, 0.02, 0.03, 0.02, 0.03 ]
    obserr_dtempf: [ 0.5, 2. 0, 4.0, 2.0, 4.0 ]
    maxvalue: 1.0e-12
    action:
      name: reject
```

IASI QC Comparison

Innovation & Observation Error after QC

Surface and Clod Sensitive Channel

UFO

Channel 1579 1039.5 cm⁻¹

GSI

IASI METOP-A 1039.500 cm⁻¹ Observation Error UFO

IASI METOP-A 1039.500 cm⁻¹ Observation Error GSI
A process of observation **Error Inflation** or called **Inverse of Error (varinv) Reduction** and *Bound (errf) Tightening* from their original values.
Section of QC YAML for Cloud/Precipitation with Strong Scattering

```yaml
# Hydrometeor Check
- Filter: Bounds Check
  filter variables:
  - name: brightness_temperature
    channels: *all_channels
  test variables:
  - name: HydrometeorChk@ObsFunction
    options:
    channels: *all_channels
    clwret_type: [ObsValue, HofX]
    clw_clr: [0.050, 0.030, 0.030, 0.020, 0.000, 0.100, 0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.030]
    clw_cld: [0.600, 0.450, 0.400, 0.450, 1.000, 1.500, 0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.200]
    obserr_clr: [2.500, 2.200, 2.000, 0.550, 0.300, 0.230, 0.230, 0.250, 0.250, 0.350, 0.400, 0.550, 0.800, 3.000, 3.500]
    obserr_cld: [20.000, 18.000, 12.000, 3.000, 0.500, 0.300, 0.230, 0.250, 0.250, 0.350, 0.400, 0.550, 0.800, 3.000, 18.000]
    maxvalue: 1.0e-12
    action:
    name: reject
```

MW QC Applications

MW QC for Surface and Hydrometeor Sensitive Channel

AMSU A N19  31.4 GHz  Innovation UFO

UFO

Innovation

GSI

Innovation
In-Core Data Assimilation – 4D H(x)

- GFS C768 (~12km) forecast model called from FV3-JEDI for 6 hour window beginning 2019-11-18 18Z.
- GFS v16 model.
- Background from operations.
- H(x) calculated in core as a post processor of the model step, no storing of 4D State anywhere.
- Interpolation is from C768 cubed sphere grid to observation locations.
In-Core Data Assimilation – 4DVar

- C768 background (from ops) and forecast.
- Native grid and resolution observer.
- Pure ensemble B matrix from C384 (25km) 40 member ensemble (from ops).
- C192 (50km) increment.
- All AMSU-A NOAA 19 (~20,000 obs).
- 3 hour window
- 2 outer loops in-core.
- BUMP for localization, interpolation etc.
Aerosol Data Assimilation

**GEOS GOCART**
- C90 (~100km) 3DEnVar
- 20 members
- 550nm Neural Network Retrieval of AOD
- ~70,000 observations
- Work done with Virginie Buhard (NASA GMAO)

**GFS GSD Chem**
- C48 (~200km) 3DEnVar
- 10 members
- CRTM simulated aerosol optical depth
- VIIRS and SUOMI-NPP
- Work done with Mariusz Pagowski (NOAA)
### Currently implemented in the EMC cycling workflow

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Forecast</th>
<th>3DVAR</th>
<th>UMD-LETKF</th>
<th>Hyb-EnVAR</th>
<th>Hyb-EnVAR + UMD-LETKF</th>
</tr>
</thead>
<tbody>
<tr>
<td>3°</td>
<td>MOM6</td>
<td>MOM6+CICE5</td>
<td>MOM6+CICE5</td>
<td>MOM6+CICE5</td>
<td>MOM6+CICE5</td>
</tr>
<tr>
<td>1°</td>
<td>MOM6</td>
<td>MOM6+CICE5</td>
<td>MOM6+CICE5</td>
<td>MOM6+CICE5</td>
<td>MOM6+CICE5</td>
</tr>
<tr>
<td>0.25°</td>
<td>MOM6</td>
<td>MOM6+CICE5</td>
<td>MOM6+CICE5</td>
<td>MOM6+CICE5</td>
<td>MOM6+CICE5</td>
</tr>
</tbody>
</table>

- **Currently implemented**
- **In progress**
- **Not implemented**

**Target system**

- **Cycling experiment**
JCSDA delivered to NOAA/NCEP the first version of JEDI-based next-generation community marine data assimilation

Cycling Experiment

- October 1, 2011 to November 2, 2011
- Coupled forecast model at ¼ degree resolution MOM6-CICE5-DataAtmosphere
- 24hr assimilation window ~1M obs per cycle
- 3DVAR with background dependent parametric B

<table>
<thead>
<tr>
<th>Variable</th>
<th>Satellite Sensor</th>
<th>In Situ</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST Infrared</td>
<td>NOAA-19, METOP-A AVHRR</td>
<td></td>
</tr>
<tr>
<td>SST Microwave</td>
<td>WindSat</td>
<td></td>
</tr>
<tr>
<td>Absolute Dynamic Topography</td>
<td>Jason-1 Jason-2</td>
<td>CryoSat-2</td>
</tr>
<tr>
<td>Ice concentration</td>
<td>F-16/F-17 SSMI, SSMI/S</td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td>Argo, CTD, XBT, TAO, PIRATA, RAMA, ...</td>
<td></td>
</tr>
</tbody>
</table>
Marine Data Assimilation

Forecast (sst, ice) 2011-10-01 to 2011-11-02

Temperature [°C]

Sea ice concentration [°C]

Preliminary Results
October 16, 2011

Forecast

Observations

sea_ice_area_fraction RMSD
region: r_gl

Date
Vision for future work
JEDI Planned Timeline

May 2020  Full set of generic QC filters
Jun 2020  Variational bias correction
July 2020  30-year ocean/sea-ice reanalysis (GODAS)
Aug 2020  Generic coupled UFO
Dec 2020  Full resolution cycling 4D-Var with outer loops
Q1 2021  Optimized ensemble (block) solvers
Q2 2021  Machine learning for QC and bias correction
Q3 2021  Continuous DA (depends on HPC resources)
Q4 2021  Coupled DA solver
2022    Coupled B matrix

2023  JEDI-GFS (global), JEDI-SAR (regional), JEDI-HAFS (hurricane), JEDI-GODAS (marine) become operational (JEDI-SFS in 2025)
Final Remarks about the JCSDA

**Conceptual leap:** More unified approach to algorithm development, observation processing, and maintenance of software.

**Streamlined processes and operations**
Unprecedented level of scientific/technical collaboration, coordination and accountability.

**Center of excellence for R2O/O2R**
Highly skilled staff committed to the success of JCSDA projects.

**JEDI system is getting mature and we are starting real-size T&E**
Read All About It!

jcsda.org/newsletters

IN THIS ISSUE

1 NEWS IN THIS QUARTER

The Joint Effort for Data Assimilation Integration (JEDI)

Joint Effort for Data Assimilation Integration (JEDI) Design and Structure

Status of Model Interfacing in the Joint Effort for Data Assimilation Integration (JEDI)

Observations in the Joint Effort for Data Assimilation Integration (JEDI)

NEWS IN THIS QUARTER

The Joint Effort for Data Assimilation Integration (JEDI)

Data Assimilation Challenges

All partners of the Joint Center for Satellite Data Assimilation (JCSDA) run data assimilation algorithms applied to their own models and applications. In 2001, the JCSDA was created to accelerate and improve the use of new satellite observing systems into each member’s data assimilation system. As Earth-observing systems constantly evolve and new systems are launched, continuous scientific developments for exploiting the full potential of the data are necessary. Given the cost and limited lifetime of new observing systems, it is important that this process happens quickly. This effort has been successful and continues to be; but, as the context evolves, new challenges emerge.
Questions?