Review of IV-TT Class4 Intercomparison Statistics for 2014-16

Fabrice Hernandez, Greg Smith, Jinshan Xu, Prasanth Divakaran, Andy Ryan, Charly Regnier and members of the IV-TT
Intercomparison and Validation Task Team (IV-TT)

Coordinates and promotes the development of a framework for the scientific validation and intercomparison of operational oceanography systems (OOFS) by

- Fostering scientific discussions on ocean physical and biogeochemical systems validation, link with CLIVAR and Climate community on common interests
- Defining metrics to assess the quality of analyses and forecasts ocean products
- Offering multi-system demonstrations and visibility to the community, link with JCOMM
- Leading to improvements of GOV’s systems

**Co-Chairs:** Fabrice Hernandez (Mercator Ocean) and Greg Smith (Environment Canada)

**Real time multi-assessment, and intercomparison started in 2013**

**Forecast accuracy of GODAE systems in 2013**

New metrics: Radar chart synthesis from the statistics of validation against observations and intercomparison of 5 operational global forecasting systems during 2013

Scores for 4 Ocean Essential Variables are provided along the four axes, normalized by the largest error

_Hernandez et al, JOO 2015_
Real time multi-assessment, and intercomparison started in 2013: demonstration and monitoring

Forecast accuracy with Class 4 metrics: comparison of 4 operational forecast against satellite altimeter sea level anomalies (correlation)
Together with the assessment of the multi-system ensemble estimates (grey)

0-100m salinity 1-day forecast: which system performs best in 2013

Ocean forecasts are compared to Argo profiles

*Ryan et al, Divakaran et al, JOO 2015*
IV-TT Class4 Intercomparison

• Brief History:
  – UK Met began providing SLA, SST and in situ T&S profile files on US GODAE ftp server in late 2012
  – January 2013 various groups started providing model equivalents
  – Several papers published demonstrating skill of GOV prediction systems for 2013 and providing a basis for the intercomparison:
    • Ryan et al. (2015); Divakaran et al. (2015), Hernandez et al. (2015)

• Here we provide an ‘annual update’ on the class4 intercomparison, present issues encountered to date and discuss future directions
Global Sea Level Anomaly
5 day lead time

- Most notable feature is change to reference dataset (AVISO)
- Update to AVISO included:
  - reduced filtering and decimation (i.e. increase in # obs)
  - 20yr reference MDT (instead of 7yr).
- Use of reference period in MDT and SLA obs needs to be the same. Lack of consistency is some feeds caused mean offsets.
- Statistics fairly consistent for 2015.
- Note that global mean SLA offset is arbitrary in systems that use a linearized free surface and thus STD should be used instead of RMSE when assessing error.
- Overall, a convergence of GOV prediction systems from 2013 to 2015 can be seen.
Global Sea Level Anomaly
5 day lead time

- Statistics fairly consistent for 2015/16.
- Overall, a convergence of GOV prediction systems from 2013 to 2015 can be seen.

May 2014: Updated AVISO feed
Global Sea Level Anomaly
5 day lead time

- Results for Sep 2015-Sep 2016
- Most notable feature is change to OMAPS
- GIOPS update on Jun. 25, 2016
- Strong consistency among 5 products
  - Potential for multi-model ensemble? (presentation by Jinshan)

May 2016: OMAPS updated
June 2016: GIOPS updated
Global Sea Surface Temperature
5 day lead time

- Strong seasonality of bias and RMSE can be seen.
- Strong sensitivity to spatial variability in observation coverage.
- Several interesting anomalies with strongly correlated errors
  - E.g. Jan-Feb 2015
- Large episodic errors in RTOFS since July 2015 symptomatic of processing issues experienced by various groups.
  - Improved quality assurance procedures and robustness required for operationalization.

Jan-Feb 2015: Errors strongly correlated between systems.
Global Sea Surface Temperature
5 day lead time

- Significant improvement in OMAPS associated with system upgrade to 0.1deg global resolution.
- Increase in RMSE for all systems in July 2016(?).
  - Needs further investigation...

Jan-Feb 2015: Errors strongly correlated between systems.
Mean annual T&S profiles

Temperature

Salinity

GOV IV-TT Workshop, Montreal, Canada, September 20-22, 2016
Mean annual T&S profiles

Temperature

Salinity

GOV IV-TT Workshop, Montreal, Canada, September 20-22, 2016
Mean annual T&S profiles

Temperature

Salinity

GOV IV-TT Workshop, Montreal, Canada, September 20-22, 2016
Summer 2016 T&S profiles
Salinity timeseries

Global In Situ Salinity, Depth Ranges: 0 to 50m

Global In Situ Salinity, Depth Ranges: 50 to 300m

Global In Situ Salinity, Depth Ranges: 300 to 1000m

Global In Situ Salinity, Depth Ranges: 1000 to 2200m

Error RMS and Bias

Num. Observations (k)
Temperature timeseries

Global In Situ Temperature, Depth Ranges: 0 to 50m

Global In Situ Temperature, Depth Ranges: 50 to 300m

Global In Situ Temperature, Depth Ranges: 300 to 1000m

Global In Situ Temperature, Depth Ranges: 1000 to 2200m
Salinity multiyear timeseries

Global In Situ Salinity, Depth Ranges:
- 0 to 50m
- 50 to 300m
- 300 to 1000m
- 1000 to 2200m

Error RMS and Bias

Num. Observations (k)
Extension to Sea ice concentration (AMSR-2)
5 day lead

- Feed in place since mid-2014.
- Currently GIOPS, PSY4 and FOAM contributing.
- Scores calculated using contingency table metrics:
  - Proportion correct total (PCT)
  - Proportion correct Ice (PCI)
  - Proportion correct Water (PCW)
  - Range [0,1]; 1 is perfect score
- Valuable contribution for Year of Polar Prediction (2017-2019)
Extension to Sea ice concentration (AMSR-2) 5 day lead

- Feed in place since mid-2014.
- Currently GIOPS, PSY4 and FOAM contributing.
- Scores calculated using contingency table metrics:
  - Proportion correct total (PCT)
  - Proportion correct Ice (PCI)
  - Proportion correct Water (PCW)
  - Range [0,1]; 1 is perfect score
- Valuable contribution for Year of Polar Prediction (2017-2019)
Summary

• Class4 Intercomparison is well-underway and showing improving robustness of data feeds and GOV prediction system skill.

• Results highlight relative strengths and weaknesses of various systems and allow a measure of progress.

• Provides valuable dataset to explore model errors.

• Extensions to sea ice and surface drift underway.
Issues to be resolved

• Common quality control of in situ profiles
  – Before distribution only, or by each centre?
• Improved quality assurance procedures needed to ensure robustness of feeds
• Communication protocols regarding updates to observations and GOV pred. systems.
• Are we ready to share the results?
Thanks!