Intercomparison and verification in operational oceanography: concepts, state-of-the-art methods and seamless approach considerations

Greg Smith and Fabrice Hernandez

Can we carry out joint activities in IV-TT and COSS-TT ?

Quality assessment objectives of operational oceanography

- Evaluate and **monitor performance** of operational system
  1. Model errors and trends
  2. Impact of the observing system
  3. Data assimilation efficiency
  4. Non constrained dynamics
- Evaluate **accuracy of products**:

  Quantify scales where, when and how close the product is to "truth"
  – Routine hindcast and forecast (skill)
  – Products derived from observation (RT or reprocessed)
  – Reanalyses: *Trends, How much better than RT products?*
- Measure **strength and weaknesses** of operated system for further improvements

Internal
Science driven

Provide timely robust and reliable products for a useful and cost effective service

- Assess **product's reliability considering user's needs**
  – Threshold, specific processes, predictability, timeliness related to accuracy

External
User driven

The validation « philosophy » adopted in our community

- **Basic principles. Defined for ocean hindcast and forecast**

  (Murphy, 1993, adopted in GODAE by Le Provost 2002, MERSEA Strand 1):

  - **Consistency: verifying that the system outputs are consistent with the current knowledge of the ocean circulation and climatologies**
  - **Quality** (or **accuracy** of the **hindcast)** **quantifying the differences between the system "best results" (analysis) and the sea truth, as estimated from observations, preferably using independent observations (not assimilated).**
  - **Performance** (or **accuracy** of the **forecast): quantifying the short term forecast capacity of each system, i.e. Answering the questions "do we perform better than persistence? better than climatology?…**

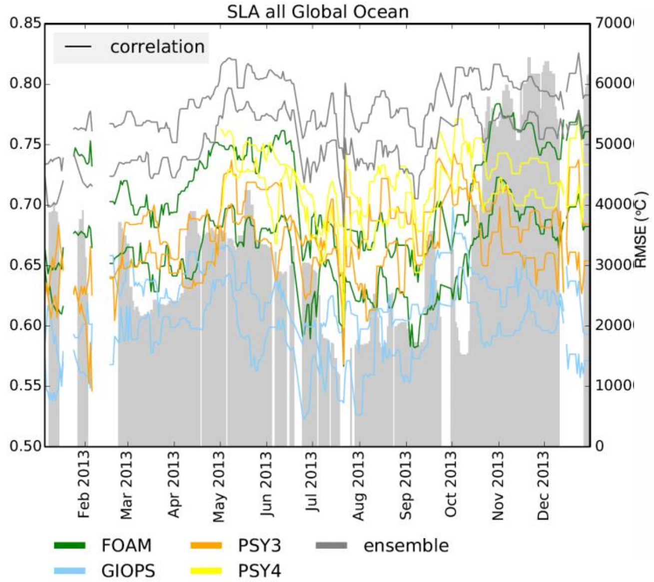
- **A complementary principle, to verify the interest for the customer (eg, Pinardi and Tonani, 2005, MFS):**

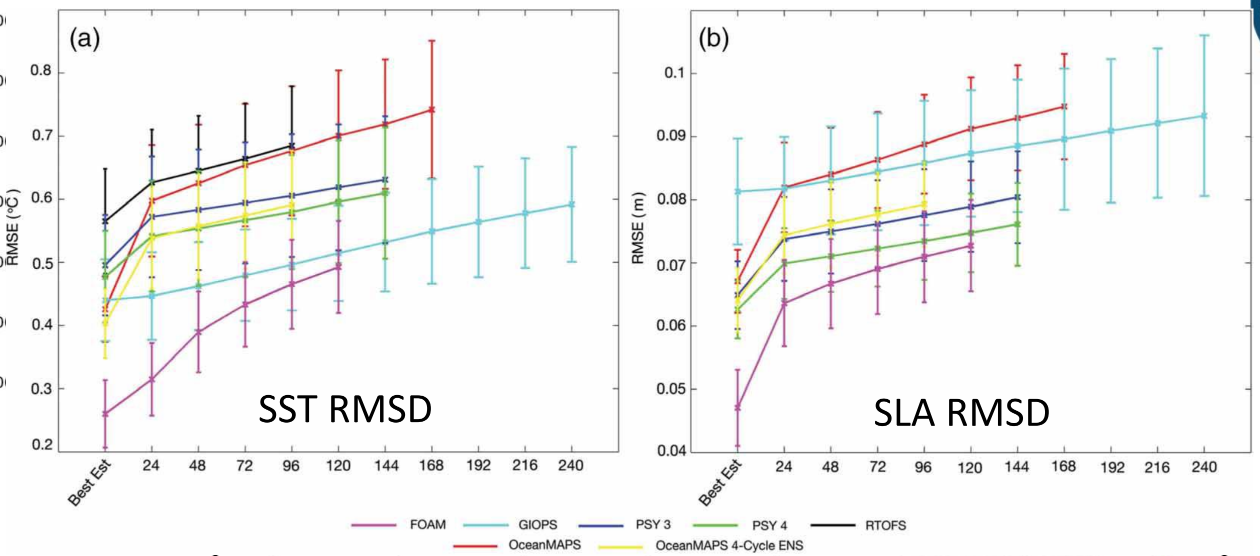
  - **Benefit: end-user assessment of which quality level has to be reached before the product is useful for an application**

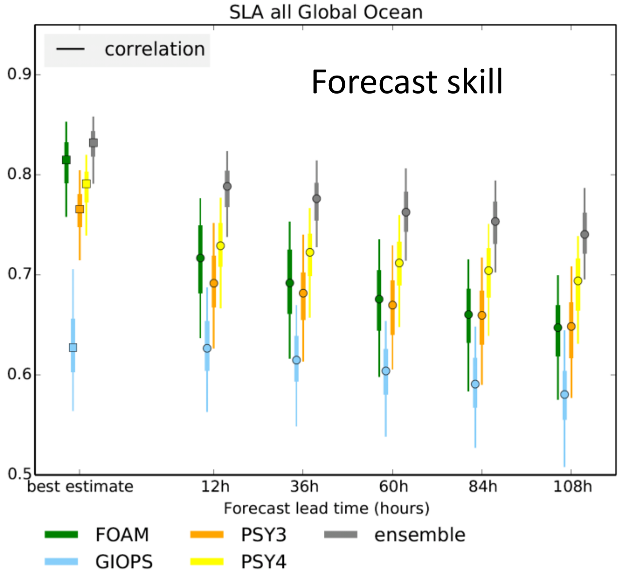
Quality assessment objectives of operational oceanography: Defining metrics:

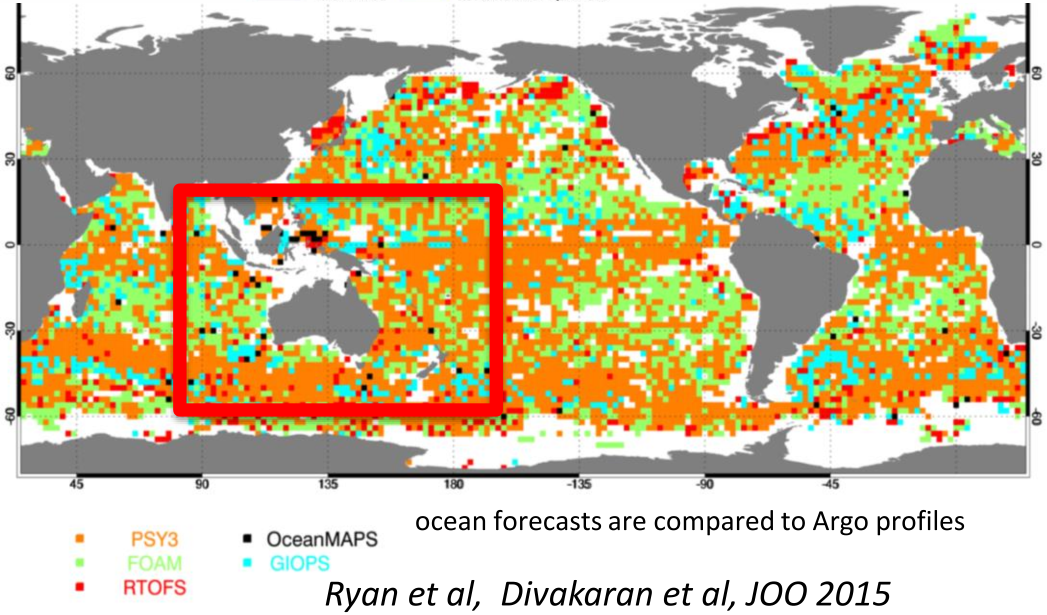
- Metrics that are **efficient for oceanographers**
- Metrics that can be **handled in similar ways** by all involved OOFS: **standardisation** (multi-system assessment, ensemble strength/weaknesses)
- Using **similar reference dataset**, where the physical content is understood (representativity)
- Metrics assessing OOFS performance that can be **computed in real-time** and off-line (reanalyses vs real time)
- Metrics issued from the assimilation scheme (misfits, residuals) OR independently
- Metrics using observations not assimilated (verification)

Example: Class4 Intercomparison









ocean forecasts are compared to Argo profiles

Ryan et al, Divakaran et al, JOO 2015

Verification challenge:

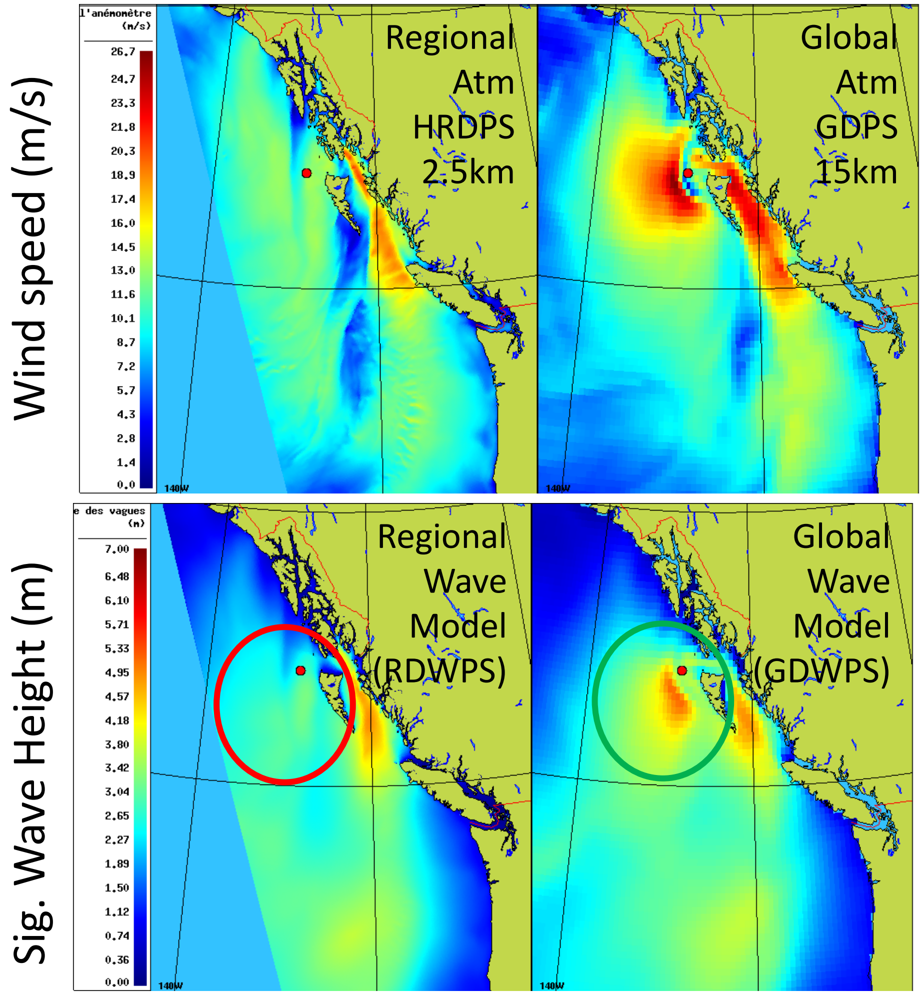
Consistency between offshore and coastal systems

- Remote effects can have important impacts
  - e.g. waves, storm surge, sea ice (conc., thick, pressure)
- Inconsistencies between atmospheric forcing
  - Higher resolution doesn`t always mean more accurate forecasts!
- OBC forcing frequency often a hinderence
  - E.g. aliasing of internal tides in daily mean forcing (ie global outputs)
- Products need consistency as well!
  - E.g. E-navigation: How do we ensure consistency in surface currents between systems?

E.g.: Canadian West Coast (2019-08-21)

Wind speed (m/s) at 27h lead

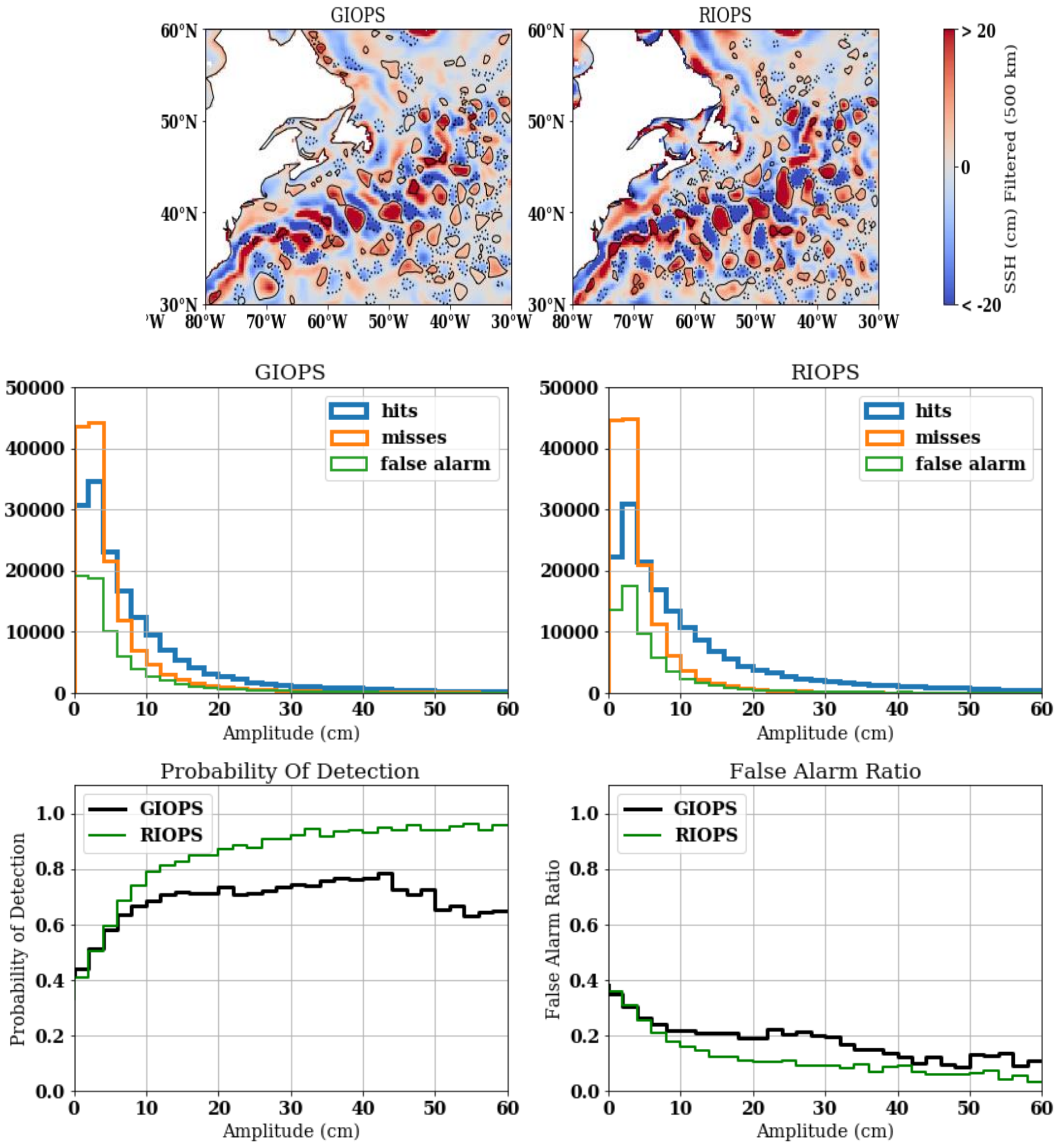
Max 26.7 m/s → storm force



Future directions:

Feature-based verification

- Significant body of knowledge and methods in NWP community (JWGFVR) in spatial verification methods (e.g. Casati et al., 2008)
  - IV-TT and JWGFVR developing links
- Spatial verification already applied to sea ice verification
  - Distance-based metrics (IIEE, SPS)
- Feature-based metrics could help to communicate strengths/weaknesses of forecasts to users
  - E.g. Eddy-tracking (Smith and Fortin, sub.)

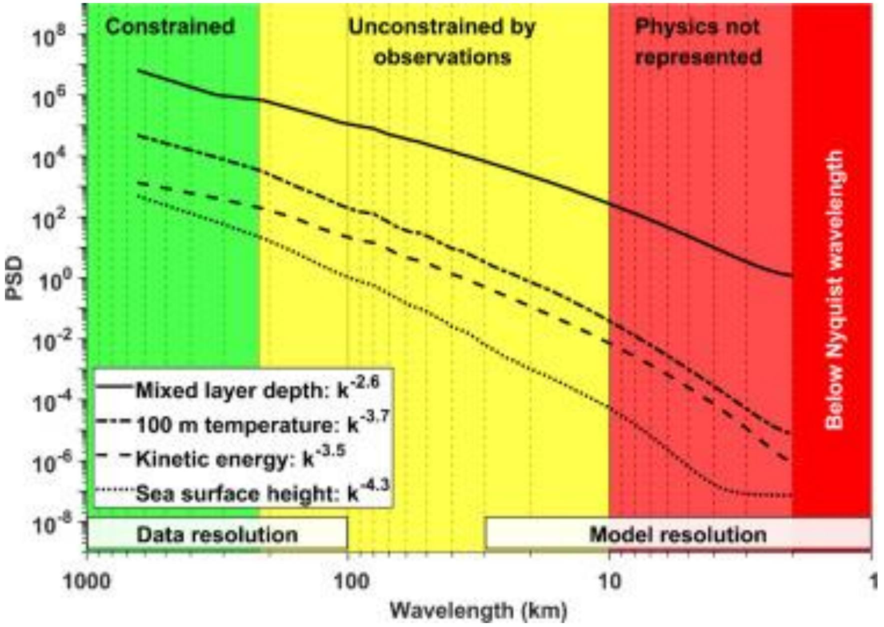


Categorical verification of eddies in Gulf Stream region in global (GIOPS, 1/4°) and regional (RIOPS; 1/12°) analysis systems. Demonstrates significant added-value of regional system in representation of eddies.

Future directions:

Scale-aware verification

- Direct comparison of fields that contain different scales of spatial and temporal variability will lead to incorrect error estimates
  - E.g. double-penalty errors
- Also need to consider representation of different processes and the scales constrained by data assimilation
  - Unconstrained variability contributes to error for many applications
  - A clear understanding of constrained scales needed to compare coastal to offshore systems to allow scale-aware verification and for reliable products
    - Jacobs et al (2021):
      - "Constrained scales contain deterministic skill"
      - "Unconstrained scales contain statistical skill"
  - If designed correctly, ensembles should provide estimate of unconstrained scales
    - Increasing resolution requires ensembles!



Jacobs et al. (Oc. Mod., 2021)

Quality assessment objectives of operational oceanography

Provide timely robust and reliable products for a useful and cost effective service

Operational oceanography is now facing same challenges and user's expectation than weather forecast, or climate assessment !

- From science driven to user driven
  - Challenges in understanding user's requests
- Operational oceanography **continuously evolving toward more complex system**
  - Coupling (atm, wave,ice, biogeochemistry, rivers) & nesting (scales)…
  - Diversity of products (blue, green and white ocean)
  - Focus: climate, seasonal, short-term predictions / open ocean, coastal
- Many applications… and decision makers
- Challenges in communicating product's reliability
  - **General public not really aware** of oceans behaviour (at the opposite of weather !)
  - Ocean intermediate **users** are **experts, with requirements in terms of quantified errors**

Quality assessment objectives of operational oceanography

Provide timely robust and reliable products for a useful and cost effective service

Operational oceanography is now facing same challenges and user's expectation than weather forecast, or climate assessment !

- These challenges are **enhanced into the regional/coastal approach**:
  - What is the paradigm in the next decade with 1/36° to 1/50° global systems, coupled with atmosphere ? What would be the coastal modeling framework ?
  - What scales are going to be observed, and constrained ?
  - What process/scales would need to be systematically validated/verified ?

Quality assessment objectives of operational oceanography: Defining metrics with a seamless approach toward the coast

- Try to apply metrics developed for **open ocean to coastal systems**
- Favour **multi-model** intercomparison and **ensemble** approaches
- Measure specifically:
  - **Benefit of nesting** approach (efficiency of enhanced physics, performance of regional/child vs global/parent system, benefit for specific applications)
  - Problems caused by **"parent" systems** (boundary forcing errors)
  - Possible benefit of **2-way nesting**
  - **Non constrained scales** on the coastal system
- Focus on **specific** coastal/regional **processes**, and quantify their representation in both regional and global systems
  - Sea Level (Rise), Mixed layer dynamics (T,S content, mixing, fronts), BottomT …

Quality assessment and seamless approach toward the coast

- Joint efforts in IV-TT and COSS-TT ?

  – Propose to evaluate a set of endorsed metrics (associated with reference obs), both in global and regional systems (in particular in OOF centres running both)

  – Establish a common synthesis by focusing on specific assessment: for example:

    - non-constrained processes and scales on the targeted OOFS?
    - Verification against a particular dataset? e.g., Sentinel-3 for SSH, toward SWOT ?