

Position Paper: GODAE OceanView Intercomparison and Validation Task Team in the context of OceanPredict : status and future activities

By F. Hernandez and G. Smith

Version v20190408

Executive summary:

- Since the beginning of GODAE, the Intercomparison and Validation Task Team (IV-TT) promotes and fosters innovative approaches in validation, verification and intercomparison activities among Ocean Operational Centres.
- The successful near real-time “Class-1&4” intercomparison activity has been on-going since 2013 and continues to grow involving several global and regional operational centres .This activity is an excellent resource for research that bring an enormous potential for innovative publishable results. It contributes directly to the scientific goals of the IV-TT and of GOV.
- The Class 1&4 activities are demanding and carried out routinely on a “best effort” basis by volunteering experts. Moreover, these quasi-operational tasks rely on non-dedicated IT resources. The reliability of human and IT resources is a significant weakness and on-going risk to the continued success of the Class 1&4 near real-time intercomparison.

In summary, the Class-1&4 intercomparison, due to its success, has become a quasi-“operational” task, performed and maintained routinely through “best efforts” of experts, with no dedicated community resources and no official support to ensure the continuity of the activity.

GOV IV-TT’s actions reach some limits that need to be revisited.

OceanPredict is an opportunity to re-define GOV IV-TT priorities, identifying and split IV-TT routine activities in a more operational framework like JCOMM/ETOOFS by:

- Transfer the “Class-1&4” intercomparison under the responsibility of a JCOMM/ETOOFS sustained framework, endorsed by all stakeholders, with structured, standardized procedures applied by all participants.
- Engage ocean forecasting centres for these routine tasks as part of JCOMM/ETOOFS, asking for dedicated resources (expertise, infrastructure) to support the activity.

This will allow the IV-TT expertise to reorient its focus on new emerging verification research topics; schedule and share research actions and projects with other GOV task teams; or initiate exchanges with other validation/verification communities.

Reminder

The GODAE OceanView Intercomparison and Validation Task Team (GOV IV-TT) is active since the beginning of GODAE, in order to **foster and promote innovative approaches in validation, verification and intercomparison** activities among Ocean Operational Centres.

Members of this task team have been developing validation frameworks in their respective operational centres. Some have been involved in a series of international projects, in particular with the CLIVAR/GSOP community (intercomparison of ocean reanalyses), the YOPP program in the Arctic etc... More important, the IV-TT is involved in the so-called “Class-4” intercomparison activity, carried on since 2013. In a lesser extent, there is also an ongoing IV-TT “Class-1” intercomparison initiative.

All these activities **rely on an active production of standardized metrics by each operational centre** involved, then a mechanism that allows all centres to share their results, bringing to an effective **comparison of global hindcasts and forecasts** among all the centres. The intercomparison synthesis has been performed over time by different volunteering experts of these centres, that eventually published some results.

This “Class-4” activity has been a success for IV-TT that could test innovative metrics for near real time verification activities since 2013, and **it appeared as a “success story”**, among others, for GOV. The number of participating centres has slightly been growing these past years. Starting with a core group producing global ocean forecasts, composed by the UK-Met (UK), the BoM (Aus.), Mercator Océan (Fr.), and NOAA/NCEP (USA), ECCO+DFO (Ca) rapidly joined, and more recently the NMEFC (China), NERSC & Met.No (No.) and CLS (Fr.) have been carried on the daily routine activity.

These activities are **performed under “best effort” resources by the different centers**. However, the “Class-4” activity is maintained with a high level of volunteer (“best-effort”) commitment in real-time by some experts in each center involved.

The technical success of the “Class-4” intercomparison activity has been spreading internally among other task teams. The OSEVal, CP, COSS and MEAP-TTs have interest in using part of the existing results or plan to mimic this kind of intercomparison framework. Externally, this pioneering activity is paving the way of more structured real time verification monitoring such as the one carried on by the CMEMS. Moreover, several ocean operational centers participating to the IV-TT activity are now using the “Class-4” results inside their own operational routine validation protocols and as part of system upgrade procedures.

Identified limits and weaknesses of the IV-TT activity

The “Class-4” and “Class-1” activities

The “Class-4” and “Class-1” activities are demanding, but carried out on best efforts. By the way these activities **have an enormous potential of innovative publishable results**. This is one of the reasons why individually, the different experts involved have enthusiastically dedicate time to maintain a high level of routine production. However, since its inception in 2013 we have seen some limits of this framework:

- The routine mechanism depends first on the volunteering center that provide the initial information. Since the beginning the UK-Met has done the effort to provide T,S, SST, SLA parameters “Class-4” information in near-real time. More recently ECCC has been providing the Sea Ice Concentration initial information. We have witnessed over time a series of outages of this production. Due to change in the operational chain at the UK-Met, with no adequate anticipation, because the UK-met expert was not aware or could not dedicate time to maintain/correct the procedures, or due to problems in the source of information gathered by the UK-Met, with no possible immediate corrective action by the expert.
- Extending the exercise to new parameters is also demanding. For several years, Mercator Océan has had some experience in velocity validation based on drifters and in 2015 there was a proposition to extend the “Class-4” framework to this new parameter. Unfortunately, the start of this activity suffered from a series of delays, due to outage in the sources of in-situ data and the time that could dedicate the expert to design, implement and test the correct production tools.
- Once initial “Class-4” files are produced they are sent to the US-GODAE server. Relation with the Program Manager of the US-GODAE server became unclear the recent years, although our technical contact there is reactive and efficient. Also, we continue to face policy issues and technical difficulties to include the contribution of our Chinese colleagues in the intercomparison. The effectiveness of US support for real-time routine activity is also under consideration with the recent political issues of the US Government “shutdown”.
- Currently, “Class-4” files are produced by participating centers. Over time, as ocean forecasting systems evolve through subsequent upgrades, the intercomparison can only be performed with a good knowledge of the way files are produced, and what are the characteristics of each system at a given time. This communication effort, by writing the relevant technical documentation is demanding, in particular if documentation needs to be revised after a system’s upgrade. Furthermore, most experts work under tight time constraints and are required to devoted time to production and scientific analyses first.
- Intercomparison is also more efficient with adequate tools where experts can communicate their results, and raise scientific issues or alerts on anomalies. Intercomparison is also more valuable in real-time if results and bulletin synthesis for monitoring the contributing operational systems can be managed with some common tools/server. Obviously, this is not the scope of the IV-TT expertise to develop such an IT framework, and IV-TT co-chairs have tried to provide some communication facilities based on existing Forum or Wiki, but this has not been particularly efficient.
- There were also several propositions for common intercomparison bulletin visualization tools. The BoM and DFO proposed softwares, and inside the CMEMS there are also similar initiatives that could be used by this purpose. However, external software are usually not adapted for internal use in operational centres, and there is no mandate of GOV IV-TT to impose some software policy to the contributing operational centres. So at this stage, there is no enthusiastic adoption of a community tool to offer a real time visualization of “Class-4” or “Class-1” results –at least externally of any forecasting centre. Such tool would imply resources to implement and maintain, and servers to host.

In conclusion this “Class-4” and “Class-1”, **due to its success, became a kind of “operational” task**, performed and maintained on routine through best efforts experts, with no dedicated community resources and official support to allow a secure activity.

Innovation and research inside IV-TT

Meanwhile, the “Class-4” and “Class-1” activities are permanent sources of research development and results, and in this sense, contribute to the scientific goals of a task team in GOV.

However, **these past years, the efforts devoted to the “Class-4” activity have slowed other initiatives**. Obviously, the part of the “Class-4” activity can nowadays be considered as “operational”, and could be re-sized in this context by operational team of ocean forecasting centers, and not any more under the task of validation experts.

These recent years, a series of new verification subjects are emerging, toward which the IV-TT scientific activity should reorient its focus:

- Extending the validation and verification approaches to the forecasting capability evolution: higher resolution and finer space and time scales ; focus from open ocean to more coastal processes ; error propagation monitoring between nested systems ; new parameters to take into account: biogeochemical, wave, or sea ice parameters ; emerging measurement technologies to be used (radiometry for surface salinity, future satellite like SKIM or SWOT giving access to fine scale and processes, like inertial waves etc...)
- Enhancing the methodological approach, from basic statistics computation of errors to more sophisticated process-oriented or user-oriented metrics, based for instance on image processing.
- Follow the path initiated by Met centres, with coupled ensemble forecasts, or multi-system forecast, where validation and verification approaches need to be revisited to account for larger exploration of the space of model errors.

In terms of community efforts, IV-TT has also some contacts with the atmospheric verification experts, and there is an obvious interest to be more involved in such framework.

In conclusion, the IV-TT is fully aware of new challenges and can dedicate time to work on it, in particular if the workload of routine “Class-4” and “Class-1” activities are reduced to scientific analyses and synthesis for publication.

Proposition to re-organize the IV-TT framework

It is identified that **“Class-4” and “Class-1” production activities are now mature**. The routine production on daily basis can be performed during forecasting centres operations, just with some supervision from the validation experts. The limitations are obviously due to lack of operational support.

Moreover, a new program context is emerging: OceanPredict and some JCOMM/ETOFS links can be considered for the operational framework. In this context, one solution could be to **identify the routine IV-TT tasks that could be considered as JCOMM/ETOFS contribution** of ocean forecasting centers. This will allow:

- To engage ocean forecasting centres for these routine tasks as part of JCOMM/ETOOFS, with a series of benefits like: production of documentation and standardized procedures, as well as providing dedicated resources to support the activity from within the different operational centres.
- To envisage with the relevant support some adequate IT infrastructure: mirroring site to the existing US-GODAE server, community tool under JCOMM rules for internal bulletin production, share of documentation
- A better visibility of this intercomparison activity toward centres not directly involved (eg, ECMWF at the global level, or many national regional ocean forecasting centres)
- A better link and/or commitment with institutions providing information (GDACs of the GOOS)

Then two issues emerge for IV-TT organization. Namely how to better interact with other GOV TT concerning verification and validation scientific issues. And how to benefit from the advances discussed and proposed by other communities, like the WWRP/WGNE Joint Working Group on Forecast Verification Research, the CMEMS Product Quality Working Group, etc..... Obviously, possible options are:

- Dedicate time to co-organize workshop/actions with other GOV task teams and discuss validation/verification approach. This is partly done, but could be enhance with specific demonstration actions
- Dedicate time to interact with other communities, by presenting our results and advances, co-organize workshops.

Priorities :

- Submit this document to shareholders : IV-TT members, GOV Patrons, decision makers of operational centres, JCOMM/ETOOFS
- Identify with IV-TT “Class-1&4” active participants what expertise, infrastructure and standards are needed to maintain a routine operational intercomparison activity
- Identify with all shareholders what are the expected outcomes of this new organization of the intercomparison (e.g., operational bulletins inside or outside each centre, visibility toward communities of users.....)
- Ensure the definition of a clear separation and responsibility among IV-TT and JCOMM-ETOOFS roles
- Redefine inside IV-TT the research priorities and actions plans