**GODAE OceanView Observing System Evaluation Task Team (OSEval-TT) Work Plan**

*Peter Oke (CSIRO) and Gilles Larnicol (CLS)*

One of the aims of GODAE OceanView is to formulate more specific requirements for observations on the basis of improved understanding of data utility. The Observing System Evaluation Task Team (OSEval-TT) is jointly formed by GODAE OceanView and the Ocean Observation Panel for Climate (OOPC). Through the task team, GODAE OceanView and OOPC partners will get organized at the international level to provide evidence-based responses to agencies and organizations in charge of sustaining the global and regional ocean observing systems used for ocean monitoring and forecasting at a range of time-scales. This activity requires consistent protocols for observation impact assessment, tools for routine production of appropriate diagnostics, common sets of metrics for inter-comparison of results, protocols for reporting the results to data providers and other stakeholders, and objective methodologies for observing system design and assessment activities.

The Task Team maintains and develops its international collaboration through regular meetings and workshops. These include the annual GODAE OceanView Science Team Meetings, and additional workshops. To date, two workshops that have focussed on OSEval-TT activities have been held under GODAE. These were held in 2007 and 2009 (see [http://www.godae.org/OSE-meetings.html)](https://docs.google.com/document/edit?id=1WREJcFF2GroYODYPeQbsQJkqCOSwWWTfZVfswDrtTKc&hl=en&pli=1) and a third workshop is planned for 2011.

GODAE OceanView activities tend to focus on short-range forecasting of the mesoscale ocean circulation, but do not exclude applications that relate to seasonal, interannual, or decadal forecasting. GODAE OceanView (and GODAE before that) has always maintained strong links with the climate and seasonal forecasting community. This has been achieved through membership of the Science Team and interactions with the CLIVAR Global Synthesis and Observations Panel (CLIVAR-GSOP). These interactions are particularly important for the OSEval-TT because of the shared dependence on the Global Ocean Observing System (GOOS); and will continue to be pursued in the coming years (e.g., through joint workshops and joint sessions at international conferences).

Traditionally, Observing System Evaluation using models has been undertaken either through Observing System Experiments (OSEs) and Observing System Simulation Experiments (OSSEs):

* OSEs involve the systematic denial (or addition) of a sub-set of observations, and the evaluation of the degradation (or improvement) in quality of the resulting analyses and forecasts. The degradation quantifies the impact of the with-held observations.
* OSSEs, sometimes referred to as twin experiments, typically use two different models. One model is used to perform a “truth” run - and it is treated as if it is the real ocean. The truth run is sampled in a manner that mimics either an existing or future observing system - yielding synthetic observations. The synthetic observations are assimilated into the second model, and the model performance is evaluated by comparing it against the truth run.

The OSEval-TT plans to coordinate a series of observing system evaluation activities, and facilitate the dissemination of OSSE activities. It is the intention of the OSEval-TT co-chairs that this work plan represents an ambitious target for our community to aspire to, and to try to fulfil. We expect that owing to limitations in time and resources, all of the planned activities will not be fully delivered. Many of the planned activities are experimental. It is not clear what level of engagement can be expected from the GODAE OceanView community. All of the planned activities require the individual groups to take the initiative to engage. By presenting a range of tasks and activities, it is anticipated that some groups will engage in some activities, and together, the GODAE OceanView community can deliver the needed information to the broader oceanographic and climate community. Details of the OSEval-TT Work Plan follow.

There are five key areas of activity that the OSEval-TT operates. These include:

1. *Routine monitoring of the Global Ocean Observing System*, delivered through:
	1. *Analysis of class 4 metrics (together with the Inter-comparison and validation task team)*; and
	2. *Performance and evaluation of Near-Real-Time OSEs*.
2. *Delayed-mode OSEs and OSSEs* for short-range and seasonal forecast systems, delivered through OSEval-TT Workshops.
3. Design and evaluation of new (e.g., SMOS) and future (e.g., new altimeter constellations) observing system components, delivered through GODAE OceanView technical workshops.
4. *Capability building*, delivered through:
	1. *OSEval-TT Workshops*; and
	2. *the OSEval-TT website.*
5. Provision and management of Observation Impact Statements (OIS) based on OSEval evidence.
	1. *Preparation and formal publication (e.g. via the OSEval-TT web site) of consensus impact statements (e.g., 1-4 pages with links to references) for observation sets tuned to data providers and other stakeholders needs;* and
	2. *Management of feedback on Impact statements and work-plan development.*

Details of each of these initiatives are presented below.

1. *Routine monitoring of the Global Ocean Observing System*

One of the outcomes of the 2nd GODAE Observing System Evaluation Workshop was agreement that we should try to undertake routine monitoring of the GOOS, and to record and disseminate how components of the GOOS are being used by forecast and reanalysis systems. The OSEval-TT will collate information on what observations are assimilated by GODAE OceanView systems; and quantify the impact of those observations on analyses and forecasts in OIS (see activity 5, below). OIS will be disseminated to OceanView patrons and observational agencies. It is anticipated that this information will help observational agencies justify and plan the maintenance, evolution and management of the GOOS.

The motivation for this activity is in recognition that the OSEval-TT has a role to play for the provision of recommendations to the observing agencies and community, particularly during and following “observing system events”. Here, we regard an observing system event to be a planned or unplanned loss or gain of instruments. Examples of observation events include data outages due to altimeter safe-hold or loss, launch of a new satellite, loss or malfunction (e.g., drift of pressure sensor) of Argo floats, etc. Reactions to these types of events are likely to be irregular, specific, short-term actions for the OSEval-TT. In these cases, the OSEval-TT will provide feedback (in the form of OIS and links to supporting material) to observational agencies to demonstrate the impact of such events on the quality of ocean analysis and forecast products generated routinely under GODAE OceanView. In practice, the OSEval-TT co-chairs will solicit input from OSEval-TT members. This input will be consolidated as an OSI by the OSEval-TT co-chairs and forwarded to the GODAE OceanView co-chairs for dissemination to observation agencies and GODAE patrons.

The GODAE OceanView OSEval-TT will monitor the GOOS and how it’s used by GODAE OceanView systems through the following activities:

1. *Analysis of Class 4 metrics*

Class 4 metrics are model-derived variables that are in observation space. This work is being pursued jointly with the Inter-comparison and Validation task team (IV-TT). The co-chairs of the IV-TT have circulated a metrics document that outlines the fields operational centers should compute and inter-compare. These include model-observation differences (innovations) for the background fields, analyses, and 1-5 day forecasts. Variables that are included in these metrics include SST, along-track SLA, and Argo T/S. A common data source has been identified, and the development of a capability to extract these metrics from every operational forecast is underway.

1. *Performance and evaluation of Near-Real-Time OSEs (NRT OSEs)*

At most operational centers, when a new forecast system is being transitioned to operations, two forecast systems are typically run in parallel - the existing system, and the new system. This suggests that these operational centers have the resources to run two versions of their forecast system in parallel. The proposal is that operational centers, in addition to performing their operational forecasts, also routinely perform OSEs for a single observation type at a time. By scheduling NRT OSEs, all operational centers that are partners in GODAE OceanView will plan on with-holding the same observation types in their NRT OSEs. This will allow the OSEval-TT to quantify the degradation in the forecast skill of multiple systems when different observation types are with-held. The schedule for NRT OSEs to be performed in the first half of 2011 follows. Clearly the OSEval-TT has no authority to insist that operational centers perform NRT-OSEs. With this in mind, the OSEval-TT has proposed the schedule below and will actively encourage operational centers to opt-in. The success of the NRT OSE program will be reviewed at the planned OSEval-TT workshop in June 2011:

January 2011: with-hold Argo data

February 2011: with-hold XBT data

March 2011: with-hold TAO data

April 2011: with-hold Jason-2 data

May 2011: with-hold all altimeter data

June 2011: with-hold AVHRR data

At the conclusion of each NRT OSE experiment, the class 4 metrics, described above, will be analysed to quantify the impact of each observation type on the analyses and forecasts of each system. Results of the NRT OSEs will be disseminated as OIS to the GODAE OceanView patrons and operational agencies. An up-to-date list of those groups planning to participate in this and other activities is maintained on the OSEval-TT website ([www.cmar.csiro.au/staff/oke/OSEval-TT.htm](http://www.cmar.csiro.au/staff/oke/OSEval-TT.htm)).

1. *Delayed-mode OSE and OSSEs*

In addition to the NRT OSEs, described above, the OSEval-TT will encourage the performance of delayed-mode OSEs and OSSEs. These experiments will be encouraged for both short-range forecast systems and seasonal forecast systems. Because the results of these OSEs and OSSEs inevitably depend on the details of the forecast system being used (e.g., model, data assimilation, initialisation, observations assimilated) it is desirable that multiple systems are used for these experiments. That way, the robust results of the OSEs and OSSEs can be identified and disseminated appropriately. This will require coordination that will be encouraged by the OSEval-TT co-chairs in the lead up to GODAE OceanView technical workshops. Because the performance and analysis of OSEs and OSSEs require a lot of resources, the OSEval-TT will not be too prescriptive on these activities. Instead, the OSEval-TT co-chairs will encourage these activities and will provide a guidance for the design of OSEs and OSSEs here, and through the OSEval website. As an example, some minimum guidance for OSEs follows:

* Observations to be assimilated (and systematically with-held) include Argo, XBT, TAO, Altimeter, and SST (other option include ocean colour/biological, sea ice and salinity data where appropriate);
* Forecast length should be 4-28 days for short-range forecast systems and 4 months or longer for seasonal forecast systems;
* Owing to the stochastic nature of seasonal forecasting, an ensemble of 10 or more forecasts should be included for seasonal applications;
* The time period for running delayed model OSEs is important; and should be chosen carefully. To perform OSEs that are relevant to today’s GOOS, a recent period is probably most relevant. To assess the impact of the number of altimeters, the period between 2003-2006 (when 4 altimeters were operating) might be most appropriate. To assess the impact of Argo, the most recent years, after Argo reached 3000 floats is preferred. In the absence of specific motivations such as those referred to above, the OSEval-TT recommends the following periods for OSEs:
	+ Mesoscale studies, including short-range forecasting: 2008
	+ Seasonal studies: 2004-2008.
* If computational resources permit, studies that span all seasons are preferable to short studies.

In addition to the traditional OSE experiments where each observation type is with-held, experiments that assess the impact of data processing (Level 2 and 3 SST; bias corrected observations; altimeter corrections etc.) could also be performed. The impact of improved error estimates could also be explored in the framework of OSEs. See [www.cmar.csiro.au/staff/oke/OSEval-TT.htm](http://www.cmar.csiro.au/staff/oke/OSEval-TT.htm) for a list of current studies.

1. *Design and evaluation of new and future observing system components*

A key objective of the OSEval-TT is its role in supporting the planning for new observation platforms and contributing to the design of the observing system. This might involve assessing the relative merits of different satellite constellations (e.g., altimetry, SST, ocean colour, SSS; revisiting the Argo design; addition of Oxygen sensors to Argo floats), different observation platforms (e.g., satellite sea surface salinity observations), or the design of integrated observing systems. The most common tool for conducting these experiments is OSSEs, but efforts in adaptive sampling are increasing in the oceanography community. These activities require significant resources, both for the computation of experiments and their analysis. It is very unlikely that an effort by the OSEval-TT to coordinate these activities, with all groups performing equivalent experiments, will be welcomed by the operational centers. Consequently, for the duration of this work plan, these activities will not be formally coordinated by the OSEval-TT, but this will be reviewed in future. Rather, GODAE OceanView technical workshops will provide a forum for sharing of results in this area, and for sharing of ideas on what key questions should be asked, and how these questions might be addressed. Each workshop will focus on a specific theme - aimed at encouraging participants to address specific questions in preparations for workshops as their resources permit. The next OSEval-TT workshop, called the GODAE OceanView and CLIVAR-GSOP Technical Workshop on Observing System Evaluation and Inter-comparisons, is being planned for 13-17 June in Monterey, CA, USA.

1. *Capability building*

a. *OSEval-TT Workshop*

In addition to providing a forum to share results that are relevant to observing system evaluation, the OSEval-TT workshops also provide a forum for identifying new (or old) and emerging techniques that could be used for observing system assessment. OSEval-TT workshops will provide a forum for sharing this information and establishing a community plan for adopting relevant methods. The outcomes from the 2nd GODAE OSEval-TT meeting (June 2009) are an example of this. At that workshop it was agreed to develop a capability to routinely assess the impact of all assimilated observations on their analysis and forecast systems. The diagnostics that spawn from this will be consolidated and disseminated to GODAE patrons and observational agencies. These diagnostics include forecast and analysis self-sensitivities. Several OceanView groups (CLS, TOPAZ, Bluelink) have developed a capability in this area following the 2nd GODAE OSEval-TT meeting.

b. *OSEval-TT website*

The OSEval-TT will establish a website that will be a reservoir for information that is relevant to observing system evaluation and design. A draft website has been established at [www.cmar.csiro.au/staff/oke/OSEval-TT.htm](http://www.cmar.csiro.au/staff/oke/OSEval-TT.htm). After this site has matured, it will be moved to the GODAE OceanView website. This website will include:

* Publication of OSEval-TT OIS;
* Announcements about upcoming and ongoing activities;
* A list of published papers that are relevant to observing system evaluation and design;
* A technical description of commonly used techniques (OSEs, OSSEs, analysis self-sensitivities, forecast sensitivities, adaptive sampling, etc.); and
* A record of past observing system events, and how the OSEval-TT responded.
1. *Provision and management of Observation Impact Statements (OIS) based on OSEval evidence.*
	1. *Preparation and formal publication (not necessarily using Journal papers) of consensus impact statements (e.g., 1-4 pages with links to references) for observation sets tuned to data providers and other stakeholders needs.*

Considerable investments of time, staff and computing resources are necessary to conduct OSEs and OSSEs. Their unique capability to provide evidence that observation types and sets have a positive impact means that the results of these experiments are particularly valuable. For data providers, such results provide strong justification for continued development and support and evidence to maintain investments for data continuity and cost effective system evolution. For management of the GOOS, OSEval results provide foundation evidence for system evolution decision-making. To maximise the utility of OSEval activities and to present consensus impact results, formal OIS will be developed. These will be short focussed reports of 1-4 pages that communicate the consensus of the OSEval-TT in a standard format focussing on key outcomes from OSEval activities, and concluding with recommendations and advice for data providers and managers. Such statements can then be used within national and regional agencies to justify continued investment for new observing systems and help to focus further research and developments funding calls.

* 1. *Management of feedback on impact statements and work-plan development.*

It is anticipated that feedback on OIS and the activities of the OSEval-TT will arise from a number of sources (e.g., users, data providers, funding agencies, space agencies etc). Feedback must be formally managed by the group in order to provide the best return on investment for the OSEval activities and to maximise the opportunity to sustain and evolve the GOOS in an informed manner. Formal workshops, points of contact and the OSEval-TT web site provide sensible means to accomplish this work. It is particularly important that feedback is managed in an open and transparent manner.