Demonstrating the complementarity of observations in an operational ocean forecasting system

D. J. Lea\textsuperscript{a\*}, M. J. Martin\textsuperscript{a} and P. Oke\textsuperscript{b}

\textsuperscript{a}Met Office, Exeter, UK
\textsuperscript{b}CSIRO Marine and Atmospheric Research, Hobart, Australia

\textsuperscript{\*}Correspondence to: Met Office, FitzRoy Road, Exeter, EX1 3PB, UK.

We have performed a series of near real-time observing system experiments with FOAM, the Met Office operational ocean forecasting system. These were conducted in parallel to the operational suite and identical to it except that certain observation types were excluded. At the start of each month the parallel system was reset to the operational restart and a run started with a different observation type excluded: in February - XBT; March - TAO/TRITON; April - Jason-2 altimeter; May - all altimeter; June - AVHRR sea surface temperature (SST) data; and July - Argo data.

We show that the existing ocean observing systems offer a good deal of complementary information. All components of the observing system offer unique, independent information that help initialise, and improve, operational ocean forecasts. Withholding XBTs causes little impact on globally averaged metrics, for example RMS innovations. However, locally we see long lasting temperature impacts ($\sim 1\degree$C) near XBT transects. Withholding TAO/TRITON data has a regional impact increasing the tropical Pacific RMS temperature and salinity innovations by 39% and 60%, respectively. Withholding Jason-2 data results in a global 4% increase in the RMS sea surface height (SSH) innovations, and small-scale changes in temperature and salinity of about 2$\degree$C at 100 m depth, and $\sim 0.2$ psu near the surface. Withholding all altimeter data leads to a 16% increase of the RMS SSH innovations. We also see changes in other model variables similar in magnitude to those from withholding Jason-2 but more widespread. Withholding AVHRR SST data produces significant changes around 1$\degree$C in model temperature at the surface and within the surface mixed layer, but there is little or no effect below this. Withholding Argo data for one month leads to a 5% increase in the temperature and salinity innovations, and changes in SSH of up to 5 cm.

Copyright © 2013 Royal Meteorological Society

\textbf{Key Words:} OSE; observing; ocean; data assimilation; GODAE OceanView

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/qj.2281

This article is protected by copyright. All rights reserved.
1. Introduction

Investigating the impact of ocean observations on operational ocean forecast systems is important for a range of applications (e.g., military, safety at sea, environmental monitoring), for funders of the current observing network, and for decisions on future observing systems. Such investigations also help developers to understand how well their data assimilation systems, that underpin each forecast system, make use of observational information. One way to quantify the impact of observations on a forecast system is to perform Observing System Experiments (OSEs; e.g., Oke and Schiller, 2007). An OSE involves running a copy of an existing assimilation run where some observations are excluded. The difference in the output of this run and the original run, assimilating all the observations, quantifies the impact of the withheld observations on the forecast system in the presence of the other observing systems.

OSEs are a powerful tool for objectively quantifying the impact of observations on a forecast system – but they are expensive. They require significant computational resources (requiring multiple versions of a forecast system to be run; and large data sets to be stored) and substantial human resources (requiring significant analysis to assess the results). As a result, most studies involving OSEs are performed for some time in the past, often several years behind real time (e.g., Vidard et al., 2007). This means that OSE studies are typically evaluating the observing system of several years ago. The relevance of such studies is therefore limited, because the observing system will have changed from the time of the study, to the present.

An initiative under GODAE OceanView, the successor of the Global Ocean Data Assimilation Experiment (GODAE) (Smith and Lefebvre, 1997; Bell et al., 2010), has been the proposal of “community near-real-time (NRT) OSEs”*. The idea is that several (ideally all) operational centres perform equivalent OSEs, withholding the same observations from their respective forecast systems, at the same time, using a set-up that is identical to their own operational system. By performing the OSEs in NRT, the results will be relevant to the present-day observing system. This relevance is important, because one of the motivations of this type of study is to identify emerging “gaps” in the observing systems, where additional observations are needed. In practice because different systems use different models, different forcing, and different assimilation systems, the impact of observations on different systems will vary – and will depend on the details of each system. If multiple systems are used for the NRT OSEs, results can be intercompared and the most robust results identified. The hope is that community NRT OSEs become an integral part of all operational centres’ activities - providing a regular, relevant and robust evaluation of the current state of the global ocean observing system. This study, where we have performed a number of NRT OSEs using FOAM (Forecasting Ocean Assimilation Model), the Met Office’s open ocean assimilation and forecasting system, is a first step in this direction. Detailed results can be found in Lea (2012).

There are a number of approaches for assessing the value and impact of observations. The OSEs mentioned above have the benefit of being reasonably straightforward to implement, if expensive. The OSSE (Observation System Simulation Experiment) is similar to the OSE but uses simulated data to test the assimilation system allowing future potential observing systems to be assessed. Another approach which is cheaper to operate than the OSEs is to use diagnostics of the assimilation to calculate observation sensitivities or observation information content (Rodgers, 2000; Cardinali et al., 2004; Desroziers et al., 2005; Chapnik et al., 2006; Todling, 2013). These can give the (linear) sensitivity of the assimilation to all the observations at the same time. The observation information content calculation (see Moore et al., 2011) requires the adjoint of the data assimilation system, something not readily


This article is protected by copyright. All rights reserved.
available for most systems. If, in future, we choose to implement such a scheme, the OSE experiments will still be useful for testing such methods.

For the moment, we perform a series of OSEs in a pseudo-operational context; running an identical copy of the operational system with the same forcing and observations with certain observations excluded. We can then compare the results with the system assimilating all the data in order to assess the sensitivity to the data excluded. These experiments are run for a month starting with the same initial model fields as the operational run but then allowing the run to evolve separately during that time. The experiments were run in NRT to properly demonstrate the impact observations were having operationally.

It is important to emphasise in cases where the observation impact seems low that this may be because of the limitations of the experiments (each covers only one month) and limitations of the ability of the FOAM system to use the information provided by the observations. Also, we have only assessed a limited range of diagnostics and it is often the case that some diagnostics suggest a bigger impact of a particular observing system than others.

In section 2 we describe the ocean analysis and forecasting system and the method used to run the OSEs. Section 3 describes the results divided into subsections for each OSE. The results of XBT withholding are in subsection 3.1, TAO/TRITON in subsection 3.2, Jason-2 altimeter in subsection 3.3, all altimeter in subsection 3.4, AVHRR SST (sea surface temperature) in subsection 3.5 and Argo in subsection 3.6. Section 4 contains the summary and conclusions.

2. Method

FOAM (Forecasting Ocean Assimilation Model) is the Met Office’s short range (out to 7 days) operational open ocean forecasting system (Storkey et al., 2010). Remotely sensed satellite SST and in situ SST data, temperature and salinity profile data, SLA (sea level anomaly) from satellite altimeters and sea ice concentration data were assimilated into the NEMO (Nucleus for European Modelling of the Ocean) (Madec, 2008) model using the Analysis Correction (AC) data assimilation scheme (Martin et al., 2007; Lorenc et al., 1991). The scheme is multivariate; made up of a 2D SST analysis where the increments are applied from the surface to the base of the mixed layer, a 2D sea surface height (SSH) analysis which has balanced temperature and salinity increments generated using the (Cooper and Haines, 1996) scheme, and separate 3D analyses of temperature and salinity profile data. The horizontal covariances are generated from the sum of two SOAR functions, with 40 km (mesoscale) and 400 km (synoptic-scale) length scales. Associated with these mesoscale and synoptic-scale horizontal covariances are vertical covariance length scales of 200 m and 100 m, respectively.

Note, though, that the AC assimilation scheme is no longer used operationally in FOAM. It was upgraded in January 2013 to NEMOVAR, a 3D-Var assimilation scheme (Mogensen et al., 2009).

In 2011, the satellite SST data included sub-sampled level 2 data from Advanced Microwave Scanning Radiometer Earth Observing System (AMSRE), Advanced Very High Resolution Radiometer (AVHRR) and Advanced Along-Track Scanning Radiometer (AATSR) data supplied by the Group for High-Resolution Sea Surface Temperature (GHRSST) project†. The in situ SST data come from drifting buoys, ship observations and moored buoys. Altimeter SLA data is along-track data provided by Aviso‡ which, in 2011, included data from the Jason-1, Jason-2 and Envisat platforms. The CNES-CLS09 v1.1 (Rio et al., 2011) mean dynamic topography (MDT) field is used to give the full dynamic topography needed for comparison with the model and data assimilation. There is an online altimeter bias correction scheme to correct for systematic errors in the MDT (Lea et al., 2008). The bulk of the profile data are now from the Argo ([Argo Science Team], 1998) array of floats providing temperature and salinity data. Additionally there are some CTD (Conductivity Temperature Depth) data and moored buoys which also measure temperature and
salinity, and XBTs (eXpendable Bathythermographs) which measure only temperature. The sea-ice concentration data is Special Sensor Microwave/Imager (SSM/I) data provided by the EUMETSAT Ocean Sea Ice Satellite Application Facility (OSI-SAF).

There is automated quality control (QC) of all data. The SST, SLA and sea-ice data are compared with the model background using a Bayesian procedure (Ingleby and Lorenc, 1997). The profile data are additionally checked using the comprehensive method described in Ingleby and Huddleston (2007). In these experiments we use the QC decisions from the operational run in the data withholding run to simplify the assessment of the results.

Data are assimilated into a global model at 1/4° resolution and various nested models at 1/12° resolution. In order to minimise the cost, the experiments described here are limited to the global model. One other important detail is that the operational system runs back 48 hours each day and assimilates all the available data in two 24 hour periods (day minus 1 and day minus 2). Running back an extra day means that data arriving late can be used to improve the analysis and forecast. The results below are all taken from the day minus 2 period where the most data are assimilated.

In this study we perform a series of OSEs; running an identical copy of the operational system with the same forcing and observations only with a particular type of observations excluded. We can then compare the results with the system assimilating all the data in order to assess the effect on FOAM of the excluded data type. First, an experiment withholding XBT data was run for all of February 2011 starting from the same initial model fields as the operational run, but then allowed to evolve separately during the month. This was followed in March by an experiment excluding TAO/TRITON data, Jason-2 altimeter data were excluded in April, all altimeter data in May, AVHRR SST in June and all Argo data in July. In all cases at the start of the month the data withholding run was initialised from the same state as the operational suite.

The one month duration of each of these OSEs is relatively short. Most OSE studies are performed for a longer period of time, from six months (e.g., Oke and Schiller, 2007) to several years (e.g., Balmaseda et al., 2007a). We recognise that a one month OSE may not give a representative impact of the longer term removal of an observation type. However, as discussed in section 1, one of the benefits of this approach is that it represents the NRT impact of observations, and is therefore more relevant for the present. In particular, we can assess the impact of “observing system events”, such as data outages (e.g., altimeter safe-hold or an outage at a data processing centre) that often only last for a short period of time. Also, with a one month OSE we can test the short term impact of the loss of a key observation platform (e.g., the end of an altimeter mission) that could happen at any time without prior warning. One main disadvantage is that the signal of withholding an observation type is likely to be smaller than for an OSE of longer duration. Recall, though, that this study is a first step to running community NRT OSEs where we may combine the results and make the impact of the OSE easier to discern.

The results will be assessed using the innovations which are the observation minus the model equivalent values before the data are assimilated (model background). We compare the operational run with the run withholding the data in question. For both runs we compare to all observations including those excluded from the data withholding run. The innovations can be thought of as a short term forecast error. We have not assessed the longer term forecast errors here but previous experience suggests the results are very similar to the innovations results for the relatively short forecasts we run (Storkey et al., 2010). Another method of assessment used is to examine the differences between the model fields which gives the spatial impact of the observing system excluded. However, to get an assessment of whether a data-type improves the model

---

1 Group for High Resolution Sea Surface Temperature: http://www.ghrsst-pp.org/

2 Aviso: http://www.aviso.oceanobs.com

This article is protected by copyright. All rights reserved.
representation of the true ocean state it is useful, also, to look at independent data. Here we do this by comparing the model to drifter derived velocities. The method used is that of Blockley et al. (2012) where the model daily mean currents are compared to the velocity calculated from the drifter positions over the day. A number of QC checks are performed and in particular any drifters which are known to have lost their drogue are excluded from the comparison.

3. Results

3.1. Withholding XBT

The “no XBT” experiment was run for all of February 2011 starting from the same initial model fields as the operational run, but then allowed to evolve separately during that month. Fig. 1 shows the locations of the XBT observations in February 2011. These form only 3.7% of the total temperature profile observations in this month.

Locally, near XBT observation locations, significant differences of 1°C or more are evident at the end of the month between the operational run and the “no XBT” run (Fig. 2). In addition, we see temperature changes in highly eddying regions, e.g. the Gulf Stream, which are not close to any XBT observations. These are most likely caused by chaotic error growth of small perturbations.

A line of XBT observations were made in the eastern Pacific starting on the 9 February 2011 near California and ending on the 21 February 2011 near Australia. The information from these observations has clearly spread further than that from another XBT line in the South Atlantic observed slightly later in the month (13 February 2011 to 21 February 2011). The spread of the observation information is affected by the time since the observation, the local currents, the Rossby wave speed and the model small scale diffusion parametrisation. The changes in temperature affect the local density gradients and the currents which have associated local changes in the model sea SSH (not shown).

Area averaged innovation statistics show very little effect from removing the XBT. If we focus on specific regions we find the Indian ocean shows the largest change with a 5% reduction in the temperature RMS error when XBT data are assimilated.

3.2. Withholding TAO/TRITON

The TAO/TRITON (Tropical Atmosphere Ocean/TRIangle Trans-Ocean buoy Network) array is a group of moored buoys measuring temperature and salinity at various depths down to 500 m in the tropical Pacific (Fig. 3). The “no TAO/TRITON” experiment where TAO/TRITON data were excluded was run for all of March 2011 starting from the same initial model fields as the operational run, but then allowed to evolve separately during that month. Of 867,922 individual subsurface temperature observations 124,530, or 14%, are TAO/TRITON observations. Out of 708,319 individual salinity observations 94,021, or 13%, are TAO/TRITON observations. In the tropical Pacific (15°S to 15°N) the array contributes the majority of the profiles (5157 out of 6508). Note, however, that the vertical sampling of the TAO/TRITON array is much sparser than say Argo and so the majority of individual observation points even in the tropical Pacific come from Argo.

In the tropical Pacific, when TAO/TRITON data are excluded, there is a dramatic increase in the RMS innovations of profile temperature (from 0.440°C to 0.612°C) and salinity (0.094 psu to 0.151 psu). The increase in the RMS innovations is also seen, to a lesser degree, in globally averaged statistics (Table 1). There is little or no effect on the SST or SSH statistics, however. There is also a much smaller impact on the profile statistics if only data other than TAO/TRITON are used in the calculation of the innovation statistics (not shown), indicating that the effect of the data is mostly localised to the area directly observed by TAO/TRITON. The region is relatively sparsely observed by Argo which may account for the large impact of TAO/TRITON, but it is also possible that the...
Figure 1. Locations of the XBT data which are excluded in the “no XBT” run, 852 profiles out of 23,243 profiles in total. The colours show the number of days into the month the observation was made.

Figure 2. Map of the temperature difference (Operational minus “no XBT”) in °C at 109.7 m depth. Calculated from daily average fields from the last day of the XBT OSE period (28 February 2011).

impact may be over emphasised because the same locations are observed every day.

Plotting the innovation statistics of temperature and salinity as a function of depth (Fig 4) shows that TAO/TRITON has an effect down to 1000 m. In temperature and salinity, there is a peak in the impact at 200 m and another smaller peak at around 500 m. In salinity there is also a substantial impact near the surface with a

<table>
<thead>
<tr>
<th>Observation Type</th>
<th>Operational</th>
<th>No TAO/TRITON</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST AATSR / °C</td>
<td>0.460</td>
<td>0.460</td>
</tr>
<tr>
<td>SSH / m</td>
<td>0.074</td>
<td>0.073</td>
</tr>
<tr>
<td>Profile T / °C</td>
<td>0.587</td>
<td>0.635</td>
</tr>
<tr>
<td>Profile S / psu</td>
<td>0.114</td>
<td>0.131</td>
</tr>
</tbody>
</table>

This article is protected by copyright. All rights reserved.
43% increase in RMS error (in the 0-15 m depth range) when TAO/TRITON data is excluded. These peaks in the impact reflect the depths of the temperature and salinity observations which are in the range of 0-300 m and at 500 m, but with information spread by both by the vertical correlations of the assimilation and by vertical model diffusion. The gap in TAO/TRITON observations at 400 m is reflected in the lack of change in the innovations at that depth. Essentially, there are no observations to see any change.

The assimilation of TAO/TRITON data has little impact near the surface in temperature, but affects the surface salinity much more. This is because the temperature is well observed by many satellite SST observations while there are few surface salinity observations to similarly constrain the salinity.

The localised effect of TAO/TRITON assimilation can be seen by examining model difference fields at 100 m at the end of the month (Fig. 5). Information generally propagates eastward at the equator at this depth. The effect of this can be seen particularly at 135°W. On the time-scale of the 1 month OSE run the impact is generally confined within the ±10° latitude range. Salinity at 100 m (Fig. 6) also shows evidence of the propagation of information particularly within a few degrees of the equator. There is a notable peak in the impact on salinity in the western Pacific. A similar surface salinity impact in magnitude and location is also seen (not shown) though the small scale details are somewhat different.

Outside of the tropics there are some differences seen in the OSE run. Some of these differences may be due to propagation of impacts away from the tropical region, perhaps by model barotropic waves or by impacts on assimilation of other data types, for example altimeter data assimilation. Though most of the differences seen, in this case, actually illustrate the difficulty of running an OSE alongside an operational system. The OSE run must as far as possible replicate the operational system (apart from excluding the observation type of interest). Unfortunately, this month there was a parallel version of the FOAM suite running forced by a new higher resolution atmosphere model. Parallel running is part of the normal operational upgrade cycle. On the 16 March 2011 this parallel suite was made operational using the model restart from the parallel run. Because the restart was switched this resulted in a sudden jump in the operational output. There was no straightforward way for the OSE to follow this leading to most of the extra-tropical differences observed. Fortunately, however, this switch appears to have had relatively little effect in the tropical oceans (assessed by comparing the
Figure 4. Tropical Pacific area innovation statistics, RMS for (a) temperature profile (in °C) and (b) salinity profile data (in psu). The black lines show results for the no TAO/TRITON run and the grey for the operational run. RMS errors as a function of depth are plotted as solid lines and mean errors are plotted as dotted lines. The maximum/minimum temperature RMS innovations are 1.03/0.07 °C in the no TAO/TRITON run compared to 0.72/0.07 °C in the operational run. The maximum/minimum salinity RMS innovations are 0.236/0.005 psu in the no TAO/TRITON run compared to 0.164/0.005 psu in the operational run.

Figure 5. Map of the temperature difference (Operational minus “no TAO/TRITON”) in °C at 109.7 m depth. Calculated from daily average fields from the last day of the TAO/TRITON OSE period (31 March 2011).

Looking at vertical sections across the Pacific along the equator (Fig. 7) the biggest effect of removing TAO/TRITON is in the thermocline region (deepest in the west at ~200 m and shallowest in the east at ~100 m). Early impacts in the OSE period are only very near the mooring locations. The information is then seen to propagate away from the observation locations. Eastward propagation is particularly notable in the eastern part of the section. Examining Hovmoller plots (for example, Fig. 8) shows the eastwards propagation at around 1 m s⁻¹, in the eastern part of the section. This is rather slower than the typical equatorial Kelvin wave speed but does match the eastward flowing Equatorial Undercurrent speed. So, it appears that
the observation information is simply carried in the mean flow rather than obviously triggering any waves on the equator.

At the end of the TAO/TRITON OSE period there are some impacts on the depth of the thermocline and also the temperature gradient in the thermocline region. Most of the thermocline depth changes are small scale and possibly random or chaotic. Generally, there is a sharpening in the gradient mostly where the thermocline is shallower. Overall the “no TAO/TRITON” run has an average $18^\circ$C to $22^\circ$C depth difference of 34.6 m compared to the operational run where it is 31.6 m. The sharpening of the gradient can also be seen in Figs. 7 and 8 where the assimilation of TAO/TRITON data tends to warm the model above the thermocline and cool it below. It is a common defect of models to have a too-diffuse thermocline so, in this regard, the assimilation of TAO/TRITON data is beneficial.

### 3.3. Withholding Jason-2

In April 2011 there were a total of 1,134,330 along track altimeter observations of which 495,429 (43%) were Jason-2 observations. The impact of excluding Jason-2 data, in the OSE this month, on the global innovation statistics is shown in Table 2. There is a small increase in the RMS innovations for all observation types. AATSR (1.3% increase), SSH (3.9%), profile temperature (1.6%) and profile salinity (2.4%). This suggests some complementarity in the data; altimeter data helps to improve the model forecast of other quantities even though other more direct observations are being made. For example, having eddies in the right place will help to improve the fit to SST data.

A time-series of the SSH innovation statistics shows that removing Jason-2 data takes some time to show its full effect (Fig. 9). This is partly due to the 10 day repeat cycle (for Jason-1 and Jason-2) which means that the local impact is only fully observed when the same altimeter track repeats. However, as the “no Jason-2” RMS error in Fig. 9 is still increasing after 10 days we conclude that the model retains...
Figure 7. Equatorial vertical section of the temperature difference (Operational minus “no TAO/TRITON”) in °C across the Pacific from 120°E to 80°W at various times throughout the OSE run. The solid lines show the longitude of the TAO/TRITON moorings on the equator. The dashed lines show moorings which are within 5 degrees latitude of the equator.

Figure 8. Hovmoller longitude-time plot showing the model temperature difference (Operational minus “no TAO/TRITON”) in °C across the Pacific at 0°N and 109.7 m depth.

information for some time after Jason-2 data is withheld. Recall that both the operational and OSE run start with the same initial conditions from the FOAM operational system which has been assimilating altimeter data continuously for some years.

The model field differences of SSH at the end of the OSE period (Fig. 10) show significant small scale impacts.
Figure 9. Global SSH innovation time series statistics in m. The black lines show the results results of the “no Jason-2” run and the grey lines the operational run. RMS errors are plotted as solid lines and mean errors as dotted lines. The maximum/minimum SSH RMS innovations are 0.082/0.071 m in the “no Jason-2” run compared to 0.079/0.068 m in the operational run.

when Jason-2 data are not assimilated. These differences are greatest in the regions with strong mesoscale variability like the Gulf Stream, Kuroshio and Antarctic Circumpolar Current. The lack of differences at high latitudes is a consequence of the weak stratification. In regions where the top to bottom temperature gradient is less than 5°C no mesoscale temperature and salinity increments (generated by Cooper and Haines, 1996) are applied. Only large scale (∼400 km) barotropic SSH increments are applied in these regions. Because the barotropic increments are large-scale it appears that removing one altimeter has very little effect. There are also some signs of coherent large scale impacts ∼+2 cm in parts of tropical Pacific and ∼−2 cm in the eastern tropical Atlantic and in the northern Indian ocean. Whether this is indicative of model bias or observation bias is not currently clear, however.

The effect of Jason-2 assimilation is also strongly evident in the 100 m temperature (Fig. 11). There are ∼2°C mesoscale differences particularly in regions of strong mesoscale variability which are coincident with the large SSH differences. However, there are also large temperature differences in the equatorial regions even though the SSH signal is quite weak, presumably a consequence of the strong vertical gradient in temperature. If we examine the surface temperature differences (not shown) we find they are quite small because this is well constrained by the assimilation of the abundant sea surface temperature data. In contrast to surface temperature we see a strong impact on surface salinity (not shown) from Jason-2 perhaps a result of the circulation changes induced by the altimeter assimilation.

Excluding Jason-2 increases the RMS current errors compared to drifters by 3% and 2% for the zonal and meridional velocity components, respectively (Table 2). We can therefore infer that the increments from Jason-2 assimilation have a positive effect on representation of the mesoscale circulation.

3.4. Withholding all altimeter

There were 1,078,113 along track altimeter observations available in May 2011. All are excluded in the “no altimeter” run. When all altimeter data are excluded there is a 16% increase in the global RMS SSH error from 7.4 cm to 8.6 cm and an increased in bias of 2 cm averaged over 1 month (Table 3). The AATSR SST, profile temperature and salinity all have increased RMS innovations, by 1.9%, 1.3% and 0.8%, respectively.

In the previous month we tested excluding Jason-2. If we compare the results, it is interesting to note that the increase in the RMS of temperature and salinity innovations was somewhat higher even though we were still assimilating the other altimeters. Allowing for the fact that we are not comparing the same month, this indicates that the other altimeters were not as beneficial to FOAM as Jason-2. Indeed, after the OSEs were run we discovered there was a problem in the upstream processing of altimeter data (from Table 3. Global summary innovation RMS for different observation types accumulated over May 2011. For the “no altimeter” and operational runs we are comparing to all observations including all altimeter data.

<table>
<thead>
<tr>
<th>Observation Type</th>
<th>Operational</th>
<th>No altimeter</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST AATSR °C</td>
<td>0.472</td>
<td>0.481</td>
</tr>
<tr>
<td>SSH / m</td>
<td>0.074</td>
<td>0.086</td>
</tr>
<tr>
<td>Profile T °C</td>
<td>0.606</td>
<td>0.614</td>
</tr>
<tr>
<td>Profile S / psu</td>
<td>0.128</td>
<td>0.129</td>
</tr>
<tr>
<td>Zonal current / m s⁻¹</td>
<td>0.225</td>
<td>0.231</td>
</tr>
<tr>
<td>Meridional current / m s⁻¹</td>
<td>0.209</td>
<td>0.216</td>
</tr>
</tbody>
</table>
AVISO) which resulted in excessive small scale filtering of real time Jason-1 and Envisat altimeter data used in FOAM. The problem did not affect the Jason-2 data. This may explain the results seen here, although we cannot disregard the possibility that it is due to problems with our assimilation method. The only way to be sure would be to perform new OSEs with more recent (corrected) data.

A time-series of the SSH innovation RMS and mean errors in Fig. 12 shows a steady increase in the RMS and a steady decrease in the mean innovations in the “no altimeter” OSE. The largest part of the RMS increase comes from an increase in the standard deviation. The increasing bias occurs because the FOAM global model evaporation, precipitation and river inflow are not exactly balanced. Without any altimeter assimilation to correct the model...
mean free surface height it rises by around 3 cm over the month.

The drift in the model SSH is obvious in the plot of differences in the model SSH between the operational and the “no altimeter” OSE at the end the month (Fig. 13). In addition there are some large (~20 cm) differences in the mesoscale due to uncorrected drifts in the mesoscale eddy field.

The impact of removing all altimeter data on model temperature at 100 m is shown in Fig. 14. At least 2°C mesoscale differences are seen. These are significantly more widespread than the equivalent changes in the “no Jason-2” OSE (Fig. 11). Also, as in the Jason-2 OSE, there are very large changes in the tropical temperatures despite the relatively small SSH signal in those regions. Again also, the surface temperature differences (not shown) are small because this is well constrained by the assimilation of sea surface temperature data.

Excluding all altimeter increases the RMS current errors compared to drifters by 3% and 2% for the zonal and meridional velocity components, respectively (Table 3). This is similar to the increase in error from only excluding Jason-2 altimeters (see section 3.3). This could relate to the problem with Jason-1 and Envisat altimeter data described above.

#### Table 4. Global summary innovation RMS for different observation types accumulated over June 2011 comparing the "No AVHRR" and operational runs. The velocity statistics do not significantly differ between the runs (not shown).

<table>
<thead>
<tr>
<th></th>
<th>Operational</th>
<th>No AVHRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST AA / °C</td>
<td>0.502</td>
<td>0.516</td>
</tr>
<tr>
<td>SSH / m</td>
<td>0.076</td>
<td>0.076</td>
</tr>
<tr>
<td>Profile T / °C</td>
<td>0.619</td>
<td>0.621</td>
</tr>
<tr>
<td>Profile S / psu</td>
<td>0.121</td>
<td>0.120</td>
</tr>
</tbody>
</table>

### 3.5. Withholding AVHRR SST

For the June 2011 period, there are 2,429,832 NOAA-AVHRR and 6,421,770 METOP-AVHRR observations assimilated into FOAM, out of a total of 26,570,218 SST observations. This means we remove 33% of the SST observations in the “no AVHRR” run. AATSR, AMSRE and in situ (from ship and drifting buoys) are the other SST datasets assimilated. The geographic distribution of the data removed is not uniform. Fig 15 shows the percentage of the SST data that is AVHRR in different locations. In the less cloud covered regions and around the coasts AVHRR is an important dataset comprising 40% or more of the SST data.

Summary global monthly average innovation statistics are shown in Table 4. The fit to AATSR SST data is improved by 2.7% when we assimilate AVHRR data. There is little impact on the SSH or profile statistics.

Even though there are many other SST data assimilated there are still significant differences, up to 1°C, in the model SST after 1 month (Fig. 16). The biggest differences are coincident with regions where the AVHRR observations are a greater part of the total SST observations (see Fig. 15). Such areas include the coastal regions, the Mediterranean, the equatorial regions and the west coast of Africa, all where over half of SST observations are AVHRR.

At 100 m depth, the temperature differences (not shown) are mostly limited to regions where the mixed layer is 100 m or deeper which in June is mostly in the southern hemisphere. As the assimilation projects the increments from SST down through the mixed layer this result is expected. There are also some deep differences seen near the equator likely due to model mixing which may...
be spuriously caused by assimilation \citep{Balmaseda2007}. This mixing may be mitigated in FOAM by the pressure bias correction scheme \citep{Bell2004} although perhaps some residual problems remain.

3.6. Withholding Argo

The “no Argo” experiment was run for all of July 2011 starting from the same initial model fields as the operational run, but then allowed to evolve separately during that time. Out of 868,576 individual subsurface temperature or salinity observations 652,758, or 75%, are Argo, and out of 744,156 individual salinity observations 634,272 or 85% are Argo.

Figure 13. Map of the SSH difference (Operational minus “no altimeter”) in m. Calculated from daily average fields from the last day of the altimeter OSE period (31 May 2011).

Figure 14. Map of the temperature difference (Operational minus “no altimeter”) in °C at 109.7 m depth. Calculated from daily average fields from the last day of the altimeter OSE period (31 May 2011).
temperature and salinity is 5% worse without Argo. The fit to AATSR SST is also slightly degraded without Argo. It should be noted that 1 month is not long enough to see the full effect of removing Argo data. Previous experience in FOAM suggests that the ocean subsurface (down to 2000 m) can take over 1 year for changes in the assimilation to take full effect.

The impact of Argo data in model temperature is very significant, particularly below the surface. In the mixed

| Table 5. Global summary innovation RMS for different observation types accumulated over July 2011. For the “no Argo” and operational runs we are comparing to all observations including Argo data. |
|-------------------------------------------------|--------|--------|
| SST AATSR / °C                                 | 0.524  | 0.527  |
| SSH / m                                        | 0.074  | 0.073  |
| Profile T / °C                                 | 0.680  | 0.728  |
| Profile S / psu                                | 0.132  | 0.139  |
| Zonal current / m s⁻¹                           | 0.232  | 0.233  |
| Meridional current / m s⁻¹                      | 0.213  | 0.215  |
layer SST data assimilation means the model is well constrained even without Argo. At 30 m, at this time of year, this means that no differences are seen in large portions of the southern hemisphere (see Fig. 17). At 100 m, mostly below the mixed layer except in some parts of the southern hemisphere, there are very significant differences between the “no Argo” and operational runs (see Fig. 18). Another notable feature is the large changes due to Argo data near Japan corresponding to a relatively high density of floats.

As we do not currently assimilate remotely sensed surface salinity data from SMOS or Aquarius, there is typically very little salinity data apart from Argo being assimilated. Consequently, the impact on the model salinity of removing Argo is significant at all depths, including at the surface. Fig. 19 shows the salinity difference at 100 m.

The removal of Argo also has an interesting effect on the model SSH. Significant large scale (∼1000 km) changes in SSH of up to 5 cm occur (Fig. 20) despite the assimilation of altimeter data. This seems to result from model biases, perhaps a drift in the water mass properties, which is not corrected by altimeter assimilation alone.

Looking at a Pacific Ocean longitude section at 40°N (Fig. 21) we see that the positive SSH anomaly is associated with a change in the temperature at around 300 m depth. The negative SSH anomaly in the western Pacific near 40°N is associated with a deeper salinity increase at 400 to 2000 m depth. This type of information may be helpful in work to reduce model biases.

Interestingly there is a small impact on the fit to drifter currents, when Argo data is removed, with the RMS error increasing by about 1% (Table 5). This is less significant than the degradation when altimeter data is removed. But, the large scale drifts in water mass properties, without Argo data assimilation, may prevent the altimeter assimilation from effectively correcting the model currents.

4. Discussion and conclusions

In order to assess the impact of different observing systems we have performed a number of observing system experiments with the operational FOAM system. These involve running a copy of the FOAM operational suite and excluding a particular observing system for a period of one month. In February 2011 XBT was excluded, in March 2011: TAO/TRITON, in April 2011: Jason-2, in May 2011: all altimeter, in June 2011: AVHRR, and in July 2011: Argo.

The main result is that the observing network provides a good degree of complementary information. For example, while XBT data have little global impact, they have considerable and persistent (at least for a month) impact locally. TAO/TRITON has a powerful effect on temperature and salinity in the Tropical Pacific, and these data are complemented by Argo data which have global but sparser coverage (in the Tropics), and consequently lower impact in the Tropics. Altimeter data have a strong impact on the mesoscale unmatched by other data types. This is evident in the small scale changes in 3D model temperature and salinity. When altimeter data are excluded there is a notable degradation in the fit to the mesoscale dominated
drifter surface currents. Such an impact on the fit to surface currents is not seen with other observation types, though there is evidence that excluding Argo data ultimately degrades the fit to drifter currents. It may be that biases in the temperature and salinity which develop without Argo assimilation prevent the model from producing a good fit to the altimeter data.

We see mesoscale structures in the AVHRR SST data, however FOAM does not use this to directly correct the circulation and so there is no effect on the fit to drifter currents. The AVHRR SST data impact and value is largely in temperature near the surface in the mixed layer. It is clear that there is some redundancy of the SST data particularly in cloud free regions which has the benefit of making the system more robust to loss of a particular satellite.

This article is protected by copyright. All rights reserved.
One conclusion of this work is that many of the impacts of removing the data take some time to become fully evident. It may, therefore, be useful to perform longer OSEs in future in order to see the full effect of removing a particular data type. Performing longer OSEs, though, is of course costly since each requires a full run of the system, so this is something that would be done only occasionally.

As the observing system and the FOAM system changes in future, we will need to rerun the OSEs since the results are specific to the observing network, model and data assimilation used at the time. In particular, we recently upgraded the assimilation scheme in FOAM to 3D-VAR (Mogensen et al., 2009). Because of this, we might expect some differences in the observation impacts compared to the results presented here.
The main weakness of the work is that it is specific to the FOAM system. This is slowly being addressed as other GODAE OceanView partners begin to perform their own OSEs. This will allow much more robust statements about the information content of particular observing systems on ocean forecasting systems in general since we will be testing them in different models and different data assimilation systems.

Acknowledgement

The authors would like to thank the GODAE OceanView community for motivating this work. We would also like to thank the anonymous reviewers for their helpful comments.

References


Mogensen K, Balmaseda M, Weaver A, Martin M, Vidard A. 2009. NEMOVAR: A variational data assimilation system for the NEMO ocean model. newsletter, ECMWF.


