DATA ASSIMILATION IN THE AUSTRALIAN BLUELINK SYSTEM

By Peter R. Oke[1,2] , Gary B. Brassington[1,3], David A. Griffin[1,2], Andreas Schiller[1,2]

[1]Centre for Australian Weather and Climate Research,
[2]Commonwealth Scientific and Industrial Research Organisation Wealth from Oceans National Research Flagship, Hobart, Australia
[3]Bureau of Meteorology, Melbourne, Australia

Abstract

The Bluelink Ocean data assimilation system (BODAS) is an ensemble-based system that underpins Australia's operational short-range ocean forecast system. The primary test-bed for the Bluelink system is the series of Bluelink ReANalysis (BRAN) experiments. Over the life of Bluelink, BRAN experiments have been used to assess the performance of the system, and to test new developments prior to integration into the forecast system and operational trials. BRAN experiments have helped identify problems with the model, assimilation system, data processing, and model initialisation. In this paper, the recent improvements of the Bluelink system are highlighted, along with some preliminary results from the application of BODAS to a relocatable coastal ocean model, also developed under Bluelink.

Introduction

Bluelink is a partnership between the Commonwealth Scientific and Industrial Research Organisation (CSIRO), the Bureau of Meteorology (BoM) and Royal Australian Navy (RAN). The primary objective of Bluelink is to develop and improve Australia's capabilities in short-range ocean forecasting and reanalysis. The Bluelink forecast system (Brassington et al. 2007) first became operational at the BoM in August 2007, and has since produced two 7-day forecasts each week. The main components of the Bluelink system are the Ocean Forecasting Australia Model (OFAM) and the Bluelink Ocean Data Assimilation System (BODAS). The primary test-bed for the Bluelink system is the series of Bluelink ReANalysis (BRAN) experiments – multi-year data assimilating model runs.

The purpose of this paper is to describe the Bluelink system, particularly BODAS and its recent enhancements, and to review some of the lessons learnt from a series of BRAN experiments. This paper is organized as follows: a short description of the Bluelink model is presented, followed by a description of BODAS. A summary of a series of BRAN experiments are described, followed by results from a recent BRAN experiment, and a demonstration of the application of BODAS to a coastal ocean forecast system.

Bluelink Ocean Model - OFAM

The global model used here is based on the Modular Ocean Model (Griffies et al., 2004) and is called the Ocean Forecasting Australia Model (OFAM). The first version of OFAM, OFAM1, used version 4p0d. OFAM2, which is still being developed, uses version 4p1. The horizontal resolution of OFAM varies from 2^o in the North Atlantic to 1/10^o in the 90^o-sector centred on Australia and south of 16^{o}N. OFAM1(2) has 47(51) levels in the vertical, with 20(24) levels in the top 200 m, and 35 levels in the top 1000 m, with a minimum of 10(5) m resolution near the surface. The horizontal grid has 1191 and 968 (1191 and 997) points in the zonal and meridional directions, respectively. The bottom topography for OFAM1 was a composite of a range of different sources, including dbdb2 (provided by the United States Naval Research Laboratory) and the General Bathymetric Charts of the Ocean (GEBCO) and AGSO2002. The OFAM2 topography is based on the Smith and Sandwell (1997) v11.1 bathymetry. The model uses the third-order quicker scheme for tracer advection (Leonard, 1979). Horizontal viscosity is resolution and state-dependent based on the Smagorinsky-scheme (Griffies and Hallberg, 2000). The turbulence closure model used by OFAM is the hybrid mixed-layer scheme described by Chen et al. (1994).

For long model runs, such as free spin-up runs and BRAN experiments, OFAM is forced by 6-hourly atmospheric fluxes from ECMWF, using fields from ERA-40 (Kallberg et al., 2004), for the period prior to August 2002, and 6-hour operational forecasts thereafter. The operational Bluelink forecast system uses 6-hourly forcing from the BoM Global Atmospheric Prediction System (GASP, e.g., Schulz et al. 2007; soon to be replaced with a version of the Unified Model, Rawlins et al. 2007).

Bluelink Ocean Data Assimilation System - BODAS

The Bluelink Ocean Data Assimilation System (BODAS; Oke et al. 2008) was initially developed for data assimilation into a global ocean forecast system. The requirements of such a system are to facilitate the assimilation of different observation types, in all possible dynamical regimes, including those of the open ocean, shelf zones and marginal seas. The assimilation of multiple observation types makes a multivariate assimilation preferable, whereby observations of one type (e.g., sea-level) influence the increments to model fields of all types (e.g., sea-level, temperature, salinity, velocity). The requirement to assimilate in a variety of different regions and dynamical regimes encourages the adoption of inhomogeneous and anisotropic background error covariance estimates, since background errors in different regions are expected to be characterised by different length-scales, and with different orientations. Multivariate, inhomogeneous, and anisotropic covariance estimates are readily obtained using ensemble data assimilation methods. It is for this reason that the Bluelink team opted to develop an ensemble-based data assimilation system. The salient aspects of BODAS are as follows:

BODAS employs an ensemble optimal interpolation (EnOI) scheme that uses a stationary ensemble of intraseasonal model anomalies, or modes, to approximate the system's background error covariance. Because we expect the background field errors of a short-range forecast system to be dominated by eddy-scale features, the ensemble is comprised of ensemble members that contain eddy-scale variability. In practice, this is achieved by computing each ensemble member by high-passing a long model run. At present, Bluelink applications use (up to) 120-ensemble members, computed from the last 10-years of a 15-year free run of OFAM. Each ensemble member is a 3-day mean minus the 3-month mean centered at the same time. The current operational system uses a 72-member ensemble.

An important feature of BODAS is covariance localisation. Using ensemble data assimilation, the influence of an observation on the model state is determined by the ensemble-based covariance between the observed state element and all other state elements. Because the ensemble is small compared to the dimension of the model subspace, the ensemble is rank-deficient and suffers from sampling error (Houtekamer and Mitchell 2001; Mitchell et al. 2002; Oke et al. 2006). The rank-deficiency means that the ensemble does not have enough degrees of freedom to adequately fit the model-data misfits (background innovations) during an assimilation step. The sampling error means that the ensemble-based covariances are noisy – particularly for long-distance covariances that are really expected to be zero. For example, sea-level errors in the Tasman Sea are not expected to be correlated with sea-level errors in the Gulf of Mexico. However, for a small ensemble, the ensemble-based covariance may be non-zero. These artificial long-distance covariances are eliminated in practice by multiplying the ensemble-based covariance by a localising correlation function (Houtekamer and Mitchell, 2001). Here, the localising function is a homogeneous, isotropic, quasi-Gaussian function with an e-folding length-scale of about 2-3 degrees. As a result, the influence of an individual observation on the model state depends on both the ensemble-based covariances and the distance between the observed location and the location of each model state element. For the covariances over short distances (less than a few hundred metres), the details of the ensemble-based covariance - including the length-scales, inhomogeneity, and the anisotropy – are retained when localisation is used (Oke et al. 2005). But the long-distance covariances are eliminated.

At present, BODAS routinely assimilates along-track sea-level anomaly (atSLA) data from all available satellite altimeters and coastal tide gauges from around Australia, plus Sea Surface Temperature (SST) observations from the Pathfinder database and AMSR-E passive microwave radiometer. Recent developments permit the assimilation of GHRSST L2P data (e.g., NAVOCEANO L2P AVHRR). In-situ temperature and salinity observations from Argo floats, the Tropical Atmosphere–Ocean (TAO) array, CTD and XBT (temperature only) surveys from a variety of different field surveys, including WOCE, Indian Ocean Thermal Archive (IOTA) and others, are also routinely assimilated. Explicit observation error estimates are assigned to each observation according to their expected instrument error, their "age" relative to the analysis time, and an estimate of their representation error – see Oke et al. (2008) for details. Representation error estimates are obtained using the method described by Oke and Sakov (2008). This method yields error estimates that depend on the model grid. For example, where the model grid spacing is 1/10^{o} the representation error is small – because the model and observations can "represent" variability of comparable scales. However, where the model is coarse, say 2^{o} in the North Atlantic, the representation error is large – because the model cannot represent all of the features and variability represented by the observations. Although somewhat counter-intuitive, this difference in representativeness is ascribed as an error to the observation, so that the model doesn't over-fit the data by "trying" to reproduce scales that are not resolvable on the model grid. That is, so the analysis step doesn't try to fit what the model regards as noise.

Despite the fact that the ocean is under-sampled, the number of discrete satellite observations is too large to be efficiently assimilated directly by BODAS. This is addressed by assimilating super-observations for SLA and SST, and by selecting only a sub-set of in situ temperature and salinity profiles to assimilate. The calculation of super-observations simply refers to the spatial averaging of SLA and SST data prior to assimilation. Super-observations are ascribed a smaller error, depending on the distribution and number of observations that are averaged. The amount of averaging and sub-sampling done by BODAS is flexible, and can readily be modified for different scenarios. Because of the spatially varying resolution of OFAM, a typical application of BODAS involves super-obing the SLA and SST data to a nominal resolution of 4/10^{o}-6/10^{o} around Australia (i.e., every 4th or 6th model grid point) and coarser elsewhere. Similarly, it is typical to select one temperature and salinity profile every degree around Australia, and coarser elsewhere. For short experiments that are focused on a particular region, or event, the resolution of the super-obing and sub-sampling can easily be modified to retain more observations in regions of particular interest.

BODAS calculates a global analysis of the model state by performing many (approximately 500) analyses on sub-domains of the model grid. For each sub-domain, observations from a halo around that sub-domain are used to influence the analysis. Provided the extent of the halo is chosen to match the distance over which the localizing function goes to zero, adjacent sub-domains produce analyses that are seamless at their point of intersection (i.e., spatially continuous), and the analysis of the full model state is equivalent to a global inversion (Figure 1). This approach differs from many ensemble-based systems (e.g., Houtekamer and Mitchel 2001; Brasseur et al. 2006; Bertino et al. 2008), who compute analyses, one grid point at a time, using observations only in the vicinity of each grid point.

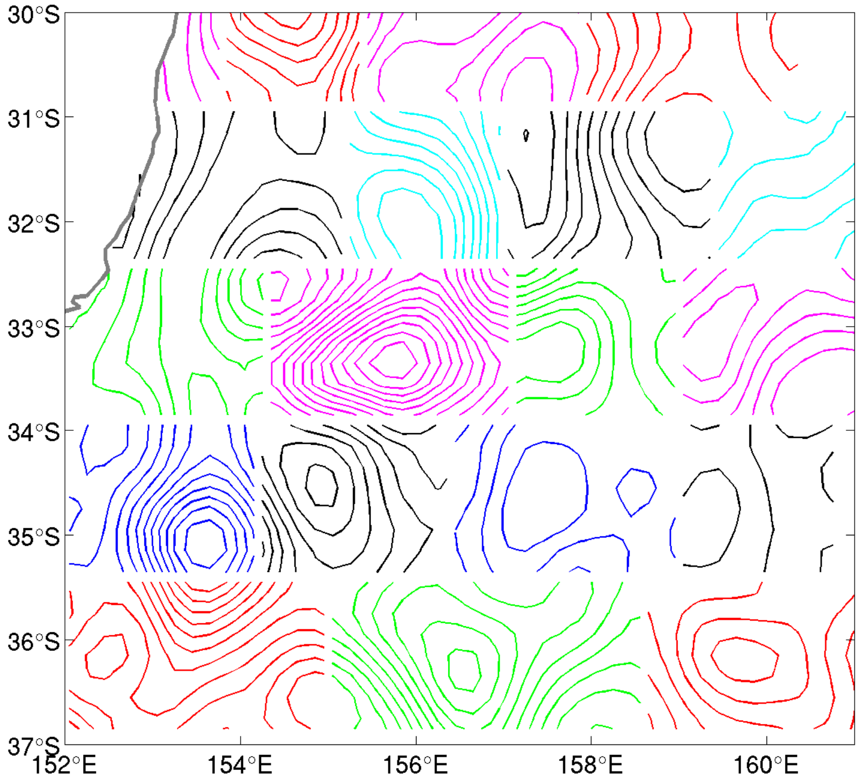

Figure 1

An example of the increments for sea-level (the contour interval is 5 cm) in the Tasman Sea produced by BODAS. Different colours represent increments computed independently. The meridional extent of each sub-domain is pre-determined, but the zonal extent of each sub-domain is adaptive, and depends on the density of the observations. Note the continuity of the increments in adjacent sub-domains.

Since it's development, BODAS has been used for many different applications, including global reanalyses (Schiller et al. 2008), operational global ocean forecasting (Brassington et al. 2007), seasonal prediction (http://poama.bom.gov.au/research/assim/ index.htm), observing system evaluation (Oke and Schiller 2007), observing system design (e.g., Oke et al. 2009), and more recently, regional (Sandery and Brassington 2008) and coastal data assimilation (see below). Some examples of these applications are described below.

Reanalysis Experiments

BRAN experiments are typically multi-year data-assimilating model runs. The purposes of BRAN experiments are twofold. Firstly, BRAN experiments are intended to facilitate testing and development of new versions of the Bluelink System prior to operational trials. Secondly, BRAN experiments are intended to provide a service to the research community for understanding ocean variability and dynamics. In this section, a review of BRAN activities is presented, along with some scientific results on ocean variability around Australia.

To date, two long (>12 years) BRAN experiments have been performed (BRAN1 and BRAN2p1), two intermediate-length (1-4 years) experiments have been performed (BRAN1p5 and BRAN2p2), and several short (3-6 month) experiments have been performed. These experiments differ in the time periods simulated, the data that is assimilated, frequency of assimilation, forcing fields, ensemble size, and the method of initialisation. A summary of the configuration of each of the main BRAN experiments is given in Table 1.

	BRAN1	BRAN1p5	BRAN2p1	BRAN2p2
Time period	10/1992-12/2004	1/2003-6/2006	10/1992-12/2006	4/2006-4/2008
Data assimilated	atSLA, T/S	atSLA, T/S, SST	atSLA, T/S, SST	atSLA, T/S, SST
Ensemble size	72	72	120	120
Assimilation interval (d)	3	7	7	7
Surface forcing	ECMWF fluxes	ECMWF fluxes	ECMWF heat/PmE fluxes & 10 m winds	ECMWF heat/PmE fluxes & 10 m winds
Rivers	none	none	seasonal	seasonal
SST (SSS) restoring	30-d (30-d)	none (none)	none (30-d)	none (30-d)
Initialisation	Updates to U, V, T, S and η in single step	Nudging to T, S & η with 1-d time-scale	Nudging to T, S & η with max(1-d,T_{in}) time-scale.	IAU to T, S, U, V, and η over 12-hours.
Known problems	Error in surface heat fluxes & bugs in BODAS	Some in situ T profiles processed incorrectly	Topographic errors in some shallow Straits	Topographic errors in some shallow Straits

Table 1

Summary of the configuration for BRAN experiments (atSLA is along-track sea-level anomaly; T/S refers to in situ temperature and salinity observations, including vertical profiles and mooring observations; SST is sea-surface temperature - to date the only SST data assimilated by BRAN is from the Pathfinder data base and from the AMSR-E mission; PmE is precipitation minus evaporation; SSS is sea-surface salinity; η is model sea-level; T_{in} is the local inertial period; IAU is Incremental Analysis Updating; ECMWF fluxes refers to ERA-40 prior to 10/2002 and ECMWF 6-hour forecasts thereafter).

Results from BRAN1 are described by Oke et al. (2005). This study demonstrates that BRAN can produce realistic mesoscale variability around Australia. However, this study also identified some problems with the Bluelink system. An error was identified in the way the surface heat flux was applied that resulted in the development of a warm bias. Some bugs were found in BODAS that meant that the salinity updates were incorrect for the first 4 years of the run. Initialization shocks, resulting from the model being updated in a single time-step, sometimes seriously degraded the reanalysis, and made the reanalyzed fields quite noisy in both space and time. All of these problems were addressed prior to operational trials of the first Bluelink forecast system at the BoM, and prior to the performance of the following BRAN experiments.

Results from BRAN1p5 are described by Oke et al. (2008). This study includes a more comprehensive assessment of BRAN, including comparisons with with-held observations. Quantitatively, it was shown that reanalyzed fields in the region around Australia in BRAN1p5 are typically within 6–12 cm of withheld atSLA observations, within 0.5–0.9^{o}C of observed SST, and within 4–7 cm of observed coastal sea-level. Comparisons with Argo profiles and surface drifting buoys show that BRAN1p5 fields are within 1^{o}C of observed sub-surface temperature, within 0.15 psu of observed sub-surface salinity, and within 0.2 m/s of near-surface currents. The fields produced by BRAN1p5 are smooth and look realistic. But it is clear from the model-data comparisons that most of the observations are under-fitted. Based on this study, initialization was identified as a key area in which the Bluelink system could be improved. Analysis of the time-mean and root-mean squared increments to sea-level also identified some biases in sea-level (Oke et al. 2008). The largest of these biases tend to be along the path of the Antarctic Circumpolar Current (ACC), indicating that perhaps the mean sea-level (MSL) field used for BRAN1p5 was inadequate. The MSL field used for all completed BRAN experiments is the time-mean of a 15-year non-assimilating run of OFAM1. A revised MSL field has recently been generated by constraining a multi-year run tightly to climatological temperature and salinity in a so-called diagnostic run using OFAM2. Research on this aspect of the Bluelink system is ongoing.

Results from BRAN2p1 are described by Schiller et al. (2008). This study includes a description of the salient features of the reanalysed circulation in the Australasian region. It included comparisons with observed and reanalyzed transport estimates for the key regions around Australia. The total (top-to-bottom) annual mean transport through the Indonesian straits, and its standard deviation, are 9.7 ± 4.4 Sv from the Pacific to the Indian Ocean with a minimum in January (6.6 Sv) and a maximum in April (12.3 Sv). The circulation of the Leeuwin Current, along the west coast of Australia, is dominated by eddy variability with a mean southward transport of 4.1 ± 2.0 Sv at 34°S. Off southern Australia, the eastward South Australian Current advects 4.5 ± 2.6 Sv at 130°E. At 32°S the East Australian Current transports 36.8 ± 18.5 Sv southward. The Coral Sea exhibits a quasi-permanent gyre between north-eastern Australia and Papua-New Guinea that is associated with the Hiri Current, which flows along the south coast of Papua-New Guinea and advects 8.2 ± 19.1 Sv into the Western Pacific Ocean. The results from BRAN2p1 are much better than BRAN1, and are very similar to those in BRAN1p5 – and like BRAN1p5, BRAN2p1 fields tend to under-fit the assimilated observations.

The latest BRAN experiment is BRAN2p2. The main difference between BRAN2p2 and BRAN2p1 and 1p5 is the initialization. The Incremental Analysis Updating (IAU) method described by Bloom et al. (1996) was adopted for BRAN2p2. Recall that BRAN2p1 and 1p5 both used nudging (Table 1). For this experiment, the increments were applied over 12 hours, with a constant weight. Examples of velocity fields at 145 m depth off south-west Western Australia, in May 2006, from BRAN1p5, 2p1, 2p2, and observations are presented in Figure 2. The observed velocity maps presented are each based on 9-days of ship-board Acoustic Doppler Current Profiler (ADCP) measurements, collected during an R. V. *Southern Surveyor* cruise in May 2006 (data provided courtesy of M. Feng) and follows a similar cruise in 2003 (Feng et al. 2007). The BRAN fields, presented for comparison, are time-averages over the sampling periods. In these example, a pair of counter-rotating eddies are evident in the observations. The eddies have a radius of only about 0.5° – approximately 50 km. This means that they are only just resolvable by a 1/10° resolution model – approximately 10 km resolution – with only about 5 grid points from the eddy center to its outer boundary. Similarly, these features are only just resolvable by the observing system. The standard GDR altimeter data, for example, comprise one estimate every 7 km along-track, and track separations are typically over 100 km. Despite these limitations, there is some evidence of these eddies in BRAN1p5 and BRAN2p1. However, they are not well reproduced. By contrast, the reproduction of these eddies in BRAN2p2 is very good, with even some of the asymmetric shapes of these eddies reproduced. There remains some errors in BRAN2p2 in the reproduction of the position of these eddies, however BRAN2p2 clearly represents an improvement in the Bluelink system in this region at this time, compared to earlier versions. The main difference between BRAN2p1 and BRAN2p2 is the initialization – nudging versus IAU.

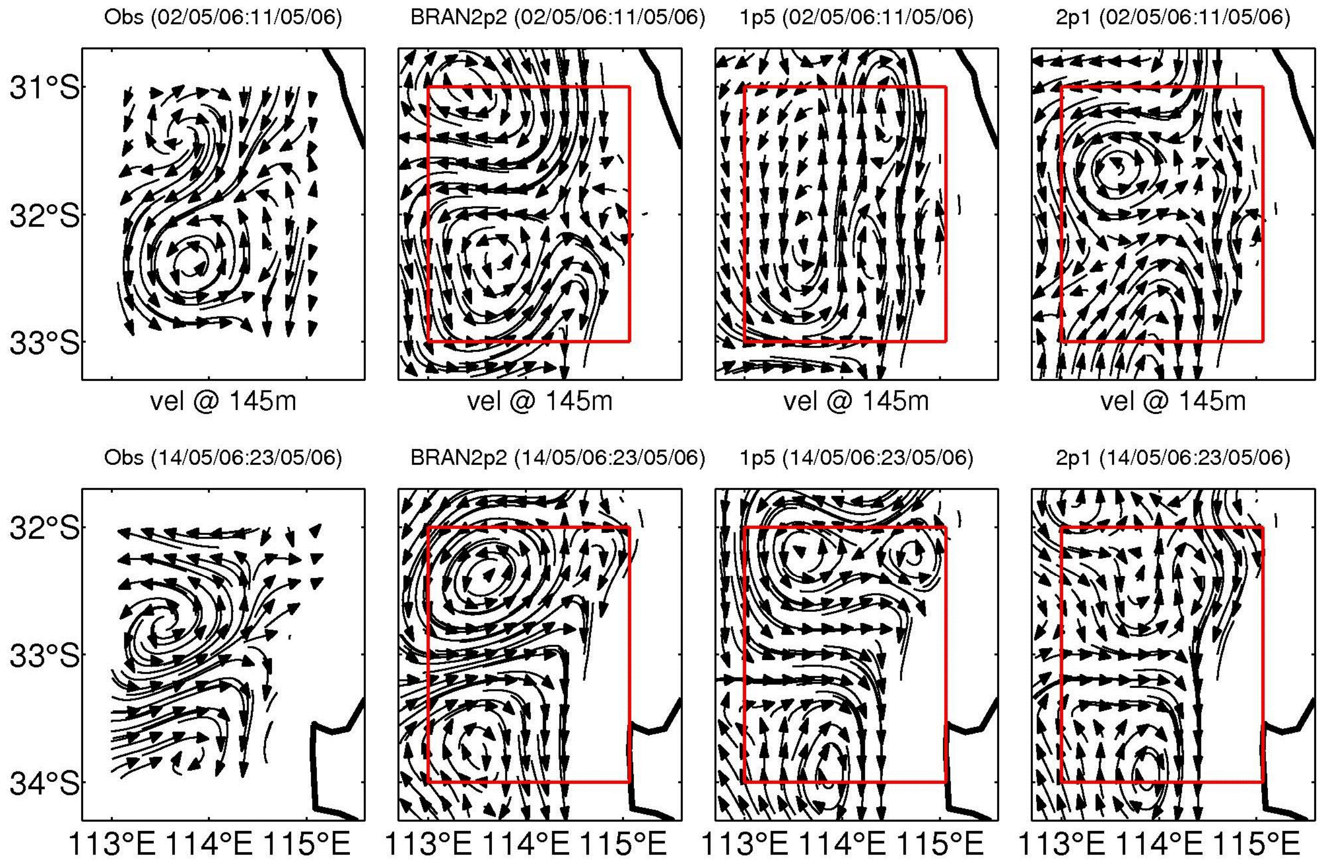

Figure 2

An example of velocities at 145 m depth from observations (left; courtesy of M. Feng) and from different versions of the BRAN for two different periods (top and bottom) off south-west Western Australia. The observed fields are mapped from ship-board ADCP measurements collected over a 9-day period. The BRAN fields are time-averages over the same period. The extent of the observation region is denoted by the red box over the BRAN fields.

Global Data Assimilation

An example of sea-level fields from the BRAN2p2 with drifter-derived velocities and trajectories overlaid is presented in Figure 3. The BRAN fields are monthly means and include the MSL. The drifter data are from the entire month. The drifter data represents the time-varying ocean circulation and is a measure of the time-integrated circulation. This is not necessarily well represented by the monthly mean sea-level fields of BRAN, but provided the variability of the circulation over each month is not too large, this comparison provides an independent assessment of the reanalyzed circulation. Note that data from the surface drifting buoys are not assimilated into BRAN. In general there is good agreement between the drifter trajectories and the sea-level contours, indicating that there is independent agreement between the reanalyzed and observed circulation. The examples in Figure 3 include situations where the drifter trajectories cross the sea-level contours. This is due on occasions to the effects of wind, or may be because a mean field (sea-level) is being compared to a Lagrangian description of the circulation (drifters). It may also be because the mesoscale features reproduced in BRAN are not precisely in the correct positions, or with the correct structures.

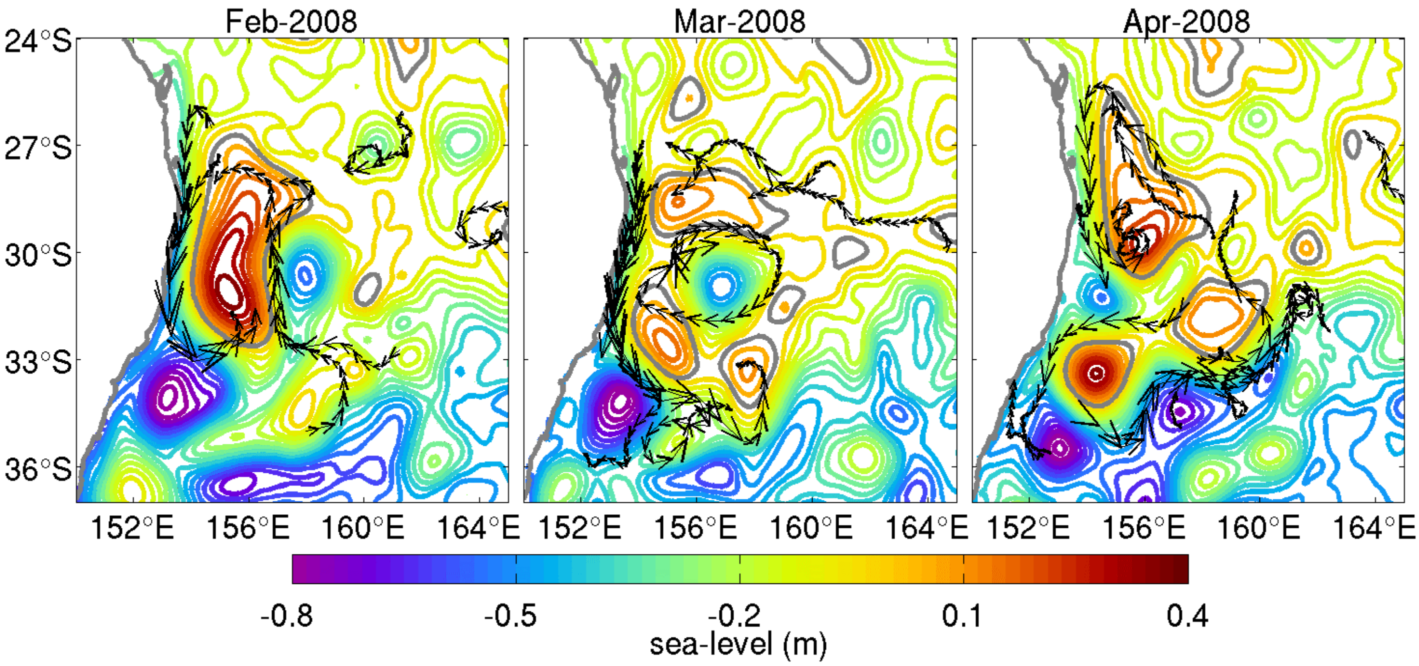

Figure 3

Monthly mean sea-level from BRAN (version 2p2), with surface drifter velocities and trajectories overlaid.

Coastal Data Assimilation

In addition to the development and application of the Bluelink global forecast and reanalysis system, the Bluelink team has developed a relocatable ocean atmosphere model (ROAM). The ROAM system is controlled by a graphical user interface that enables a non-expert user to quickly define the extent of a model domain, a forecast period (e.g., 1-7 days), and the key model components (i.e., ocean, atmosphere, waves), and execute a forecast independently in near-real-time. The intension is for ROAM to be applied by an operator for domains of around 100 to 500 km in extent. The resolution of the ocean component of ROAM is typically 1-10 km, and the model is nested within either the Bluelink ReANalysis (BRAN) system for delayed-mode applications, or the operational Bluelink forecast system for near-real-time forecasts. A recent development under Bluelink is the incorporation of BODAS into the ROAM control system and the addition of ocean data assimilation to the user's choice of specifications. The benefits of the addition of ocean data assimilation to ROAM is demonstrated here through an example to the Bonney coast, a region of frequent wind-driven upwellings, off South Australia.

The ocean model used in ROAM is the Sparse Hydrodynamic Ocean Code (SHOC; Herzfeld 2009). SHOC is a z-level primitive equation model that has been developed at CSIRO over many years. For this application, the horizontal resolution of SHOC is 5 km – twice the resolution of the Bluelink model. The surface wind stress is the same for BRAN and SHOC, and is from ERA40.

Within ROAM, SHOC is typically integrated for up to a 7-day forecast. For the examples presented here, daily mean fields of velocity, temperature, salinity, and sea-level from BRAN2p1 are used to construct the initial and boundary fields for SHOC. Each integration of SHOC includes a 4-day spin-up period, followed by a 7-day forecast.

The data assimilation in BRAN is sequential, and is performed on a 7-day update cycle. BRAN can therefore be considered to be a series of 7-day forecasts. For the examples described here, the SHOC forecasts are synchronised with the BRAN update cycle so that BRAN forecasts can be directly compared to SHOC forecasts, with and without data assimilation - hereafter denoted as SHOC and SHOC+DA, respectively. A sequence of 8 forecasts cycles are reported here, including a 4-day spin-up and 7-day forecast for each cycle. The period chosen for this comparison is the 2-month period spanning February/March 1995. This period corresponds to a series of wind-driven upwelling events, and is the focus of a detailed study by Griffin et al. (1997), who sought an explanation for a massive Pilchard die-off that occurred off southern Australia at this time.

The version of BODAS that is applied to SHOC is the same as that used for both the reanalysis and operational Bluelink systems. For the SHOC+DA runs, all of the assimilation calculations are performed on a sub-domain of the global model, with $1/10^o$ resolution. For each day of the 4-day spin-up period, the BRAN fields are modified by BODAS. For each day, BODAS treats the daily mean BRAN fields as the background field, and combines these fields with SST observations from 3-day composite AVHRR fields produced by CSIRO. An analysis field is generated for each day of the 4-day spin-up ($t<0$), and the increments from the last day of the spin-up ($t=0$) are used to update the “analysis” fields for the forecast period ($t>0$). This is intended to reduce the discontinuities in time that may occur in the transition from the spin-up period to the forecast period. In practice, on day 1 of the forecast, 80% of the day 0 increment is applied to the BRAN fields. On day 2, 60% of the day 0 increment is applied, and so on. This aspect of the assimilation has not yet been tuned properly.

Note that the assimilation performed here is all done on a sub-domain of the global model grid – not on the grid of the coastal model. One advantage of this is that the ensemble from the global model can readily be used for coastal data assimilation. Of course, this assumes that the statistics of both the global and coastal models are comparable. Another advantage is that the coastal model, SHOC, is integrated in almost the same manner for both the free run without data assimilation, and the run with data assimilation.

An example of the SST field from independent (un-assimilated) observations and from 5-day forecasts from BRAN, SHOC and SHOC+DA is presented in Figure 4. This figure shows a strong signature of wind-driven upwelling in the observations, with very cold waters upwelled to the surface and becoming advected offshore. The wind stress prior to this period is moderate and upwelling favourable (Figure 4f). Despite the coarse resolution of the ERA40 forcing fields used here ($2x2^o$), BRAN produces an upwelling, but it is weaker than the observed event. Similarly, SHOC produces an upwelling, but is also too weak. The SHOC+DA run produces a stronger upwelling that is in better agreement with the observations.

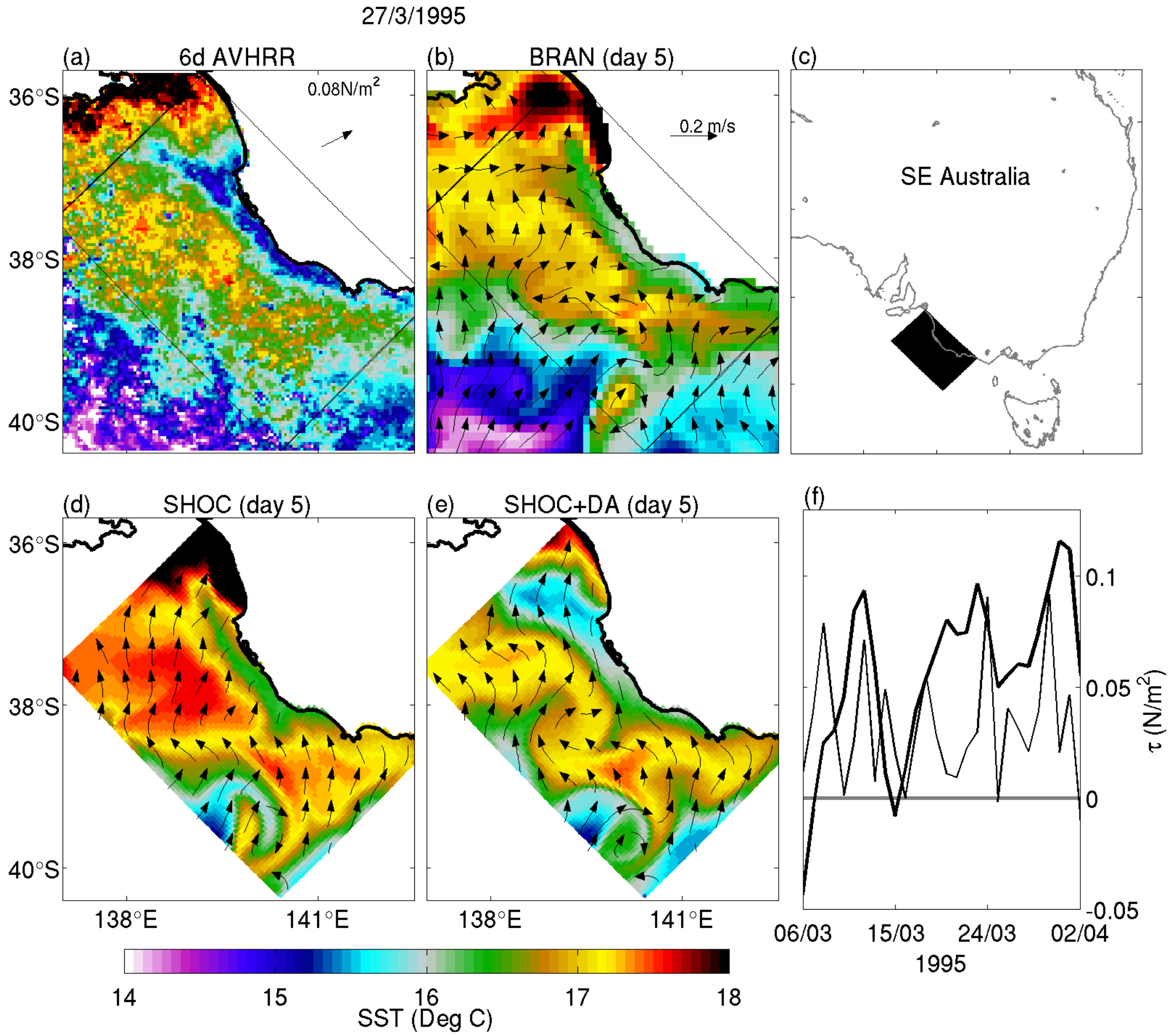

Figure 4

An example of SST from (a) 6-d composite AVHRR, (b) daily mean BRAN (version 2p1), (d) daily mean SHOC, and (e) daily mean SHOC plus data assimilation. The model fields are valid 5 days after initialisation. The arrow in panel (a) shows the daily mean wind stress along with the magnitude. The region of the SHOC domain is shown in panel (c) and the time series of zonal (bold) and meridional (thin) wind stress is plotted in panel (f). The arrows in panels (b, d, and e) show the daily mean surface velocities.

The model fields are compared to 1-day composite AVHRR SST observations across 8 consecutive 7-day forecasts. These statistics are summarized in Figure 5, showing the root-mean-squared difference (RMSD) fields, presented as a function of the forecast lead time. The forecast lead time is negative during the spin up period and positive during the forecast period. Figure 5 shows that the RMSD is greatest for SHOC without assimilation, and is smallest for SHOC+DA. The difference between these runs is greatest during the spin-up period, when SST data are assimilated, but remains significant out to 7-day forecasts.

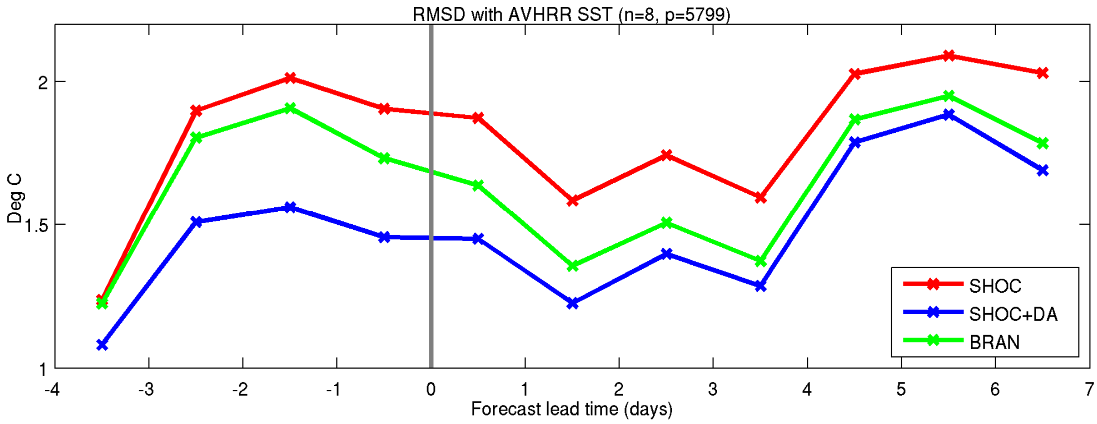

Figure 5
Root-mean-squared difference between 1-day composite AVHRR SST and daily mean SST from SHOC (no assimilation), SHOC+DA (with assimilation of AMSR-E SST), and BRAN (version 2p1).

Because SHOC has higher resolution than BRAN, one might expect SHOC, even without data assimilation, to out-perform BRAN. However, this is not the case here. This is probably because the difference in resolution is only small (5 km compared to 10 km). SHOC also has an additional source of error through the open boundaries. The fields are well behaved at the boundaries (Figure 4) with incoming features retaining their structure and out-going features leaving the domain with no obvious artifacts. However, the boundaries certainly remain a source of error.

Despite the main event considered here, a wind-driven upwelling, being due to surface forcing, rather than initialization, there is still a significant benefit of updating the initial conditions and boundary fields to better match reality. This is one demonstration of the benefit of data assimilation in coastal models.

Conclusion

BODAS was initially developed under Bluelink for global ocean data assimilation. BODAS was initially developed for short-range forecasting of the mesoscale circulation in the open ocean. But since its development, BODAS has also been incorporated into the operational Bluelink forecast system, run at the BoM, and has routinely been used for reanalysis experiments. Many aspects of the Bluelink system have been improved as a result of the BRAN experiments and the system has demonstrated measureable improvements over the lifetime of Bluelink. In addition to global data assimilation, BODAS has also been used for observing system evaluation, observing system design, and coastal data assimilation. Development of BODAS is ongoing. Specific challenges ahead include the application of BODAS to a global 1/10° model that is planned for Bluelink. Better use of observations is also an important ongoing challenge and the problem of model initialisation remains an issue. Research in these areas continues under Bluelink

Acknowledgements

Financial support for this research is provided by CSIRO, the Bureau of Meteorology, and the Royal Australian Navy as part of the Bluelink project, and the US Office of Naval Research (Grant No. N00014-07-1-0422). The author also acknowledges the many contributions of the Bluelink science team. Satellite altimetry is provided by NASA, NOAA, ESA and CNES. Drifter data are provided by NOAA-AOML and SST observations are provided by NASA, NOAA and Remote Sensing Systems.

References

Bertino, L., and Lisaeter, K.A. 2008: The TOPAZ monitoring and prediction system for the Atlantic and Arctic oceans, *Journal of Operational Oceanography*, **1**, 15–19, 2008.

Bloom, S.C., Takas, L.L., Da Silva, A.M., Ledvina, D. 1996: Data assimilation using incremental analysis updates. *Monthly Weather Review*, **124**, 1256–1271.

Brassington, G.B., Pugh, T., Spillman, C., Schulz, E., Beggs, H., Schiller, A., Oke, P.R. 2007: BLUElink> development of operational oceanography and servicing in Australia. *Journal of Research Practice Information Technology*, **39**, 151–164.

Brasseur, P., Bahurel, P., Bertino, L., Birol, F., Brankart, J.-M., Ferry, N., Losa, S., Remy, E., Schrter, J., Skachko, S., Testut, C.-E., Tranchant, van Leeuwen, P.-J., Verron, J. 2006: Data assimilation in operational ocean forecasting systems: the MERCATOR and MERSEA developments. *Quarterly Journal of the Royal Meterological Society*, **131**, 3561-3582.

Chen, D., Rothstein, L.M., Busalacchi, A.J. 1994: A hybrid vertical mixing scheme and its application to tropical ocean models. Journal of Physical Oceanography, **24**, 2156-2179.

Feng, M., Majewski, L.J., Fandry, C.B., and Waite, A.M. 2007: Characteristics of two counter-rotating eddies in the Leeuwin Current system off the Western Australian coast. *Deep-Sea Research*, ***54***, 961–980.

Griffies, S.M., Hallberg, R.W. 2000: Biharmonic friction with a Smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models. *Monthly Weather Review*, **128**, 2935-2946.

Griffies, S.M., Pacanowski, R.C., Rosati, A. 2004: A technical guide to MOM4. GFDL Ocean Group Technical Report No. 5. NOAA/Geophysical Fluid Dynamics Laboratory, 371pp.

Griffin, D.A., Thompson, P.A., Bax, N.J., Bradford, R.W., Hallegraff, G.M. 1997: The 1995 mass mortality of pilchard: no role found for physical or biological oceanographic factors in Australia, *Marine and Freshwater Research*, **48**, 27-42.

Herzfeld, M. 2009: Improving stability of regional numerical ocean models. Ocean Dynamics **59**, 21-46.

Houtekamer, P.L.,Mitchell, H.L. 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, **129**, 123–137.

Kallberg, P., Simmons, A., Uppala, S., Fuentes, M., 2004: The ERA-40 archive, European Centre for Medium-range Weather Forecasts (ECMWF), ECMWF Re-Analysis Project (ERA), ERA-40 Project Report Series 17, 31 pp.

Leonard, B.P. 1979: A stable and accurate convective modelling procedure based on quadratic upstream interpolation. *Computational Methods of Applied Mechanical Engineering*, **19**, 59–98.

Mitchell, H.L., Houtekamer, P.L., and Pellerin, G. 2002: Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Monthly Weather Review*, **130**, 2791–2808.

Oke, P.R., Schiller, A., Griffin, D.A., Brassington, G.B. 2005: Ensemble data assimilation for an eddy-resolving ocean model of the Australian region. *Quarterly Journal of the Royal Meteorological Society*, **131**, 3301–3311.

Oke, P. R., P. Sakov, and S. P. Corney 2006: Impacts of localisation in the EnKF and EnOI: Experiments with a small model. *Ocean Dynamics*, **57**, 32-45.

Oke, P.R., and Schiller, A. 2007: Impact of Argo, SST and altimeter data on an eddy-resolving ocean reanalysis. *Geophysical Research Letters*, **34**, L19601, doi:10.1029/2007GL031549.

Oke, P.R., Brassington, G.B., Griffin, D.A. and Schiller, A. 2008: The Bluelink Ocean Data Assimilation System (BODAS), *Ocean Modelling*, **21**, 46-70, doi:10.1016/j.ocemod.2007. 11.002.

Oke, P. R., and P. Sakov 2008: Representation error of oceanic observations for data assimilation. *Journal of Atmospheric and Oceanic Technology*, **25**, 1004-1017.

Oke, P. R., P. Sakov and E. Schulz, 2009: A comparison of shelf observation platforms for assimilation into an eddy-resolving ocean model. *Dynamics of Atmospheres and Oceans*, doi:10.1016/j.dynatmoce.2009.04.002, in press.

Rawlins, F., Ballard, S.P., Bovis, K.J., Clayton, A.M., Li, D., Inverarity, G.W., Lorenc, A.C., and Payne T.J. 2007: The Met Office global four-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, **133**, 347–362.

Sandery, P.A., and Brassington, G.B. 2008: Preliminary evaluation of a coupled ocean-atmosphere prediction system. *CAWCR Research Letters*, **1**, 19-23.

Schiller, A., Oke, P.R., Brassington, G.B., Entel, M., Fiedler, R., Griffin, D.A., and Mansbridge, J.V. 2008: Eddy-resolving ocean circulation in the Asian-Australian region inferred from an ocean reanalysis effort. *Progress in Oceanography*, **76**, 334-365.

Schulz, E.W., Kepert, J.D. and Greenslade, D.J.M. 2007: An Assessment of Marine Surface Winds from the Australian Bureau of Meteorology Numerical Weather Prediction Systems, *Weather and Forecasting*, **22**, 613-636.

Smith, W.H.F., and Sandwell, D.T. 1997: Global seafloor topography from satellite altimetry and ship depth soundings, *Science*, **277**, 1957-1962.