

## Combining ensemble and variational data assimilation

Dr Peter R. Oke

CSIRO Marine and Atmospheric Research and Wealth from Oceans Flagship Program  
GPO Box 1538 Hobart TAS 7001 Australia  
phone: (+61 3) 6232-5387 fax: (+61 3) 6232-5123 email: [peter.oke@csiro.au](mailto:peter.oke@csiro.au)

Dr. Alexandre Kurapov

College of Oceanic and Atmospheric Sciences, Oregon State University  
Corvallis, Oregon, USA  
Phone (+1 541) 737-2865 fax: (+1 541) 737-2064 Email: [kurapov@coas.oregonstate.edu](mailto:kurapov@coas.oregonstate.edu)

Award Number: N62909-10-1-7068,  
including partial support from ONR Global's Naval International  
Cooperative Opportunities Program (NICOP)  
<http://www.cmar.csiro.au/staff/oke/onr.htm>

### LONG-TERM GOALS

The long-term goal of this project is to develop and apply practical methods for data assimilation to improve the short-range prediction of mesoscale ocean variability.

### OBJECTIVES

The primary objective of this work is to develop an ocean data assimilation system that exploits the strengths of both the ensemble-based (e.g., Evensen 2003; Houtekamer and Mitchell 1998; Tippett et al. 2003) and variational (e.g., Bennett 2002) approaches to data assimilation. The first step in this project is to perform a comprehensive inter-comparison of an ensemble-based data assimilation system with a 4d-var system for a suite of coastal model configurations. The second step is to identify the strengths and weaknesses of each system and to improve both systems by *borrowing* components from the other system. Ultimately, a single ensemble-var system will be developed. We will investigate the extent to which the ensemble-var system can outperform both the ensemble-based and variational approaches, both in terms of forecast skill (accuracy) and computational efficiency (throughput).

### APPROACH

An ensemble-based data assimilation system, based on an Ensemble Optimal Interpolation (EnOI) scheme (Oke et al. 2002; Evensen 2003), has been developed and tested for a range of applications by Dr. Oke (Oke et al. 2005; 2008; 2009; 2010; Yin et al. 2010; O'Kane et al. 2010), and a four-dimensional variational (4d-var) system, based on the representer method (Chua and Bennett 2001; Bennett 2002), has been developed and tested for coastal ocean applications by Dr. Kurapov (Kurapov

# Report Documentation Page

Form Approved  
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE <b>30 SEP 2011</b>		2. REPORT TYPE		3. DATES COVERED <b>00-00-2011 to 00-00-2011</b>	
4. TITLE AND SUBTITLE <b>Combining ensemble and variational data assimilation</b>				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Oregon State University, College of Oceanic and Atmospheric Sciences, Corvallis, OR, 97331</b>				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			

et al. 2007; 2009 2010).

Conceptually, both the ensemble and variational approaches perform analogous tasks. Both methods implicitly generate estimates of the system's background error covariance, both interpolate and extrapolate background innovations (model-observation differences) onto the full model state (including all variables at all model grid points), and both seek to minimize some norm of the model-observation misfits. The ensemble-approach has an advantage that it uses the non-linear model operator to generate and evolve the system's error covariance, but has the disadvantage that it is rank deficient (i.e., usually uses too few ensemble members to completely span the error sub-space of the system). The variational representer-based approach, by contrast, is full rank, but it uses linear operators to derive the system's error covariance, or representers. The ensemble-approach is typically implemented as a three-dimensional interpolation (e.g., Oke et al. 2005; 2008) and the representer-approach is fundamentally four-dimensional (Kurapov et al. 2007, 2009). There are pros and cons of both approaches. Under this project, we will exploit the strengths of both approaches, with the goal of developing a superior ensemble-var data assimilation system.

A key hypothesis in this study is that the adjoint and tangent linear model developed for a particular ocean general circulation model – in this case the Regional Ocean Modelling System (ROMS) – can successfully be applied to a different ocean model – in this case the Sparse Hydrodynamic Ocean Code (SHOC; Herzfeld 2006). If this can be demonstrated, the technical developments of the ocean modelling community can potentially be consolidated significantly – and the progress of this community could be accelerated. We have made steps towards testing this hypothesis by configuring both SHOC and ROMS for the same region, on the same grid.

The EnOI system to be used under this project is routinely applied to regional configurations of SHOC. The technical aspects of applying the 4dVar system to SHOC have been achieved. This has involved porting the Advanced Variational Regional Ocean Representer Analyzer (AVRORA; Kurapov et al., 2007; 2009; 2020) to the same computer platform as SHOC, and enabling AVRORA to assimilate data from the same database that the EnOI system draws on. These steps have been largely technical - but have enabled the Australian-based researchers involved in this project to acquire a good working knowledge of AVRORA. We are now well-placed to fulfill the objectives of this project by first inter-comparing EnOI with 4dVar, then bringing together.

A key characteristic of the 4dVar approach is the state-dependence of the background error covariance estimates - quantified by the representers. Enabling state-dependent estimates of the background error covariance using the EnOI is one way of bringing these approaches together.

Dr. Peter Oke is the P.I. on this project and leads the ocean data assimilation activities at CSIRO Marine and Atmospheric Research. Dr. Oke is working closely with Dr. Kurapov, from OSU, under this project, and also in collaboration with Dr. Terrence O'Kane and Dr. Chaojiao Sun from CSIRO.

## **WORK COMPLETED**

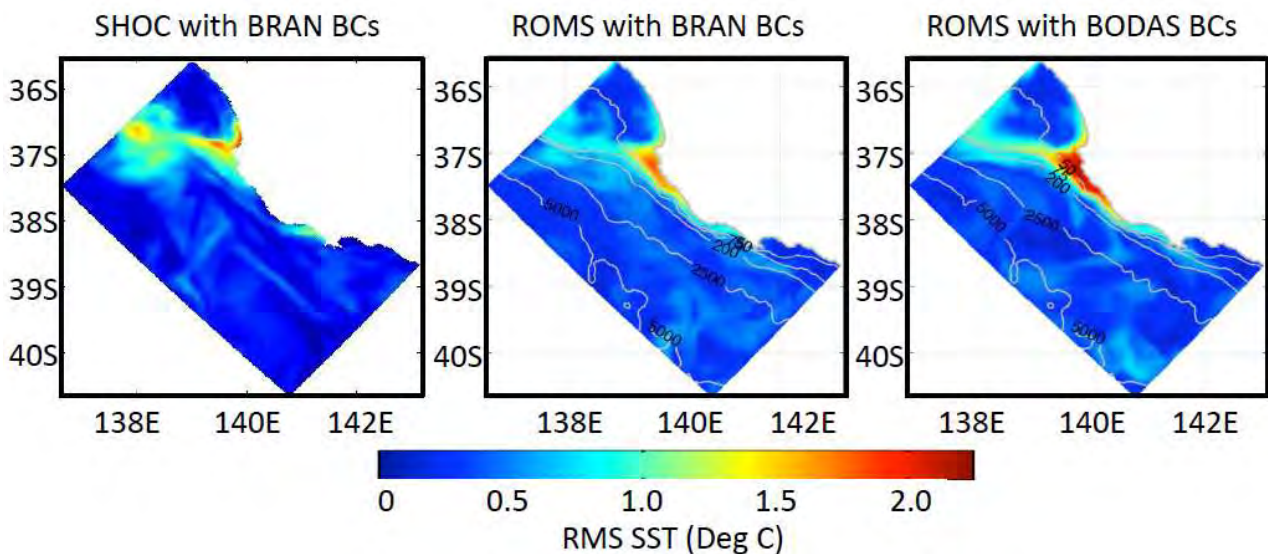
A series of comparisons between ROMS and SHOC, configured for the Bonney coast off southern Australia have been conducted. Results show that both models produce comparable wind-driven circulation. The experiments include both ROMS and SHOC nested within fields from a global ocean reanalysis, and fields from a regional EnOI-based analysis. Comparisons demonstrate that both ROMS and SHOC produce more realistic circulation when they are nested within the EnOI-based analyses.

The next step of applying the 4dVar system to both models is underway. Modules have been written to allow the 4dVar system to assimilate observations from the database used by the EnOI system. This will facilitate a direct inter-comparison of methods.

State-dependent estimates of the background error for the EnOI system have been produced by establishing a simple EnOI-breeding capability (O’Kane et al. 2010). The EnOI-breeding approach has been applied to a regional model of the Tasman Sea. Results demonstrate that this efficient approach yields state-dependent estimates of the system’s background errors that are meaningful.

## RESULTS

A comparison between ROMS and SHOC, both configured for the Bonney coast off southern Australia, has been performed. The results show that both models reproduce the variability associated with wind-driven upwelling (Figure 1). The intensity of the upwelling is similar in both models. Both models benefit from being nested within the EnOI-based regional analysis (labelled BODAS BCs). In these experiments the EnOI-based regional analyses (BODAS BCs) are computed by assimilating observed sea-surface temperature, along-track sea-level anomalies, and in situ temperature and salinity from Argo profiles every day. The background field for each analysis is derived from version 2.2 of the Bluelink ReANALYSIS (BRAN2.2; Oke and Griffin 2010).

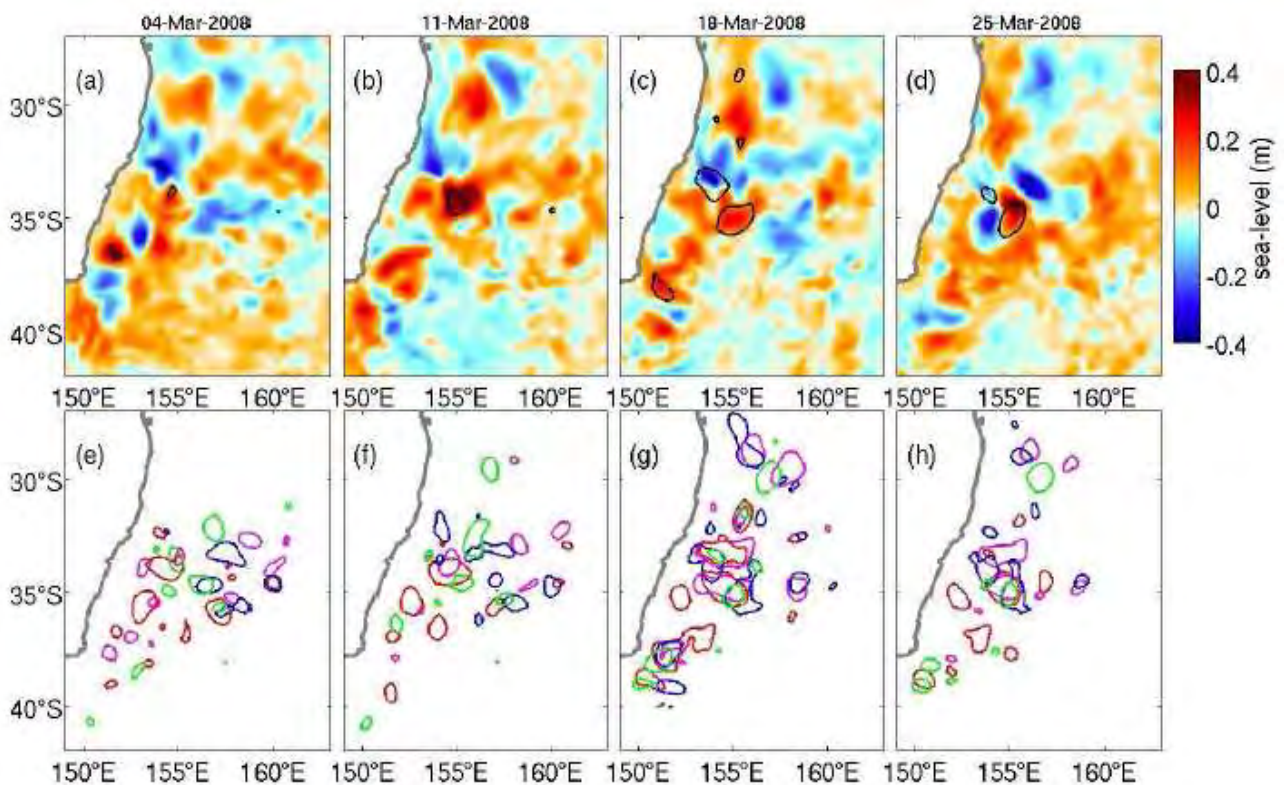


**Figure 1: Standard deviation of sea surface temperature for February 2008 from SHOC with BRAN boundary conditions (left), ROMS with BRAN boundary conditions (middle), and ROMS with BODAS(i.e., regional EnOI-based analyses) boundary conditions.**

*[Figure showing the similarity between two different ocean models off the Bonney coast, and the benefit of assimilated within EnOI-based analysis fields.]*

Breeding is a simple method that involves the perturbations of an analysis field, and the sequential rescaling of the perturbed forecast anomalies with respect to a specified norm. In theory, breeding generates the perturbations fields that project most strongly onto the fastest growing modes. Bred vectors can therefore provide an estimate of the likely growth of the background error over a given

assimilation cycle. The background errors are state-dependent. An example of breeding, applied together with an EnOI system (Oke et al. 2008), is presented in Figure 2 (from O’Kane et al. 2010). Figure 2 shows the mean of 4-bred vectors for three different times, along with an estimate of the background error for sea-level (computed by differencing a forecast with a verifying analysis). In each case shown in Figure 2 the bred vectors have large amplitude in regions that correspond to large background errors (the mean of the bred vectors is contoured in black over the estimates of forecast error). Note that the bred-vectors are computed without prior knowledge of the actual background errors. The growth of each bred vector is therefore due to growing instabilities in the model that correspond to the development or evolution of mesoscale eddies and the associated sub-mesoscale circulation. In practice, such instabilities are poorly predicted owing to their chaotic nature, so the background errors around these growing instabilities are often relatively large.



**Figure 2: Examples of (a-d) forecast error for sea-level (colour) and the ensemble averaged bred vector (contours, showing plus/minus 0.35 m contours) in the Tasman Sea; and (e-h) four bred vectors overlaid. Each bred vector is a different colour.**

*[Figure showing the alignment of the bred vectors and the state-dependence of background errors from a data assimilating ocean model.]*

## IMPACT/APPLICATIONS

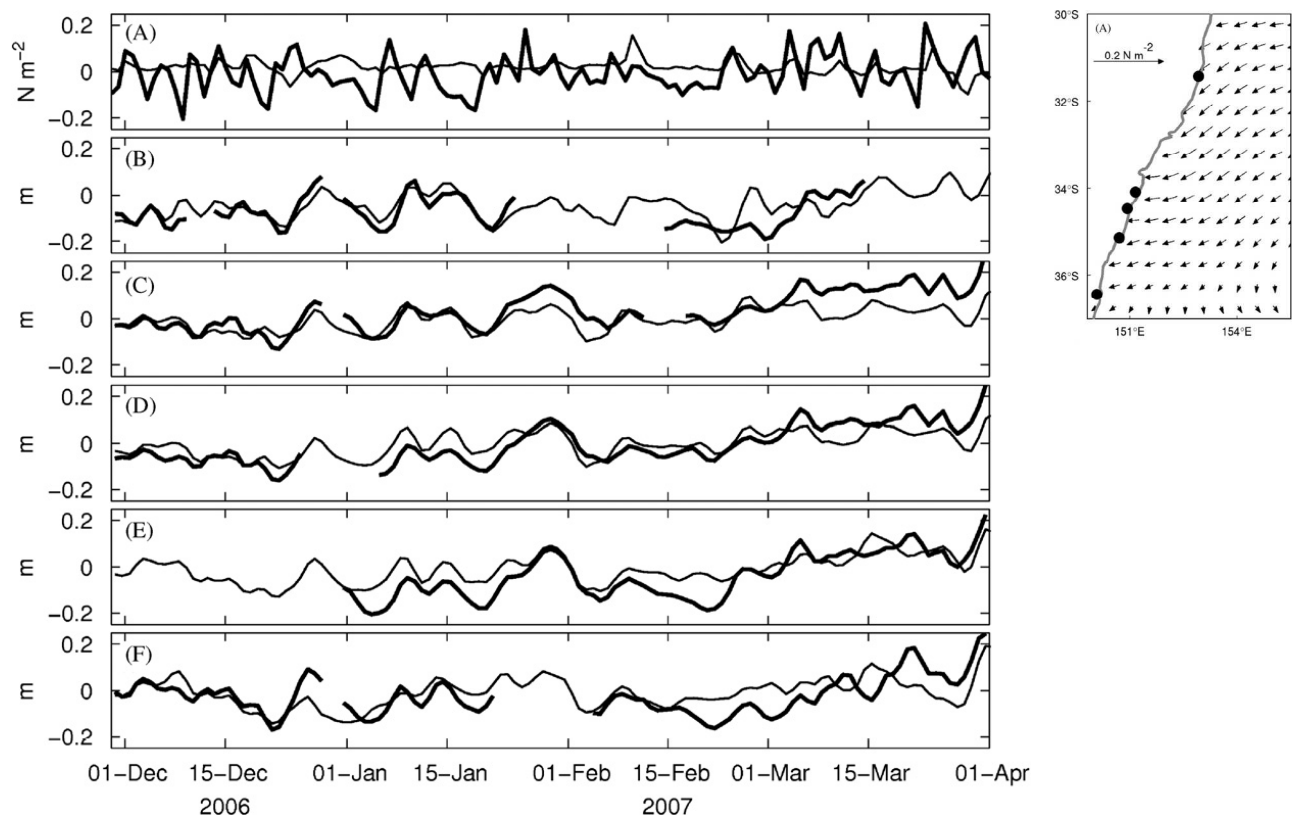
We are a step closer to evaluating the hypothesis that an adjoint and tangent-linear model from one model (in this case, ROMS) can be used, together with a 4dVar system, to assimilate observations into a second independent model (in this case, SHOC). If this hypothesis can be verified, the technical

challenges of developing an adjoint model and a tangent linear model for every new version of an ocean model will be significantly simplified.

The breeding method employed here (see Figure 2) reliably predicts the state-dependence of the system's background error. When combined with EnOI, or a 4dVar system, breeding should ultimately make these systems more efficient. Breeding is also suitable for adaptive sampling, where additional observations (e.g., gliders) could be deployed in regions of large errors or growing instabilities.

## RELATED PROJECTS

Bluelink is a partnership between CSIRO, the Bureau of Meteorology and the Royal Australian Navy. Many of the research activities undertaken in Bluelink have strong synergies with the project that is the subject of this annual report. The main objective of Bluelink is the development and application of an ocean forecast system for the mesoscale circulation around Australia. Applications of the Bluelink system are well documented (e.g., Oke et al. 2005; 2008; 2009; 2010; Schiller et al. 2008). An indication of the performance of BRAN is shown in Figure 3, where observed and modelled sea-level is compared for several unassimilated tide gauges off south-eastern Australia (Oke and Griffin 2010).



**Figure 1: Time series of (A) meridional (bold) and zonal (thin) wind stress at 35.1S; and (B–F) sea-level anomaly from BRAN (thin) and unassimilated coastal tide gauges off south-eastern Australia. Tide gauge locations are denoted by the dots in the panel to the right.**

*[Figure showing the good agreement between observed and modelled sea-level off eastern Australia from the Bluelink Reanalysis system.]*

The most recent activities under Bluelink involve the development of a near-global (excluding the Arctic) eddy-resolving ocean model. The new Bluelink model will underpin future ocean reanalyses and Australia's future operational short-range ocean forecast system. Application of the EnOI system referred to throughout this annual report is underway.

## REFERENCES

- Bennett, A. F., 2002: Inverse Modeling of the Ocean and Atmosphere, 234 pp., Cambridge Univ. Press, New York.
- Chua, B. and A. F. Bennett, 2001: An inverse ocean modeling system. *Ocean Modelling*, **3**, 137-165.
- Evensen, G., 2003: The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53**, 343–367.
- Herzfeld, M., 2006: An alternative coordinate system for solving finite difference ocean models. *Ocean Modelling*, **14**, 174-196.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, **126**, 796–811.
- Kurapov, A. L., G. D. Egbert, J. S. Allen, R. N. Miller, 2007: Representer-based variational data assimilation in a nonlinear model of nearshore circulation, *Journal of Geophysical Research*, **112**, C11019, doi:10.1029/2007JC004117.
- Kurapov, A. L., G. D. Egbert, J. S. Allen, R. N. Miller, 2009: Representer-based analyses in the coastal upwelling system, *Dynamics of Atmospheres and Oceans*, **48**, 198-218, doi:10.1016/j.dynatmoce.2008.09.002.
- Kurapov, A. L., D. Foley, P. T. Strub, G. D. Egbert, and J. S. Allen, 2010: Variational assimilation of satellite observations in a coastal ocean model off Oregon, *J. Geophys. Res.*, submitted.
- O’Kane, T. J., P. R. Oke, P. A. Sandery, 2010: Predicting the East Australian Current. *Ocean Modelling*, **38**, 251-266.
- Oke, P. R., J. S. Allen, R. N. Miller, G. D. Egbert and P. Michael Kosro, 2002: Assimilation of surface velocity data into a primitive equation coastal ocean model. *Journal of Geophysical Research*, **107**(C9), 3122-3147.
- Oke, P. R., A. Schiller, D. A. Griffin and G. B. Brassington, 2005: Ensemble data assimilation for an eddy-resolving ocean model, *Quarterly Journal of the Royal Meteorology Society*, **131**, 3301-3311.
- Oke, P. R., G. B. Brassington, D. A. Griffin and A. Schiller, 2008: The Bluelink Ocean Data Assimilation System (BODAS). *Ocean Modelling*, **20**, 46-70.
- 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*, **48**, 121-142, doi:10.1016/j.dynatmoce.2009.04.002.
- Oke, P. R., G. B. Brassington, D. A. Griffin and A. Schiller, 2010: Ocean Data Assimilation: a case for ensemble optimal interpolation. *Australian Meteorological and Oceanographic Journal*, **59**, 67-76.
- Oke, P. R., and D. A. Griffin, 2010: The cold-core eddy and strong upwelling off the coast of New South Wales in early 2007. *Deep Sea Research*, doi:10.1016/j.dsr2.2010.06.006.
- Schiller, A., P. R. Oke, G. B. Brassington, M. Entel, R. Fiedler, D. A. Griffin, and J. V. Mansbridge, 2008: Eddy-resolving ocean circulation in the Asian-Australian region inferred from an ocean reanalysis effort. *Progress in Oceanography*, **76**, 334-365.
- Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Monthly Weather Review*, **131**, 1485–1490.

Yin, Y., O. Alves, P. R. Oke, 2010: An ensemble ocean data assimilation system for seasonal prediction. *Monthly Weather Review*, **139**, 786-808.

## PUBLICATIONS

O'Kane, T. J., P. R. Oke, P. A. Sandery, 2010: Predicting the East Australian Current. *Ocean Modelling*, **38**, 251-266. [published, refereed]

Oke, P. R., T. J. O'Kane, 2010: Observing system design and assessment. *Operational Oceanography in the 21<sup>st</sup> Century*, Brassington, G. B., A. Schiller, Eds., Springer, 123-154. [published, refereed]

Yin, Y., O. Alves, P. R. Oke, 2010: An ensemble ocean data assimilation system for seasonal prediction. *Monthly Weather Review*, **139**, 786-808. [published, refereed]

Oke, P. R., M. Balmaseda, M. Benkiran, J. A. Cummings, E. Dombrowsky, Y. Fujii, S. Guinehut, G. Larnicol, P.-Y. Le Traon, and M. J. Martin, 2010: Observational requirements of GODAE Systems. In *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2)*, Venice, Italy, 21-25 September 2009, Hall, J., Harrison, D. E. & Stammer, D., Eds., ESA Publication WPP-306. [published, refereed]

Oke, P. R., and D. A. Griffin, 2010: The cold-core eddy and strong upwelling off the coast of New South Wales in early 2007. *Deep Sea Research*, doi:10.1016/j.dsr2.2010.06.006. [published, refereed]

Schiller, A., S. E. Wijffels, J. Sprintall, R. Molcard and P. R. Oke, 2010: Pathways of Intraseasonal Variability in the Indonesian Throughflow Region. *Dynamics of Atmospheres and Oceans*, **50**, 174-200. [published, refereed]

Langlais, C., A. Schiller, P. R. Oke, 2010: Southern Ocean fronts in the Bluelink reanalysis. *Mercator Quarterly Newsletter*, **36**, 50-57. [published, refereed]

Oke, P. R., G. B. Brassington, D. A. Griffin and A. Schiller, 2010: Ocean Data Assimilation: a case for ensemble optimal interpolation. *Australian Meteorological and Oceanographic Journal*, **59**, 67-76. [published, refereed]

Andreu-Burillo, I., G. B. Brassington, P. R. Oke, H. Beggs, 2010: Including a new data stream in the Bluelink Ocean Data Assimilation System. *Australian Meteorological and Oceanographic Journal*, **59**, 77-86. [published, refereed]

Preprints of submitted manuscripts are available from <http://www.cmar.csiro.au/staff/oke/>.