Check for updates

AGU100 ADVANCING EARTH AND SPACE SCIENCE

JGR

Journal of Geophysical Research: Oceans

RESEARCH ARTICLE
10.1029/2017JC013685

Observation Impact in a Regional Reanalysis of the East Australian Current System

Colette Kerry[1], **Moninya Roughan**[1], **and Brian Powell**[2]

[1]School of Mathematics and Statistics, Faculty of Science, UNSW Australia, Sydney, New South Wales, Australia, [2]Department of Oceanography, School of Earth and Ocean Sciences and Technology, University of Hawai'i at Mānoa, Honolulu, HI, USA

Key Points:
- Combining observations with a numerical model of the EAC, we reveal which observations contribute most to changes in the modeled estimate
- For the metrics used, observations taken in regions with greater natural variability contribute the most
- Using model physics to compute analysis increments, observation impact is far reaching: upstream and downstream, and forward and backward in time

Correspondence to:
C. Kerry,
c.kerry@unsw.edu.au

Citation:
Kerry, C. G., Roughan, M., & Powell, B. S. (2018). Observation impact in a regional reanalysis of the East Australian Current system. *Journal of Geophysical Research: Oceans*, *123*, 7511–7528. https://doi.org/10.1029/2017JC013685

Received 5 DEC 2017
Accepted 4 OCT 2018
Accepted article online 11 OCT 2018
Published online 23 OCT 2018

Abstract The East Australia Current dominates the circulation along the east coast of Australia; therefore, identifying observations that best constrain its transport and eddies may help improve circulation estimates. Observational data sets are sparse in time and space and numerical models are unable to predict the timing and location of eddies due to their chaotic nature. Data assimilation combines observations with a numerical model such that the model better represents the observations and provides the dynamic context. This study uses variational methods to quantify how oceanic observations from various platforms impact model estimates of transport and eddy kinetic energy in the East Australia Current. The most influential observations are, in this order, satellite-derived sea surface temperature; radial components of sea surface velocity from an high-frequency radar array midway along the coast; satellite-derived sea surface height, temperature, salinity, and velocity observations from a full-depth mooring array in the upstream portion of the domain; and subsurface hydrographic data measured by ocean gliders. Not only do the high-frequency radar observations have high impact on transport estimates at the array location, but also they have significant impact both upstream and downstream. Likewise, the impact of the mooring array observations is far reaching, contributing to transport estimates hundreds of kilometers downstream. The observation impact of deep gliders deployed into eddies is particularly high. Significantly, we find that observations taken in regions with greater natural variability contribute most to constraining the model estimates.

Plain Language Summary Estimating the ocean state requires a combination of numerical modeling and ocean observations. This study quantifies how different observation types and locations contribute to estimates of transport and eddy variability in the East Australian Current.

1. Introduction

Predicting the ocean's eddying mesoscale ocean circulation requires a combination of numerical modeling and ocean observations. Observations are typically sparse in time and space relative to the dimensionality of the ocean and do not capture the complete, three-dimensional structure of the circulation. Numerical models allow us to solve the physical equations that govern the ocean circulation on a discritized grid, but the mesoscale circulation is chaotic and models alone are unable to correctly predict the timing and location of the eddies. Using data assimilation techniques to combine the observations with the model allows the model to interpolate in time and space between observations and provide a complete ocean state estimate that represents the observations. Both the observations and the model are subject to uncertainties that are taken into account in the assimilation process. Data assimilation for state estimation is important for ocean forecasting as it allows regular updates of the initial conditions with an ocean state estimate that has reduced uncertainty.

In this work, we use four-dimensional variational data assimilation (4D-Var) to solve for increments in the model initial conditions, boundary forcing, and surface forcings such that the difference between the model solution of the time-evolving flow and all available observations is minimized over an assimilation interval. 4D-Var uses the linearized model equations and their adjoint to solve for the increments, and the length of the interval is limited by the validity of the linear assumption. In performing these calculations we compute the sensitivity of the ocean state to each individual observation, and this information can be used to quantify how particular observation platforms contribute to estimates of specific circulation metrics.

©2018. American Geophysical Union. All Rights Reserved.

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Mesoscale eddies are particularly energetic in western boundary currents (WBCs; Beal & Elipot, 2016; Jia et al., 2011; Zhai et al., 2010), and these dynamic systems are typically challenging regions to predict (e.g., Metzger et al., 2014) and observe (e.g., Roughan et al., 2014, 2017). The southward-flowing EAC is the WBC of the South Pacific subtropical gyre; it dominates the ocean circulation along the southeast coast of Australia and is characterized by high mesoscale eddy variability (Mata et al., 2006). The current strengthens and typically separates from the coast between 31° and 33° S (the EAC separation point) forming large warm- and cold-core eddies that make up the energetic eddy field in the Tasman Sea (Godfrey et al., 1980; Oke & Middleton, 2000; Wilkin & Zhang, 2007). The EAC presents an ideal laboratory for examining how ocean observations can inform state estimates of an eddy-dominated boundary current. In particular, the changing regime as the current flows poleward allows us to assess the relative impact of observing variables from different platforms in different dynamical regimes.

We have developed a numerical model of the southeastern Australia oceanic region that extends from north of where the EAC is most coherent and encompasses the eddy field in the Tasman Sea (Figure 1a). We use 4D-Var data assimilation to assimilate observations from a variety of platforms over a 2-year period (2012–2013; Kerry et al., 2016) and, in this paper, quantify the impact of specific data streams on estimates of the EAC. The EAC and its associated eddy field dominate the circulation along the east coast of Australia. We define measures of volume transport and eddy kinetic energy (EKE) for the observation impact calculations to identify observations that best constrain EAC transport and the intensity of the eddy field. As the current dynamics vary latitudinally from a more coherent current in the north to an eddy-dominated regime to the south, we consider volume transport through four shore normal sections spanning the EAC-dominated region. Because eddies dominate the circulation in the Tasman Sea after the EAC separates from the coast, we chose the metric to describe eddy intensity as the spatially averaged EKE over a region of the Tasman Sea. The results identify the most influential observation types and locations and may help toward improving observing system design for better state estimation and prediction.

2. Methodology

2.1. Data Assimilative Model Configuration

For the EAC region, we have configured a high-resolution (2.5–6 km) numerical ocean model using the Regional Ocean Modelling System (ROMS 3.4). The model takes boundary forcing from the BlueLink ReANalysis (BRAN3, Oke et al., 2013) and uses atmospheric forcing from the Australian Bureau of Meteorology's ACCESS reanalysis (Puri et al., 2013). This configuration provides a good representation of the mean and variability of the mesoscale EAC circulation (Kerry et al., 2016) and does not include the tides. We use 4D-Var data assimilation to constrain the model with 2 years of observational data (2012–2013), from a variety of traditional and newly available observation platforms, and assimilate over 5-day time windows to ensure that the linear assumption remains reasonable. The background error covariances are estimated by factorization following Weaver and Courtier (2001). We only prescribe univariate covariance; the dynamics are coupled through the use of the tangent-linear and adjoint models in the assimilation but not in the statistics of the background error covariances. The data assimilation system achieves a significant reduction in the difference between the modeled solution and both assimilated and independent observations, providing a dynamically consistent *best estimate* of the ocean state over the 2-year period.

The assimilated observations include

- AVISO SSH: daily, gridded (1/4° × 1/4°) satellite-derived sea surface height (SSH) data from Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO; Centre National d'Etudes Spatiales, 2015);
- NAVO SST: sea surface temperature (SST) from the U.S. Naval Oceanographic Office's Global Area Coverage Advanced Very High Resolution Radiometer level-2 product (NAVOCEANO's GAC AVHRR L2P SST). Data have a resolution of 4 km and are available two to three times per day;
- Aquarius SSS: daily, gridded (1° by 1°) satellite-derived sea surface salinity (SSS) from the National Aeronautics and Space Administration's Aquarius satellite (www.aquarius.umaine.edu).
- Argo floats: temperature and salinity from 1,229 profiles of the upper 2,000 m from Argo floats (www.argo.ucsd.edu);
- XBT: temperature profiles from expendable bathythermographs (XBTs) along repeat lines; PX34, which is the Sydney-Wellington route, and PX30, which is the Brisbane-Fiji route;

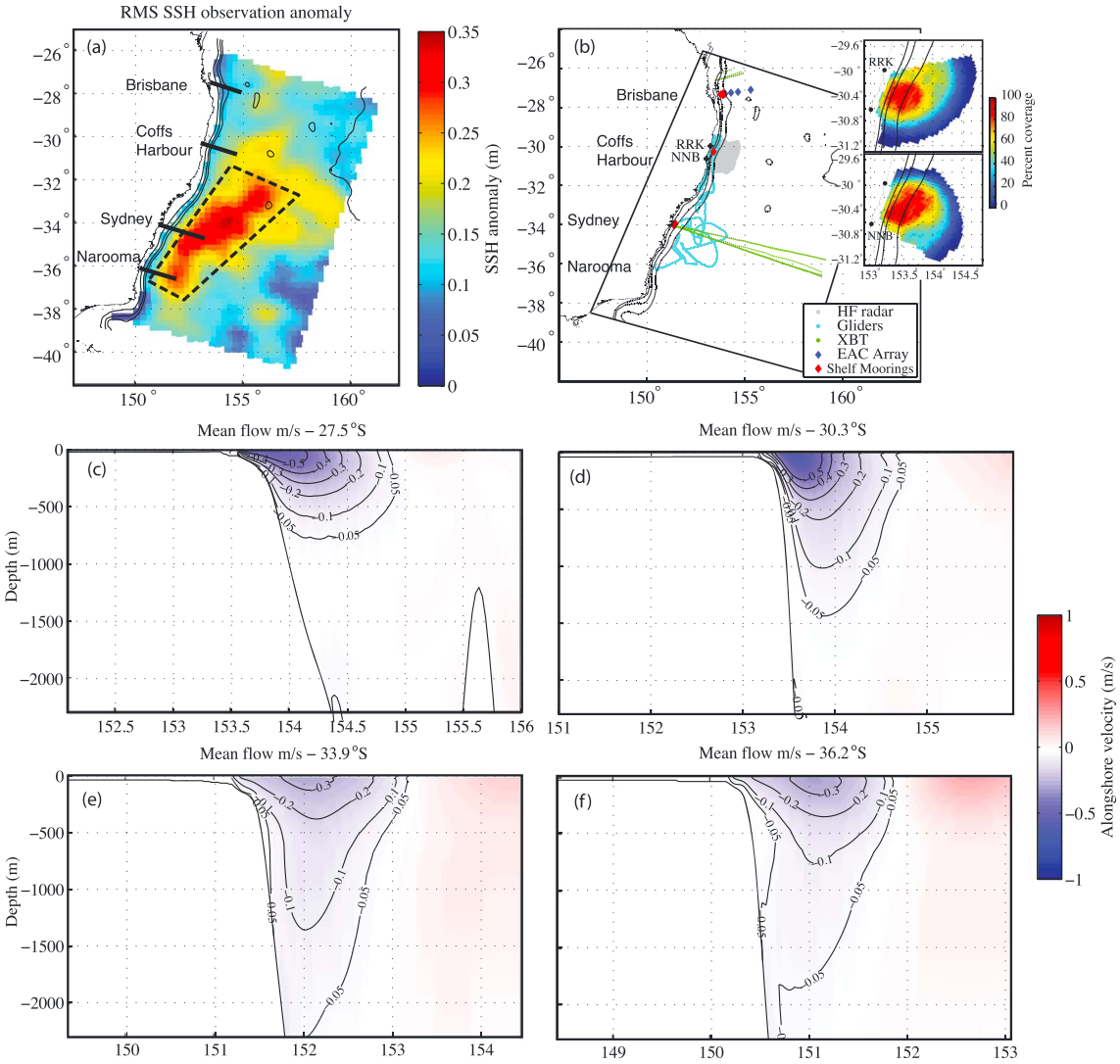

Figure 1. (a) Model domain with RMS SSH anomalies from 10 years of AVISO data. (b) Observations used in the assimilation (SSH, sea surface temperature, sea surface salinity, and Argo observations are not shown). Insets in (b) show the percentage coverage of the surface radial data from the two high-frequency radar stations. Mean alongshore velocity from the 2-year reanalysis through shore normal sections crossing the coast at Brisbane (27.5° S, [c]), Coffs Harbour (30.3° S, [d]), Sydney (33.9° S, [e]), and Narooma (36.2° S, [f]). The black solid lines in (a) are the sections across which the alongshore volume transport is computed; the black dashed line shows area over which spatially averaged eddy kinetic energy is computed. EAC = East Australia Current; RMS = root-mean-square; SSH = sea surface height; XBT = expendable bathythermograph.

- HF Radar: radial component of surface currents (hereafter referred to as *radial velocities*) from a high-frequency (HF) radar WERA phased array system with two sites north and south of Coffs Harbour (Red Rock, 29.98° S, 153.23° E, and North Nambucca, 30.62° S, 153.011° E). The instruments and data are managed by the Australian Coastal Ocean Radar Network (http://imos.org.au/acorn.html), and the radial velocities are assimilated directly following the method described by De Souza et al. (2015);
- NSW shelf moorings: temperature and velocity data from three moorings located along the New South Wales (NSW) continental shelf, located off Coffs Harbour, 30° S (CH100), and Sydney, 33.9° S (SYD100 and SYD140);
- SEQ shelf moorings: temperature, salinity, and velocity data from two moorings located on continental shelf and shelf slope off Brisbane in South East Queensland (SEQ), 27.5° S, (SEQ400 and SEQ200) in approximate water depths of 400 and 200 m, respectively;
- EAC transport array: an array of five deep water moorings (EAC 1–5), which measure temperature, salinity, and velocities throughout the water column. The array is the offshore extension of the SEQ shelf moorings off Brisbane, was positioned where the EAC is predicted to be most coherent and was designed to measure the mean and time-varying EAC transport (Sloyan et al., 2016);

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

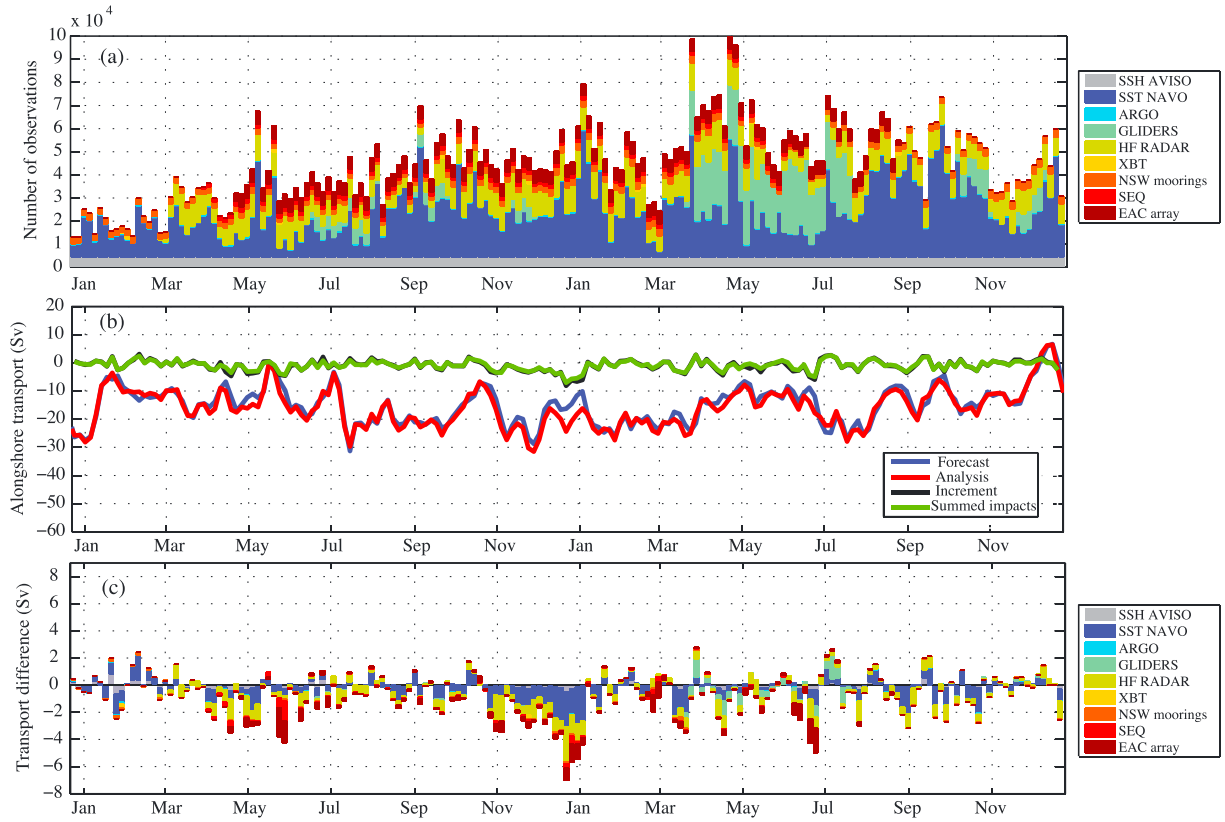

Figure 2. (a) Number of observations from each observation platform in each 5-day assimilation window over the 2-year reanalysis. (b) Volume transport through the shore normal section crossing the coast at Brisbane (27.5° S) for the forecast and the analysis, averaged over each 5-day window, the difference in transport between the forecast and the analysis (the increment), and the summed observation impacts. Note that the increment and summed impacts should match closely if the linear assumption is valid. (c) The impact of each observation platform on the transport increment for each 5-day window. AVISO = Archiving, Validation and Interpretation of Satellite Oceanographic data; EAC = East Australia Current; HF = high frequency; SEQ = South East Queensland; SSH = sea surface height; SST = sea surface temperature; NAVO = Naval Oceanographic Office; NSW = New South Wales; XBT = expendable bathythermograph.

- Ocean gliders: temperature and salinity profiles from autonomous ocean gliders deployed by the Australian National Facility for Ocean Gliders (http://imos.org.au/anfog.html) mostly on the NSW continental shelf, between 29.5° S and 32.3° S, with two missions between 20 March 2013 and 22 July 2013 extending offshore, further south, and to depths of 900 m.

The locations of the glider missions, XBT lines, shelf and offshore moorings, and the HF radar array are shown in Figure 1b, with the insets showing the percentage coverage of the surface radial velocity data from the HF radar. The 1,229 Argo floats are distributed fairly evenly across the model domain over the 2-year period (see Figure 5a of Kerry et al., 2016). The temperature and velocity observations from the moorings are collected at high sampling frequencies (every 5–10 min) and are processed prior to assimilation to remove high frequency variability not resolved by the model. The temperature observations are low-pass filtered to remove variability at periods shorter than the inertial period and the observations are applied 6 hourly in the assimilation. Velocity observations are low-pass filtered at 30 hr to remove variability due to tides and inertial oscillations and applied 6 hourly. It is important to remove the tidal signal from velocity observations as the barotropic tidal velocities are of a similar order of magnitude to the subtidal velocities. The speeds and angles of the radial velocities from the HF radar are spatially averaged onto the model grid and a 24-hr boxcar-averaging filter is used to remove tides and inertial oscillations that are not resolved by the model. The number of observations (after processing) from each observation platform for each 5-day assimilation window is shown in Figure 2a. The SST data are patchy with a varying number of observations per assimilation window as the satellites do not provide observations through clouds.

The reader is referred to Kerry et al. (2016) for a thorough description of the model configuration, data assimilation scheme, observations, and the reanalysis performance.

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.2. Observation Impacts

We are interested in understanding how the observations impact our estimate of the ocean circulation. In solving the state estimation problem with 4D-Var, we compute the dynamical error covariance between the observations and the model that allows us to directly compute the impact of each observation on the circulation estimate. This paper does not aim to provide a thorough mathematical description of 4D-Var data assimilation, which can be found in Bennett (2002), Courtier et al. (1993), Courtier et al. (1994), Le Dimet and Talagrand (1986), Moore et al. (2004), Moore, Arango, Broquet, Edwards, et al. (2011), Stammer et al. (2002), Talagrand and Courtier (1987), Thépaut and Courtier (1991), and Weaver et al. (2003). The formulation of the observation impact calculations is summarized below; for more detail the reader is directed to Moore, Arango, Broquet, Powell, et al. (2011) and Powell (2017).

We use 4D-Var to adjust the forecast model initial conditions, boundary conditions, and surface forcing such that the difference between the new model solution (the analysis) and the observations is minimized, in a least squares sense, over the assimilation time window. The state estimation problem can be written as

$$\mathbf{x}_a = \mathbf{x}_f + \mathbf{K}\mathbf{d}, \tag{1}$$

where $\mathbf{x}_f$ is the model forecast state vector and $\mathbf{x}_a$ is the resulting adjusted model state, referred to as the analysis. The innovation vector, $\mathbf{d}$, describes the residuals between the observations ($\mathbf{y}$) and the model and is given by $\mathbf{d} = \mathbf{y} - \mathcal{H}(\mathbf{x}_f)$, where $\mathcal{H}$ transforms the model forecast state vector $\mathbf{x}_f$ to observation space. $\mathbf{K}$ is the Kalman gain matrix that encompasses the observation and background error covariance matrices and, in 4D-Var, the tangent linear and adjoint models (refer to Powell, 2017, Equation (2) and associated discussion). The $\mathbf{K}$ matrix acts on the innovation vector to compute the analysis increments (in model space).

In order to quantify the observation impacts we must define a scalar measure of the ocean circulation that is a function of the model state variables, $J = Q(\mathbf{x})$, where $\mathbf{x}$ is the model state vector (all model state variables, boundary, and surface forcings). As such, we are interested in how our scalar measure of the circulation J changes over the assimilation interval,

$$\Delta J = Q(\mathbf{x}_a) - Q(\mathbf{x}_f). \tag{2}$$

It follows that

$$\Delta J = Q(\mathbf{x}_f + \mathbf{K}\mathbf{d}) - Q(\mathbf{x}_f). \tag{3}$$

A first-order Taylor Expansion gives

$$\Delta J = \mathbf{d}^T \mathbf{K}^T \frac{\partial Q}{\partial \mathbf{x}_f}. \tag{4}$$

$\mathbf{K}^T$ is the adjoint of the assimilation procedure, and $\frac{\partial Q}{\partial \mathbf{x}_f}$ is the derivative of Q about the model forecast state vector. Using equation (4) we can quantify the contribution of each individual observation to the change in J between the forecast and the analysis. $\mathbf{K}^T$ acts to project the sensitivity of the measure to the forecast onto the observation locations, and the impact is found by scaling this sensitivity by the innovation. Some scalar functions require second-order expansion, and following Gelaro et al. (2007) the solution is then given by

$$\Delta J = \frac{1}{2}\mathbf{d}^T \mathbf{K}^T \left[\frac{\partial Q}{\partial \mathbf{x}_a} + \frac{\partial Q}{\partial \mathbf{x}_f}\right]. \tag{5}$$

Here we use these methodologies to understand how observations impact estimates of five different scalar measures of the ocean circulation. We look at estimates of volume transport through four shore normal sections and of spatially averaged EKE over a specified region in the Tasman Sea. The circulation measures are defined below.

2.2.1. Alongshore Volume Transport

The four shore normal sections are shown in Figure 1a overlain on a plot of the root-mean-square (RMS) SSH observation anomaly from 10 years of AVISO data, providing a measure of SSH variability. The sections are located shore normal from the coast originating at

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

- Brisbane: 27.5° S,
- Coff Harbour: 30.3° S,
- Sydney: 33.9° S, and
- Narooma: 36.2° S.

The cross-sectional area over which we compute the transport is based on the mean alongshore velocity, v, sections from the 2-year reanalysis, shown in Figures 1c to 1f. The transport is computed for the region within the 0.05 m/s contour. The mean current weakens and deepens as it moves poleward, and so the cross-sectional area for the transport computation is different for each section.

For each of the four sections, our chosen measure, J, is the mean volume transport through the section over the 5-day assimilation window. The transport through a rectangular shore normal section from cross-shore distance x_0 to x_i, from the surface down to a chosen depth D, averaged over the time interval from t_0 to $t_0 + T$, is given by

$$J = \frac{1}{T \times 10^6} \int_{t_0}^{t_0+T} \int_{-D}^{0} \int_{x_0}^{x_i} (\mathbf{v}) \mathrm{d}x\mathrm{d}z\mathrm{d}t, \tag{6}$$

where $\mathbf{v}$ is the alongshore velocity and J has units of sverdrups. Computing the transport through the chosen cross-sectional area (within the 0.05 m/s contour in the v mean), we write $J = (1/10^6)\overline{\Sigma_{i_0}^{i_{\mathrm{end}}}(\mathbf{v}_i \mathrm{d}x_i \mathrm{d}z_i)}$, where i_0 to i_{end} are the grid cell indices within the area, the overbar represents the time mean over the 5 days, $\mathbf{v}$ is the alongshore velocity, and dxdz is the cell area. For volume transport we use the first-order formulation (equation (4)) for the impact calculation.

2.2.2. Eddy Kinetic Energy

We investigate the observation impacts on a scalar measure describing EKE over the region of elevated SSH variability in the Tasman Sea. Our chosen measure, J, is the mean EKE over each 5-day assimilation window, spatially averaged over the area shown in Figure 1a and from the surface to a depth of 450 m. The upper 450 m was chosen as the EKE is highest in this depth range. The EKE averaged over the area from x_0 to x_i and y_0 to y_i, from the surface down to a chosen depth $D = 450$ m, averaged over the time interval from t_0 to $t_0 + T$, is given by

$$J = \frac{1}{TDA \times 10^4} \int_{t_0}^{t_0+T} \int_{-D}^{0} \int_{x_0}^{x_i} \int_{y_0}^{y_i} \left((u - \bar{u})^2 + (v - \bar{v})^2 \right) \mathrm{d}y\mathrm{d}x\mathrm{d}z\mathrm{d}t, \tag{7}$$

where $\mathbf{u}$ and $\mathbf{v}$ are the cross-shore and alongshore velocities, respectively; the overbar denotes the time mean, A is the total area, and J has units square centimeters per square second. Computing this from the model output for grid cell indices i_0 to i_{end} inside the chosen area and to a depth of 450 m, we write $J = (1/V \times 10^4)\overline{\Sigma_{i_0}^{i_{\mathrm{end}}}((u_i - \bar{u}_i)^2 + (v_i - \bar{v}_i)^2)}\mathrm{d}x_i\mathrm{d}y_i\mathrm{d}z_i$, where d$xdydz$ is the cell volume and V is the total volume. Because the EKE calculation has squared terms, we use the second-order Taylor expansion formulation (equation (5)) for the impact calculation.

3. Results

3.1. Alongshore Volume Transport

The EAC flows poleward and is most coherent off Brisbane (27.5° S), also the location of the EAC mooring array (Sloyan et al., 2016). At Coffs Harbour (30.3° S), just upstream of where the EAC typically separates from the coast, the time-mean EAC transport is greater than that at off Brisbane due to recirculation (Kerry et al., 2016). Downstream of Coffs Harbour (30.3° S) the current, having strengthened on the narrowing shelf, separates from the coast forming an energetic eddy field in the Tasman Sea (Mata et al., 2006). The Sydney cross section (33.9° S), downstream of the typical separation point, traverses this region of high eddy variability and the Narooma section is downstream of it at 36.2° S (see Figure 1a). For the 2-year reanalysis, time series of volume transport through the four shore normal sections, averaged over each 5-day window, for the forecasts and the analyses are shown in Figures 2b, 3a, 4a, and 5a. The transport varies least off Brisbane, where the EAC is most coherent, and most off Sydney, where the eddy variability is the greatest.

For each 5-day window, we are interested in how the observations impact the change in 5-day averaged transport estimate between the forecast (before data assimilation) and the analysis, ΔJ (equation (2)). The

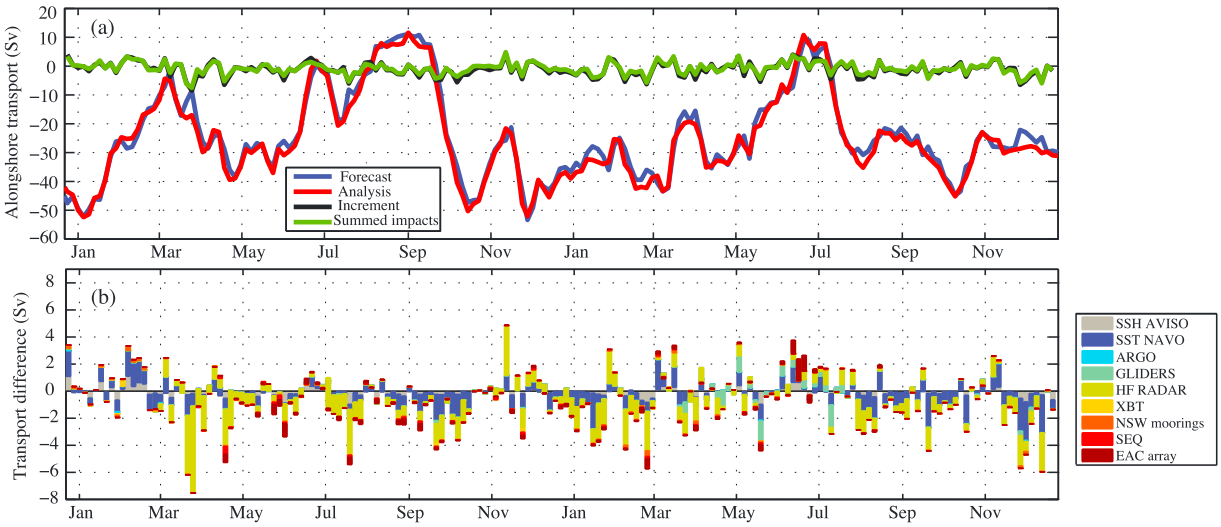

Figure 3. (a) Volume transport through the shore normal section crossing the coast at Coffs Harbour (30.3° S) for the forecast and the analysis, averaged over each 5-day window, the difference in transport between the forecast and the analysis (the increment), and the summed observation impacts. (b) The impact of each observation platform on the transport increment for each 5-day window. AVISO = Archiving, Validation and Interpretation of Satellite Oceanographic data; EAC = East Australia Current; HF = high frequency; SEQ = South East Queensland; SSH = sea surface height; SST = sea surface temperature; NAVO = Naval Oceanographic Office; NSW = New South Wales; XBT = expendable bathythermograph.

observation impacts are computed using the linearized model equations and, for the impacts to be valid, the linear assumption must remain reasonable. If this is the case the increment ΔJ, which represents the transport difference between the two nonlinear models, should match closely the sum of the observation impacts for all observations. These are shown by the black and green lines in Figures 2b, 3a, 4a, and 5a and match well for all four sections.

Each observation may increase or decrease the volume transport depending on how that observation changes the state estimate. The sum of the observation impacts for each observation platform is presented to illustrate the combined effect, shown in Figures 2c, 3b, 4b, and 5b for transport through the four sections. A negative (positive) transport increment corresponds to an increase (decrease) in poleward transport in the

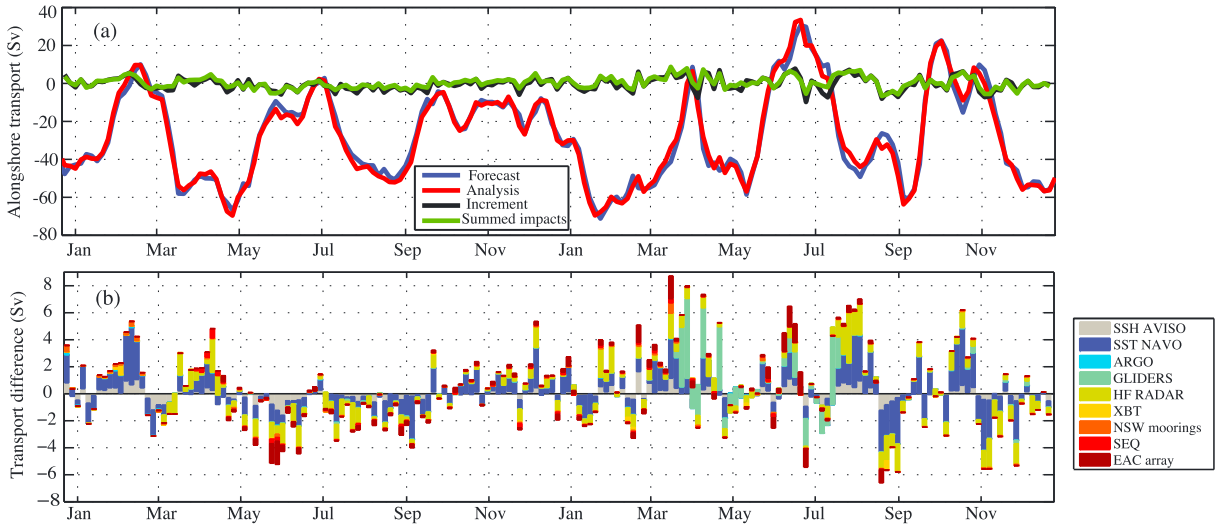

Figure 4. (a) Volume transport through the shore normal section crossing the coast at Sydney (33.9° S) for the forecast and the analysis, averaged over each 5-day window, the difference in transport between the forecast and the analysis (the increment), and the summed observation impacts. (b) The impact of each observation platform on the transport increment for each 5-day window. AVISO = Archiving, Validation and Interpretation of Satellite Oceanographic data; EAC = East Australia Current; HF = high frequency; SEQ = South East Queensland; SSH = sea surface height; SST = sea surface temperature; NAVO = Naval Oceanographic Office; NSW = New South Wales; XBT = expendable bathythermograph.

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

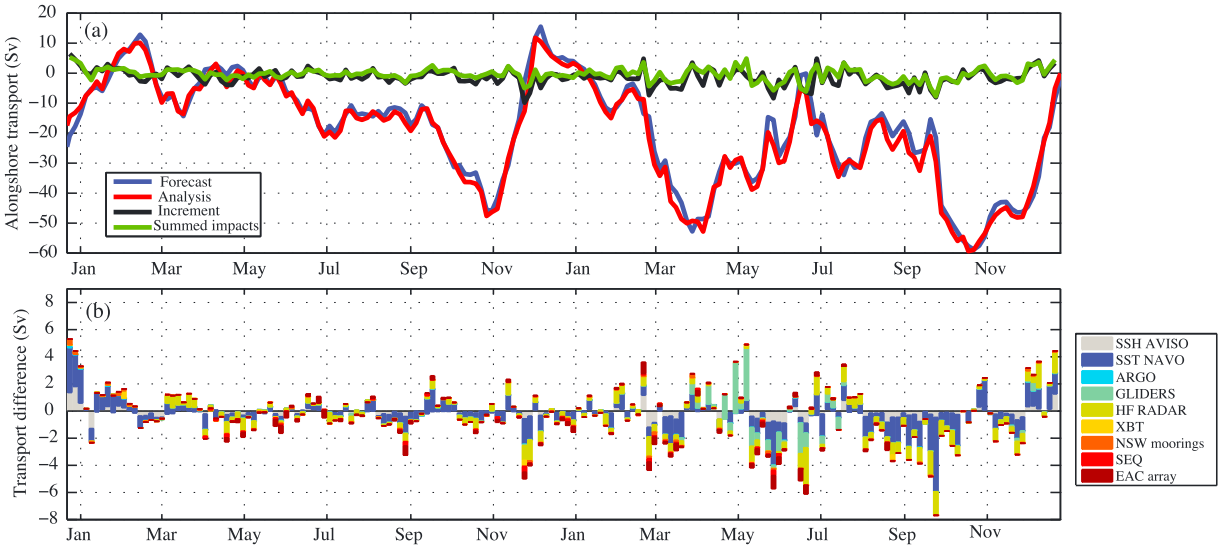

Figure 5. (a) Volume transport through the shore normal section crossing the coast at Narooma (36.2° S) for the forecast and the analysis, averaged over each 5-day window, the difference in transport between the forecast and the analysis (the increment) and the summed observation impacts. (b) The impact of each observation platform on the transport increment for each 5-day window. AVISO = Archiving, Validation and Interpretation of Satellite Oceanographic data; EAC = East Australia Current; HF = high frequency; SEQ = South East Queensland; SSH = sea surface height; SST = sea surface temperature; NAVO = Naval Oceanographic Office; NSW = New South Wales; XBT = expendable bathythermograph.

analysis. The number of observations from each observation platform is shown in Figure 2a. The transport increments range from about ±8 Sv and are typically a factor of 2 smaller than the tendency term (the change in 5-day averaged transport from one 5-day window to the next). For the four sections, the mean of the absolute value of the increments is 8–11% of the mean transport.

The contribution of each observing platform to changes in the model estimates varies considerably over the 2-year period, as it depends on the evolving flow and the observation coverage for each specific assimilation window. It should be noted that, because the impacts are scaled by the innovations, high impacts often occur when the model-data differences are large; for example, for observations in regions where the model state more rapidly evolves in the forecast. Furthermore, because we are summing the individual impacts over all observations in each platform, observation platforms with greater data volume are often more dominant. Observation impact differs from observation sensitivity (see Moore, Arango, Broquet, Powell, et al., 2011), which describes the sensitivity of circulation estimates to changes in the observations and is independent of the actual values of the observations. Interpretation of the impact results with respect to these considerations is discussed in more detail in section 4 below.

3.1.1. Time-Averaged Observation Impacts

Averaging the impacts of each platform over all assimilation windows over the 2-year period provides an overview of the average dominance of the various platforms. This is presented in Table 1, which shows percentage impacts of each observation platform over the 2 years. The percentages are calculated as the absolute values of the impact, I, for each platform, p, as a percentage of the total summed absolute impacts across all platforms, averaged over the 2 years. The percentage impact of each platform is given by

$$\Sigma_{i=1}^{i=n}\left(\frac{\mid I_{i,p} \mid}{\Sigma_{p=1}^{p=N} \mid I_{i,p} \mid}\right) \times \frac{100}{n}, \tag{8}$$

where $i = 1 : n$, n is the number of assimilation intervals, and $p = 1 : N$, where N is the number of platforms, and $I_{i,p}$ represents the impact of platform p for assimilation interval i (in sverdrups).

The SST observations dominate the number of observations and have the greatest impact on average over 2 years for the Brisbane, Sydney, and Narooma sections. For the Coffs Harbour section, the HF radar observations have comparable impact to SST on average. The other platforms that have significant impacts at all four sections are the HF radar, the satellite-derived SSH, the EAC transport array, and the gliders. The influence of

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
Average Percentage Impact of Each Observation Platform Over the 2-Year Period, Calculated as the Time Mean of the Absolute Values of the Impact for Each Platform as a Percentage of the Total Summed Absolute Impact (Equation (8))

	Mean % no. obs.	Mean % impact				
		27.5° S	30.3° S	33.9° S	36.2° S	EKE
SSH AVISO	11.77	10.67	10.36	12.46	13.76	11.76
SST NAVO	43.62	38.48	35.86	43.11	42.97	43.49
SSS Aquarius	0.35	0.005	0.005	0.005	0.006	0.006
Argo	1.03	0.85	0.86	0.84	1.01	0.93
Gliders	8.85	6.10	5.48	6.87	5.68	7.24
HF radar	17.22	28.62	36.41	23.81	23.80	23.90
XBT	0.19	0.37	0.52	0.64	0.58	0.61
NSW moorings	6.73	2.74	2.83	2.48	2.43	2.61
SEQ	2.40	2.80	1.62	2.08	1.84	2.20
EAC array	7.83	9.38	6.06	7.71	8.69	7.26

Note. AVISO = Archiving, Validation and Interpretation of Satellite Oceanographic data; EAC = East Australia Current; HF = high frequency; SEQ = South East Queensland; SSH = sea surface height; SSS = sea surface salinity; SST = sea surface temperature; NAVO = Naval Oceanographic Office; NSW = New South Wales; XBT = expendable bathythermograph.

the HF radar array and the EAC mooring array is far reaching. The HF radar observations dominate the impacts for the Coffs Harbour section, where the instruments are located, but also have significant impact on transport estimates both upstream and downstream. Likewise, the EAC array has the greatest impact at its location at 27.5° S but also impacts estimates as far as Narooma, 1,000 km downstream.

The percentage impacts of SSH and SST observations are greatest for the two poleward sections (Sydney and Narooma) that traverse the more eddy-dominated region, compared to the Brisbane and Coffs Harbour sections where the EAC flow is more coherent. Conversely, the HF radar observations are more dominant off Coffs Harbour and Brisbane relative to the two poleward sections, as the radar array constrains the EAC flow. The Brisbane cross section is located where the EAC is most coherent (Sloyan et al., 2016) and off Coffs Harbour the EAC is observed as a coherent jet in the HF radar footprint where it flows on average within 50 km of the coast (Archer et al., 2017). Further south, eddies dominate and influence the transport through the Sydney and Narooma sections, and the satellite observations are particularly useful in constraining the eddies.

From the 20 March 2013 to the 22 July 2013, two glider missions extended into the eddy field offshore of Sydney reaching depths of 900 m. Over this period, the glider observations had a significant impact on transport estimates off Sydney and Narooma, as well as at the cross sections to the north. Indeed, if we compute the percentage impact in the same way as for Table 1 but only over this time period, the impact of the glider observations is 29.89%, 28.24%, 35.97%, and 29.48% for the transport through 27.5°S, 30.3°S, 33.9°S, and 36.2°S, respectively.

3.1.2. Latitudinal Dependence

The EAC flow evolves from north to south with a deepening and weakening of the mean flow (Figure 1) as the circulation regime transitions from a more coherent boundary current to a more eddy-dominated regime poleward of the separation point. Because of this latitudinal variation in current dynamics it is particularly interesting to analyze how the observation impacts vary with latitude. To do this we present the contribution of observations from each observation platform on transport through the four shore normal sections summed in latitude bins (Figure 6).

Using bins of 0.25°, the absolute values of the impacts of all observations that fall within each bin over all assimilation windows are grouped by observation platform and summed (Figures 6a to 6d). The *x* axes represent the summed absolute value of the (positive and negative) change in transport through the particular cross-shore section that is a result of all of the observations that fall in that particular latitudinal bin over 2 years. For example, referring to Figure 6a, all of the observations that fall between 30.30°S and 30.55°S con-

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

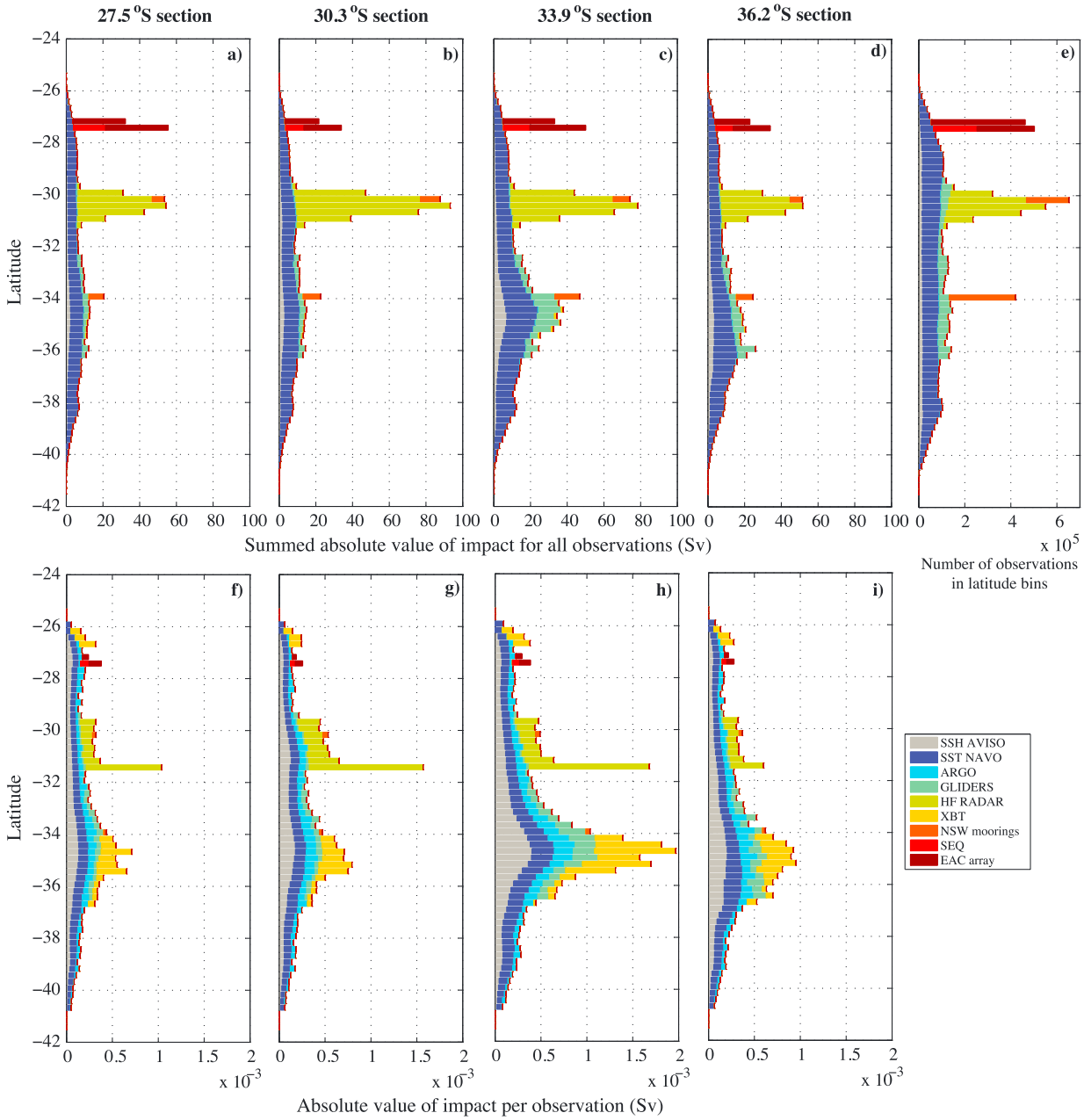

Figure 6. The observation impacts on transport through the four shore normal sections grouped into latitude bins of 0.25°. Summed absolute value of impacts for observations from each observation platform in each bin over the 2 years (a–d), the number of observations in each bin (e), and as in (a–d) but normalized by the number of observations (f–i). AVISO = Archiving, Validation and Interpretation of Satellite Oceanographic data; EAC = East Australia Current; HF = high frequency; SEQ = South East Queensland; SSH = sea surface height; SST = sea surface temperature; NAVO = Naval Oceanographic Office; NSW = New South Wales; XBT = expendable bathythermograph.

tribute to a total change in transport of 54.2 Sv at the Brisbane (27.5°S) section over the entire 2-year period. This *summed impact* provides an indication of the relative impact of observations in each latitudinal bin. The number of observations in each bin is shown in Figure 6e. In Figures 6f to 6i the summed values for each platform in each bin are normalized by the number of observations from that observation platform in the specific latitude bin. Note that this analysis includes all assimilation windows over the 2-year period to investigate the average latitudinal dependence; however, as seen in Figures 2c, 3b, 4b, and 5b, the impacts are time variable.

From the summed impacts (Figures 6a to 6d), the dominance of SST, observations from the EAC transport array, HF radar, and glider data is evident. The impact of the SEQ moorings at 27.5°S and the NSW moorings

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

at 30°S and 34°S can also be seen in these plots. South of about 32°S, where the EAC has typically separated from the coast, SST and SSH have an increased impact with respect to the number of observations per bin, which remains relatively constant (Figure 6e). Glider observations are taken along the coast from 29°S to 36°S (Figure 6e), but the glider data between 34° and 36°S have a disproportionally large impact on transport off Sydney (Figure 6c) because these particular observations correspond to offshore missions that specifically targeted mesoscale eddies in the Tasman Sea.

The observation impact relative to the number of observations is more easily interpreted from Figures 6f to 6i. The impact of SSH and SST observations normalized by the number of observations is greatest for observations in the region of elevated eddy activity from about 32–37°S (Figure 1a), compared to observations in the other, less eddy-dominated regions to the north and south. Measurements from Argo floats in this latitudinal range also have a more considerable impact per observation; although the summed impact is small due to the relatively small number of observations. Likewise, the impact per observation for measurements from the Sydney-Wellington XBT line (34–36°S) is large.

This latitudinal analysis shows that observations over the eddy field make the greatest contribution to changes in the model estimates of volume transport along the coast. SSH, SST, and Argo observations are evenly distributed across all latitudes, and observations taken in the region of high eddy variability (32–37°S) have more impact than the same observations taken elsewhere as they provide information to constrain the highly variable region. Even where the jet is mostly coherent at 27.5° S (Sloyan et al., 2016) and 30.3° S (Archer et al., 2017), satellite and Argo observations of the eddy field from 32–37° S have greater impact on transport than the same observation types upstream of the separation point (Figures 6f to 6g).

3.1.3. Depth Dependence

In a similar manner to the above analysis in latitude bins, we can examine the observation impacts in depth bins for the observation platforms providing subsurface data. In Figure 7 we present the observation impacts with depth for three chosen circulation metrics; transport through 27.5°S (Brisbane), transport through 33.9°S (Sydney), and spatially averaged EKE in the Tasman Sea (presented in the next section). Grouping the observations into 50-m bins, the absolute values of the impacts of all observations that fall within each bin over all assimilation windows are grouped by observation platform and summed (Figures 7a to 7c). The number of observations in each bin are shown in Figure 7d, and Figures 7e to 7g show the summed values normalized by the number of observations of each type in each bin.

Of the subsurface observations, the EAC array dominates the impacts for transport off Brisbane (Figure 7a), while for transport off Sydney both the gliders and the EAC array make contributions of similar magnitude (Figure 7b). The shelf moorings off Sydney, Coffs Harbour, and Brisbane (SEQ) also have considerable summed impacts on transport through both sections. This analysis reveals that the greatest summed impacts occur in the upper 250 m of the water column where data density is much higher.

Accounting for data density, the impacts normalized by number of observations are largest for measurements in the upper 400 m. Below 400 m the impacts taper off for all subsurface platforms except the gliders, which continue to have a considerable impact for measurements across all depths (the gliders sample down to 1000 m). The upper 400 m represents the depth region of greatest variability and uncertainty as it typically contains the mixed layer and pycnocline—an undersampled region with a key influence on current transport and eddy structure, so observations that help reduce this uncertainty make an important contribution to constraining these circulation measures. Presenting the observation impact relative to the number of observations is useful as it reveals the value of observing the upper 400 m of the water column.

Note that the XBT data have not been included in these figures as the summed impacts are so small relative to the other platforms but their impact per number of observations are large (of similar magnitude to all other platforms added) making Figures 7e to 7g difficult to interpret. The large impact per observation for the XBT observations indicates that these data are providing important new information about the subsurface ocean, thus contributing to large changes in the circulation measures relative to the number of observations made.

3.2. Eddy Kinetic Energy

EKE describes the energy associated with the velocity perturbations about the mean flow (equation (7)). It provides a measure of the circulation variability about the mean which, in a mesoscale setting, is typically associated with the eddies. While the free-running ocean model is configured to correctly represent the spatial distribution and intensity of the mesoscale variability on average, correct representation of the timing,

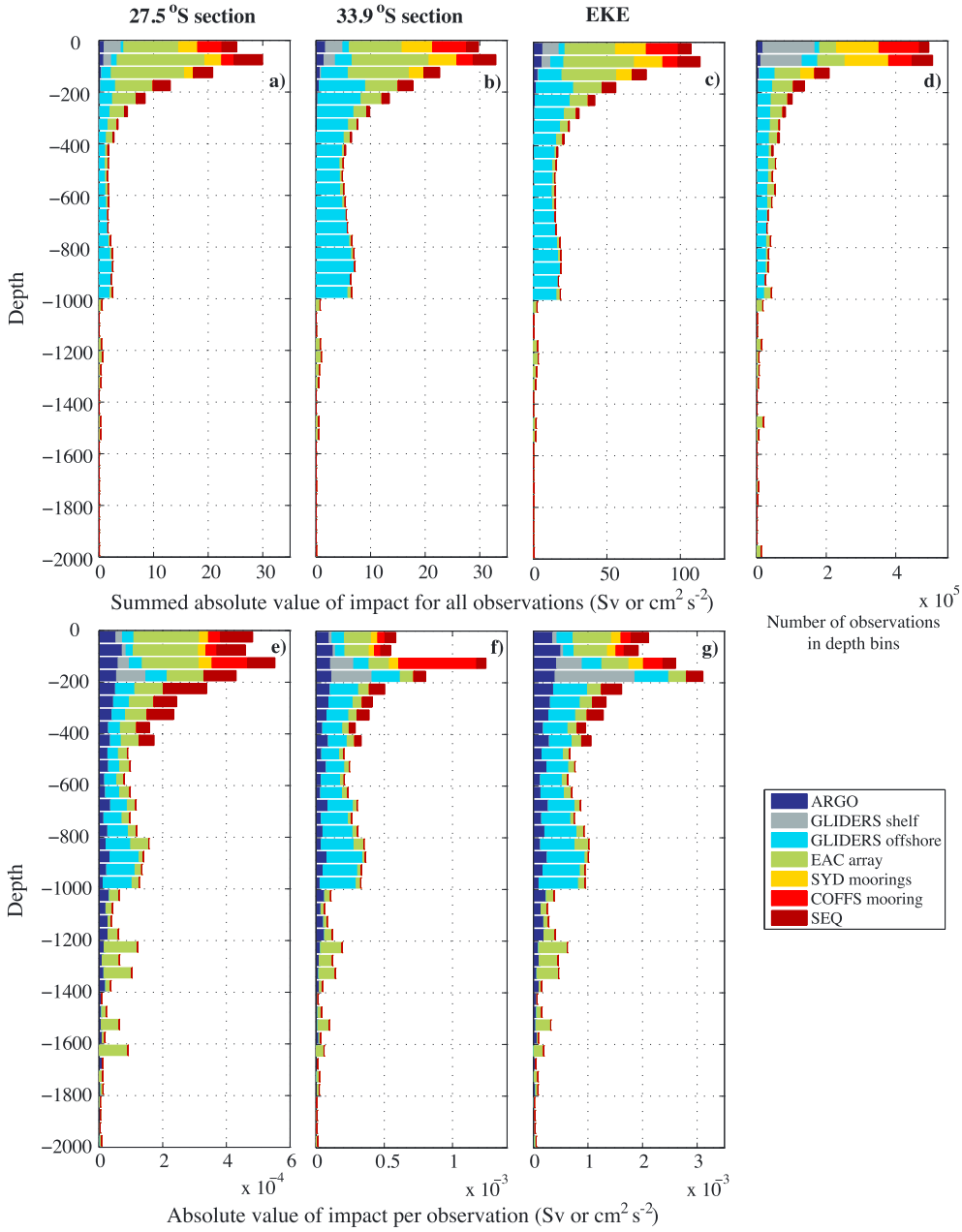

Figure 7. The observation impacts on transport through the 27.5° S and 33.9° S sections and spatially averaged eddy kinetic energy over the Tasman Sea, grouped into depth bins of 50 m. Summed absolute value of impacts for observations from each observation platform in each bin over the 2 years (a–c), the number of observations in each bin (d), and as in (a)–(c) but normalized by the number of observations (e–g). EAC = East Australia Current; SEQ = South East Queensland.

locations, and strength of individual eddies requires data assimilation. Over the 5-day forecast, the eddies may evolve differently and the purpose of the analysis is to adjust the model such that the eddies' evolution better represents the observations. The observation impacts on spatially averaged EKE describe how specific observations contribute to the adjustments affecting eddy intensity in the Tasman Sea region.

Time series of the averaged EKE computed from the forecast and the analysis are shown in Figure 8a. For the impact calculations we use the second-order formulation (equation (5)) and the increment ΔJ, representing the transport difference between the two nonlinear models, matches closely the sum of the observation impacts for all observations (black and green lines in Figure 8a). The EKE increments range from −17 cm^2/s^2

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

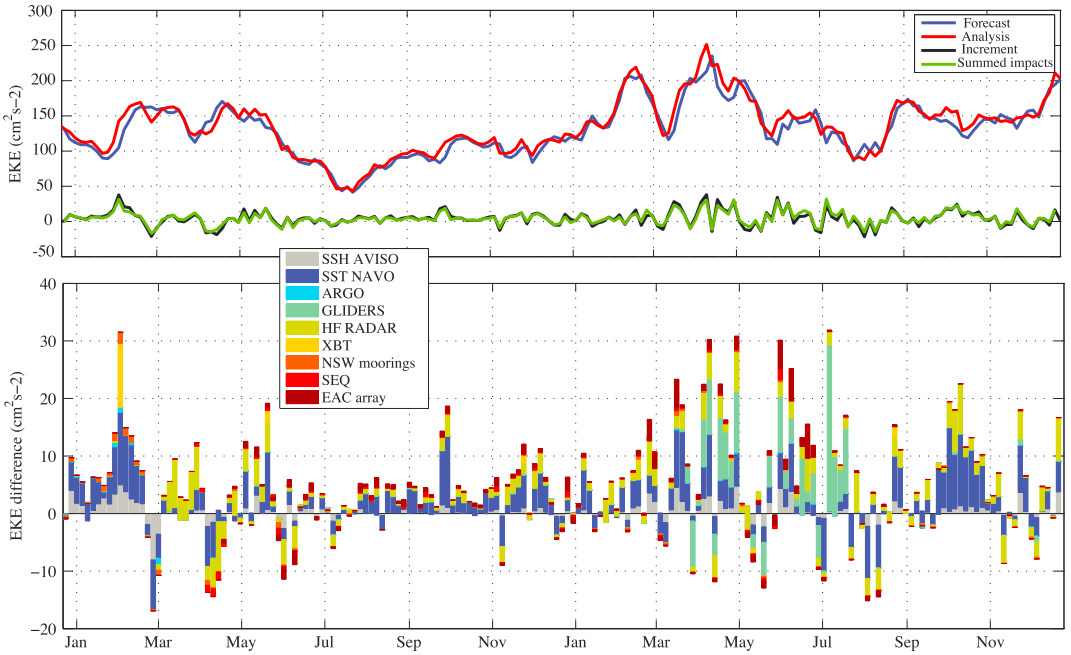

Figure 8. (a) Eddy kinetic energy averaged over the area shown in Figure 1 and down to 450m, for the forecast and the analysis, averaged over each 5-day window, the difference in eddy kinetic energy between the forecast and the analysis (the increment) and the summed observation impacts. Note the increment and summed impacts should match closely if the linear assumption is valid. (b) The impact of each observation platform on the EKE difference for each 5-day window. AVISO = Archiving, Validation and Interpretation of Satellite Oceanographic data; EAC = East Australia Current; HF = high frequency; SEQ = South East Queensland; SSH = sea surface height; SST = sea surface temperature; NAVO = Naval Oceanographic Office; NSW = New South Wales; XBT = expendable bathythermograph.

to +32 cm^2/s^2, and the mean of the absolute value of the increments is 6.5% of the mean EKE value over the 2 years (135 cm^2/s^2).

The contribution to the EKE increments for each observation platform are shown in Figure 8b, and the average percentage impacts over the 2-year period are presented in Table 1. As for volume transport, the gliders also have large impacts on the EKE over the time period when they sampled eddies offshore of Sydney (20 March 2013 and 22 July 2013). Computing the average percentage impacts as in Table 1 but only over that time period, the percentage impacts are 8.9%, 24.93%, 38.34%, 14.36%, and 9.98% for SSH, SST, gliders, HF radar, and the EAC transport array, respectively. The percentage impacts are calculated in the same way as those for alongshore volume transport, described in section 3.1.1 and by equation (8).

Grouping the observation impacts on EKE into depth bins (Figures 7c and 7g) reveals that the gliders and the EAC array are the most valuable subsurface observations in constraining the EKE. Similar to the volume transport metrics, the greatest summed impacts occur in the upper 250 m of the water column while the impacts normalized by number of observations are largest for measurements in the upper 400 m.

4. Discussion

4.1. Interpreting Observation Impact

To understand how we might improve sampling strategies, it is useful to understand why certain observations might have greater impacts on the circulation estimates than others and how the observations adjust the model. There are several advantages of the methodology used in this study to compute observation impacts, which was also used by Moore, Arango, Broquet, Powell, et al. (2011) and Powell (2017). This method considers the complete assimilation system with the full suite of assimilated observations and, because we compute the impact of each specific observation, we can quantify the relative impact of different data streams, observed variable types, and observation locations. For example, in this study we quantified the impact of a number of data streams without the need to run several simulation experiments, we studied the time variable impacts, and we quantified how the observation impacts vary in space (both latitudinally and with depth). However, several considerations must be made when interpreting the observation impact results.

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

First, as shown in equation (4), the impact is scaled by the forecast-data difference (the innovation). Observation impact quantifies the contribution of specific observations to changes in a specific circulation estimate, and larger forecast-data differences will typically result in larger changes to the model state (with the associated model background and observation errors accounted for) resulting in larger changes to the chosen circulation metric. Higher innovations are likely to occur in regions where the forecast skill degrades more quickly, resulting in high impacts for observations taken in these regions. High innovations may also result from bad data, but careful preprocessing of the observations in this study minimizes the risk of assimilating bad data. As such, observations taken where the model forecast innovations are small will have low impact, yet these observations may still be important to the observing system.

Observation sensitivity (see Moore, Arango, Broquet, Powell, et al., 2011) is independent of the actual values of the observations and is more suited to asses the value of a particular type of observation independent of the forecast errors. However, for a specific region and assimilation system, observation impacts are useful as they show the value of observing regions where forecast errors are likely to grow more quickly. For example, this study highlights the importance of observing regions of high variability, where the model may deviate the most over the forecast period; however, clearly observations of the less variable, slowly evolving regions are still important. Understanding the optimal spatial coverage and frequency of observations in the more variable compared to less variable regions requires more study.

Second, observations that are not independent of each other will have less impact per observation. For SST, the decorrelation length scales are typically large, so we are assimilating a large number of observations that are not linearly independent; that is, additional observations may provide very little additional information about the ocean state. However, because there are so many SST observations they can have a large total impact. On the other hand, the XBT observations taken between 34° and 36°S have a large impact per observation as the observations provide new information about the subsurface structure of the ocean; yet the summed impact of these observations is very small. As seen in Figures 6a to 6d, SST dominates the summed impacts at most latitudes, but its impact per observation (Figures 6f to 6i) is less than that of the SSH observations and of similar magnitude to the Argo floats. The observation impact normalized by the number of observations is dependent on the number of observations of a certain type that we choose to assimilate. For example, we could subsample the SST observations to include less linearly dependent observations and get a greater impact per observation. This is why we present (in Figures 6 and 7) both the integrated impacts and the impacts per observation. The impacts normalized by the number of observations are useful to show the value of low density observations such as those from Argo and XBTs and to reveal the relative impacts of observations in the latitudinal range of greatest eddy variability and in the upper 400 m of the water column.

Finally, the impacts of observations in our specific assimilation system are dependent on our prior choices of observation and model background error covariances, which are encompassed in the Kalman gain matrix, **K** (equations (4) and (5)). For example, observations with high errors are penalized in the assimilation and so will have low impact. While we have put considerable effort into making reasonable estimates of these error covariances and provided checks of their consistency (refer to section 4.1 of Kerry et al., 2016) they are not the true error covariances. Indeed, estimating the error covariances is one of the largest areas of continued research in data assimilation.

4.2. Constraining the Transport and EKE

The impact of each specific observation on our transport and EKE estimates depends on how the information carried by the observation is projected into model space through the data assimilation technique to alter the model state estimate. As the EAC is a geostrophic flow, there is a corresponding surface displacement and isopycnal tilt associated with the flow. Observations that provide information about the surface displacement and the isopycnal tilt across the current will constrain the transport.

The geostrophic balance gives rise to the thermal wind equations that describe the proportionality between the vertical derivative of velocity and the horizontal components of the density gradient (refer to Gill, 1982, equations 7.7.9). Integrating twice from the surface to the depth of no motion, assuming a constant horizontal density gradient, gives the alongshore transport,

$$Q = -\frac{g}{2f\rho_o} z^{*2} \frac{\partial \rho}{\partial x}, \tag{9}$$

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

where z^* is the depth of no motion and x represents the cross distance. So the transport increment between the forecast and the analysis is proportional to the difference in the cross-shore density gradient.

Both barotropic and baroclinic instabilities are thought to drive eddy formation and evolution in the EAC (Mata et al., 2006). Studies of the EAC suggest that north of and at the separation zone, barotopic instabilities allow perturbations to grow leading up to an eddy shedding event, and once the eddies shed their evolution is likely driven by baroclinic instability (Mata et al., 2006; Macdonald et al., 2016). Increments in the EKE in the Tasman Sea region may result from adjustments to the current flow north of its separation, driving changes to the eddy shedding dynamics (typically driven by barotropic instabilities), or changes to the eddies' evolution (due to baroclinic instabilities). Lindzen and Farrell (1980) show that the growth rate of the fastest growing baroclinic instability is strongly dependent on the horizontal density gradient. As such, EKE also relies strongly on subsurface structure.

The subsurface ocean is typically undersampled. Although we make use of several subsurface observation platforms (gliders, moorings, Argo floats, and XBTs), the observations remain temporally and spatially sparse, or highly localized. By quantifying the impacts we have shown the relative influence of various observation platforms, with the dominant platforms being the SST, the HF radar array, the SSH, the EAC mooring array, and the ocean gliders. Note that the three most dominant platforms sample only the sea surface. These relative impacts represent an average over the 2-year period (the impacts for each 5-day assimilation window are flow dependent and vary considerably from one window to the next) and are summed across all observations in each platform so are influenced by data volume. Satellite-derived surface observations (SST and SSH) are widespread across the model domain and regularly available. The SST data in particular have a high spatial resolution of 4 km and is available two to three times a day in the absence of clouds. These observations are important in defining the boundaries of the mesoscale eddies and constraining their surface expression, key to representing their temporal and spatial evolution. The surface tilt associated with mesoscale eddies is projected into subsurface, to alter the isopycnal tilt. Surface current observations from the HF radar prove particularly useful in constraining the EAC flow where it is mostly coherent (Brisbane and Coffs Harbour), while satellite observations are more useful, on average, where volume transport is eddy driven (Sydney and Narooma, Table 1).

In the assimilation we adjust the initial and boundary conditions and the surface forcing, and adjustments to initial conditions are almost entirely responsible for the volume transport and EKE increments. The greatest adjustments in the initial conditions are made to SSH and temperature. The spatially averaged normalized increments (increment magnitudes) for initial conditions, averaged over the 2 years, are 89% (2.6 cm), 77% (0.18 ° C), 51% (0.038 m/s), 44% (0.0087), and 38% (0.027 m/s) for SSH, temperature, alongshore velocity, salinity, and cross-shore velocity, respectively. The normalized increments are represented as percentages that describe the mean increment adjustment relative to the typical variability of the variable over a 5-day period and allow comparison of the relative adjustments of different variables. For example, for SSH the spatially averaged mean increment adjustment to the initial conditions is 2.61 cm, and the square root of the spatially averaged mean 5-day variances of SSH for a long (10-year) model simulation is 2.93 cm, giving a normalized increment of 89%.

It is notable that the SST observations impact the alongshore transport much more than SSH (Table 1), despite the direct dynamical link between SSH slope and the surface geostrophic current. In the assimilation, we exclude SSH observations over water depths less than 1,000 m, because the observations are noisy on the continental shelf and the AVISO-gridded product is not able to resolve the processes that occur here. This equates to a distance of between 35 and 100 km offshore (given the model bathymetry). The four sections through which the transport is computed extend 156–200 km offshore and, given the relatively low resolution (1/4° $\times$ 1/4°) of the mapped SSH product and the exclusion of SSH data along the continental shelf, the cross-shore slope along each section is described by a maximum of four to six data points. Therefore, the SSH observations may not provide a faithful estimate of the slope across the transport sections. The SST observations dominate the total number of observations; we assimilate a total of 3.72 million SST observations over the 2 years, compared to 1.03 million subsurface temperature observations, 0.84 million SSH observations, and 0.82 million subsurface velocity observations. Over all assimilation windows, the mean percentage number of SST observations relative to the total number of observations is 43.6%, compared to 11.6% for subsurface temperature observations, 12.1% for SSH, and 10.5% for subsurface velocity. As such, the SST observations make the great-

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

est contribution to constraining temperature, and thus density, in the initial conditions, thus impacting the EAC transport estimates.

The EAC array and the gliders are the most dominant subsurface observations (Figure 7). The EAC array traverses the EAC off Brisbane, providing important information to constrain the flow where it tends to be most coherent. When the glider observations become available they have high temporal and spatial density therefore providing detailed information on the depth and structure of the mixed layer and pycnocline. We found that the glider measurements had large impacts on EAC transport and EKE for the time periods that they sampled eddies offshore of the continental shelf off Sydney. Averaged over the assimilation periods for which they were sampling, the offshore gliders had percentage impacts of 28–36% on transport estimates through the four sections, and 38% for EKE (computed as in equation (8)). This is a similar finding to that of Powell (2017) who shows that, in the Hawaiian region, the gliders have a large impact on accurate representation of Hawaiian Lee Counter Current transport. The profiling subsurface observations from Argo floats and XBTs have high observation impacts relative to the number of observations (as seen in the latitudinal analysis, Figures 6f to 6i). These observations have small total impacts as they are temporally and spatially sparse, but their impact relative to the number of observations is high as they provide new information about the subsurface ocean.

Because, in 4D-Var, the model error covariances are flow-dependent, time-varying fields, the impacts of observations taken at a specific location can be far reaching as information captured by the observation propagates in the ocean. Information from an assimilated observation can propagate by various oceanic processes resolved by the model including advection, barotropic waves, baroclinic waves, and mixed barotropic-barolinic waves (such as coastally trapped waves). For example, Kurapov et al. (2011) showed that assimilated information was propagated by coastally trapped waves in a variational assimilation study of the Oregon coast. Zhang et al. (2010) and Powell (2017) use a representer-based approach to reveal how information is propagated from a specific observation to other locations in the ocean. The observation information propagates both forward and backward in time because both past and future times must be consistent with the observation occurrence at its particular time and place.

In this study, the HF radar and the EAC array observations have significant impacts on transport at all four cross sections. It would take 23 days for a current of 0.5 m/s to travel the 1,000 km from Brisbane (the EAC array location) to Narooma, so over the 5-day assimilation windows this information cannot be carried by advection alone. Coastally trapped waves propagate with the coast to the left in the Southern Hemisphere and, in the southeast Australian region, typically have phase speeds of 2.1–4.0 m/s (Maiwa et al., 2010; Woodham et al., 2013); meaning, they could travel 1,000 km in 2.9–5.5 days. Internal waves could also potentially propagate information in this region at speeds fast enough to traverse long distances in the 5-day windows. Thus, it is possible that information is translated by coastally trapped waves and internal waves; however, the propagation patterns of perturbations have not been presented in this work.

5. Conclusion

Mesoscale eddies dominate the circulation in most WBCs, and these highly dynamic regions are inherently difficult to model and observe. This study combines observations with a numerical model to predict the EAC, the WBC of the South Pacific subtropical gyre, and quantifies the value of various data streams on circulation estimates of volume transport and EKE. Averaged over the 2-year reanalysis period we find that the most influential observations are the satellite derived SST, the radial velocities from the HF radar off Coffs Harbour, the satellite derived SSH, the data from the full depth EAC mooring array, and the ocean glider observations (in that order). These relative impacts are given the data volumes that were assimilated in this system; for example, the dominance of SST is due to the large number of SST observations assimilated.

Overall, we find that observations taken in regions with greater natural variability cause the greatest change in our circulation estimates. SSH and SST observations of the region of elevated eddy energy between 32° and 37° S have more impact per observation on transport estimates along the coast than the same observations taken elsewhere. Observations taken in the upper 400 m of the water column contribute more to changes in the circulation estimates than deeper observations, as they sample the depth region of greatest uncertainty and reveal information about the structure of the mixed layer and pycnocline. The observation impact of gliders deployed into EAC eddies off Sydney is particularly high.

We show the far-reaching effect of some localized observation platforms as, with 4D-Var, information captured by the observations propagates in the ocean in space and time. The HF radar observations have high impact on transport estimates at the instrument location, but they also have significant impact hundreds of kilometers both upstream and downstream. Likewise, the impact of the EAC transport array is far reaching, contributing to transport estimates up to 1,000 km downstream of its location.

This study reveals the relative impacts of the different observation platforms that inform estimates of our chosen circulation measures, given our rigorously tuned data assimilative model of the southeast Australian region. These circulation measures were chosen because hindcasting and forecasting EAC transport and eddy intensity is of fundamental importance. Different circulation measures may be impacted differently. Understanding which observations are useful in informing model estimates of the EAC is a key step toward providing improved state estimates and predictions.

While this study is specific to the EAC region and the specific observing system, its findings are widely relevant to data assimilation systems of other regions. The results are particularly relevant to other WBCs in which the flow typically evolves from a more coherent jet adjacent to the coast to an eddy field. More work is required to understand the relative importance of observing the variable versus less variable regions of the ocean to achieve optimum state estimates and predictions.

Acknowledgments

This research and C. Kerry were supported by an Australian Research Council Discovery Project DP1401. Model initial conditions and boundary forcing comes from the Bluelink ReANalysis version 3p5 (BRAN3, Oke et al., 2013). Surface forcing is provided by the Australia Bureau of Meteorology's (BOM) Australian Community Climate and Earth-System Simulation (ACCESS) 12-km product (Puri et al., 2013). We thank Holly Simms from BOM for making it available to us. The observations used for assimilation were primarily sourced from the the Australian Integrated Marine Observing System's (IMOS) data portal (https://portal.aodn.org.au)—IMOS is a national collaborative research infrastructure, supported by the Australian Government. The reanalysis output over the 2-year period (2012–2013) is archived at UNSW Australia and can be made available for research purposes (contact the corresponding author of this paper). We sincerely thank Peter Oke, CSIRO Hobart, for providing data in the preliminary phases of this work and for his invaluable discussions.

References

Archer, M., Roughan, M., Keating, S., & Schaeffer, A. (2017). On the variability of the East Australian Current: Jet structure, meandering, and influence on shelf circulation. *Journal of Geophysical Research: Oceans*, *122*, 1–18. https://doi.org/10.1002/2017JC013097

Beal, L. M., & Elipot, S. (2016). Broadening not strengthening of the Agulhas Current since the early 1990s. *Nature*, *540*, 570–573.

Bennett, A. F (2002). *Inverse modeling of the ocean and atmosphere*. Cambridge: Cambridge University Press.

Centre National d'Etudes Spatiales (2015). SSALTO/DUACS user handbook: (M)SLA and (M)ADT near-real time and delayed time products AVISO sateliitle altimetry data.

Courtier, P., Andersson, E., Heckley, W. A., Kelly, G., Pailleux, J., Rabier, F., et al. (1993). Variational assimilation at ECMWF (Tech. Rep. 194). Reading: European Centre for Medium-Range Weather Forecasts.

Courtier, P., Thépaut, J. N., & Hollingsworth, A. (1994). A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, *120*, 1367–1387.

De Souza, J. M. A. C., Powell, B., Castillo-Trujillo, A. C., & Flament, P. (2015). The vorticity balance of the ocean surface in Hawaii from a regional reanalysis. *Journal of Physical Oceanography*, *45*, 424–440.

Gelaro, R., Zhu, Y., & Errico, R. M. (2007). Examinations of various-order adjoint-based approximations of observation impact. *Meteorologische Zeitschrift*, *16*(6), 685–692.

Gill, A. E (1982). *Atmosphere-ocean dynamics*. San Diego: Academic Press.

Godfrey, J. S., Cresswell, G. R., Golding, T. J., Pearce, A. F., & Boyd, R. (1980). The separation of the East Australian Current. *Journal of Physical Oceanography*, *10*, 430–440.

Jia, F., Wu, L., & Qiu, B. (2011). Seasonal modulation of eddy kinetic energy and its formation mechanism in the southeast Indian Ocean. *Journal of Physical Oceanography*, *41*, 657–665.

Kerry, C. G., Powell, B. S., Roughan, M., & Oke, P. R. (2016). Development and evaluation of a high-resolution reanalysis of the East Australian Current region using the Regional Ocean Modelling System (ROMS 3.4) and Incremental Strong-Constraint 4-Dimensional Variational (IS4D-Var) data assimilation. *Geoscientific Model Development*, *9*, 3779–3801. https://doi.org/10.5194/gmd-9-3779-2016

Kurapov, A. L., Foley, D., Strub, P. T., Egbert, G. D., & Allen, J. S. (2011). Variational assimilation of satellite observations in a coastal ocean model off Oregon. *Journal of Geophysical Research*, *116*, C05006. https://doi.org/10.1029/2010JC006909

Le Dimet, F., & Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, *38A*, 97–110.

Lindzen, R. S., & Farrell, B. (1980). A simple approximate result for the maximum growth rate of baroclinic instabilities. *Journal of the Atmospheric Sciences*, *37*, 1648–1654.

Macdonald, H. S., Roughan, M., Baird, M. E., & Wilkin, J. (2016). The formation of a cold-core eddy in the East Australian Current. *Continental Shelf Research*, *114*, 72–84. https://doi.org/10.1016/j.csr.2016.01.002

Maiwa, K., Masumoto, Y., & Yamagata, T. (2010). Characteristics of coastal trapped waves along the southern and eastern coasts of Australia. *Journal of Oceanography*, *66*, 243–258.

Mata, M. M., Wijffels, S. E., Church, J. A., & Tomczak, M. (2006). Eddy shedding and energy conversions in the East Australian Current. *Journal of Geophysical Research*, *111*, C09034. https://doi.org/10.1029/2006JC003592

Metzger, E., Smedstad, O., Thoppil, P., Hurlburt, H., Cummings, J., Wallcraft, A., et al. (2014). US Navy operational global ocean and Arctic ice prediction systems. *Oceanography*, *27*(3), 32–43.

Moore, A. M., Arango, H. G., Broquet, G., Edwards, C., Veneziani, M., Powell, B. S., et al. (2011). The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part III—Observation impact and observation sensitivity in the California Current System. *Progress in Oceanography*, *91*, 74–94. https://doi.org/10.1016/j.pocean.2011.05.005

Moore, A. M., Arango, H. G., Broquet, G., Powell, B. S., Zavala-Garay, J., & Weaver, A. T. (2011). The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part I—System overview and formulation. *Progress in Oceanography*, *91*, 34–49. https://doi.org/10.1016/j.pocean.2011.05.004

Moore, A. M., Arango, H. G., Di Lorenzo, E., Cornuelle, B. D., Miller, A. J., & Neilson, D. J. (2004). A comprehensive ocean prediction and analysis system based on the tangent linear and adjoint of a regional ocean model. *Ocean Modelling*, *7*, 227–258.

Oke, P. R., & Middleton, J. H. (2000). Topographically induced upwelling off eastern Australia. *Journal of Physical Oceanography*, *30*, 512–530.

Oke, P., Sakov, P., Cahill, M. L., Dunn, J. R., Fiedler, R., Griffin, D. A., et al. (2013). Towards a dynamically balanced eddy-resolving ocean reanalysis: BRAN3. *Ocean Modelling*, *67*, 52–70.

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Powell, B. S. (2017). Quantifying how observations inform a numerical reanalysis of Hawaii. *Journal of Geophysical Research: Oceans*, *122*, 1–18. https://doi.org/10.1002/2017JC012854

Puri, K., Dietachmayer, G., Steinle, P., Dix, M., Rikus, L., Logan, L., et al. (2013). Operational implementation of the ACCESS Numerical Weather Prediction system. *Australian Meteorological and Oceanographic Journal*, *63*, 265–284.

Roughan, M., Keating, S., Schaeffer, A., Heredia, P. C., Rocha, C., Griffin, D., et al. (2017). A tale of two eddies: The bio-physical characteristics of two contrasting cyclonic eddies in the East Australian Current. *Journal of Geophysical Research: Oceans*, *122*, 2494–2518. https://doi.org/10.1002/2016JC012241

Roughan, M., Schaeffer, A., & Suthers, I. (2014). Sustained ocean observing along the coast of southeastern Australia: NSW-IMOS 2007–2014. In *Coastal Ocean Observing Systems*. Cambridge: Academic Press.

Sloyan, B. M., Ridgway, K. R., & Cowley, R. (2016). The East Australian Current and property transport at 27S from 2012–2013. *Journal of Physical Oceanography*, *46*(3), 993–1008.

Stammer, D., Wunsch, C., Giering, R., Eckert, C., Heimbach, P., Marotzke, J., et al. (2002). The global ocean circulation during 1992–1997, estimated from ocean observations and a general circulation model. *Journal of Geophysical Research*, *107*, 3118.

Talagrand, O., & Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quarterly Journal of the Royal Meteorological Society*, *113*, 1311–1328.

Thépaut, J. N., & Courtier, P. (1991). Four dimensional variational data assimilation using the adjoint of a multilevel primitive equation model. *Quarterly Journal of the Royal Meteorological Society*, *117*(502), 1225–1254. https://doi.org/10.1002/qj.49711750206

Weaver, A., & Courtier, P. (2001). Correlation modelling on the sphere using generalized diffusion equation. *Quarterly Journal of the Royal Meteorological Society*, *127*, 1815–1846.

Weaver, A. T., Vialard, J., & Anderson, D. L. T. (2003). Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part I: Formulation, internal diagnostics, and consistency checks. *Monthly Weather Review*, *131*, 1360–1378.

Wilkin, J. L., & Zhang, W. G. (2007). Modes of mesoscale sea surface height and temperature variability in the East Australian Current. *Journal of Geophysical Research*, *112*, C01013.

Woodham, R., Brassington, G. B., Robertson, R., & Alves, O. (2013). Propagation characteristics of coastally trapped waves on the Australian Continental Shelf. *Journal of Geophysical Research: Oceans*, *118*, 4461–4473. https://doi.org/10.1002/jgrc.20317

Zhai, X., Johnson, H. L., & Marshall, D. P. (2010). Significant sink of ocean-eddy energy near western boundaries. *Nature Geoscience*, *3*, 608–612.

Zhang, W. G., Wilkin, J. L., & Levin, J. C. (2010). Towards an integrated observation and modeling system in the New York Bight using variational methods. Part II: Representer-based observing strategy evaluation. *Ocean Modelling*, *35*, 134–145.

21699291, 2018, 10, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2017JC013685 by University of New South Wales, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License