**Peter R. Oke and Andreas Schiller**

CSIRO Marine and Atmospheric Research

# Observing system design using ensemble optimal interpolation:

## application to the tropical Indian Ocean mooring array

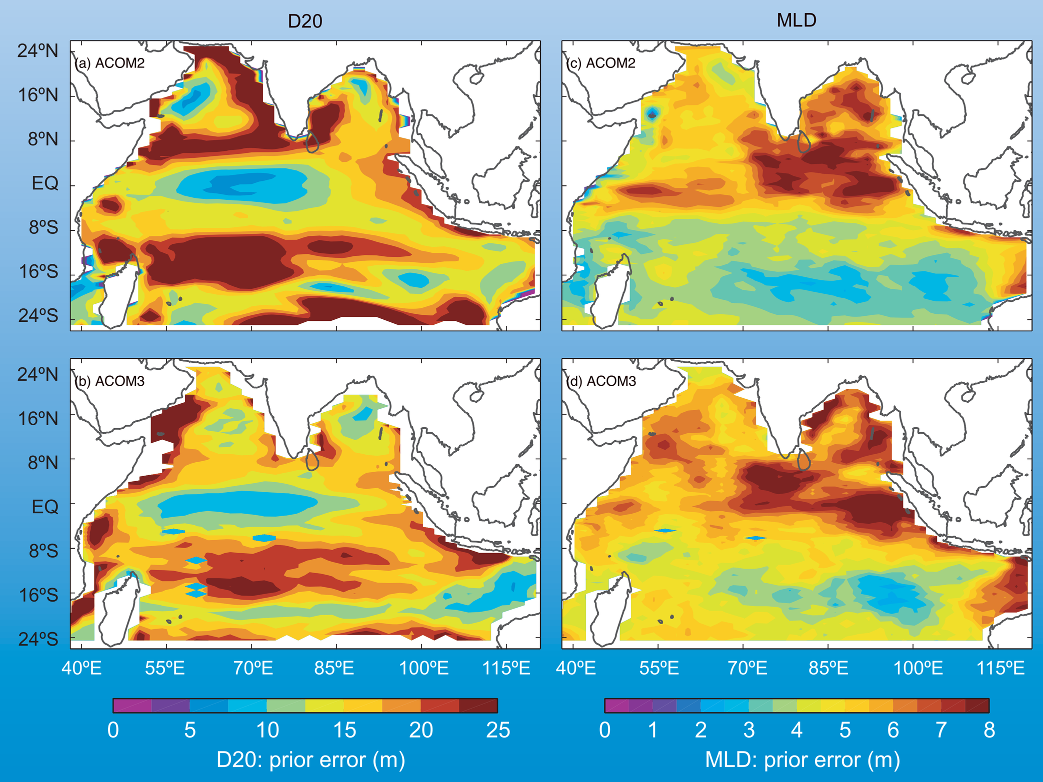


**Figure 1** Standard deviation of D20 (left) and MLD (right) from ACOM2 (top) and ACOM3 (bottom).

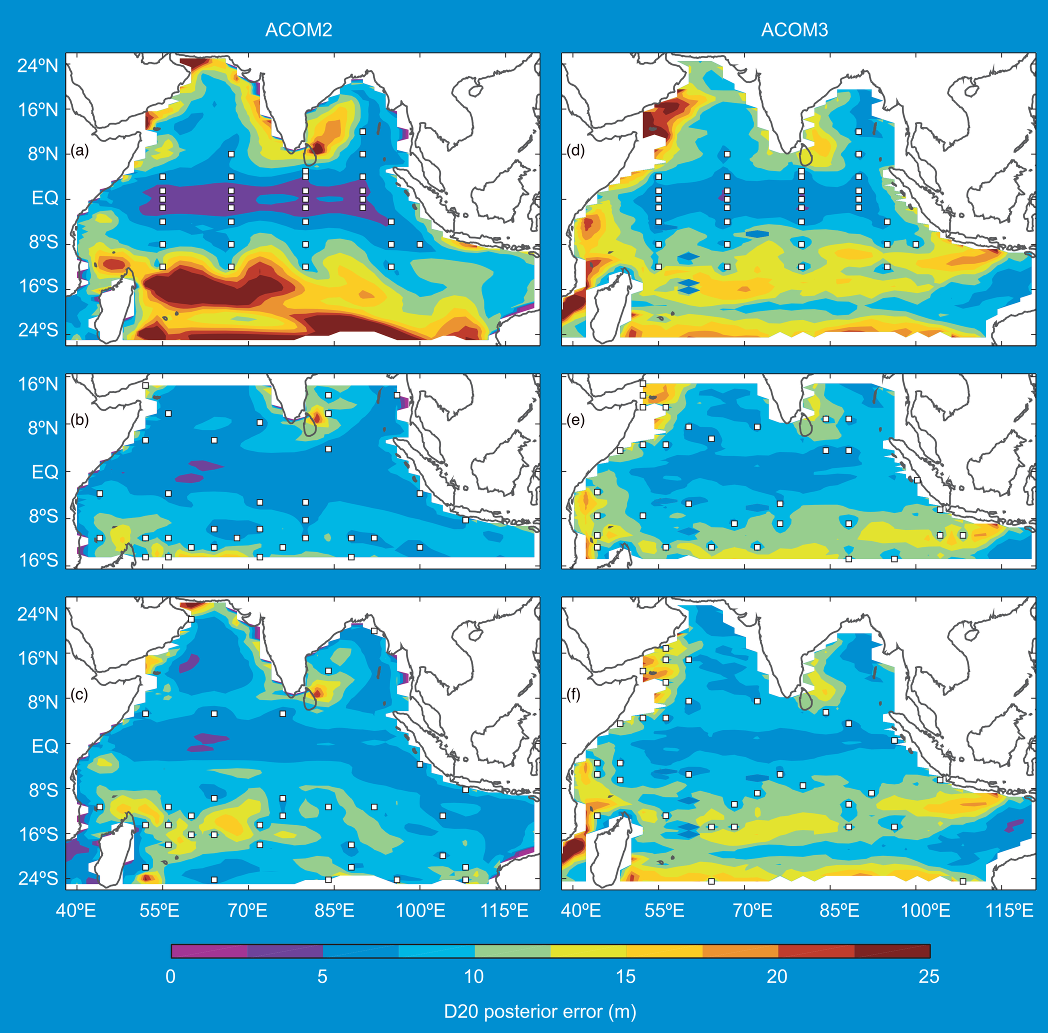


**Figure 3** Theoretical posterior error for D20 from ACOM2 (left) and ACOM3 (right) using the proposed array (top) and an optimal array on the ±15° (middle) and ±25° (bottom) domain.

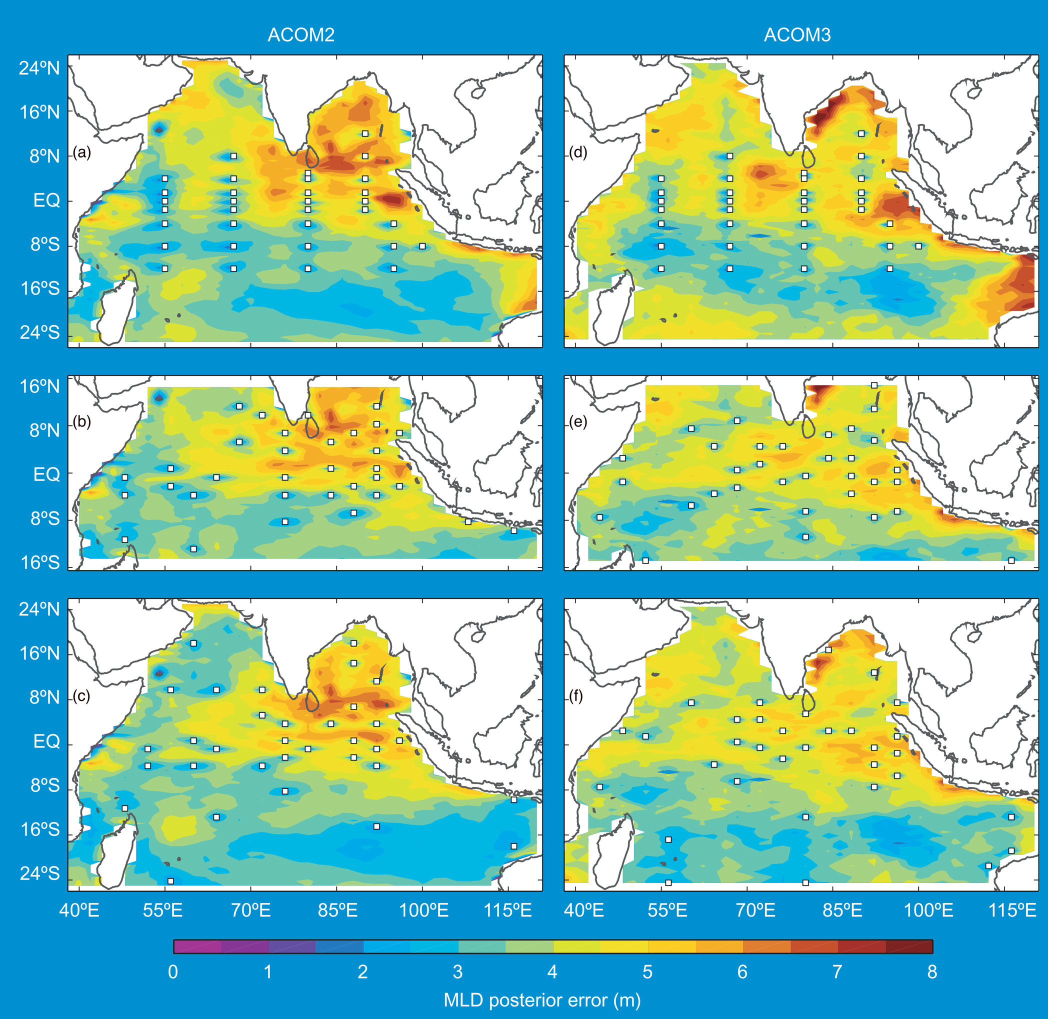


**Figure 4** As for Figure 3, except for MLD.

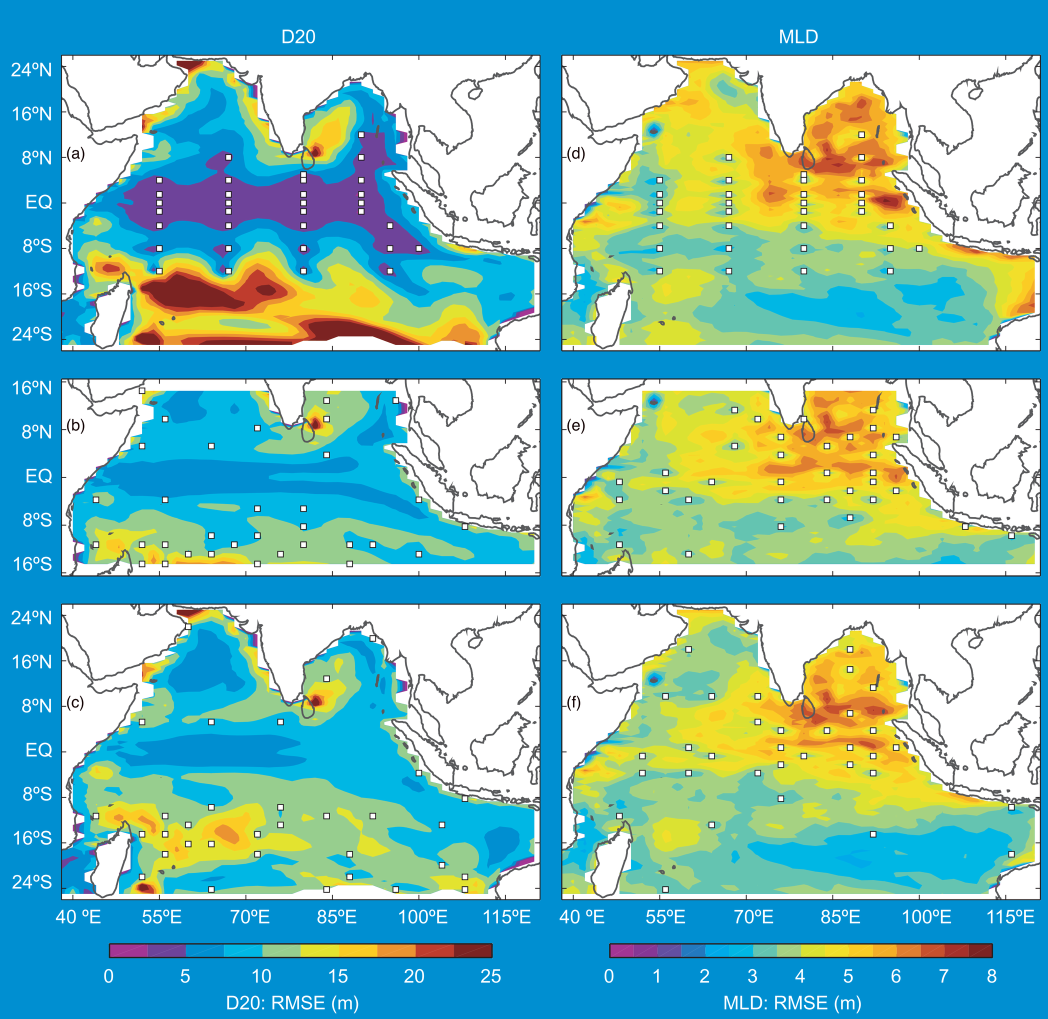


**Figure 5** RMSE for D20 (left) and MLD (right) for OSSEs using the proposed array (top) and an optimal array on the ±15° (middle) and ±25° (bottom) domain for ACOM2. These fields should be compared to the Figure 3 and 4.

## Objective

A series of Observing System Simulation Experiments (OSSEs) are performed using Ensemble Optimal Interpolation (EnOI) to design an improved mooring array for the tropical Indian Ocean. We apply a procedure to determine the optimal array of observations that minimises the analysis/posterior error variance. We apply the system to the depth of the 20° isotherm (D20), representing interannual variability, and high-pass filtered mixed layer depth (MLD), representing intraseasonal variability.

## Method

We consider fields from two different global ocean models with different resolution, surface fluxes and run for different periods. The details of these models are summarised in **Table 1**. EnOI produces gridded analyses $\mathbf{w}^a$, by solving

$$\mathbf{w}^a = \mathbf{w}^B + \mathbf{K}\,(\mathbf{d}-\mathbf{H}\mathbf{w}^B) \quad (1)$$

$$\mathbf{K} = \mathbf{P}^B\mathbf{H}^T\,(\mathbf{H}\mathbf{P}^B\mathbf{H}^T + \mathbf{R})^{-1}$$

where $\mathbf{w}^B$ is the gridded background field that is here defined as the temporal mean from a long model run; $\mathbf{K}$ is the Kalman gain; $\mathbf{d}$ is a vector of observations; $\mathbf{H}$ interpolates from grid- to observation-space; $\mathbf{R} = \varepsilon^2\mathbf{I}$ is the observation error covariance matrix, where $\varepsilon = 4$m is the assumed observation error;

$$\mathbf{P}^B = \mathbf{A}\mathbf{A}^T/(t-1) \quad (2)$$

quantifies the background error covariances, where $t$ is the ensemble size (here equivalent to the number of realizations from a long model run); $\mathbf{A} = [\mathbf{w}'_1\ \ \mathbf{w}'_2\ \ldots\ \mathbf{w}'_t]$ is a matrix of anomalies; and $\mathbf{w}'_i$ is the *ith* model anomaly from the background field/long-term mean. The theoretical analysis error covariance matrix $\mathbf{P}^a$ resulting from (1) is given by

$$\mathbf{P}^a = \mathbf{P}^B - \mathbf{K}\mathbf{H}\mathbf{P}^B \quad (3)$$

The prior and posterior error variances are given by the diagonals of $\mathbf{P}^B$ and $\mathbf{P}^a$, respectively; and the basin-averaged background and expected analysis errors (EAE) are given by $\sqrt{\mathrm{trace}(\mathbf{P}^B)/n}$ and $\sqrt{\mathrm{trace}(\mathbf{P}^a)/n}$, where $n$ is the number of grid points.

We seek to define $\mathbf{H}$, such that the EAE is minimised. In practice, we start with locations at every model grid point, we eliminate the location that, when withheld, gives the smallest EAE. We recursively repeat the procedure until the desired number of locations remain.

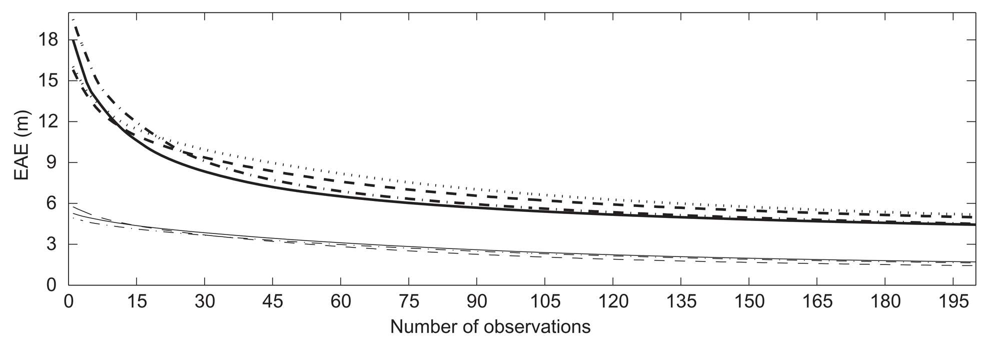


**Figure 6** Plots of the basin-averaged posterior error as a function of the number of observations for D20 (bold) and MLD (thin) for ACOM2, ACOM3 on the ±15° and ±25° domains.

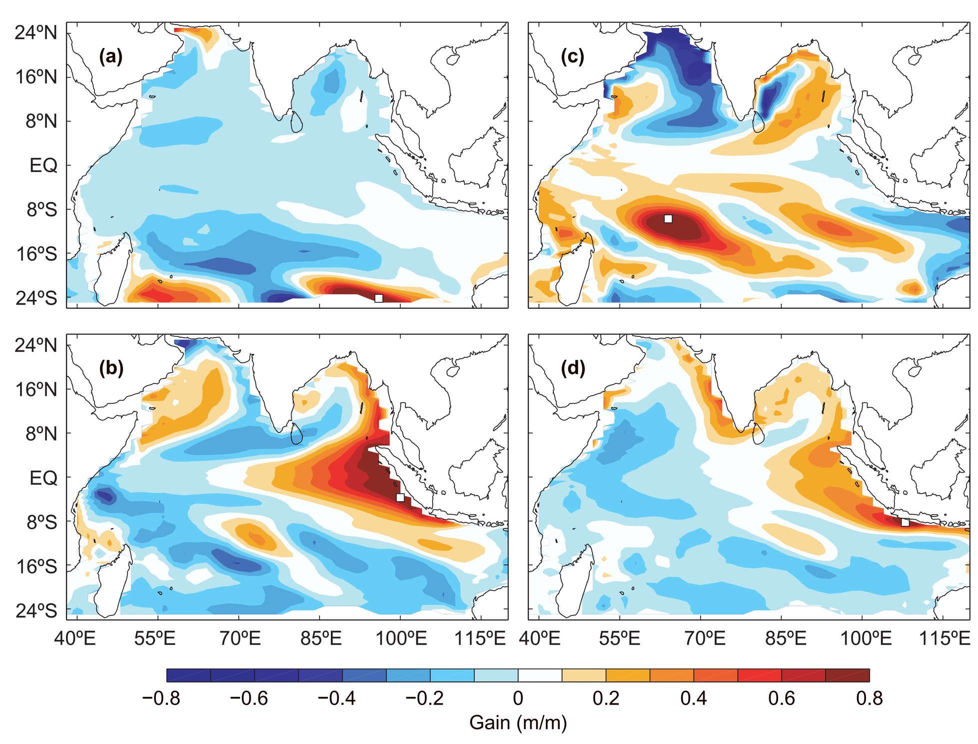


**Figure 7** Examples of the gain matrix for D20 using the ACOM2 ensemble, for the best 4 observation locations (denoted by the squares).

**Table 1** Details of model configurations

| | ACOM2 | ACOM3 |
|---|---|---|
| Model code | MOM2 | MOM3 |
| Resolution (x, y, z) | (2°, 0.5-1.5°, 25 levels) | (0.5°, 0.33°, 33 levels) |
| Wind forcing | NCEP-NCAR + FSU | ERS1/2 |
| Non-shortwave heat | ABL + flux correction | ABL + flux correction |
| Shortwave heat flux | as above | OLR + NCEP |
| Freshwater flux | as above | Monthly analyses |
| Simulated period | 1982-1994 | 1992-2000 |

## Results

The standard deviations of D20 and MLD are shown in **Figure 1**. Under the assumptions of (1-2), these fields can be regarded as prior error estimates.

The optimal observation locations for D20 and MLD are presented in **Figure 2**, for experiments using the ACOM2 and ACOM3 ensemble for two domains that extend within ±15° and ±25° of the equator. We note that while the details of the arrays differ, the general features are quite similar. For example, for D20, the optimal arrays tend to have many observations between 5-15°S, with very few observations along the equator. By contrast for MLD, the optimal arrays tend to have many observations within 5° of the equator; and particularly to the east.

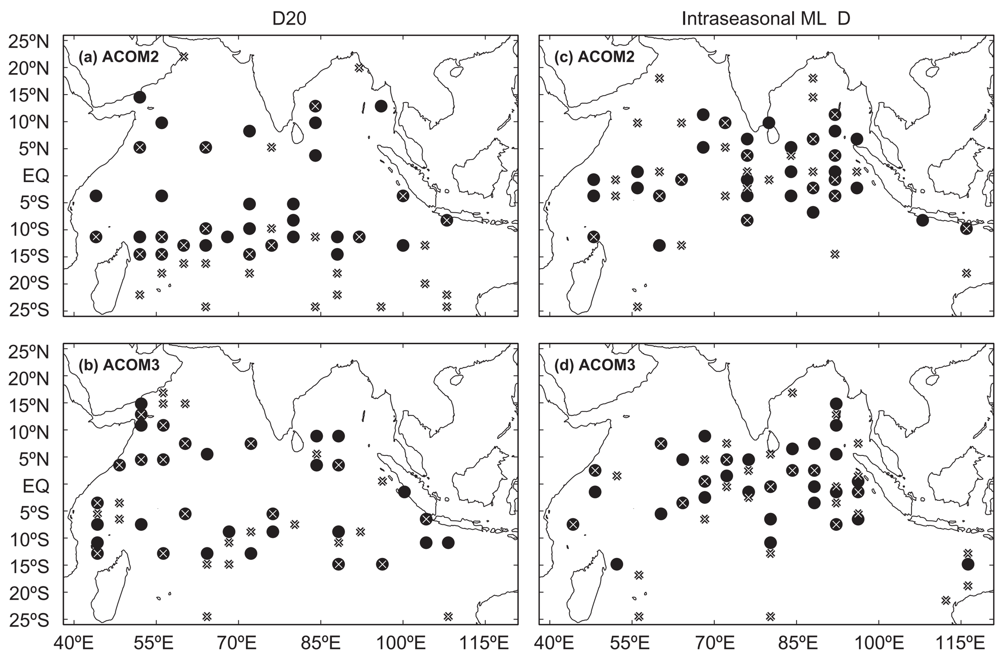


**Figure 2** Map showing the optimal observation locations for D20 (left) and MLD (right) using a 152-member ensemble from ACOM2 (top) and a 101-member ensemble from ACOM3; Crosses (circles) are for the ±25° (±15°) domain.

The theoretical posterior error for D20 and MLD are presented for all cases considered here in **Figure 3** and **4** respectively. For comparison, we present the root-mean-squared error (RMSE) for OSSEs spanning the period 1982-1988 using the ACOM2 ensemble for the period 1989-1994 in **Figure 5**. For each of these OSSEs, we contaminate the observations with normally distributed white noise with a standard deviation of 4 m. We note that the RMSE fields in **Figure 5** correspond well to the theoretical error estimates in the corresponding panels in **Figures 3** and **4**.

An indication of the number of observations the EnOI systems require to adequately represent D20 and MLD variability is given in **Figure 6**, showing the basin-averaged analysis error as a function of array size. For comparison, the basin-averaged prior error (**Figure 1**) for D20 and MLD is about 20 m and 5 m respectively; and recall that the observations are assumed to have an error of 4 m. These calculations suggest that an array of 30 moorings is probably a good choice for resolving interannual variability; and that intraseaonal variability is probably poorly constrained by this EnOI system, even when a much larger array is used; owing to its small spatial scales.

Some insight into why the EnOI system prefers observations at particular locations is gained by considering fields from the gain matrix in **Figure 7**. If only one observation is available and the background innovation $(\mathbf{d}-\mathbf{H}\mathbf{w}^B)$ is 1 m, the gain is equivalent to the EnOI-derived increments.

## Conclusions

We find that in general, observations south of 8°S and off the Indonesian coast are most important for resolving interannual variability; while observations within 5° of the equator; and particularly to the east, are important for resolving intraseasonal variability.

Wealth from Oceans National Research Flagship

**Acknowledgements:** This research is funded by Australia's CSIRO through appropriation funding and the Wealth from Oceans Flagship Program; and by the US Office of Naval Research Ocean Modelling Program through Grant N000140410345.

Poster designed by Lea Crosswell, CSIRO Marine and Atmospheric Research Communication Group