18/06/2018

I would like to take the opportunity to thank the reviewers for their constructive feedback. Please find the response to their comments below. The following format has been adopted:

Reviewer comment

Author response

Text within the manuscript

Anonymous Referee #1

General comments

This paper investigates the impact of data assimilation window length on a short-range ocean forecasting system. Current operational ocean forecasting systems use a range of different assimilation windows from 6 hours to 10 days, and yet there has been little work to look at the impact of different assimilation windows in one systems. This makes the topic of the paper novel and relevant for the ocean modelling community. As we move towards coupled ocean-atmosphere data assimilations the length of the time windows used in the ocean are likely to reduce to be consistent with atmosphere assimilation windows. This study may therefore be of particular interest to those developing coupled data assimilation systems. However, the results in the study would be more significant if the experiments were not using a synchronous data assimilation method. The author should address this explicitly earlier in the paper. It's not clear that you couldn't just achieve similar improvements through asynchronous data assimilation.

The 1-day cycle is neither synchronous or asynchronous as we use daily binned observations. Over the years we have run experiments using asynchronous FGAT on a 3-day cycle and never found it to yield any improvements over synchronous DA. I have added a statement to the methods section regarding this.

An asynchronous 3-day cycle FGAT (First Guess Appropriate Time) system was not compared with the 1-day or 3-day cycle systems as FGAT did not provide any significant improvements over the synchronous 3-day cycle. Mean increments and forecast errors from FGAT were comparable to the 3-day cycle (not shown).

The innovation statistics (Figure 6) from the 7 day forecasts are a significant result and strongly support the use of a 1 day window over a 3-day window in this system. However, the paper over emphasises the results from the mean and absolute mean increments. As alluded to in the background section, it is difficult to draw conclusions from comparisons of mean increments alone and the current organisation of the paper puts too much weight on the increment results. It would strengthen the interpretation of the assimilation increments if they were discussed within the context of the forecast statistics. I think that the paper could be substantially improved by presenting the innovation statistics first, as the key result, and providing the increments as supplementary evidence. Throughout the paper the author states that the differences in mean absolute increments suggest observations are having a greater impact with a one day window. I think that you need to be careful with how this statement is used. A reader may misinterpret this as meaning that larger increments automatically lead to an improved system. Presenting this result within the context of improved forecast innovations would make the statement more robust. The author should also clarify that larger increments do not necessarily mean a better data assimilation system.

The forecast statistics are now presented first in the results section. There were statements in the abstract and manuscript that did explicitly say that greater observation impact and smaller mean increments do not imply a better system. Nonetheless, the abstract has been changed to

A brief examination of the relationship between data assimilation cycle length and observation impact in a practical global mesoscale ocean forecasting setting is provided. Behind real-time reanalyses and forecasts from two different cycle length systems are compared and skill is quantified using all observations typically available for ocean forecasting. A 1-day Ensemble Optimal Interpolation (EnOI) cycle is compared to a 3-day cycle. The mean analysis increments for the 1-day system are significantly smaller suggesting a less biased system. Comparison of mean absolute increments identifies observations have greater impact in the 1-day system. Whilst smaller mean increments and greater observation impact do not guarantee a better forecast system, analysis of 7-day parallel forecasts show that the 1-day cycle system delivers improvement in predictability when compared to all available independent observations. The results are dependent on region, model and observing system, however, show the 1-day cycle provides an overall improvement in predictability, particularly in the subsurface. This appears to mainly come from less biased initial conditions and suggests greater retention of memory from observations and improved balance in the model.

The conclusion went on to say that

“The 1-day cycle system was shown to provide less biased initial conditions and improved forecasts, suggesting the relatively smaller mean analysis increments were reliably sampled.”

This has been removed and the conclusions have been rewritten

In places the paper seems to lack details or justification. For example, the choice of forecast period for assessment or the choice of assimilation windows for the experiments. And in some places the paper seems to make contradictory conclusions about the results (particularly in relation to the mean increments). The paper should be modified to give a clearer narrative.

Justification for assimilation windows is based on the 3-day window behind real-time cycle and the 7 day forecasts that the current operational system uses and that the 3-day window is about the maximum length window that would be appropriate for assimilation into mesoscale eddy resolving system.

The following text has been changed to address this

“OceanMAPS is global eddy resolving, forced by Numerical Weather Prediction (NWP) and runs on a 3-day data assimilation cycle.”

to

“OceanMAPS is global eddy resolving, forced by Numerical Weather Prediction (NWP), runs on a 3-day data assimilation cycle and carries out 7-day forecasts”

In order to address some contradictions the following text in the abstract has been changed from

“This is thought to come from less biased initial conditions, greater observation impact and improved consistency with respect to the timing of model and observations.“

to

“This appears to mainly come from less biased initial conditions.“

I think that there are some errors in the interpretation and description of the results. More details are given in the specific comments.

Specific comments

Abstract:
Page 1, line 4-5. The mean increments look to be approximately 1/3 smaller in the 1-day experiment, which is what you would expect for linear error growth. I don’t think that you can make any statements about bias here without consideration of the error growth throughout the assimilation window. This statement is also inconsistent with your discussion of the mean increment results on page 4, line 8. Page 1, line 9. I don’t remember seeing any statistics which showed that the biggest improvements were in the Western Boundary currents.

The following statement regarding bias is in the context of error growth and is not inconsistent with other statements.

“… the mean increments are much smaller in the 1-day than the 3-day system. This is a natural result of shorter cycle length as model error growth is more constrained with more frequent analyses…”

This has been changed to

“.. mean increments are about one third smaller in the 1-day than the 3-day system, which can be expected for approximate linear error growth.”

Table 3 has robust statistics that show improvement in the bias in the Tasman Sea – where error is dominated by dynamical instabilities in the East Australian Current. Regardless, this is just one WBC region so I have removed the claims to better forecast skill for WBCs from the text.

Background:
Page 1, line 21-22. Over fitting is not just a problem for long data assimilation windows. In fact a long data assimilation window with good super obing or thining could produce smoother increments and be less influence by noise in the observations.

Point taken, however, in synchronous DA the longer the window, the more time averaging goes into the super-observation and mesoscale eddies become smeared out, less balanced, less accurate, less realistic.

Page 2, lines 11. This paragraph is a bit confusing. It seems to argue that the mean increments are not a good indicator of bias, which contradicts your result on line 5 in the abstract. I didn't really understand what the purpose of this paragraph was. To justify the use of mean absolute increments?

The paragraph sets the context for the problem, which is the reliability of the mean increment as a proxy for bias. Statements are made on the limitations then the following statement is made which sets the qualification.

"Provided observation coverage is sufficient, observation bias is minimal and background error covariances are physically meaningful, well sampled mean analysis increments can be a reasonable indication of model bias"

Page 2, line 29. What is the forecast range of OFAM3? This might give more context to your choice of forecasts.

OFAM3 was a 20-year spin-up run (free-model with atmospheric reanalysis forcing). Forecast range is not relevant to OFAM3.

Page 2, line 30. Have you specified the horizontal resolution anywhere? This is important since the focus of the paper is mesoscale forecasting.

This has now been specified

Page 3, line 10. "EnKF-C (Sakov, 2014) with Ensemble Optimal Interpolation (EnOI)" is not general terminology for a data assimilation scheme. This name is too specific to be used without context. In reality, I think you are actually using EnOI?

The sentence has been changed for clarification

"For data asimilation the EnKF-C software (Sakov, 2014) is used in Ensemble Optimal Interpolation (EnOI) (Evensen, 2003) mode"

Page 3, line 14. More details about the observation operator would be useful. You should also define the linear observation operator in equation 1.

This has been addressed

Page 3, line 15. More details could be given on the data assimilation system, e.g clarifying that this is a synchronous data assimilation scheme, defining when in the time window the increments are applied (presumably the middle).

It has been clarified that the 3-day cycle is synchronous and the scheme is clearly shown in Figure 2.

Page 3, line 23. You discuss the impact of super obing before introducing that you have used super obing. I think the order should be switched round.

This has been rearranged and improved

Results:
Page 4, line 8. Seems contradictory to the abstract (page1, line 4-5)

This has been clarified

Page 4, line 11. The MAI from the 2 experiments are only directly comparable if the forecast error growth is linear. It is worth discussing this here. Your results in Figure 7 should give some indication of the forecast error growth. From these figures it looks like the forecast error growth in the first day is a bit larger than subsequent days.

Have added the following to address this

"MAI between the two systems is only directly comparable if the forecast error growth is linear. Forecast error growth in the two systems is largest on the first day (shown later in Figure 7)."

Page 4. It could also be useful to consider the variability in the increments.

The temporal RMS of the increments were initially looked at, and these were ~1/3 smaller in one day system similar to the mean increments. This was not considered to add anything and prompted looking for another metric (MAI).

Page 4, line 13-14. Would you expect the fact that you are assimilating more observation in the 1-day experiment to also impact on the magnitude of the increments?

Not necessarily as with the shorter cycle there is less error growth and generally smaller increments.

Page 4, line 16. What is the temporal resolution of the kinetic energy outputs in Figure 7.

This is 6 hourly. I have added this to the text

Page 4, line 27. But also the model is only free running for 1 day before the next increments is applied, so less time to drift.

OK

Page 4, line 28. Wouldn't the eddy kinetic energy be a better representation of the mesoscale energy?

Yes, the total kinetic energy in the systems represents this well as most of the kinetic energy in the ocean, and in the eddy-resolving ocean model used here, is in the mesoscale.

Page 4, line 29. If you are going to claim that the model kinetic energy is closer to the observations, you should also show the observation kinetic energy. Comparing the results to the observations would also give more context for the difference between the two data-assimilation experiments. From the current Figure it's not clear how significant this increase in Kinetic Energy is.

There are no observations of kinetic energy that would represent the kinetic energy in the model. One might derive KE from a geostrophic current product based on altimeter tracks (GSLA) but this has issues and is not representative of the model total KE. For example, GSLA does not project subsurface accurately, is unreliable on the shelf and can have over fitting of sparse altimeter tracks. The data assimilation system tells us that the analysis, which is closer to the observations, and satisfies more than just altimetry, has greater total kinetic energy every cycle, regardless of cycle length. The significance of the increase is important. We have not studied this in detail. The point been made here is that the signal is there and it makes a difference. Ultimately the improved forecast errors signify the 1-day system to be closer to observations.

Page 5, line 10. "mean forecast bias is more significantly reduced for SLA, SST and sub-surface temperature" - the mean forecast bias for SLA is actually slightly larger in the 1-day experiment in Table 3.

This statement is referring to global stats in Table 1. The mean forecast bias for SLA in Table 3 is for the Tasman Sea region, not really a problem. There is a sentence near the end that acknowledges this

"It is noted that, whilst an overall improvement in global performance was detected, in some regions the 1-day scheme may not perform better than the 3-day system"

Page 5, line 13. Are the increments applied in the middle of the time window? Could you clarify this.

The increments are applied at the analysis time. The observation window is centred as illustrated in Figure 2.

Page 5, line 10-11. Are the MAD statistics in Figure 6 calculated in the same way (using the same forecasts) as those presented in Table 2? Why do look so different? For example, the subsurface temperature MAD at day 7 looks to have a value of approximately 0.625, but in table 2 it's given as 0.603.

Table 2 provides the average of the MAD for the 7-day forecast. It should check out as the average of days 1-7 in Figure 6. The existing statement says we include day 0 (position of star in figure 2) in Figure 6 for illustration but do not include this in forecast stats as day 0 is not independent.

"For day zero, statistics are included that represent the errors in the initial conditions and the observation window partially overlaps half of this day in both systems so the statistics for day zero cannot be regarded as independent"

Conclusion:
Page 5, line 25-26. "Further 1 year runs of the two systems with an improved model using renanalysis bulk flux forcing have confirmed (not shown) that the 1-day cycle provides improvements in forecasting the mesoscale circulation in the western boundary current regions." Is this the evidence for the statement in the abstract that the biggest improvements are in the western boundary current region? You should show this result if it forms part of your main conclusions. It would, in general, be good to see more results focused on the mesoscale region given that the focus of the paper is mesoscale forecasting.

I have changed all references to improvements in WBCs to improvements in forecasting the mesoscale circulation.

Technical Corrections

Page 9, Figure 1. There is quite a lot of irrelevant information on this plot which makes the key information difficult to see. The current vectors make the figure appear noisy in print, and they are not discussed anywhere in the paper. It would be best if they were removed.

I would prefer that the original of Figure 1 is supplied to the reviewers and included in the final PDF. This unfortunately did not carry through to the reviewed manuscript. There is a wealth of information that can be clearly seen by zooming in.

Page 10, Figure 4. It would be better to use a sequential colour bar for Mean Absolute Increments.

Thanks for the suggestion. Figure 4 still shows what was intended.

Page 11, Figure 6 caption. typo throughout.

Fixed

Page 11, Figure 7, (d). It's very difficult to see the results from the 3-day experiment for salinity.

Figure 7, which is now Figure 3 has been improved.

Anonymous Referee #2

Specific Comments: 1) Figure 1 is poorly displayed, at least in the paper version that will be used by most readers. "Unassimilated forward independent super-observations are shown with coloured circles and grey outline on the same colorscale": the grey outline is barely visible, while it is virtually impossible to see the coloured circles. Maybe C1 a second panel can be shown?

The original submitted figure is high resolution and did not translate to as good resolution in the GMDD generated manuscript. My preference is the original be embedded in the final electronic version allowing readers to see the many interesting details within this figure. Most readers should be able to refer to the electronic version.

2) Figure 3 is confusing with three colors, blue, red and orange (?). I am guessing the red color is simply the overlap between blue and orange. This plot is relatively simple, and can be displayed by black-and-white lines (solid vs. dotted) showing a small bar or shaded lines representing the +/- 1 standard deviation.

Thanks for the suggestion The figure shows what was intended.

3) p1, line 18, "...project into unobserved variables", replace "into" with "onto"

Changed

4) p2, line 5, replace "~~" with "approximately"

Done

5) in the "Data and Methods" section, there is no place to mention the model grid size, which should be explicitly stated. I have to read the Oke et al., 2013a to find out this information as 1/10 degree.

Have explicitly now stated the horizontal resolution

6) p3, line 8, replace "3 hourly" with "3-hourly"

Done

7) some acronyms not commonly used by the community are not necessary, e.g., MAI, MAD

These have been defined once only at the first occurrence as appropriate for scientific writing.

8) Figure 2, only one star (representing forecast base time) is shown, should be dis- played at the center of every 3-day cycle, right?

This could be done, however, the message is more effectively communicated the way it is displayed as it shows that both cycles need to propagate the model past the last analysis to the end of the observation window.

9) the font for some figures should be somewhat larger, particularly if multiple figures are printed on the same page

Compare Results

Old File:	versus	New File:
gmd-2017-298-manuscript-version2.pdf		**gmd-2017-298-AC3-supplement.pdf**
12 pages (9.70 MB)		**14 pages (16.41 MB)**
5/3/18, 9:24:52 pm		31/7/18, 10:10:36 am

Total Changes

159

Text only comparison

Content

75 Replacements

52 Insertions

32 Deletions

Styling and Annotations

0 Styling

0 Annotations

Go to First Change (page 1)

Data assimilation cycle length and observation impact in mesoscale ocean forecasting

Paul Sandery[1]

[1]CSIRO Oceans and Atmosphere, Castray Esplanade Battery Point TAS 7008

Correspondence: Dr Paul A Sandery (paul.sandery@csiro.au)

Abstract. A brief examination of the relationship between data assimilation cycle length and observation impact in a practical global mesoscale ocean forecasting setting is provided. Behind real-time reanalyses and forecasts from two different cycle length systems are compared and skill is quantified using all observations typically available for ocean forecasting. A 1-day Ensemble Optimal Interpolation (EnOI) cycle is compared to a 3-day cycle. The mean analysis increments for the 1-day system are significantly smaller suggesting a less biased system. Comparison of mean absolute increments identifies observations have greater impact in the 1-day system. Whilst smaller mean increments and greater observation impact do not guarantee a better forecast system, analysis of 7-day parallel forecasts show that the 1-day cycle system delivers improvement in predictability, particularly for the subsurface. This improvement appears to mainly come from less biased initial conditions and suggests greater retention of memory from observations and improved balance in the model.

1 Background

Cycle length in sequential data assimilating forecasting systems is an important setting that relates to dynamical scales resolved by the numerical model and the observation system. Many ocean forecasting systems, for example those described in Cummings and Smedstad (2014); Martin et al. (2007); Chassignet et al. (2009); Ferry et al. (2010); Bertino et al. (2008), make different choices around cycle length. Shorter cycle-length implies more frequent analyses and initialisation of the dynamical model. This may not necessarily lead to a better forecast system. In multivariate systems, observed variables project onto unobserved variables and systems tend to perform best when model error covariances are adequately sampled and there is reasonable coverage of multiple observation types. Longer cycles favour better coverage, however, can introduce larger analysis increments, temporal representation errors and overfitting of observational data. Bias is a fundamental problem in atmospheric and ocean forecasting affecting system performance. Bias arises within an assimilation cycle shared by issues related to the assimilation system, the model and observations. Identifying the cause of bias can be almost impossible (Houtekamer and Zhang, 2016). Mean analysis increments are sometimes used to detect model bias (Houtekamer and Mitchell, 2005; Oke et al., 2013b) and some bias correction schemes are based on this (Zhang et al., 2016; Takacs et al., 2016; Ha and Snyder, 2014).

Some care must be taken when using mean analysis increments as a proxy for model bias as they are dependent on the structure of the background covariances and also contain observation bias (Dee, 2005). Furthermore they can be approximately zero and relatively meaningless in regions of few or no observations or when large errors of opposite sign cancel out over time. Provided observation coverage is sufficient, observation bias is minimal and background error covariances are physically meaningful, well sampled mean analysis increments can be a reasonable indication of model bias. Dee and Da Silva (1998) illustrated that mean analysis increments tend to underestimate forecast bias. This is because they depend on the rate and period of growth of perturbations, i.e. model error growth, so they are forecast lead-time and cycle length dependent. This questions the use of mean analysis increments to estimate and compare the bias of forecasting systems with different cycle lengths. It appears that the cycle length, however, should be based on that which is best for predictability. Aspects of this are touched on by running twin experiments with a global ocean forecasting system using cycle lengths of 1 and 3-days. The system used in this study is the current Bureau of Meteorology Ocean Model Analysis and Prediction System (OceanMAPS) version 3. Previous versions of this system are documented in Brassington (2013) and Brassington et al. (2007). OceanMAPS is global eddy resolving, forced by Numerical Weather Prediction (NWP), runs on a 3-day data assimilation cycle and carries out 7-day forecasts. It is able to constrain aspects of the mesoscale variability to the available real-time observations. It produces forecasts of synoptic features of the ocean circulation, such as the locations of eddies and fronts, daily changes in sea surface temperature and mixed layer depth, wind driven surface flows and coastal trapped waves. As typical for ocean forecast systems like OceanMAPS, the largest errors tend to occur in regions of most rapidly growing dynamical instabilities (O'Kane et al., 2011), such as western boundary currents and along the Antarctic Circumpolar Current (ACC). Some of these features and the characteristic spatio-temporal scales resolved by the model are captured in Figure 1, which presents a snapshot of Sea Level Anomaly (SLA) for the 9th September 2013. The behind real-time forecasted SLA is shown with unassimilated forward independent super-observations for the same day from the 1-day cycle system. Information regarding the use of forward super-observations for forecast verification, as used in this study, can be found in Sakov and Sandery (2015) and Sandery and Sakov (2017).

2 Data and Methods

The Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model version 4.1 (MOM4p1) (Griffies et al., 2009) is used. This is a Boussinesq three-dimensional primitive equation volume conserving ocean model. The OceanMAPS grid has 0.1° horizontal resolution and is the same as the Ocean Forecasting Australia Model version 3 (OFAM3) (Oke et al., 2013a), which is based on bathymetric data from Smith and Sandwell (1997). The grid has 51 vertical levels and the top cell approximates quantities at 2.5 m depth with the average resolution in the upper 200 m being approximately 10 m. The physical model settings include the use of a 4th-order Sweby advection method and a scale dependent isotropic Smagorinksy biharmonic horizontal mixing scheme as described in Griffies and Halberg (2000). The General Ocean Turbulence Model (GOTM) κ-ϵ scheme is used for vertical mixing. Note that tides are not explicitly modelled, rather a parameterisation of tidal mixing is implemented using the scheme of Lee et al. (2006).

Initial conditions for both systems are the same and taken from the multi-year OFAM3 spin-up for the 1st January 2012. The 1 and 3-day cycle systems are spun-up with data assimilation over a 1-year period to the 1st Janaury 2013. Hindcasts are continued throughout 2013 and a series of 7-day forecasts, 3-days apart, with identical base dates as illustrated in Figure 2 are carried out from 3rd January 2013. The forecast experiments were done behind real-time, therefore observations in the 12-24 hours prior to forecast base time were available to both systems, whereas in practice they would not be available in this period in a real-time system. The model is forced by 3-hourly prescribed surface fluxes of momentum, heat and salt from the Bureau of Meteorology operational global NWP system version 1 which is known as ACCESS-G APS1 (Australian Community Climate and Earth System Simulator). For data asimilation the EnKF-C software (Sakov, 2014) is used in Ensemble Optimal Interpolation (EnOI) (Evensen, 2003) mode. The analysis equation and background error covariances can be written as

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{B}\mathbf{H}^T\left[\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right]^{-1}\left[\mathbf{y} - \mathcal{H}(\overline{\mathbf{x}^f})\right], \tag{1a}$$

$$\mathbf{B} \equiv \mathbf{A}\mathbf{A}^T\left[(\mathbf{m}-\mathbf{1})\right]^{-1}, \tag{1b}$$

where $\mathbf{x}^a$ and $\mathbf{x}^f$ are analysis and forecast state vectors respectively; $\mathbf{y}$ is an observation vector; $\mathbf{H}$ is a linear observation operator, i.e. $\mathbf{H} = \nabla\mathcal{H}(\mathbf{x})$, where $\mathcal{H}$ is a linear affine observation operator; $\mathbf{B}$ is background error covariance; $\mathbf{R}$ is observation error covariance; $\mathbf{A}$ represents ensemble anomalies; $\mathbf{m}$ is ensemble size and ^{T} denotes matrix transposition. $\mathbf{x}^f$ is taken to be an instantaneous model state, whereas $\overline{\mathbf{x}^f}$ is a 1 day mean and 3 day mean in the respective systems. The system uses no nudging or incremental analysis updating (Ourmières et al., 2006), rather the model is directly initialised to the analysis. This approach allows the model to run the complete length of each cycle as a dynamical forecast without being influenced by forcing from nudging terms in the model equations. It also includes any initialisation shock from imbalance in the analysis in order to assess the impact of this on forecasts. $\mathbf{B}$ is based on a 144 member ensemble of intra-seasonal (1-day minus bimonthly mean) anomalies generated from an 18 year run of OFAM3. A source of time filtering is implicit in the innovation vector $[\mathbf{y} - \mathcal{H}(\overline{\mathbf{x}^f})]$ from the fact that the super-observations tend to represent averages over the time window, particulary for observations with relatively larger coverage, such as SST. An asynchronous 3-day cycle FGAT (First Guess Appropriate Time) system was not compared with the 1-day or 3-day cycle systems as FGAT did not provide any improvements over the synchronous 3-day cycle. Mean increments and forecast errors from FGAT were comparable to the 3-day synchronous cycle (not shown).

Both the 1-day and 3-day systems assimilate the same original observations only once. The following observations are converted to super-observations weighted by inverse error variance. Altimetric SLA is taken from from the Radar Altimeter Database System (RADS) (Schrama et al., 2000) using tide, mean dynamic topography and inverse barometer corrections. SLA observations are limited to water depths greater than 200 m. Sea surface temperature (SST) retrievals from the NAVO-CEANO (May et al., 1998) and WindSat (Gaiser et al., 2004) databases are used. All available in-situ temperature and salinity observations on the Global Telecommunications System (GTS) are used. These include Argo profiles (Roemmich et al., 2009), Conductivity Temperature Depth (CTD) and eXpendable BathyThermograph (XBT) profiles. The EnOI systems are run in a cycle scheme that centres the observation window as shown in Figure 2. The amount of super-observations generated by the system from the original observations for the 1-day system is larger than the 3-day system. The total number of super-observations used in 2013 is shown in Table 1. In the data assimilation, a 250 km localisation radius is used for all observation types. The mean sea-level from OFAM3 (Oke et al., 2013a) is used for the model's mean dynamic topography to assimilate along track SLA observations.

3 Results

Global forecast innovation errors for the 1 year behind real-time period for 2013 are provided in Table 1. These are based on forward unassimilated observations, which can be regarded as independent. The 1-day cycle benefits statistically from a shorter forecast lead-time. These errors suggest an improvement in performance in constraining SLA, SST and sub-surface temperature and salinity. In order to determine whether this result is only dependent on forecast lead-time, a series of 44 parallel 7-day forecasts using identical base dates from 3rd January 2013 are analysed. These forecasts are compared to unassimilated observations. The global 7-day mean forecast errors are shown in Table 2 and Table 3 repeats this for the Tasman Sea region. It's interesting to note that whilst mean absolute deviation (MAD) global forecast errors are marginally smaller in the 1-day system that mean forecast bias is more significantly reduced for SLA, SST and sub-surface temperature. Figure 3 shows the global MAD forecast error growth as a function of lead time. Note that in order to ensure genuine forecasts are made the model is propagated to the end of the respective observation window in both systems, which is the position of the star in Figure 2. Daily mean forecast fields are saved and these are compared to the observations. For day zero, statistics are included that represent the errors in the initial conditions and the observation window partially overlaps half of this day in both systems so the statistics for day zero cannot be regarded as independent. The results suggest the 1-day system is better overall as a forecast system with improvements in lead time of about 1 day in surface variables and up to 7-days in sub-surface variables. The errors for salinity are relatively high for both systems as no restoring to salinity is used, however, the relative improvement is apparent.

The mean analysis increments for SST and SLA are shown in Figure 4. Three key features emerge regarding this estimate of model bias in the mean increments. There is an equatorial eastern Pacific Ocean cold bias, a southern ocean high latitude warm bias and mesoscale warm and cold biases in the western boundary current and ACC regions. Without speculating on the source of these systematic model errors it is noted that the first two aformentioned bias features have been detected in the CSIRO Climate Analysis Forecast Ensemble (CAFE) System, which is a configuration of the GFDL coupled model version 2.1 (CM2.1) run under an ensemble Kalman filter data assimilation framework. Figure 4 shows that mean increments are about one third smaller in the 1-day than the 3-day system, which can be expected for approximate linear error growth. The spatial patterns are very similar with the main difference being amplitude. Another way to compare increments over a period of time is to calculate the Mean Absolute Increment (MAI) (Figure 5). This is done in the following way. In each 3-day period the 1-day increments are summed and then the absolute values calculated. The mean of the absolute values over the 1 year period are then calculated. MAI for the two systems is only directly comparable if the forecast error growth is linear. The difference in mean increments between the two systems suggests this, however, error growth in the two systems is largest on the first day (as seen in Figure 3) and becomes mainly linear after this. Regardless, the differences in spatial distribution of MAI for SLA and SST, shown in Figure 6, indicate the 1-day system has generally larger MAI. It can bee seen there is a greater impact from the observing system. The 1-day system projects more information from observed variables into unobserved variables through the background error covariances due to the relatively smaller observation coverage per analysis. For instance, in-situ observations from the Tropical Atmosphere Ocean - Triangle Trans-Ocean Buoy Network (TAO-TRITON) moored array in the equatorial Pacific Ocean produce larger MAI on SLA and SST in the 1-day system. It is also evident that the 1-day system has larger MAI in the western boundary currents and ACC. Figure 6 shows, as expected, that SST projects more into SLA in the 1-day system. SST observations in the 1-day system appear to be having greater impact in the regions of fastest growing dynamical instabilities. Interestingly, the relatively smaller MAI for SST in the 1-day system in the Inter-Tropical Convergence Zone (ITCZ), in the tropical warm pool in the western Pacific Ocean and at high latitudes in the Southern Ocean indicate the observing system is having less impact in these areas in the 1-day system.

Data assimilation typically injects energy into a forecast model as the observed fronts can be sharper than what can be supported by the model. In each cycle we usually see a jump in total kinetic energy with subsequent diminishing until the end of the forecast. This can be caused by factors such as insufficient horizontal and vertical resolution and imbalances in the analysis. Figure 8 shows total kinetic energy at 6 hourly temporal resolution. Here it can be seen that data assimilation in the 1-day system renders the state at a higher kinetic energy level, with smaller amplitude temporal fluctuations between cycles. The latter reflects the smaller increments per cycle in the 1-day system, however, the larger kinetic energy state indicates that more energy is retained in the mesoscale eddies, which infers that the gradients in SLA are maintained closer to observations. The larger MAI for SLA and SST in the 1-day system in the western boundary current and ACC regions reflects that observations are having a larger impact in these regions. The total kinetic energy dissipation for both systems in 2013 was calculated by summing the dissipation within each cycle and removing the trend. The 1-day system total kinetic energy dissipation is 8.4×10^{18} J and that for the 3 day system is 9.6×10^{18} J. The relative total kinetic energy dissipation, estimated by subtracting the mean dissipation from the respective systems, shows that the 1-day system has approximately 17% less relative kinetic energy dissipation than the 3-day system suggesting it is more effective at preserving SLA gradients and may be more dynamically balanced.

4 Conclusions

Global errors from a set of 44 parallel 7-day forecasts over a 1 year period in 2013 showed the 1-day cycle system delivered improvements in predicting sea surface temperature, sea level anomaly, subsurface temperature and salinity. The difference in mean absolute increments between the two cycle length systems indicated that the same observations had a greater impact on the 1-day system, with a larger degree of observed variables projecting onto unobserved variables. Greater observation impact does not necessarily lead to an improved forecast system as overfitting observations can produce dynamical imbalances, which can have deleterious effects on forecasts. The results, however, indicate that the 1-day cycle takes greater advantage of the observations and, compared to the 3-day cycle, is less biased in initial conditions and forecasts. This suggests also that the background error covariances are a reasonable estimate of model error. With the shorter cycle length data assimilation introduces a larger amount of kinetic energy from the observations into the state, bringing the model closer to a realistic representation of ocean's kinetic energy. The 1-day cycle introduced a larger amount of information from the observations into the model with more frequent smaller adjustments at finer scales. The overall improvement in predictability, particularly in the subsurface, suggests greater retention of memory from observations and improved balance in the model. It is noted that, whilst an overall improvement in global performance was detected, in some regions the 1-day scheme may not perform better than the 3-day system. The results are a practical example of the influence of cycle length in global mesoscale ocean forecasting with the current observation network. The 1-day cycle is closer to asynchronous data assimilation and appears to be an improvement over the First Guess Appropriate Time (FGAT) approach (Cummings, 2005; Lee, 2005; Atlas et al., 2011) as our FGAT experiments did not yield as significant an improvement.

Acknowledgements. This work was carried out within the Bluelink Project with financial support from the Australian Bureau of Meteorology, Commonwealth Scientific and Industrial Research Organisation and Royal Australian Navy. Numerical simulations were undertaken using the Raijin supercomputer at the National Computational Infrastructure.

5 Additional Information

5.1 Code availability

The ocean model is available at https://github.com/mom-ocean/MOM4p1 and the data assimilation code can be found at https://github.com/sakov/enkf-c. These codes are documented within. The OceanMAPS3 system and observation processing scripts are intellectual property of the Bureau of Meteorology.

References

Atlas, R., Hoffman, R. N., Ardizzone, J., Leidner, S. M., Jusem, J. C., Smith, D. K., and Gombos, D.: A cross-calibrated, multiplatform ocean surface wind velocity product for meteorological and oceanographic applications, Bulletin of the American Meteorological Society, 92, 157, 2011.

Bertino, L., Lisæter, K., and Scient, S.: The TOPAZ monitoring and prediction system for the Atlantic and Arctic Oceans, Journal of Operational Oceanography, 1, 15–18, 2008.

Brassington, G. B.: Multicycle ensemble forecasting of sea surface temperature, Geophysical Research Letters, 40, 6191–6195, 2013.

Brassington, G. B., Pugh, T., Spillman, C., Schulz, E., Beggs, H., Schiller, A., and Oke, P. R.: BLUElink> Development of Operational Oceanography and Servicing in Australia, Journal of Research and Practice in Information Technology, 39, 151–164, 2007.

Chassignet, E. P., Hurlburt, H. E., Metzger, E. J., Smedstad, O. M., Cummings, J. A., Halliwell, G. R., Bleck, R., Baraille, R., Wallcraft, A. J., Lozano, C., et al.: US GODAE: global ocean prediction with the HYbrid Coordinate Ocean Model (HYCOM), Tech. rep., DTIC Document, 2009.

Cummings, J. A.: Operational multivariate ocean data assimilation, Quarterly Journal of the Royal Meteorological Society, 131, 3583–3604, 2005.

Cummings, J. A. and Smedstad, O. M.: Ocean Data Impacts in Global HYCOM*, Journal of Atmospheric and Oceanic Technology, 31, 1771–1791, 2014.

Dee, D. P.: Bias and data assimilation, Quarterly Journal of the Royal Meteorological Society, 131, 3323–3343, 2005.

Dee, D. P. and Da Silva, A. M.: Data assimilation in the presence of forecast bias, Quarterly Journal of the Royal Meteorological Society, 124, 269–295, 1998.

Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, Ocean dynamics, 53, 343–367, 2003.

Ferry, N., Parent, L., Garric, G., Barnier, B., and Jourdain, N. C.: Mercator global Eddy permitting ocean reanalysis GLORYS1V1: description and results, Mercator-Ocean Quarterly Newsletter, 36, 15–27, 2010.

Gaiser, P. W., St Germain, K. M., Twarog, E. M., Poe, G. A., Purdy, W., Richardson, D., Grossman, W., Jones, W. L., Spencer, D., Golba, G., et al.: The WindSat spaceborne polarimetric microwave radiometer: Sensor description and early orbit performance, Geoscience and Remote Sensing, IEEE Transactions on, 42, 2347–2361, 2004.

Griffies, S. M. and Halberg, R. W.: Biharmonic Friction with a Smagorinsky-Like Viscosity for Use in Large-Scale Eddy-Permitting Ocean Models, Monthly Weather Review, 128, 2935–2946, 2000.

Griffies, S. M., Schmidt, M., and Herzfeld, M.: Elements of MOM4p1, Tech. rep., 2009.

Ha, S.-Y. and Snyder, C.: Influence of surface observations in mesoscale data assimilation using an ensemble Kalman filter, Monthly Weather Review, 142, 1489–1508, 2014.

Houtekamer, P. and Zhang, F.: Review of the ensemble Kalman filter for atmospheric data assimilation, Monthly Weather Review, 144, 4489–4532, 2016.

Houtekamer, P. L. and Mitchell, H. L.: Ensemble kalman filtering, Quarterly Journal of the Royal Meteorological Society, 131, 3269–3289, 2005.

Lee, H.-C., Rosati, A., and Spelman, M. J.: Barotropic tidal mixing effects in a coupled climate model: Oceanic conditions in the Northern Atlantic, Ocean Modelling, 11, 464 – 477, https://doi.org/http://dx.doi.org/10.1016/j.ocemod.2005.03.003, http://www.sciencedirect.com/science/article/pii/S1463500305000326, 2006.

Lee, M.-S.: Preliminary tests of first guess at appropriate time (FGAT) with WRF 3DVAR and WRF model, Asia-Pacific Journal of Atmospheric Sciences, 41, 495–505, 2005.

Martin, M., Hines, A., and Bell, M.: Data assimilation in the FOAM operational short-range ocean forecasting system: A description of the scheme and its impact, Quarterly Journal of the Royal Meteorological Society, 133, 981–995, 2007.

May, D. A., Parmeter, M. M., Olszewski, D. S., and McKenzie, B. D.: Operational processing of satellite sea surface temperature retrievals at the Naval Oceanographic Office, Bulletin of the American Meteorological Society, 79, 397–407, 1998.

O'Kane, T. J., Oke, P. R., and Sandery, P. A.: Predicting the East Australian Current, Ocean Modelling, 38, 251–266, 2011.

Oke, P. R., Griffin, D. A., Schiller, A., Matear, R. J., Fiedler, R., Mansbridge, J., Lenton, A., Cahill, M., Chamberlain, A., M., and Ridgway, K.: Evaluation of a near-global eddy-resolving ocean model, Geoscientific Model Developments, 6, 591–615, 2013a.

Oke, P. R., Sakov, P., Cahill, M. L., Dunn, J. R., Fiedler, R., Griffin, D. A., Mansbridge, J. V., Ridgway, K. R., and Schiller, A.: Towards a dynamically balanced eddy-resolving ocean reanalysis: BRAN3, Ocean Modelling, 67, 52–70, 2013b.

Ourmières, Y., Brankart, J.-M., Berline, L., Brasseur, P., and Verron, J.: Incremental analysis update implementation into a sequential ocean data assimilation system, Journal of Atmospheric and Oceanic Technology, 23, 1729–1744, 2006.

Roemmich, D., Johnson, G. C., Riser, S., Davis, R., Gilson, J., Owens, W. B., Garzoli, S. L., Schmid, C., and Ignaszewski, M.: The Argo Program: Observing the global ocean with profiling floats, 2009.

Sakov, P.: EnKF-C user guide, version 0.41, 2014.

Sakov, P. and Sandery, P.: Comparison of EnKF and EnOI regional ocean reanalysis systems, Ocean Modelling 89, 45-60, 2015.

Sandery, P. A. and Sakov, P.: Ocean forecasting of mesoscale features can deteriorate by increasing model resolution towards the submesoscale, Nature Communications, 8, 1566, 2017.

Schrama, E. J., Scharroo, R., and Naeije, M.: Radar Altimeter Database System (RADS): Towards a generic multi-satellite altimeter database system, Netherlands Remote Sensing Board (BCRS), Programme Bureau, Rijkswaterstaat Survey Department, 2000.

Smith, W. H. F. and Sandwell, D. T.: Global Sea Floor Topography from Satellite Altimetry and Ship Depth Soundings, Science, 277, 1956–1962, 1997.

Takacs, L. L., Suárez, M. J., and Todling, R.: Maintaining atmospheric mass and water balance in reanalyses, Quarterly Journal of the Royal Meteorological Society, 142, 1565–1573, 2016.

Zhang, B., Tallapragada, V., Weng, F., Sippel, J., and Ma, Z.: Estimation and correction of model bias in the NASA/GMAO GEOS5 data assimilation system: Sequential implementation, Advances in Atmospheric Sciences, 33, 659–672, 2016.

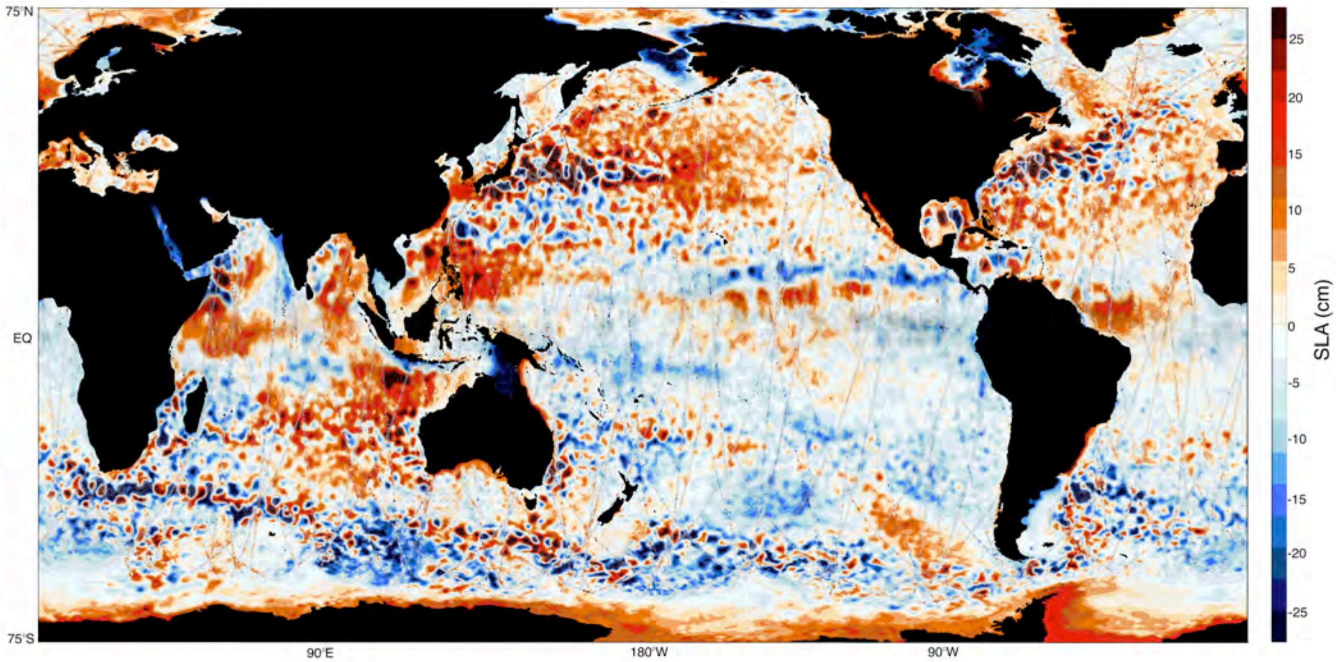

Figure 1. Forecast sea-level anomaly (SLA) from the 1-day cycle system for the 3rd of September 2013. Unassimilated forward independent super-observations are shown with coloured circles and grey outline on the same colorscale. The figure is high resolution and may be zoomed in for a detailed inspection of any region in the electronic version. Also shown are surface current vectors (black arrowheads) and surface wind vectors (blue arrowheads).

Table 1. Global mean behind real-time forecast innovation mean absolute deviation (MAD) and bias for sea-level anomaly (SLA), sea surface temperature (SST), sub-surface temperature (T) and salinity (S) from 1 year behind real-time period for 2013. See Figure 2 for cycle scheme. Total number of super-observations used in 2013 shown. † 1-day system ‡ 3-day system.

Variable (units)	MAD†	Bias†	MAD‡	Bias‡	Observations†	Observations‡
SLA (cm)	5.14	0.05	5.48	0.08	27070422	26033356
SST (K)	0.277	0.014	0.330	0.03	210063788	175258730
T (K)	0.517	-0.0877	0.539	-0.0934	6125208	5964116
S (psu)	0.13	0.0096	0.14	0.0104	5562515	5380711

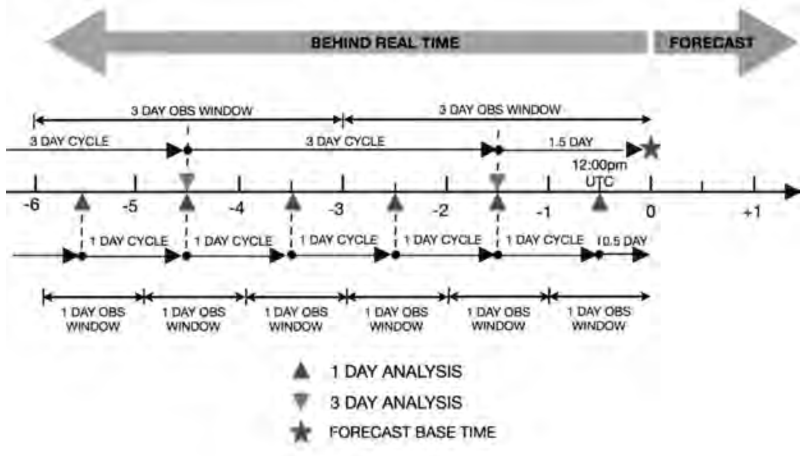

Figure 2. The analysis-forecast scheme used to compare the 3-day with the 1-day cycle system.

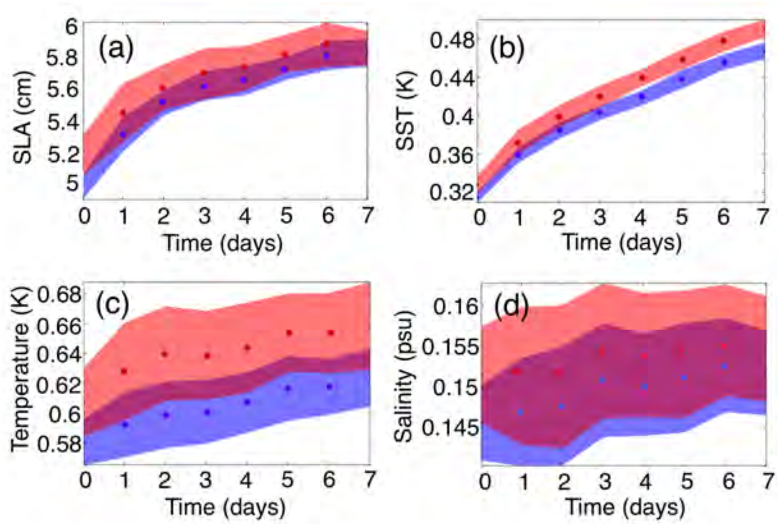

Figure 3. Global 7-day forecast innovation error statistics from series of identical base dates for (a) sea-level anomaly, (b) sea surface temperature, (c) sub-surface temperature and (d) sub-surface salinity. 1-day system shown in blue and 3-day system shown in red. The envelopes represent $\pm$ 1 standard deviation in forecast error. See Figure 2 for forecast scheme.

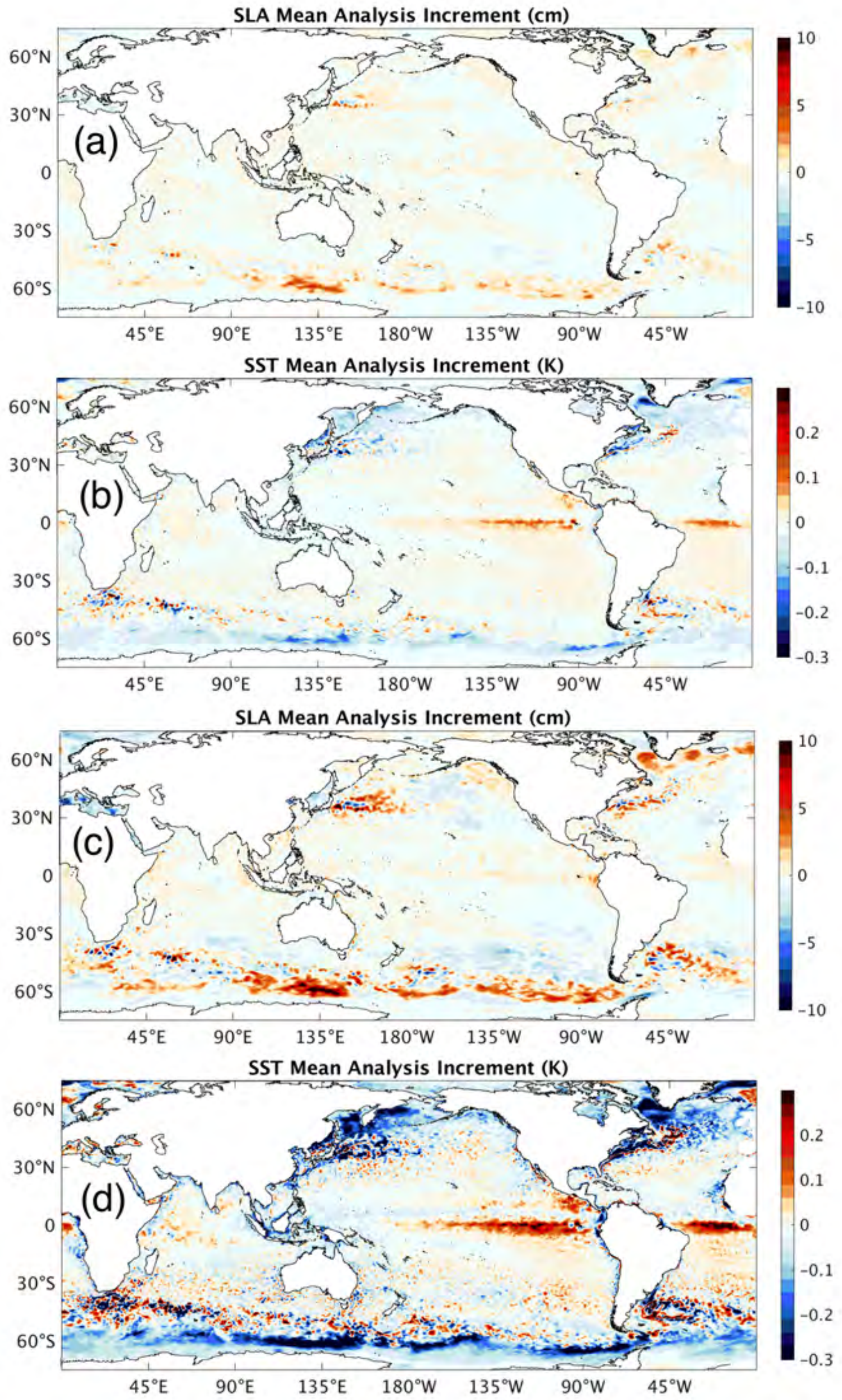

Figure 4. Mean analysis increments for sea-level anomaly (SLA) and sea surface temperature (SST) for the 1-day (a,b) and 3-day system (c,d).

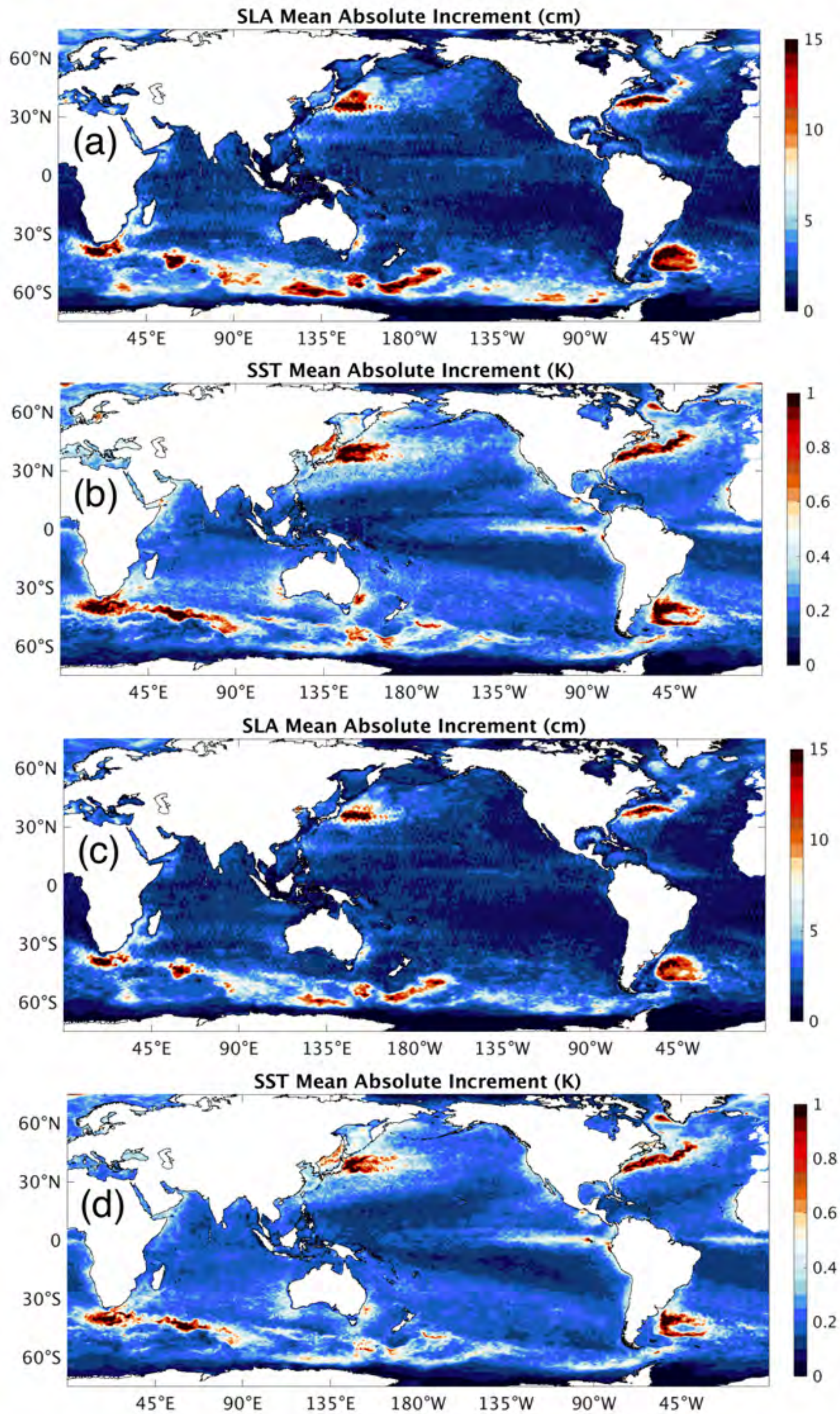

Figure 5. Mean absolute increments for sea-level anomaly (SLA) and sea surface temperature (SST) for the 1-day (a,b) and 3-day system (c,d).

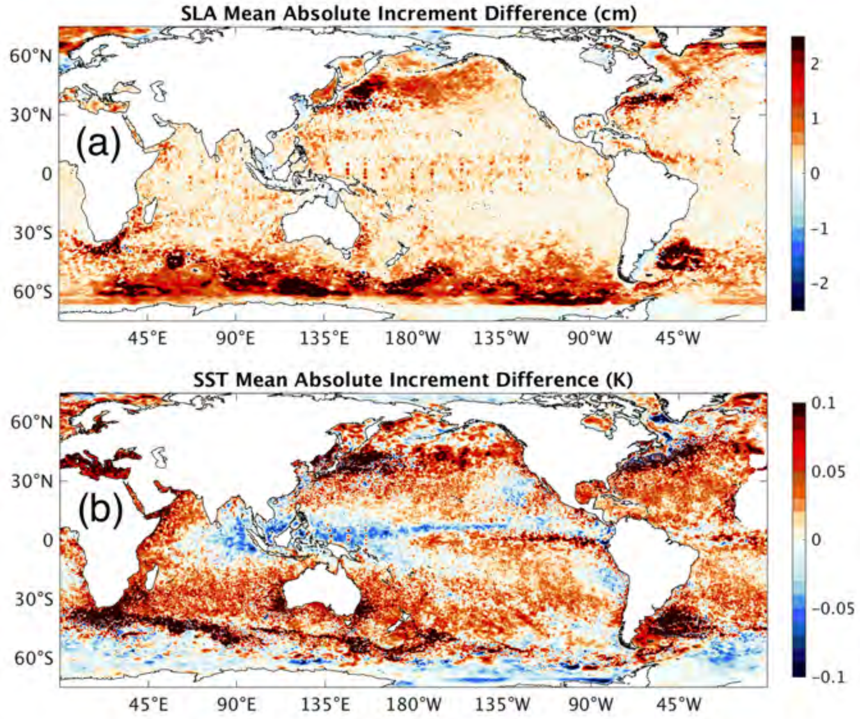

Figure 6. Difference (1-day minus 3-day) in mean absolute increment (a) for sea-level anomaly (SLA) and (b) sea surface temperature (SST).

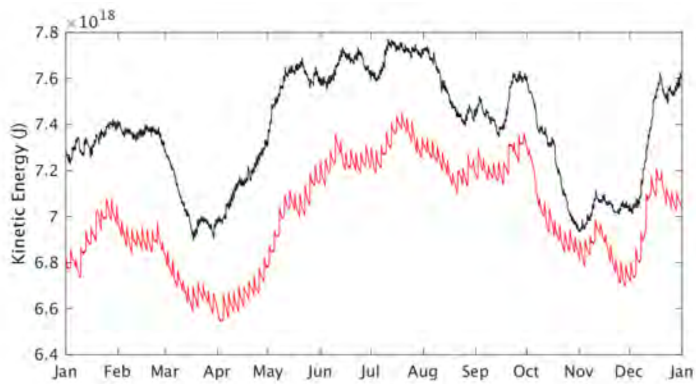

Figure 7. Total kinetic energy (Joules) for the 1-day (black) and 3-day systems (red) throughout 2013.

Table 2. Global mean and 7-day mean forecast innovation mean absolute deviation (MAD) and bias for sea-level anomaly (SLA), sea surface temperature (SST), sub-surface temperature (T) and salinity (S) from series of 44 7-day forecasts, 3-days apart from 3rd January 2013. See Figure 2 for information on how the base dates are aligned. ⋆ Total number of super-observations used to verify the 44 7-day forecasts shown. † 1-day system ‡ 3-day system.

Variable (units)	MAD†	Bias†	MAD‡	Bias‡	Observations⋆
SLA (cm)	5.51	0.0152	5.60	0.197	21272458
SST (K)	0.417	0.0151	0.435	0.0457	237176982
T (K)	0.603	-0.0979	0.616	-0.136	5276357
S (psu)	0.153	0.0349	0.155	0.0341	4974538

Table 3. As for Table 2, except for Tasman Sea region. † 1-day system ‡ 3-day system.

Variable (units)	MAD†	Bias†	MAD‡	Bias‡	Observations⋆
SLA (cm)	7.11	0.0674	7.21	0.0667	202548
SST (K)	0.478	-0.0634	0.488	-0.124	3145463
T (K)	0.573	-0.067	0.617	-0.119	48056
S (psu)	0.104	0.0674	0.098	0.0667	51144