

The impact of hybrid oceanic data assimilation in a coupled model: a case study of a tropical cyclone

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open access

Leung, T. Y. ORCID: <https://orcid.org/0000-0003-0056-284X>,
Lawless, A. S. ORCID: <https://orcid.org/0000-0002-3016-6568>,
Nichols, N. K. ORCID: <https://orcid.org/0000-0003-1133-5220>,
Lea, D. J. and Martin, M. J. (2022) The impact of hybrid
oceanic data assimilation in a coupled model: a case study of
a tropical cyclone. Quarterly Journal of the Royal
Meteorological Society, 148 (746). pp. 2410-2430. ISSN 0035-
9009 doi: 10.1002/qj.4309 Available at
<https://centaur.reading.ac.uk/105720/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1002/qj.4309>

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

RESEARCH ARTICLE

The impact of hybrid oceanic data assimilation in a coupled model: A case study of a tropical cyclone

Tsz Yan Leung^{1,2,3}  | Amos S. Lawless^{1,2}  | Nancy K. Nichols^{1,2}  |
Daniel J. Lea⁴  | Matthew J. Martin⁴ 

¹School of Mathematical, Physical and Computational Sciences, University of Reading, Reading, UK

²National Centre for Earth Observation, Reading, UK

³MetOffice@Reading, Reading, UK

⁴Met Office, Exeter, UK

Correspondence

T. Y. Leung, MetOffice@Reading, Meteorology Building, University of Reading, Whiteknights Road, Earley Gate, Reading, RG6 6ET, Berkshire, UK.
Email: tsz-yan.leung@metoffice.gov.uk

Funding information

Newton Fund, Grant/Award Number: P107915; Natural Environment Research Council, Grant/Award Number: PR140015

Abstract

Tropical cyclones tend to result in distinctive spatial and temporal characteristics in the upper ocean, which suggests that traditional, parametrisation-based background-error covariances in oceanic data assimilation (DA) may not be suitable. Using the case study of Cyclone *Titli*, which affected the Bay of Bengal in October 2018, we explore hybrid methods that combine the traditional covariance modelling strategy used in variational methods with flow-dependent estimates of the ocean's error covariance structures based on a short-range ensemble forecast. This hybrid approach is investigated in the UK Met Office's state-of-the-art system. Single-observation experiments in the ocean reveal that the hybrid approach is capable of producing analysis increments that are time-varying, more anisotropic and vertically less uniform. When the hybrid oceanic covariances are incorporated into a weakly coupled DA system, the sea-surface temperature (SST) in the path of the cyclone is changed, not only through the different specifications of background-error covariances used in the SST assimilation, but also through the propagation of subsurface temperature differences to the surface as a result of vertical mixing associated with the cyclone's strong winds. The coupling with the atmosphere then leads to a discrepancy in the cyclone's central pressure, which brings forth further SST differences due to the different representations of the cyclone's emerging cold wake.

KEYWORDS

air–sea coupling, background-error covariance modelling, data assimilation, tropical cyclones

1 | INTRODUCTION

Skilful numerical weather prediction (NWP) relies on a good data assimilation (DA) scheme that combines prior information of the forecast model with new observations

(Daley, 1991; Kalnay, 2003). Since both the prior information—often known as “background”—and the observations contain errors, some knowledge of these errors needs to be specified for the DA scheme to work well. This article focuses on the former category of errors,

that is, background errors. The covariance structure of such errors is an important topic in DA, because it is crucial in determining how information from observations is spread to the unobserved parts of the system (Bannister, 2008a). This means that the accuracy of such covariances is crucial to the success of DA schemes (Cardinali *et al.*, 2004) and therefore of forecasts that their outputs initiate.

How background-error covariances are specified traditionally depends on the assimilation scheme. In variational DA methods such as the three- and four-dimensional variational (3DVar/4DVar) schemes (Courtier *et al.*, 1998; Rabier, 2005), which are used by many NWP centres across the world, they are typically modelled under the assumption that background errors follow a multivariate Gaussian distribution (Bannister, 2008a). This Gaussian distribution is centred at zero, since model biases are often treated separately (see, e.g., Dee, 2005). In terms of its covariances, balance relationships inherent to the fluid (Bannister, 2008b; Weaver *et al.*, 2005) are exploited. Such relationships decompose the fluid flow into several dynamical components that are assumed to be mutually uncorrelated (Bannister, 2008b). Within each component, variances and spatial correlations are parametrised (Bannister, 2008b), using techniques that vary across NWP and ocean forecasting centres and models (e.g., Weaver and Courtier, 2001; Purser *et al.*, 2003; Deckmyn and Berre, 2005). The fact that physical balances are preserved by variational DA schemes makes these schemes a popular choice among NWP and ocean forecasting centres. Traditionally, the parametrisations of error variances and correlations are based largely on crude climatological estimates. While they can be improved by being allowed to vary with the local flow, it is unlikely that covariances modelled in this way can fully capture the complex nature of inhomogeneity and anisotropy of the fluid (Bannister, 2008b). This could result in suboptimal analyses in regions where the dynamics is substantially different from the climatology, such as tropical cyclones, although the implicit propagation of background-error covariances by the tangent-linear model in 4DVar could help alleviate the problem.

On the other hand, ensemble DA methods such as the ensemble Kalman filter (Evensen, 1994) estimate background-error covariances based on a short-range forecast ensemble. Unlike variational methods, ensemble methods are by their very nature flow-dependent. However, the limited ensemble size (typically $\mathcal{O}(10)$ to $\mathcal{O}(100)$) in comparison with the high dimensionality of NWP and ocean forecasting models (typically $\mathcal{O}(10^8)$ to $\mathcal{O}(10^9)$ degrees of freedom) means that analysis increments would only be able to explore a very limited part of the state space if the empirical covariances were not processed

further. This is because the theory of linear algebra dictates that the rank of a raw ensemble-estimated covariance matrix must be limited by the ensemble size (Evensen and van Leeuwen, 1996). Moreover, the undersampling could lead to “filter divergence”—the eventual collapse of the ensemble into a single trajectory (Jazwinski, 1970). In view of these shortcomings, localisation (Houtekamer and Mitchell, 2001; Hamill *et al.*, 2001) is applied to the empirically estimated covariance matrix to taper spurious long-distance correlations to zero and improve the matrix rank. Additionally, the ensemble is inflated at every assimilation cycle through an *ad hoc* procedure that injects variance (Anderson and Anderson, 1999; Houtekamer and Zhang, 2016).

There has been a growing interest over the past couple of decades in DA methods that combine features of both ensemble and variational methods (e.g., Hamill and Snyder, 2000; Lorenc, 2003). Indeed, it has been shown in the NWP context that a hybrid between the two could draw on the strengths of both and produce better analyses than either of them on its own (Clayton *et al.*, 2013; Kuhl *et al.*, 2013; Wang *et al.*, 2013). Bannister (2017) provides a review of the many hybrid formulations proposed to date. Following the nomenclature of Lorenc (2013), here we consider hybrid-3DEnVar, a variational-based scheme that is the same as 3DVar except that it allows modelled background-error covariances to be combined with localised, empirical estimates from a short-range forecast ensemble. At the UK Met Office, a four-dimensional version of this (hybrid-4DVar) has already been running operationally in the atmospheric DA system (Clayton *et al.*, 2013). Pre-operational trials of the ocean system with hybrid-3DEnVar show promising results, including a substantial improvement in global background innovation statistics compared with standard 3DVar or 3DEnVar (Lea *et al.*, 2022).

As the Met Office and other NWP centres move towards an Earth-system modelling and DA approach where the coupling between model components is expected to enable a more consistent exchange of information across interfaces (Williams *et al.*, 2017; Browne *et al.*, 2019), it is worth considering the impact of hybrid oceanic background-error covariances in a coupled framework. Coupled modelling is known to benefit the prediction of tropical cyclones (Magnusson *et al.*, 2019), since air–sea interaction is crucial to their development (Chen and Zhang, 2019; Chen *et al.*, 2021). Recent work has also shown that initialising coupled systems with coupled DA methods has the potential to improve tropical cyclone analyses and forecasts further (Chen and Zhang, 2019). Using coupled DA, therefore, it can be expected that the benefits, felt in the ocean, of introducing hybrid background-error covariances in an oceanic DA system

could also lead to better tropical cyclone predictions and hence potentially could save many lives.

To highlight the physical mechanisms through which the improved ocean state can benefit the atmosphere, a pilot case study is carried out during the active period of a tropical cyclone. This will be the focus of our article, which is structured as follows. Section 2 introduces Cyclone *Titli* (RSMC-Tropical Cyclones New Delhi, 2019), the subject of this case study. Section 3 provides a few details of the Met Office's coupled DA system, which is a weakly coupled system according to the definition of Penny *et al.* (2017). Attention is given to its oceanic component, where hybrid background-error covariances are introduced. In Section 4, results of a series of single-observation experiments within the uncoupled ocean model are presented. This enables assessment of background-error covariance structures under a range of hybrid combinations and demonstrates the time-varying nature of the hybrid covariances. Returning to the coupled model in Section 5, differences in physical features of the cyclone seen in cycled DA experiments with and without hybrid oceanic covariances are discussed, in light of the covariance structures revealed in the previous section. Finally, Section 6 summarises the article and draws a few conclusions.

2 | SYNOPSIS OF THE TROPICAL CYCLONE

The tropical cyclone chosen for this study is Cyclone *Titli*, which led to tens of deaths and brought significant damage to eastern India in October 2018. The track of *Titli* is shown in Figure 1. It formed at 0300 UTC on October 8 in the middle of the Bay of Bengal and gradually crossed the Bay on a northwesterly track over the following three days. It exhibited rapid intensification during October 10 and reached the classification of Very Severe Cyclonic Storm (118–165 $\text{km} \cdot \text{hr}^{-1}$ sustained winds) by the time of its landfall at approximately 0000 UTC on October 11, at 18.8° North, 84.5° East, near the border between the Indian states of Andhra Pradesh and Odisha. Following its landfall, the storm quickly weakened and later changed its course to northeasterly while crossing Odisha. It eventually dissipated over the Indian state of West Bengal at 0000 UTC on October 13.

Figure 2 shows various atmospheric and oceanic analysis fields in the run-up to *Titli*'s landfall. The mean sea-level pressure and wind fields in Figure 2a indicate that the cyclone was highly organised and symmetric. The sea-surface temperature (SST), which had been at least

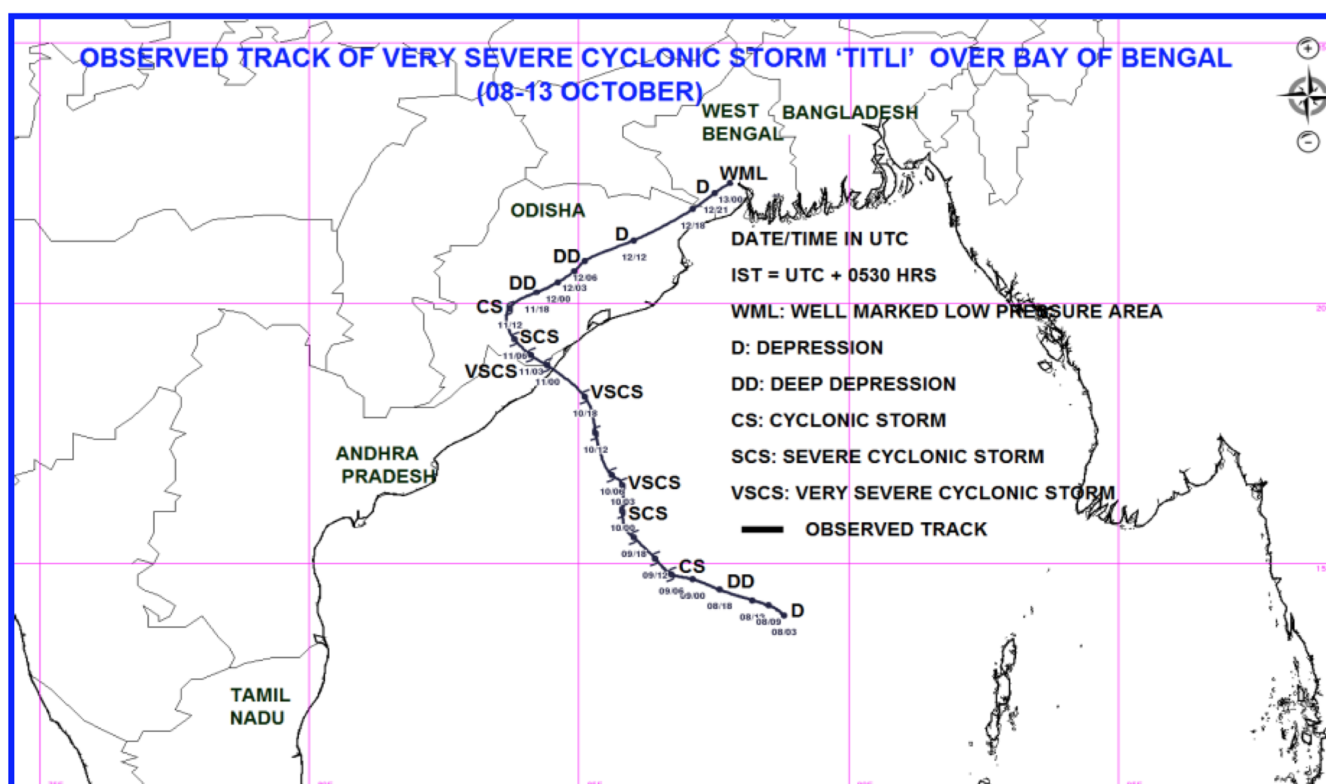


FIGURE 1 The track of Cyclone *Titli* (courtesy of the India Meteorological Department). The intensity categories are defined by maximum three-minute sustained surface winds: Depression ($31\text{--}50 \text{ km} \cdot \text{hr}^{-1}$), Deep Depression ($51\text{--}62 \text{ km} \cdot \text{hr}^{-1}$), Cyclonic Storm ($63\text{--}88 \text{ km} \cdot \text{hr}^{-1}$), Severe Cyclonic Storm ($89\text{--}117 \text{ km} \cdot \text{hr}^{-1}$), Very Severe Cyclonic Storm ($118\text{--}165 \text{ km} \cdot \text{hr}^{-1}$), Extremely Severe Cyclonic Storm ($166\text{--}220 \text{ km} \cdot \text{hr}^{-1}$), and Super Cyclonic Storm ($\geq 221 \text{ km} \cdot \text{hr}^{-1}$) [Colour figure can be viewed at wileyonlinelibrary.com]

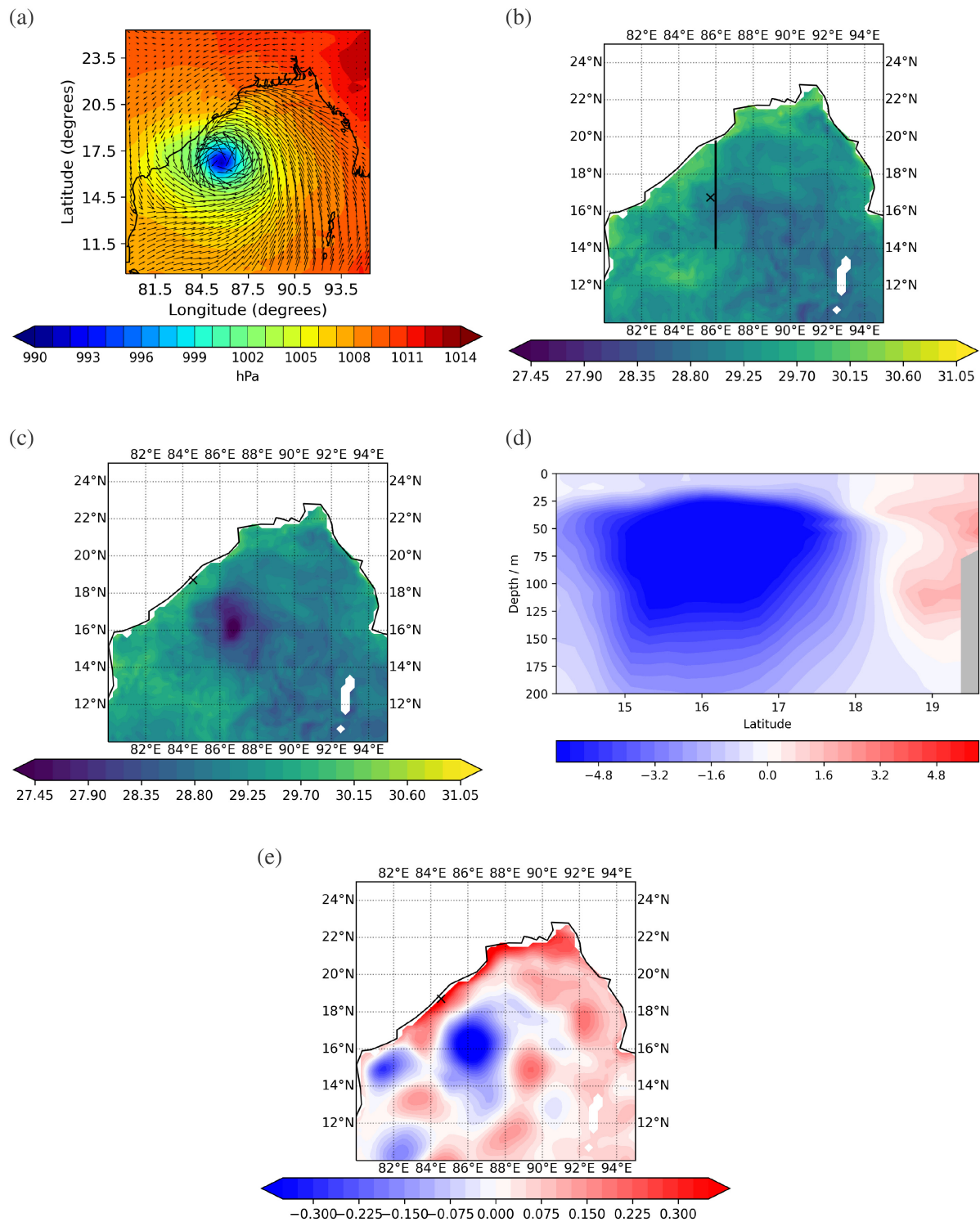


FIGURE 2 (a) Mean sea-level pressure (colours) and surface wind (arrows) analyses at 0600 UTC on October 10, 2018. An arrow of unit length on this latitude–longitude grid represents a wind speed of 10 m s^{-1} . (b, c) SST analyses ($^{\circ}\text{C}$) averaged over the 3-hr periods ending at (b) 0600 UTC on October 10, 2018 and (c) 0000 UTC on October 11, 2018, respectively. (d) Temperature anomaly analysis ($^{\circ}\text{C}$) along a meridional cross-section at 86.0°E averaged over the 3-hr period ending at 0000 UTC on October 11, 2018. The horizontal extent of the cross-section is marked by the black line in panel (b). The anomaly is relative to the 2011–2015 climatology of October. (e) SLA analysis (metres) averaged over the 3-hr period ending at 0000 UTC on October 11, 2018. All panels of this figure are taken from the Control coupled experiment described in Section 5. The black crosses in panels (b), (c), and (e) indicate the positions of the centre of the cyclone in the middle of the respective 3-hr periods [Colour figure can be viewed at wileyonlinelibrary.com]

28 °C throughout the Bay of Bengal (Figure 2b), decreased as the cyclone intensified (Figure 2c). This is associated with the cyclone's dynamical entrainment and air–sea heat exchange (Price, 1981). Beneath the ocean mixed layer resided a mesoscale cold-core eddy (Figure 2d), which had been a persistent feature in the region for about two weeks (not shown). The cooling of the sea surface in Figure 2c can therefore be seen as a temporary extension of the cold-core eddy to the surface. The increased seawater density associated with the lower temperatures of the eddy is manifest in the negative sea-level anomaly¹ (SLA) shown in Figure 2e.

3 | THE COUPLED NWP SYSTEM

A development version of the Met Office's coupled NWP system is used in this study. This development version is the Global Coupled 3 (GC3) configuration of the Unified Model (Williams *et al.*, 2017), which consists of an atmospheric component (Walters *et al.*, 2019), a land-surface component (Walters *et al.*, 2019), an oceanic component (Storkey *et al.*, 2018), and a sea-ice component (Ridley *et al.*, 2018). The atmospheric and land-surface components are considered to be “tightly coupled” in the sense that they run on the same model grid and share the same executable files; likewise, the oceanic and sea-ice components are tightly coupled. Across the air–sea interface, fields such as heat and momentum fluxes are exchanged on an hourly basis using the Ocean Atmosphere Sea Ice Soil (OASIS) coupler (Craig *et al.*, 2017). DA for this development version is weakly coupled (according to the definition of Penny *et al.*, 2017), which means that the components share the same background trajectory from an integration of the coupled model but compute analyses from the background innovations independently.

As this article is focused on the impact of a development in the oceanic DA part of the system, its incorporation into the coupled model, and the introduction of air–sea coupled DA, only these aspects of the coupled NWP system are elaborated in the following subsections. The reader is advised to consult the references provided above for a description of the other components of the system.

3.1 | NEMO–CICE and NEMOVAR

The global Forecasting Ocean Assimilation Model (FOAM: Blockley *et al.*, 2014) is the operational short-range ocean

forecasting system at the Met Office. It consists of the ocean model Nucleus for European Modelling of the Ocean (NEMO: Madec and NEMO Team, 2016) and the Los Alamos sea-ice model (CICE: Hunke *et al.*, 2015), which are tightly coupled together, as well as the variational DA system NEMOVAR (Waters *et al.*, 2015).

The current configuration of NEMO–CICE, known as ORCA025, runs on a tripolar grid with an equivalent horizontal resolution of 0.25° latitude and longitude. This translates to about 25 km in the Bay of Bengal. For the oceanic component NEMO, the state variables are potential temperature (T ; hereafter “temperature”), practical salinity (S ; hereafter “salinity”), sea-surface height (SSH), and zonal and meridional currents. Of these, SSH is a two-dimensional field, whereas the rest are three-dimensional with 75 levels resolved in the vertical. The top level is approximately 1 m thick and the vertical resolution decreases with depth. An ensemble version of FOAM has recently been developed (Lea *et al.*, 2022), consisting of 37 ensemble members. To account for model uncertainty, stochastic parametrisation is applied on all ensemble members except one (the control member). A mix of stochastic parametrisation schemes are used, and the strengths of these schemes are tuned so that the ensemble is reliable in most regions (based on our assumptions about the observation-error statistics). Details of the stochastic parametrisation are described in Storto and Andriopoulos (2021) and Lea *et al.* (2022). The ocean surface is forced by an atmospheric and land-surface ensemble (the global version of the Met Office Global and Regional Ensemble Prediction System, MOGREPS-G: Bowler *et al.*, 2008), which itself features stochastic model perturbations and an ensemble of data assimilations with perturbed observations. Further details of the ocean model configuration used in this study can be found in Storkey *et al.* (2018).

NEMOVAR takes a background ensemble forecast generated by NEMO–CICE and runs a corresponding 37-member ensemble of DAs using the incremental formulation (cf. Courtier *et al.*, 1994) of 3DVar with first guess at appropriate time (FGAT: Fisher and Andersson, 2001). Apart from the control ensemble member described above, each ensemble member generates perturbed initial conditions for the model by assimilating perturbed observation values and locations to account for observational uncertainty. The control member assimilates the same observations, but without perturbations. The assimilation cycle is 24 hr, running from 0000 UTC to 0000 UTC of the following day. Analysis increments are added to the background fields incrementally over this 24-hr window, using the incremental analysis updating (IAU) procedure of Bloom *et al.* (1996).

¹SLA is the difference between sea-surface height and the mean dynamic topography, which is fixed in time but spatially varying.

The standard 3DVar background-error covariance matrix used in NEMOVAR, \mathbf{B}_{mod} , is modelled through a balance transformation that maps the state variables to several control variables that are assumed to be uncorrelated. For the oceanic component, these control variables are temperature, unbalanced salinity, and unbalanced SSH. The general formulation of NEMOVAR also includes unbalanced horizontal currents as an additional oceanic control variable, but this is not currently used, as no velocity data are assimilated at present. Details of the balance relationships are given in Weaver *et al.* (2005). A notable change in the Met Office implementation is that the SSH–density² balance (cf. their equation 19),

$$\delta\eta_B = - \int_{-z_{\text{ref}}}^{-z_0} \frac{\delta\rho}{\rho_0} dz, \quad (1)$$

defines z_0 to be the mixed-layer depth instead of $z_0 = 0$. (Here, z is the vertical coordinate, z_{ref} is a reference depth, ρ_0 is a reference density, and $\delta\rho$ and $\delta\eta_B$ are analysis increments for the density and balanced SSH, respectively.) This change effectively prevents SSH and mixed-layer variables from influencing each other through the balance relationship.

Waters *et al.* (2015) describe the parametrisations used to model the standard deviations and correlation length-scales of \mathbf{B}_{mod} for each control variable. Generally speaking, these parametrisations are rather crude and are largely based on seasonally varying climatological estimates, though they also depend on the local mixed-layer depth, as the dynamics within the mixed layer are very different from the dynamics in the deeper ocean (Schiller and Ridgway, 2013). The correlation length-scales are used to define approximately Gaussian spatial correlations through an implicit diffusion operator (Weaver and Courtier, 2001; Weaver and Mirouze, 2013; Weaver *et al.*, 2016), which now allows multiple length-scales to be specified (Mirouze *et al.*, 2016). For temperature and unbalanced salinity, two length-scales are used in the horizontal direction: a length-scale based on the Rossby radius of deformation (about 87.5 km, or equivalently 0.79° latitude, at the point in the Bay of Bengal chosen for the single-observation experiments in Section 4), and a fixed scale of 444 km (equivalent to 4° latitude). As for SSH, the length-scales for the balanced part of its increments are associated with baroclinic errors and are determined by the temperature and salinity length-scales just described. The unbalanced part of the SSH increments is associated with barotropic errors (Waters *et al.*, 2015), which propagate very quickly and therefore have a long scale of 444 km.

The new hybrid DA scheme (hybrid-3DEnVar) for the oceanic component uses a different background-error covariance matrix $\mathbf{B} = (1 - \alpha)\mathbf{B}_{\text{mod}} + \alpha\mathbf{B}_{\text{ens}}$, which is a weighted combination of \mathbf{B}_{mod} and a matrix of localised error covariances estimated by the ensemble of the day, \mathbf{B}_{ens} . Here, α is a number between 0 and 1 that defines the weighting between the two matrices. It should be noted that the 3DVar scheme described above corresponds to the special case where $\alpha = 0$. The matrix \mathbf{B}_{ens} is computed by taking the elementwise product between a localisation matrix and the empirical covariance matrix, but in practice this is achieved through an equivalent sequence of matrix–vector multiplications (Weaver *et al.*, 2018). As in Lea *et al.* (2022), univariate localisation is used in this study, meaning that localisation is only applied to individual control variables and does not change the prescribed balance relationships. The ensemble correlation information between different control variables is therefore discarded, and the localisation matrix becomes block-diagonal. Each block of the localisation matrix is defined through the same family of approximately Gaussian diffusion operators as the one used to construct spatial correlations in \mathbf{B}_{mod} . However, there is no localisation in the vertical. The horizontal localisation length-scale is set to be twice the Rossby radius of deformation, which is about 175 km (equivalent to 1.58° latitude) at the point in the Bay of Bengal chosen for the single-observation experiments in Section 4. This is between the short and long length-scales of \mathbf{B}_{mod} at that location (87.5 and 444 km, respectively).

Observations used by the oceanic component of NEMOVAR include remotely sensed SST and SLA, as well as in situ temperature and salinity profiles such as those from Argo floats. SST data from buoys, surface drifters, and ships are also used. In terms of observational coverage, since most of the SST measurements come from infrared satellite instruments (Lea *et al.*, 2014), the availability of observed SST data is linked to the cloudiness of the region: SST observations are often spatially dense in clear-sky regions, whereas cloudier regions, relying on other SST data sources, are less observed. On the other hand, SLA can only be measured directly beneath the tracks of satellite altimeters, so observations follow satellite trajectories. Profile observations measure temperature and salinity beneath the ocean surface, usually to 2000 m depth, but are horizontally very sparse. On an average day, only two to three observation profiles are available across the Bay of Bengal. Lea *et al.* (2014) demonstrated the complementarity of these different observation types. Broadly speaking, temperature and salinity profiles set the large-scale ocean density structure, whereas SLA observations constrain the mesoscale and SST observations constrain the temperature within the mixed layer.

²Density is a diagnostic variable derived from temperature and salinity through the ocean's equation of state (see Equation 2 in Section 4.2.2).

3.2 | Incorporation of NEMOVAR into the coupled data assimilation system

At the Met Office, the oceanic and sea-ice model NEMO–CICE described in Section 3.1 is coupled to the atmospheric and land-surface model UM–JULES (which is named after the respective component models, the Unified Model (Cullen, 1993; Brown *et al.*, 2012) and the Joint UK Land Environment Simulator (Best *et al.*, 2011; Clark *et al.*, 2011)). UM–JULES runs on a latitude–longitude grid. For this study, a lower-resolution version of UM–JULES than operational forecasts is used, namely N320, which has approximate zonal and meridional resolutions of 60 and 40 km respectively in the Bay of Bengal. The OASIS coupler (Craig *et al.*, 2017) is responsible for the hourly exchange of information across the air–sea interface.

In light of the faster time-scales of physical processes in the atmosphere than in the ocean (Bauer *et al.*, 2015), the coupled model runs on the same assimilation cycles as UM–JULES. These cycles are six-hourly: the nominal analysis times are 0000, 0600, 1200, and 1800 UTC each day, and are therefore more frequent than the assimilation cycles of FOAM by a factor of four. Each cycle assimilates observations over the 6-hr period centred at the nominal analysis time. Because of the reduction in cycle length of the oceanic and sea-ice components, the background-error covariance matrix of NEMOVAR is due to be remodelled to represent six-hourly forecast errors instead of daily ones. However, since oceanic and sea-ice fields vary slowly, the background-error covariance matrix based on daily forecast errors described in Section 3.1 is considered to be a reasonable approximation for the purposes of this study.

DA for the atmospheric component uses an ensemble of hybrid-4DVars, that is, an ensemble of four-dimensional variational assimilations with hybrid background-error covariances (Clayton *et al.*, 2013). For the land-surface component, the simplified extended Kalman filter is used (Gómez *et al.*, 2020). A weakly coupled DA system using NEMOVAR and the atmospheric and land-surface DA systems will be used in the experiments in Section 5. The weak coupling means that NEMOVAR and the atmospheric and land-surface DA system share the same background information generated by the coupled model, but run the assimilations independently. The analyses for different components are then combined to initialise the coupled forecast, and this forecast provides the background information for the next assimilation cycle. The IAU time period for the oceanic and sea-ice components of the coupled model is set to be the first half of the 6-hr assimilation window, so that the analysis increments are added fully by the nominal analysis time.

4 | SINGLE-OBSERVATION EXPERIMENTS IN THE OCEAN

We first seek to understand how hybrid background-error covariances can change analysed ocean fields, by means of single-observation experiments within the context of FOAM. Assimilating a single direct synthetic observation of a state variable allows the underlying background-error covariance structure to be revealed, since the analysis increment field is proportional to the field of background-error covariances between the observed state variable and all state variables (Lawless, 2013). By comparing covariance structures directly across various hybrid weightings α , we will examine in what ways flow dependence can be incorporated into the DA system through hybridising the background-error covariance matrix. The methodology of the experiments will be discussed in greater detail in Section 4.1, and Section 4.2 will report on the results.

4.1 | Experimental design

To illustrate the time-varying nature of the ensemble-based covariances, we perform two sets of single-observation experiments: (a) the observation being assimilated before the passing of the cyclone, and (b) the observation being located near the centre of *Titli's* cold wake. The point 86.7°E, 16.2°N and the date October 11, 2018 are chosen for the second set of experiments, as the cold wake is strongest there (see Figure 2c). For fair comparison, the first set of experiments assimilate observations at the same location but on an earlier date of October 9, 2018. The centre of the cyclone is at about 170 km south of this point at the beginning of that day when the ensemble covariances for \mathbf{B}_{ens} are computed.

Each set of experiments consists of 12 separate assimilation runs, a combination of four values of the hybrid weight α (0.0, 0.2, 0.8, and 1.0) and three types of observation (SST, SLA, and temperature at a depth of 100 m). The runs are performed under identical settings except for α and/or the observation. A separate 37-member ensemble of NEMO–CICE that runs cycled assimilations using the standard 3DVar–FGAT scheme is used to determine the parametrisations for \mathbf{B}_{mod} and to construct \mathbf{B}_{ens} . This ensemble has been running for more than seven months before the date of Cyclone *Titli* and is fully spun up when the cyclone hits. A selection of this ensemble's statistics for temperature on both October 9 and 11 is shown in Figure 3. It is evident that the presence of the cold wake on October 11 changes the ensemble's temperature spread substantially, which, as we shall see, has a significant impact on how the analysis responds to observations when \mathbf{B}_{ens} is

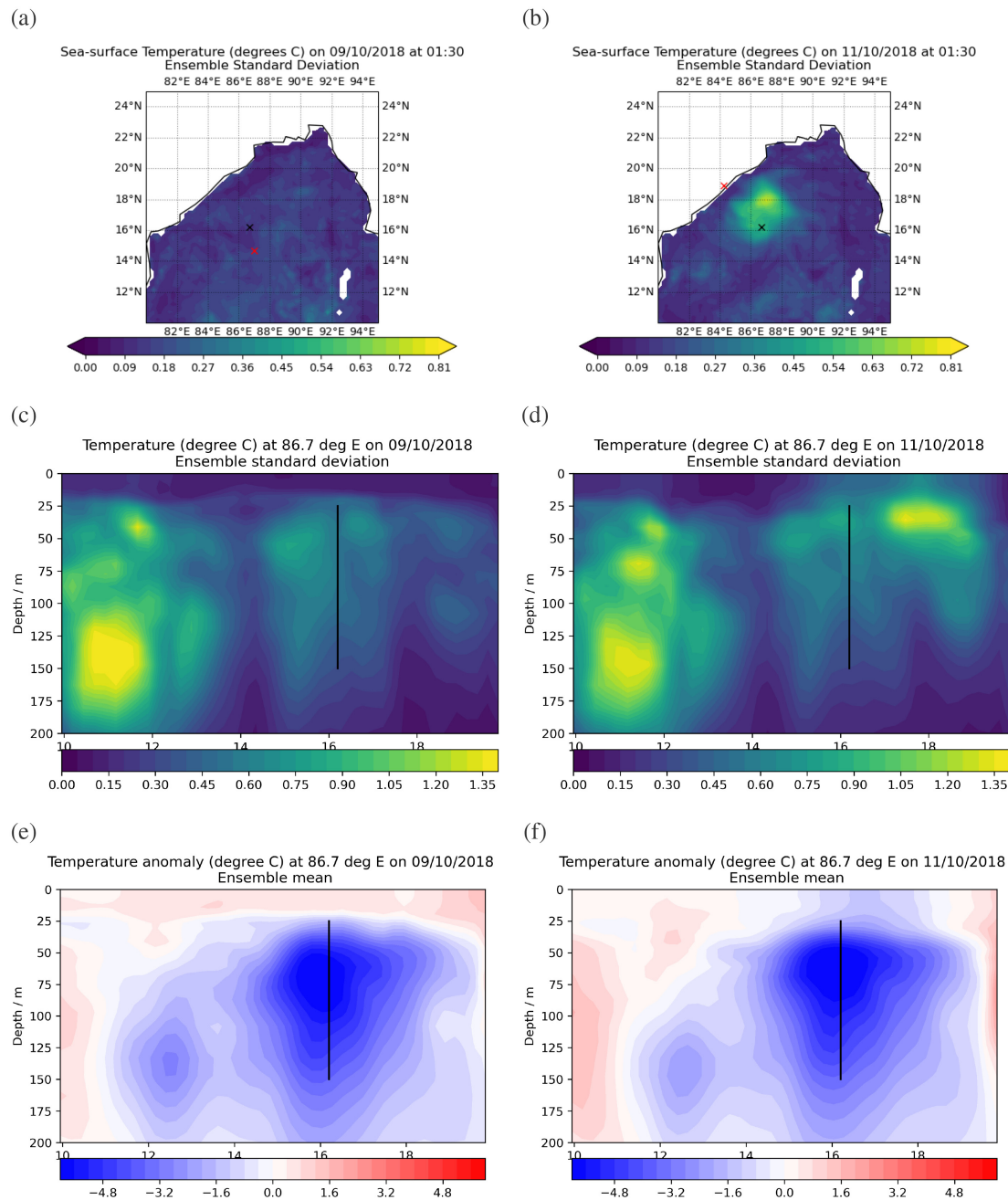


FIGURE 3 (a,b) Ensemble standard deviation of SST, (c,d) ensemble standard deviation of temperature along a meridional cross-section at 86.7°E, and (e,f) ensemble mean of temperature anomaly along the same cross-section, on (a,c,e) October 9, 2018 and (b,d,f) October 11, 2018. The surface fields (a,b) are temporally averaged over 0000–0300 UTC of the respective day, whereas the cross-sectional fields (c–f) are temporally averaged over the whole UTC day. The anomaly is relative to the 2011–2015 climatology of October. The positions of the centre of Cyclone *Titli* at 0130 UTC on the respective day (the centre of the averaging window) are marked by the red crosses in panels (a) and (b). The horizontal position of the observation for the single-observation experiments (86.7°E, 16.2°N) is marked by the black crosses in panels (a) and (b), and by the black vertical lines in panels (c–f) [Colour figure can be viewed at wileyonlinelibrary.com]

given a heavy weight. Moreover, a comparison between Figure 3c,e (or Figure 3d,f) shows that local maxima of the ensemble spread are generally associated with not the centres but the edges of oceanic eddies. The larger spread is therefore a result of the ensemble members disagreeing on eddy positioning.

The background innovations are 0.5°C for the SST observation, 0.08 m for the SLA observation, and 1.0°C for the temperature observation 100 m below the surface, which are of the same order as the global root-mean-square values of the innovations for the respective variables. The choice of 100-m depth is motivated by

the fact that it usually falls within the thermocline and, since the parametrisations for \mathbf{B}_{mod} depend greatly on the mixed-layer and thermocline depths, one may expect that the assimilation increments will be sensitive to α when the observation is placed there.

4.2 | Results

4.2.1 | An SST observation

The balance relationships of \mathbf{B}_{mod} assume that the effect of assimilating SST observation(s) is limited to modulating near-surface temperatures. This is because the design of its vertical correlation length-scales effectively disables propagation of surface information into the deeper ocean (Waters *et al.*, 2015), and the temperature–salinity balance is turned off within the mixed layer (Weaver *et al.*, 2005). For this reason we shall restrict our discussion to the SST increment, which is shown in Figure 4 for both dates and the various values of α .

It is clear from the figure that the magnitude of the increment decreases as the weight given to \mathbf{B}_{ens} increases. This is because of a large discrepancy between the standard deviations in \mathbf{B}_{mod} and \mathbf{B}_{ens} . At the grid point closest to the observation location, the SST background-error standard deviation in \mathbf{B}_{mod} is approximately 0.8 °C (not shown), compared with just about 0.1 °C in \mathbf{B}_{ens} (Figure 3a) on October 9, which is significantly smaller than the observation-error standard deviation (about 0.6 °C). This leads to the lack of any substantial SST increment in Figure 4g when \mathbf{B}_{ens} is given full weight. The discrepancy is considerably smaller on October 11, with the standard deviation being close to 0.5 °C in \mathbf{B}_{ens} . The larger ensemble spread is due to the effect of the cyclone's passing, as Figure 3b suggests. The fact that the background- and observation-error standard deviations have comparable magnitudes on October 11 prevents the apparent vanishing of the SST increment in the $\alpha = 1.0$ experiment that day (Figure 4h).

Figure 4 also demonstrates the effect of using a dual correlation length-scale in \mathbf{B}_{mod} and a separate localisation length-scale in \mathbf{B}_{ens} . In Figure 4a,b, the closely packed contours near the point of observation suggest a decorrelation of modelled background errors over short distances associated with the length-scale equal to the Rossby deformation radius, which is about 87.5 km (equivalent to 0.79° latitude) at that location. On the other hand, the more separated contours further afield show the effect of the longer, 444-km (equivalent to 4° latitude) length-scale. Since \mathbf{B}_{ens} has only one localisation length-scale, which is twice the Rossby radius, it is no surprise that the increment for the $\alpha = 1.0$ experiment on October 11 is concentrated

within about 2° latitude and longitude of the observation location (Figure 4h), although the structure of ensemble correlations and the spatial variation of the \mathbf{B}_{ens} standard deviations (Figure 3b) also contribute to this. The fact that such a localisation length-scale is shorter than the 444-km length-scale of \mathbf{B}_{mod} provides another explanation for the reduction of the increment's spatial extent as α increases.

The spatial structure of the increment remains more or less isotropic as α is varied. This is due to the relative uniformity of \mathbf{B}_{ens} standard deviations within the immediate vicinity of the observation location, within the localisation radius. However, the flow dependence and anisotropy associated with \mathbf{B}_{ens} are not obscured completely. For example, the southern half of the outermost contour in Figure 4f bears some resemblance to the contours of the ensemble standard deviation plot in Figure 3b.

4.2.2 | An SLA observation

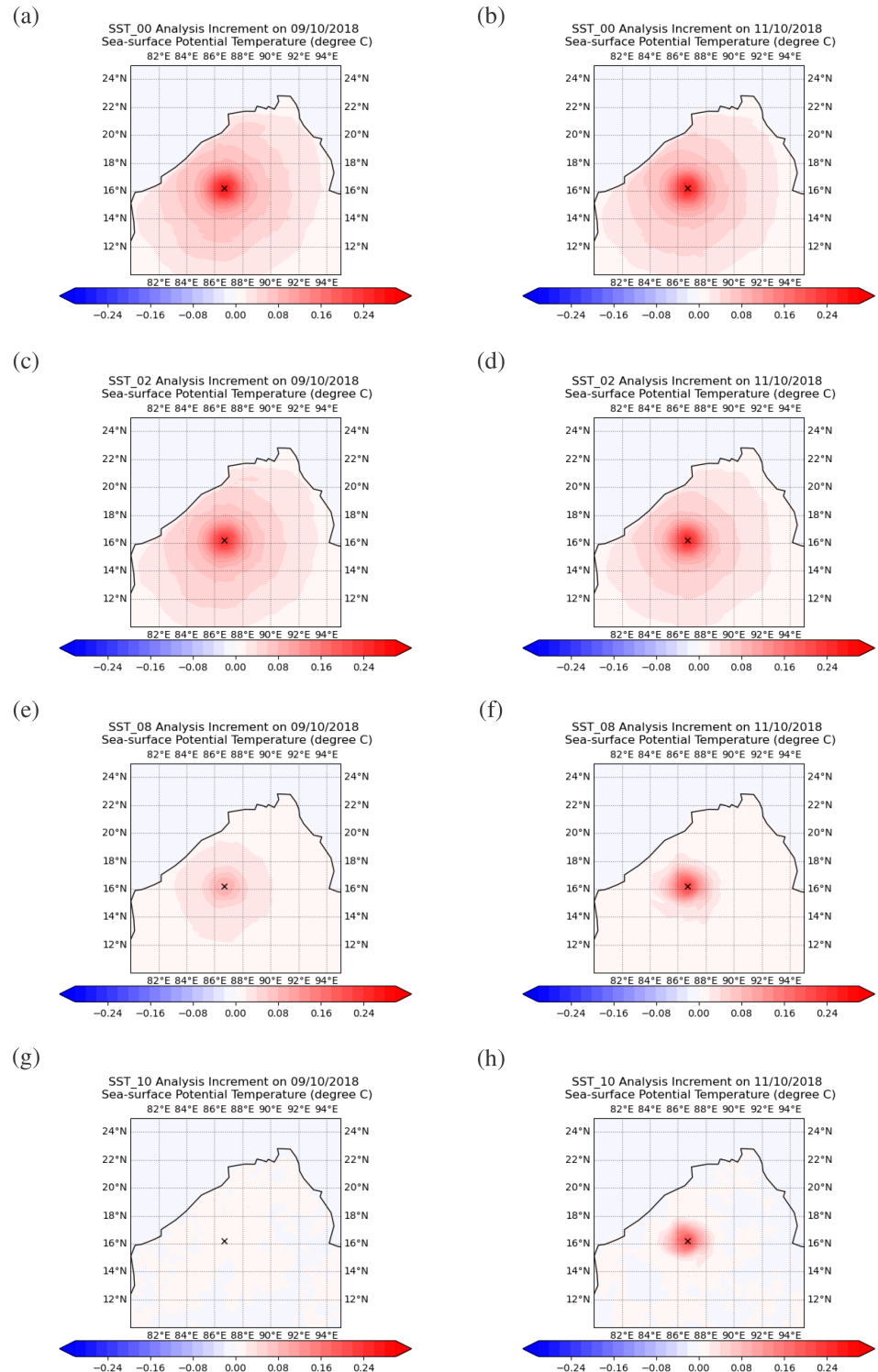
The assimilation of SLA observations is associated with the lifting and lowering of water columns (Cooper and Haines, 1996). This causes the density to change in a way described by the linearised balance given in Equation 1. Since the z_0 in that equation is set to be the mixed-layer depth, SLA observations can have an impact on the density in the deeper ocean only, but not within the mixed layer. Density increments can then be related to temperature and salinity increments (δT and δS respectively) through the linearised equation of state (Weaver *et al.*, 2005):

$$\delta\rho = -f(T, S) \delta T + g(T, S) \delta S, \quad (2)$$

where $f(T, S)$ and $g(T, S)$ are non-negative linearisation coefficients. From Equations 1 and 2, it can be deduced that an increase in SSH is associated with an increase in subsurface temperature and a decrease in subsurface salinity, and vice versa.

Figure 5 shows the temperature increment along a meridional cross-section through the location of the observation, for the different values of the weighting coefficient α . For the October 9 experiments, it can be seen that the overall magnitude of the increment decreases as α increases. This can again be attributed to a discrepancy in the standard deviations of \mathbf{B}_{mod} and \mathbf{B}_{ens} . It is also interesting to see in the $\alpha = 0.8$ and 1.0 experiments (Figure 5e,g), where the ensemble is given more weight in the background-error covariance matrix, that the subsurface temperature increments are tilted and the maximum temperature increment does not fall directly beneath the observation point. The increments are more concentrated to the south of the observation than to the north. This can be related to the fact that the ensemble has a larger spread

FIGURE 4 SST analysis increments in response to an SST observation at 86.7°E, 16.2°N on (a,c,e,g) October 9, 2018 and (b,d,f,h) October 11, 2018, for $\alpha =$ (a,b) 0.0, (c,d) 0.2, (e,f) 0.8, and (g,h) 1.0. The location of the observation is marked by the black crosses [Colour figure can be viewed at wileyonlinelibrary.com]



to the south of the observation (Figure 3c). However, the latitude of the maximum increment (about 15.7°N) does not coincide with the latitude of the local maximum of the ensemble spread (about 15.2°N), despite there being some resemblance between the contours of these fields. This is because the magnitude of the increment associated with \mathbf{B}_{ens} is not only determined by the ensemble's standard deviation, but also influenced by the structure

of ensemble correlations, as well as the horizontal localisation that makes the correlations decay away from the observation point.

On October 11, the magnitude of the temperature increment is larger when \mathbf{B}_{ens} is given more weight (Figure 5b,d,f,h). This shows that the temperature variances in \mathbf{B}_{ens} are likely to be larger than those in \mathbf{B}_{mod} somewhere along the water column in the vicinity of the

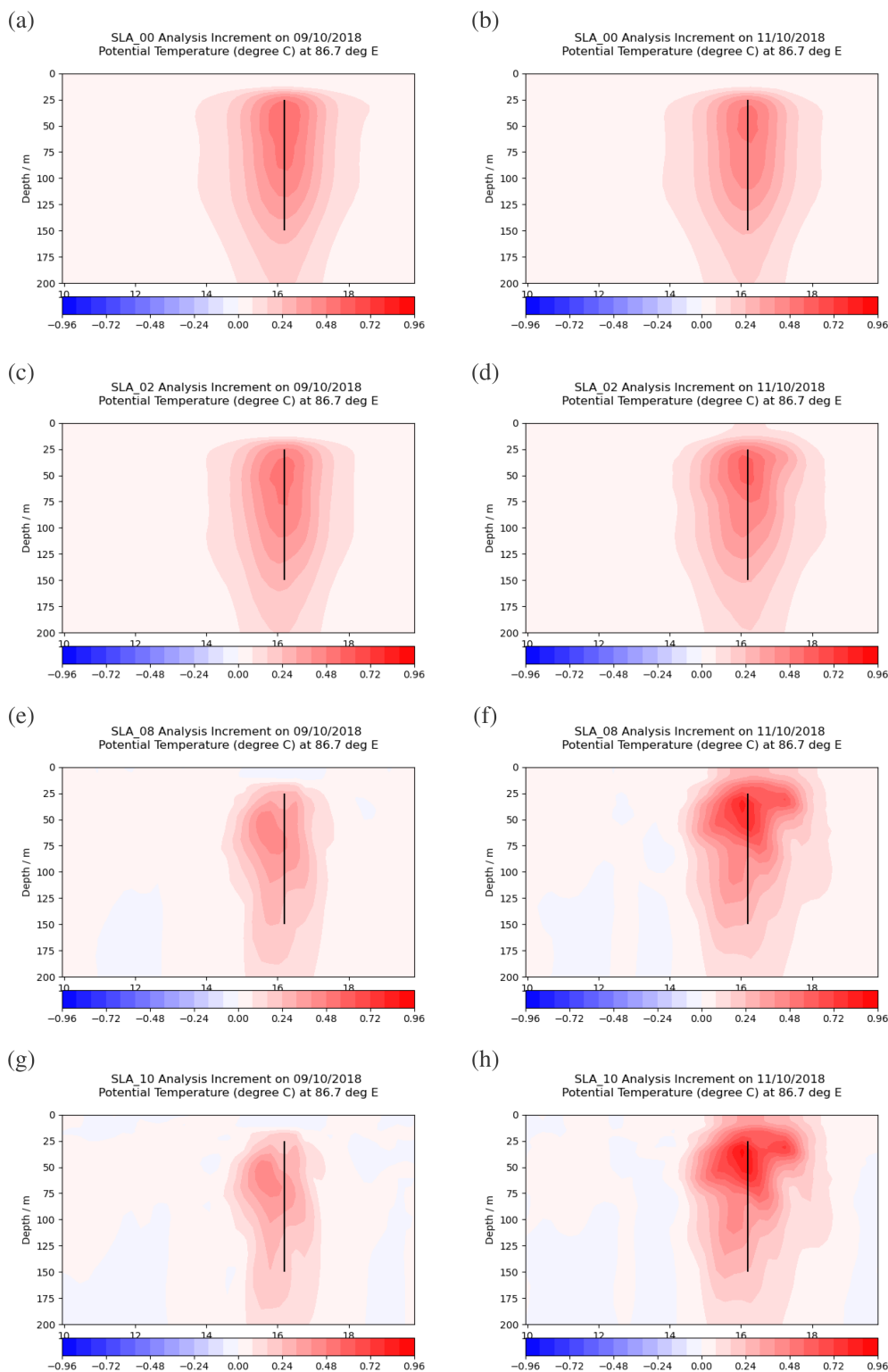


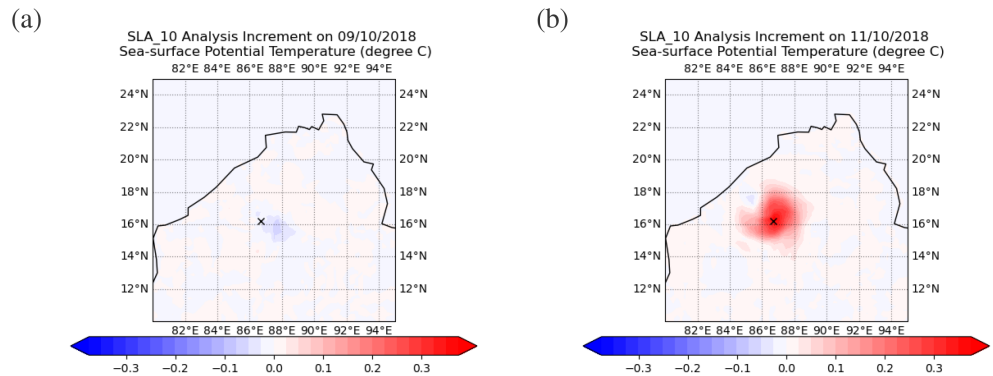
FIGURE 5 Temperature increments along a meridional cross-section at 86.7°E in response to an SLA observation at 86.7°E, 16.2°N on (a,c,e,g) October 9, 2018 and (b,d,f,h) October 11, 2018, for $\alpha =$ (a,b) 0.0, (c,d) 0.2, (e,f) 0.8, and (g,h) 1.0. The latitude of the observation is marked by the black vertical lines [Colour figure can be viewed at wileyonlinelibrary.com]

observation location. Also, the magnitude of the increment associated with B_{ens} is larger on October 11 than on October 9 (Figure 5g,h), due to the larger ensemble spread in the region (Figure 3c,d). The large spread, together with the ensemble correlations and lack of vertical localisation, allows the increment to propagate to the ocean surface, even in experiments where B_{ens} is given a weight as small

as 0.2 (Figure 5d). The separation between the mixed layer and the deeper ocean, assumed in the design of B_{mod} , no longer holds in these experiments when a tropical cyclone has just induced a cold wake in the region.

In Figure 5g,h, negative increments appear in regions beyond the localisation radius in the $\alpha = 1.0$ experiment on both dates. This indicates that negative temperature

FIGURE 6 SST increments in response to an SLA observation at 86.7°E, 16.2°N on (a) October 9, 2018 and (b) October 11, 2018, for $\alpha = 1.0$. The location of the observation is marked by the black crosses [Colour figure can be viewed at wileyonlinelibrary.com]



correlations exist between those regions and parts of the water column beneath the point of the SLA observation. While these negative correlations could simply be an artefact of sampling error, Figure 3e,f suggests that the local peak of temperature spread at 12°–14°N, 100–200 m depth in Figure 3c,d is associated with the edge of a secondary cold-core eddy extended southwards from the main one centred at about 16°E. It is likely that the negative ensemble correlations and increments arise from a disagreement among ensemble members over the positioning of either or both of these eddies.

The contrasting behaviour of \mathbf{B}_{ens} between the October 9 and 11 cases can be seen better in Figure 6, which shows the surface temperature increment when $\alpha = 1.0$. On October 9 before the cyclone passes, this increment is very weak and is probably no more than noise arising from the lack of vertical localisation. However, with the presence of the cold wake on October 11, the surface increment becomes substantial. The horizontal structure of the surface increment can be associated with the structure of the ensemble variance in Figure 3b.

Figure 7 shows the corresponding SSH increments. It is clear that the increment is increasingly anisotropic as \mathbf{B}_{ens} is given more weight, on both October 9 and 11. However, unlike the case of temperature increments, the shape of the \mathbf{B}_{ens} -based SSH increment does not seem to bear any resemblance to the ensemble's SSH spread (not shown). This is because temperature, and not SSH, is the lead variable in the balance relationships (which are used in both \mathbf{B}_{mod} and \mathbf{B}_{ens}). The SSH increment is the aggregated effect of the density increment across many vertical levels (Equation 1), plus a contribution associated with the unbalanced component of SSH.

4.2.3 | A subsurface temperature observation

Finally, we consider the effect of assimilating a single temperature observation 100 m below the surface. Information from observations below the mixed layer such

as this cannot propagate into the mixed layer and the ocean surface by virtue of \mathbf{B}_{mod} alone, because of the mixed-layer dependence of the parametrisation of vertical correlation length-scales (Waters *et al.*, 2015). Indeed, the cross-sectional temperature increments in Figure 8 show that \mathbf{B}_{ens} can also capture robustly the separation between the mixed layer and the waters beneath it on October 9. Since there is no vertical localisation in \mathbf{B}_{ens} , the lack of increment in the mixed layer can most likely be explained by the smallness of ensemble spread in that part of the ocean (Figure 3c), together with the structure of ensemble correlations. On the other hand, the increment on October 11 associated with \mathbf{B}_{ens} reaches to the surface, thanks to the large ensemble spread associated with the cold wake. Indeed, the point of largest increment in the $\alpha = 1.0$ experiment on October 11 (Figure 8h) is not anywhere close to the point of the observation, but instead in a region of large temperature variance in the ensemble (Figure 3d).

It is striking to see a significant variation in the horizontal and vertical extents of the increments as α is varied. Vertically, the impact of the observation can be felt further away in experiments where \mathbf{B}_{ens} is given more weight, because of the absence of vertical localisation in \mathbf{B}_{ens} and the short vertical correlation length-scales implied by \mathbf{B}_{mod} . However, the latter should be interpreted with some caution, since in practice profiles of observations are assimilated throughout a water column, rather than single isolated observations, which could mean that the vertical correlation model of \mathbf{B}_{mod} might not be well-tuned.

One can see the effect of the different horizontal correlation/localisation length-scales in \mathbf{B}_{mod} and \mathbf{B}_{ens} in Figure 8. The Rossby-radius-dependent correlation length-scale of \mathbf{B}_{mod} (about 87.5 km, or equivalently 0.79° latitude) is reflected in the closely packed contour lines in Figure 8a,b close to the point of observation. The longer length-scale of \mathbf{B}_{mod} (444 km, or equivalently 4° latitude) enables a response at locations further away, beyond the region of influence of \mathbf{B}_{ens} which has a localisation length-scale of about 175 km (equivalent to 1.58° latitude).

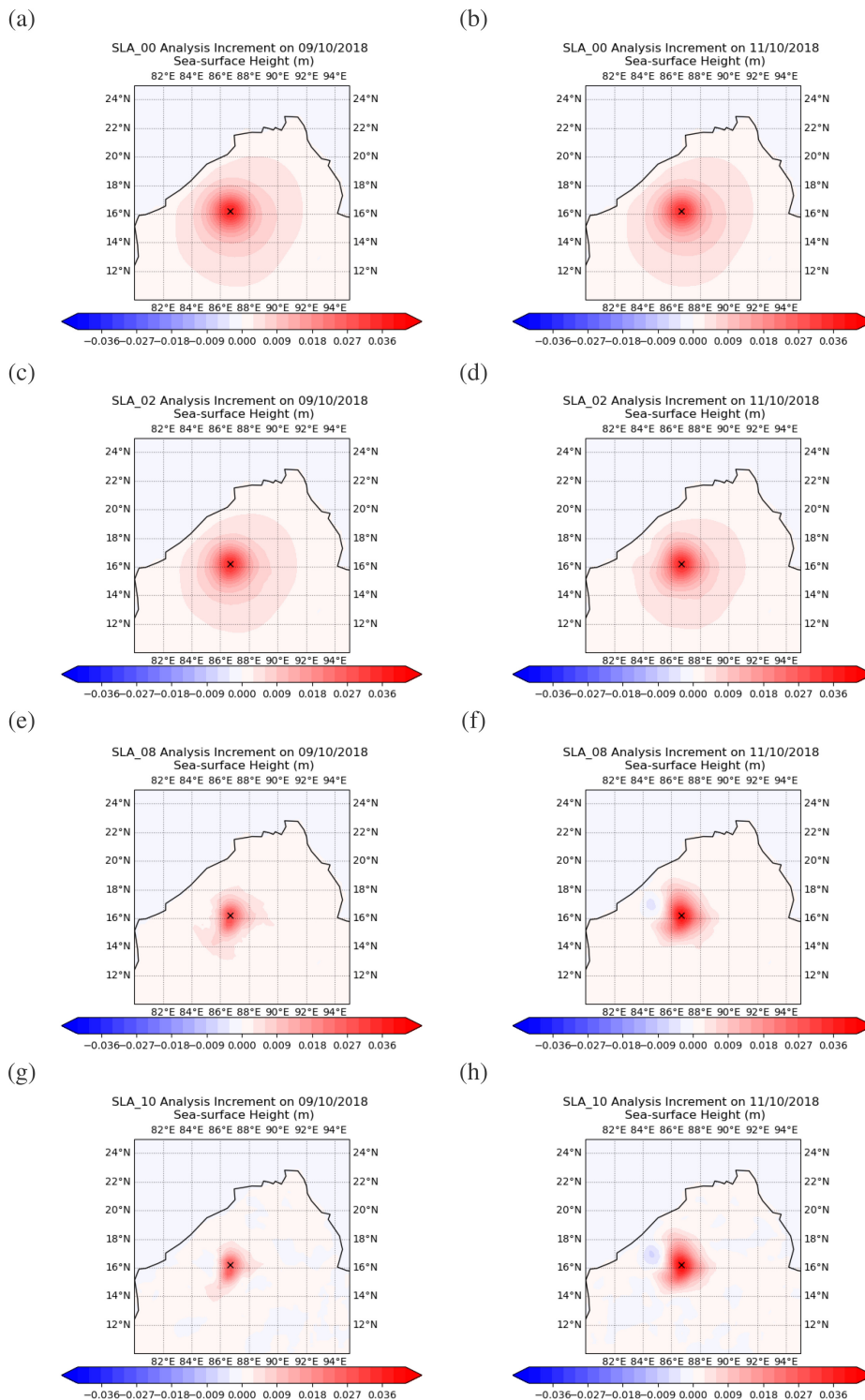


FIGURE 7 SSH increments in response to an SLA observation at 86.7°E, 16.2°N on (a,c,e,g) October 9, 2018 and (b,d,f,h) October 11, 2018, for $\alpha =$ (a,b) 0.0, (c,d) 0.2, (e,f) 0.8, and (g,h) 1.0. The location of the observation is marked by the black crosses [Colour figure can be viewed at wileyonlinelibrary.com]

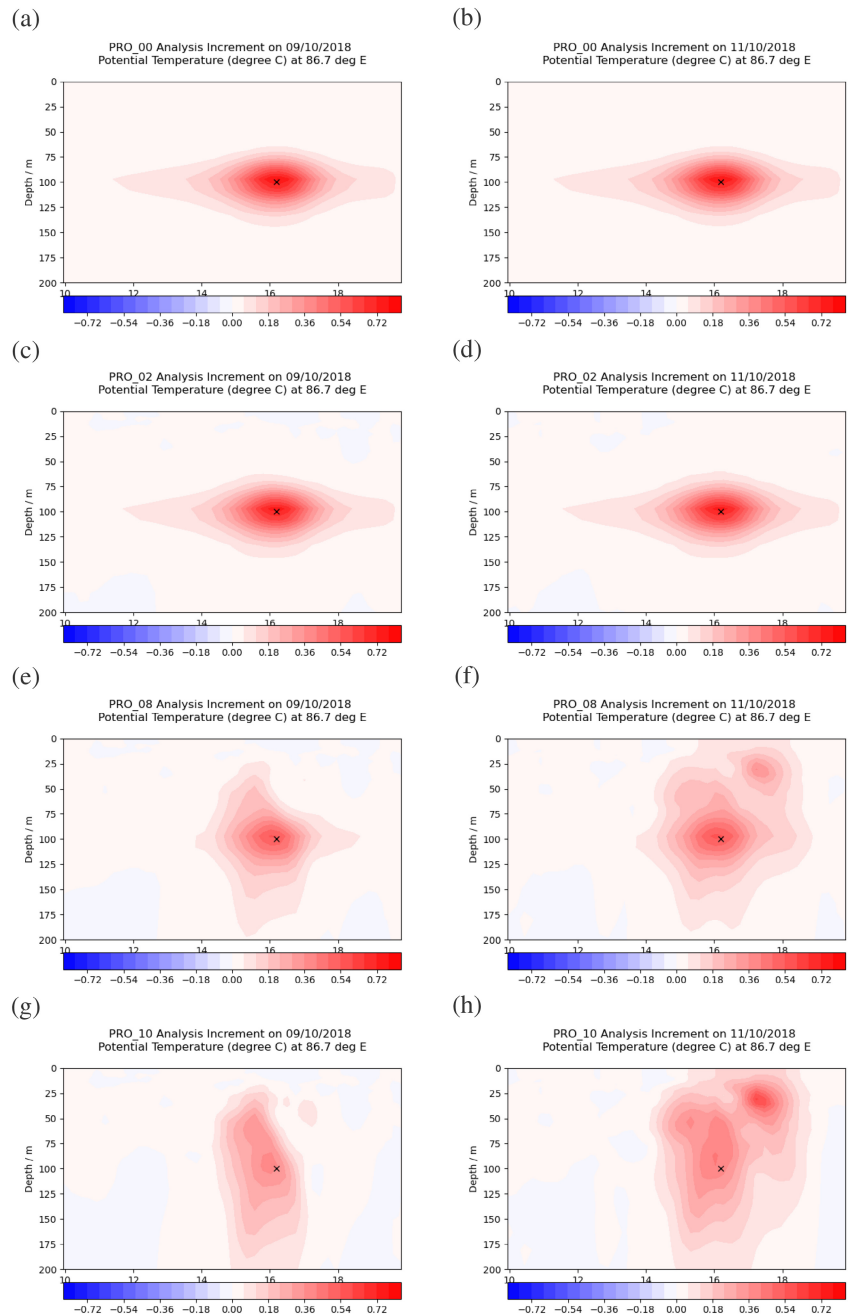
The subsurface temperature increments in large- α experiments (Figure 8e–h) are tilted, with similar characteristics to the corresponding increments in response to an SLA observation (Figure 5c,d). Like Figure 7, the SSH response to the subsurface temperature observation (not shown) is also anisotropic in these experiments, although the magnitudes are smaller here.

5 | IMPACTS OF HYBRID OCEANIC COVARIANCES IN THE COUPLED SYSTEM

5.1 | Experimental design

Having examined the structures of hybrid background-error covariances, we turn our attention to the coupled

FIGURE 8 Temperature increments along a meridional cross-section at 86.7°E in response to a temperature observation 100 m below the surface at 86.7°E, 16.2°N on (a,c,e,g) October 9, 2018 and (b,d,f,h) October 11, 2018, for $\alpha =$ (a,b) 0.0, (c,d) 0.2, (e,f) 0.8, and (g,h) 1.0. The location of the observation is marked by the black crosses [Colour figure can be viewed at wileyonlinelibrary.com]



DA system and investigate how such hybridisation in the ocean changes the representation of Tropical Cyclone *Titli* through cycling of the DA system. Two cycled assimilation experiments are considered. They are run under identical settings, except that the weight α given to \mathbf{B}_{ens} in the oceanic component is 0.0 in one of the experiments (“Control”) and 0.8 in the other (“Hybrid”). Hence, only the impact of oceanic DA changes is assessed here; the hybrid DA weight in the atmosphere is the same. As a simplification, the experiments are run for a single deterministic member instead of a full ensemble, and the matrices \mathbf{B}_{ens} for the atmosphere and ocean are derived from offline ensembles. The atmospheric ensemble is

taken from an operational run of MOGREPS-G, for which DA is performed using the ensemble transform Kalman filter (which was the operational ensemble DA scheme used at the time of the cyclone). On the other hand, the oceanic \mathbf{B}_{ens} is derived from a FOAM research ensemble that features hybrid-3DEnVar with FGAT, $\alpha = 0.8$, and ensemble inflation (note that this ensemble is not the same as the one used for the construction of \mathbf{B}_{ens} in the single-observation experiments of Section 4). The inflation scheme is return-to-prior-spread (RTPS; Whitaker and Hamill, 2012) and the inflation factor is 0.8. Lea *et al.* (2022) show that this particular combination of α and the inflation factor produces the best results globally, and it is

on this basis that an oceanic ensemble with these parameters is chosen to construct \mathbf{B}_{ens} for the oceanic component of the coupled DA system used here. It should be noted that this offline ensemble runs on daily assimilation cycles and is thus only available to be incorporated into the coupled DA system once a day. As a result, the ensemble information at 0000 UTC each day is re-used for the 0600, 1200, and 1800 UTC cycles later that day. This can be justified by the slow evolution of oceanic fields and ensemble statistics compared with the cycle length of the DA system (6 hr).

The two experiments are run for a 19-day period, from the 0600 UTC cycle on September 25, 2018 to the 0000 UTC cycle on October 14, 2018 inclusive. The initial conditions for the atmospheric and land-surface components are from a separate coupled run, whereas the initial conditions for the oceanic and sea-ice components are from the unperturbed member of the above-mentioned FOAM ensemble. The start date of the coupled experiments is selected in such a way that both the Control and Hybrid experiments are given nearly two weeks to adjust before Cyclone *Titli* becomes active on October 8, 2018. Both experiments assimilate the full set of observations available operationally at the time.³

Since a low-resolution atmospheric model component is used, the detailed structure of Cyclone *Titli* is not expected to be resolved fully in this experiment. In particular, one may expect the intensity of the cyclone to be substantially underestimated (Davis, 2018). Hence, rather than verifying the Control and Hybrid runs against independent data and determining which run is better, we shall focus on explaining how the analyses produced by the two runs can be qualitatively different from each other. In this way, we may understand better the physical mechanisms by which hybrid DA in the ocean can influence tropical cyclone development. A few interesting features seen in this case study are presented in the following subsections.

5.2 | Impacts within the ocean

Figure 9 shows the temperature difference between the Hybrid and Control analyses along a meridional cross-section at 86°E (the same cross-section as Figure 2d) at various times throughout the experiment. A subsurface dipole of temperature difference can be found during and after the passage of Cyclone *Titli* (Figure 9c,d), thus

indicating that the analysis temperatures in the Hybrid run are warmer than the Control run at 15°–16°N, but cooler at 17.5°–19°N. The positioning of this dipole means that the Hybrid run has a larger northward extent of the cold-core eddy shown in Figure 2d. We are interested in this dipole because the subsurface temperature difference is brought to the ocean surface as the cyclone passes through the region (Figure 9d). The propagation of subsurface (hence cooler) waters to the surface during the passage of tropical cyclones is a well-known phenomenon resulting from vertical mixing associated with the cyclone's strong winds (Price, 1981). It leads to a cold wake of surface waters behind the cyclone centre, as Figure 2c illustrates for the case of Cyclone *Titli*.

The origin of this dipole of temperature difference can be traced back to October 3, 2018, well before the formation of the cyclone. The part of the dipole with positive values of Hybrid-minus-Control temperature difference (at approximately 16°N and near the surface at that time) first started to develop during the 1200 UTC assimilation cycle that day, whereas the development of the negative part (at 18°–19°N at that time) began one cycle later, during the 1800 UTC cycle (Figure 9a,b). Further investigation indicates that the formation of the dipole is linked to a couple of altimeter tracks measuring SLA that day. These are shown in Figure 10. During the 1200 UTC cycle, data from an altimeter track in the region 85°–87°E, 16°–18°N are assimilated with substantially negative innovation values in the Control experiment. The innovation values in this region are very similar in the Hybrid experiment (not shown). In light of the different analysis increment structures seen in the single-observation experiments (e.g., Figure 5), the formation of the positive part of the dipole can be attributed to the relevant portion of the altimeter track where the innovations are far from zero. Similarly, the formation of the negative part can be attributed to the large SLA innovations at approximately 86°E, 18.5°N during the 1800 UTC cycle the same day.

The discussion here about the temperature dipole highlights the fact that the impact of using different assimilation schemes can be long-lasting. In this case, the description of vertical mixing in the upper ocean caused by a tropical cyclone, and therefore of the SST in the cyclone's wake, can be influenced by different treatments of SLA observations that pre-date even the formation of the cyclone. The present discussion also highlights that the role of the ocean in representing tropical cyclones in coupled models is not limited to SST or heat content in the mixed layer. Instead, subsurface oceanic fields can also contribute.

The evolution of SST difference between the two runs in the run up to *Titli*'s landfall is shown in Figure 11. It can be seen that, at the time of landfall, a characteristically

³The only exception is that clear-sky radiance data from the *GOES-16* satellite, which the Met Office operational system only began assimilating during our experimental period, are not used in our experiments. This should not concern us, as *GOES-16* is a geostationary satellite far away from the Bay of Bengal.

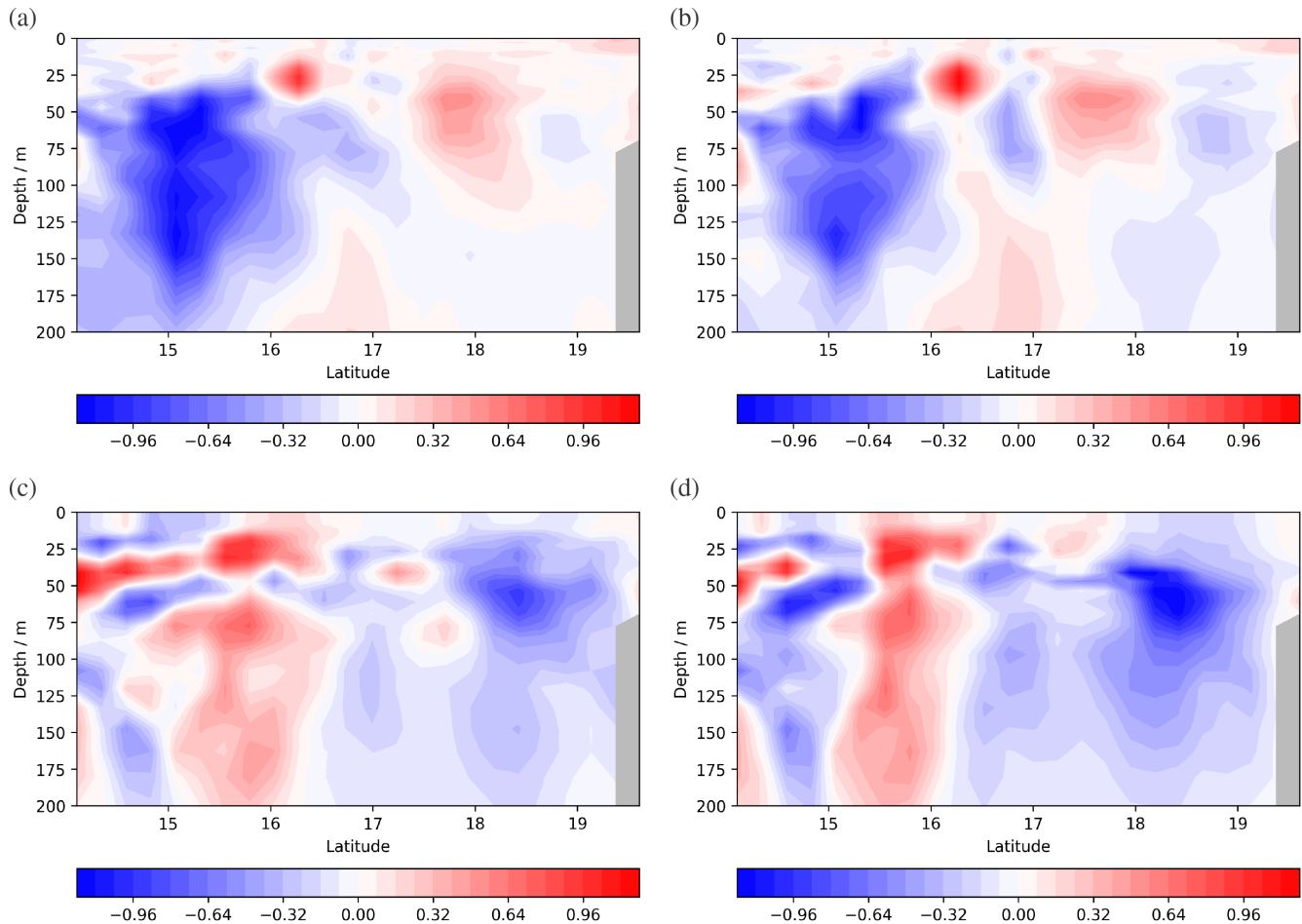
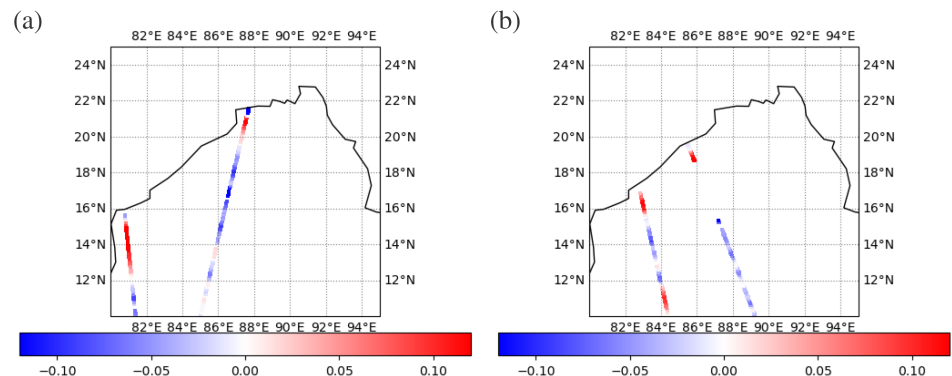


FIGURE 9 Hybrid-minus-Control differences of temperature analyses ($^{\circ}\text{C}$) along a meridional cross-section at 86.0°E averaged over the 3-hr periods ending at (a) 1200 UTC on October 3, 2018, (b) 1800 UTC on October 3, 2018, (c) 0600 UTC on October 10, 2018, and (d) 0000 UTC on October 11, 2018, respectively [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 10 Background innovations of SLA observations (m) in the Bay of Bengal assimilated in the Control experiment during the (a) 1200 UTC and (b) 1800 UTC cycles on October 3, 2018 [Colour figure can be viewed at wileyonlinelibrary.com]



larger-scale dipole of SST difference compared with the surroundings (indicated by the black box in Figure 11d) sits at a location consistent with the subsurface temperature difference dipole shown in Figure 9. Moreover, the timing of formation of the surface dipole is consistent with the known phenomenon that vertical mixing associated with the cyclone brings subsurface temperature differences to the surface. That being said, a couple of

other factors may also contribute to the resulting spatial pattern of SST difference (although it is not straightforward to untangle them). It could be a direct result of the different ways that SST observations are assimilated in the Control and Hybrid runs. On the other hand, it may also be explained by interaction between the atmosphere and the ocean, as will be discussed in the next subsection.

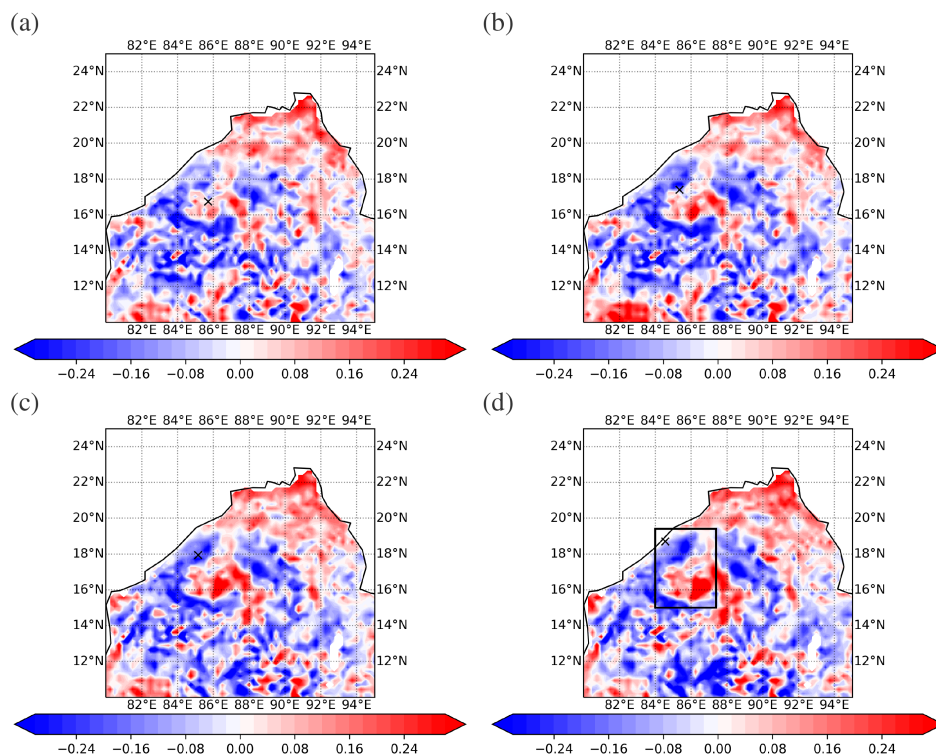


FIGURE 11 Hybrid-minus-Control differences of SST analyses ($^{\circ}\text{C}$) averaged over the 3-hr periods ending at (a) 0600 UTC, (b) 1200 UTC, and (c) 1800 UTC on October 10, 2018, and (d) 0000 UTC on October 11, 2018. The black crosses indicate the positions of the centre of the cyclone in the middle of the respective 3-hr periods. The black box in panel (d) identifies the SST difference dipole mentioned in the text [Colour figure can be viewed at wileyonlinelibrary.com]

5.3 | Impacts on the atmosphere and follow-on impacts on the SST

While the Control and Hybrid analyses agree on *Titli*'s track in the run-up to the cyclone's landfall (not shown), the analysed central pressures are found to be quite different at 0600 UTC on October 10, which is 18 hr before the landfall. The discrepancy is shown in Figure 12a. (Due to the low resolution of the atmospheric component, both runs significantly underestimate the strength of the cyclone in terms of central pressure. The minimum central pressure achieved by *Titli*, according to an independent source, is 972 hPa (RSMC-Tropical Cyclones New Delhi, 2019).) Figure 12b confirms that, at this time, the mean sea-level pressure analysis of the Hybrid run is higher not only at the cyclone's centre but also throughout the core of the cyclone (the blue region in Figure 2a). Moreover, Figure 12a reveals that the difference in central pressure is absent in earlier analyses and is much smaller in later analyses. Hence it could be deduced that during the period 0000–0600 UTC on October 10 the cyclone intensifies faster in the Control run than in the Hybrid run, yet by the time of the landfall the cyclones in the two runs have similar strength. Further investigation (not shown) confirms that the intensification during 0000–0600 UTC is not seen in the background pressure field in either run. Since atmospheric observations are assimilated in the same way, it is the coupling with the ocean during the oceanic IAU period that leads to the different descriptions of the cyclone's intensification.

The fact that Cyclone *Titli* is stronger in the Control run on October 10 can have a further impact on the Hybrid-minus-Control SST difference. The stronger winds associated with the Control-run cyclone enhance mixing in the upper ocean, thus bringing cooler subsurface waters to the surface (Price, 1981). This leads to lower SSTs in the Control run than in the Hybrid run in the region where such mixing occurs, and therefore contributes towards the positive part of the Hybrid-minus-Control temperature difference dipole in Figure 11d. This might explain why the positive part of that dipole is of a larger magnitude than its negative part.

6 | DISCUSSION AND CONCLUSIONS

In variational DA, background-error covariance structures determine how information from observations is spread to the unobserved parts of the system. In oceanic models, these covariances have traditionally been modelled by parametrisations that depend mainly on macroscopic properties of the ocean and have limited dependence on local conditions. As such, these parametrisations may be unable to represent the ocean's mesoscale error structure accurately. Moreover, they may not be suitable in the region of tropical cyclones, because of the cyclones' effect on the spatial and temporal characteristics in the upper ocean. In this work, we have illustrated how incorporating ensemble information of flow dependence

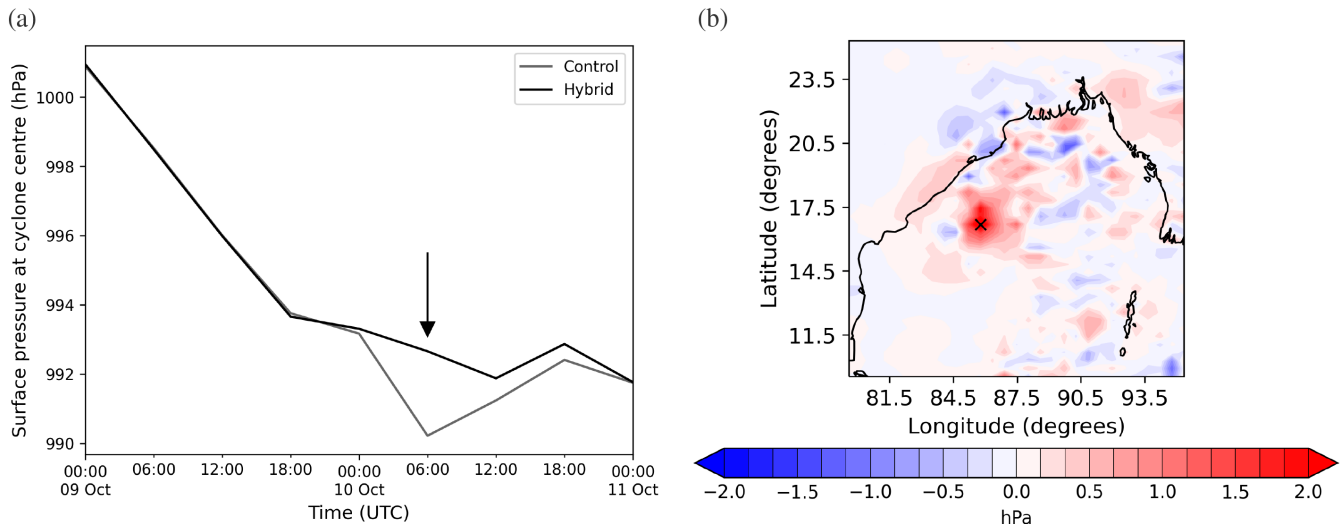


FIGURE 12 (a) Evolution of *Titli*'s central pressure analysis in the 48 hr before landfall, with a temporal resolution of 6 hr. The darker shade is for the Hybrid run and the lighter shade is for the Control run. (b) Hybrid-minus-Control difference of mean sea-level pressure analyses at 0600 UTC on October 10, 2018 (marked by the arrow in panel a). The black cross indicates the position of the centre of the cyclone, which, in this case, is the same for the Hybrid and Control runs [Colour figure can be viewed at wileyonlinelibrary.com]

into the modelling of oceanic background-error covariances in a coupled DA system changes the way that the atmosphere and ocean interact in the context of a case study of a tropical cyclone.

We have systematically carried out single-observation experiments to examine the structure of background-error covariances for the ocean. The analysis increment in response to assimilating a single direct observation of a model variable is proportional to a column of the background-error covariance matrix, \mathbf{B} . Here, \mathbf{B} is a weighted combination between \mathbf{B}_{mod} , which consists of traditional parametrisations of error covariances used in standard 3DVar, and \mathbf{B}_{ens} , a matrix of localised covariances estimated by the ensemble of the day. Hybrid weights of $\alpha = 0.0, 0.2, 0.8$, and 1.0 are considered, where α is the relative weight given to \mathbf{B}_{ens} . The higher α is, the more flow-dependent information is included in the estimation of background-error covariances.

The single-observation experiments are performed using the Met Office's FOAM system during the passage of Tropical Cyclone *Titli* across the Bay of Bengal. Three types of observation, namely an SST observation, an SLA observation, and a temperature observation at a depth of 100 m, are tested at a point near the centre of the cyclone's cold wake, both before and after its formation. A variety of structures are seen in the increments corresponding to different observation types as well as different values of α . The matrices \mathbf{B}_{mod} and \mathbf{B}_{ens} play their roles in controlling the shape and spatial extent of the increments, through their correlation and localisation length-scales respectively. In addition, the overall magnitude of the

increment can be quite sensitive to the background-error standard deviation. Such sensitivity is seen particularly in the single-SST-observation experiments, when the SST spread is very small compared with the observation-error standard deviation, thereby highlighting the particular importance of modelling SST ensemble spread accurately when hybrid DA is used in coupled models. The increments are generally more anisotropic and vertically less uniform, as more weight is given to \mathbf{B}_{ens} . Moreover, their spatial structures can be quite different before and after the cold wake forms, although the shapes of these irregular and time-varying structures can always be attributed to the ensemble spread of the day. Compared with \mathbf{B}_{mod} , which is used in standard 3DVar and has only limited dependence on the local flow, incorporating \mathbf{B}_{ens} into the assimilation system allows a response to observations that is more strongly linked to the local flow.

These single-observation experiments provide valuable insights into understanding how hybridisation of background-error covariances in the ocean might change the analysed representation of *Titli* in the Met Office's coupled model. We have seen qualitatively that the SST field, which is often crucial to the accurate modelling of tropical cyclones, can be affected through several intertwining mechanisms. Most notably, subsurface temperature differences can propagate to the ocean surface through vertical mixing associated with the cyclone's strong winds. The long memory of the subsurface ocean means that the origins of such subsurface differences can even pre-date (and therefore be unrelated to) the cyclone. Apart from the role

of the different representations of the subsurface ocean, hybrid covariances can also change the assimilation of SST observations directly. The relatively few SST observations constraining the assimilation systems (due to the presence of clouds near the cyclone's centre) make it easy for SST analyses to diverge from each other. Moreover, air–sea interaction means that the different SST fields can lead to a discrepancy in the simulated central pressure of the cyclone. The fact that the cyclone is stronger in one simulation than in the other can then contribute to further differences in the SST field, due to the different representations of the cyclone's cold wake. All in all, these illustrations suggest that the effects of hybrid oceanic covariances on SST can be very complex when tropical cyclones are modelled in a coupled DA framework.

In light of the low resolution of the atmospheric model component, we have not attempted to investigate whether the hybridisation of oceanic covariances improves analyses and forecasts of the cyclone's track and strength when compared against independent data. Changes in background innovation statistics for our study period and region are too small to be statistically significant. In order to understand whether hybrid covariances in the ocean generally improve tropical cyclone analyses and forecasts, a comprehensive assessment involving more cyclones and a higher-resolution model is required. We hope to carry out such a study in the future and report its results in a separate article.

Further developments to the weakly coupled NWP system with hybrid atmospheric and oceanic DA are underway at the Met Office. The new system will enable the running of an ensemble that produces flow-dependent background-error covariances online, which are to be used in the next assimilation cycle. This ensemble system will be useful for understanding further the characteristics of background-error covariances in tropical cyclones. In particular, cross-fluid error covariances produced by this weakly coupled system will be examined in order to assess whether a future upgrade to more strongly coupled DA methods can be justified.






AUTHOR CONTRIBUTIONS

Tsz Yan Leung: investigation; methodology; writing – original draft. **Amos S. Lawless:** conceptualization; funding acquisition; methodology; project administration; supervision; writing – review and editing. **Nancy K. Nichols:** conceptualization; funding acquisition; methodology; supervision; writing – review and editing. **Daniel J. Lea:** investigation; resources; software; writing – review and editing. **Matthew J. Martin:** conceptualization; methodology; resources; software; writing – review and editing.

ACKNOWLEDGEMENTS

This work was carried out under the Newton Fund project “Novel Assimilation Methods for Coupled Atmosphere–Ocean Prediction” (Grant Agreement P107915) under the “Weather and Climate Science for Service Partnership India” programme coordinated by the Met Office, with additional funding from the Natural Environment Research Council's support for the National Centre for Earth Observation (Contract Number PR140015). The authors thank Anthony Weaver for developing the hybrid DA code in NEMOVAR, the three anonymous reviewers for their valuable comments on an earlier version of the manuscript, and the India Meteorological Department for granting permission to reproduce Figure 1.

ORCID

Tsz Yan Leung  <https://orcid.org/0000-0003-0056-284X>
Amos S. Lawless  <https://orcid.org/0000-0002-3016-6568>
Nancy K. Nichols  <https://orcid.org/0000-0003-1133-5220>
Daniel J. Lea  <https://orcid.org/0000-0003-1736-628X>
Matthew J. Martin  <https://orcid.org/0000-0003-0293-3106>

REFERENCES

- Anderson, J.L. and Anderson, S.L. (1999) A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, 127, 2741–2758.
- Bannister, R.N. (2008a) A review of forecast error covariance statistics in atmospheric variational data assimilation. I: characteristics and measurements of forecast error covariances. *Quarterly Journal of the Royal Meteorological Society*, 134, 1951–1970.
- Bannister, R.N. (2008b) A review of forecast error covariance statistics in atmospheric variational data assimilation. II: modelling the forecast error covariance statistics. *Quarterly Journal of the Royal Meteorological Society*, 134, 1971–1996.
- Bannister, R.N. (2017) A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 143, 607–633.
- Bauer, P., Thorpe, A. and Brunet, G. (2015) The quiet revolution of numerical weather prediction. *Nature*, 525, 47–55.
- Best, M.J., Pryor, M., Clark, D.B., Rooney, G.G., Essery, R.L.H., Ménard, C.B., Edwards, J.M., Hendry, M.A., Porson, A., Gedney, N., Mercado, L.M., Sitch, S., Blyth, E., Boucher, O., Cox, P.M., Grimmond, C.S.B. and Harding, R.J. (2011) The Joint UK Land Environment Simulator (JULES), model description – Part 1: energy and water fluxes. *Geoscientific Model Development*, 4, 677–699.
- Blockley, E.W., Martin, M.J., McLaren, A.J., Ryan, A.G., Waters, J., Lea, D.J., Mirouze, I., Peterson, K.A., Sellar, A. and Storkey, D. (2014) Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new global FOAM forecasts. *Geoscientific Model Development*, 7, 2613–2638.

- Bloom, S.C., Takacs, L.L., da Silva, A.M. and Ledvina, D. (1996) Data assimilation using incremental analysis updates. *Monthly Weather Review*, 124, 1256–1271.
- Bowler, N.E., Arribas, A., Mylne, K.R., Robertson, K.B. and Beare, S.E. (2008) The MOGREPS short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 134, 703–722.
- Brown, A., Milton, S., Cullen, M., Golding, B., Mitchell, J. and Shelly, A. (2012) Unified modeling and prediction of weather and climate: a 25-year journey. *Bulletin of the American Meteorological Society*, 93, 1865–1877.
- Browne, P.A., de Rosnay, P., Zuo, H., Bennett, A. and Dawson, A. (2019) Weakly coupled ocean–atmosphere data assimilation in the ECMWF NWP system. *Remote Sensing*, 11, 234.
- Cardinali, C., Pezzulli, S. and Andersson, E. (2004) Influence-matrix diagnostic of a data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 130, 2767–2786.
- Chen, X., Nystrom, R.G., Davis, C.A. and Zarzycki, C.M. (2021) Dynamical structures of cross-domain forecast error covariance of a simulated tropical cyclone in a convection-permitting coupled atmosphere–ocean model. *Monthly Weather Review*, 149, 41–63.
- Chen, X. and Zhang, F. (2019) Development of a convection-permitting air–sea-coupled ensemble data assimilation system for tropical cyclone prediction. *Journal of Advances in Modeling Earth Systems*, 11, 3474–3496.
- Clark, D.B., Mercado, L.M., Stith, S., Jones, C.D., Gedney, N., Best, M.J., Pryor, M., Rooney, G.G., Essery, R.L.H., Blyth, E., Boucher, O., Harding, R.J., Huntingford, C. and Cox, P.M. (2011) The Joint UK Land Environment Simulator (JULES), model description – Part 2: carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, 4, 701–722.
- Clayton, A.M., Lorenc, A.C. and Barker, D.M. (2013) Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Quarterly Journal of the Royal Meteorological Society*, 139, 1445–1461.
- Cooper, M. and Haines, K. (1996) Altimetric assimilation with water property conservation. *Journal of Geophysical Research: Oceans*, 101, 1059–1077.
- Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljević, D., Hamrud, M., Hollingsworth, A., Rabier, F. and Fisher, M. (1998) The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: formulation. *Quarterly Journal of the Royal Meteorological Society*, 124, 1783–1807.
- Courtier, P., Thépaut, J.N. and 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.
- Craig, A., Valcke, S. and Coquart, L. (2017) Development and performance of a new version of the OASIS coupler, OASIS3-MCT_3.0. *Geoscientific Model Development*, 10, 3297–3308.
- Cullen, M.J.P. (1993) The unified forecast/climate model. *Meteorological Magazine*, 122, 81–94. https://digital.nmla.metoffice.gov.uk/IO_faa7dfb4-d31a-4a1b-92ba-16a145be1509/ [Accessed 24th June 2021].
- Daley, R. (1991) *Atmospheric Data Analysis*. Cambridge: Cambridge University Press.
- Davis, C.A. (2018) Resolving tropical cyclone intensity in models. *Geophysical Research Letters*, 45, 2082–2087.
- Deckmyn, A. and Berre, L. (2005) A wavelet approach to representing background error covariances in a limited-area model. *Monthly Weather Review*, 133, 1279–1294.
- Dee, D.P. (2005) Bias and data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 131, 3323–3343.
- Evensen, G. (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99, 10143–10162.
- Evensen, G. and van Leeuwen, P.J. (1996) Assimilation of Geosat altimeter data for the Agulhas Current using the Ensemble Kalman Filter with a quasigeostrophic model. *Monthly Weather Review*, 124, 85–96.
- Fisher, M. and Andersson, E. (2001) *Developments in 4D-Var and Kalman filtering*. Reading, UK: European Centre for Medium-Range Weather Forecasts. Technical Memorandum 347.
- Gómez, B., Charlton-Pérez, C.L., Lewis, H. and Candy, B. (2020) The Met Office operational soil moisture analysis system. *Remote Sensing*, 12, 3691.
- Hamill, T.M. and Snyder, C. (2000) A hybrid ensemble Kalman filter-3D variational analysis scheme. *Monthly Weather Review*, 128, 2905–2919.
- Hamill, T.M., Whitaker, J.S. and Snyder, C. (2001) Distance-dependent filtering of background error covariance estimates in an Ensemble Kalman Filter. *Monthly Weather Review*, 129, 2776–2790.
- Houtekamer, P.L. and Mitchell, H.L. (2001) A sequential Ensemble Kalman Filter for atmospheric data assimilation. *Monthly Weather Review*, 129, 123–137.
- Houtekamer, P.L. and Zhang, F. (2016) Review of the Ensemble Kalman Filter for atmospheric data assimilation. *Monthly Weather Review*, 144, 4489–4532.
- Hunke, E.C., Lipscomb, W.H., Turner, A.K., Jeffery, N. and Elliott, S. (2015) *CICE: the Los Alamos sea ice model: documentation and software user's manual, version 5.1.* Los Alamos National Laboratory. <http://www.ccpo.odu.edu/~klinck/Reprints/PDF/cicedoc2015.pdf> [Accessed 22nd June 2021].
- Jazwinski, A.H. (1970) *Stochastic Processes and Filtering Theory*. London: Academic Press.
- Kalnay, E. (2003) *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge: Cambridge University Press.
- Kuhl, D.D., Rosmond, T.E., Bishop, C.H., McLay, J. and Baker, N.L. (2013) Comparison of hybrid ensemble/4DVar and 4DVar within the NAVDAS-AR data assimilation framework. *Monthly Weather Review*, 141, 2740–2758.
- Lawless, A.S. (2013). Variational data assimilation for very large environmental problems. In M. Cullen, M.A. Freitag, S. Kindermann, and R. Scheichl (Eds.), *Large Scale Inverse Problems: Computational Methods and Applications in the Earth Sciences*, pp. 55–90. Göttingen, Germany: De Gruyter.
- Lea, D.J., Martin, M.J. and Oke, P.R. (2014) Demonstrating the complementarity of observations in an operational ocean forecasting system. *Quarterly Journal of the Royal Meteorological Society*, 140, 2037–2049.
- Lea, D.J., While, J., Martin, M.J., Weaver, A., Storto, A. and Chrast, M. (2022) A new global ocean ensemble system at the Met Office: Assessing the impact of hybrid data assimilation and inflation settings. *Quarterly Journal of the Royal Meteorological Society*, 148, 1996–2030.

- Lorenc, A.C. (2003) The potential of the ensemble Kalman filter for NWP – a comparison with 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, 129, 3183–3203.
- Lorenc, A.C. (2013) *Recommended nomenclature for EnVar data assimilation methods*. https://www.wcrp-climate.org/WGNE/BlueBook/2013/individual-articles/01_Lorenc_Andrew_EnVar_nomenclature.pdf [Accessed 21st June 2021].
- Madec, G. and NEMO Team (2016) *NEMO ocean engine (version 3.6 stable). Note du Pôle de modélisation* Vol. 27. Paris, France: Institut Pierre-Simon Laplace. ISSN 1288-1619.
- Magnusson, L., Bidlot, J.-R., Bonavita, M., Brown, A.R., Browne, P.A., de Chiara, G., Dahoui, M., Lang, S.T.K., McNally, T., Mogensen, K.S., Pappenberger, F., Prates, F., Rabier, F., Richardson, D.S., Vitart, F. and Malardel, S. (2019) ECMWF activities for improved hurricane forecasts. *Bulletin of the American Meteorological Society*, 100, 445–457.
- Mirouze, I., Blockley, E.W., Lea, D.J., Martin, M.J. and Bell, M.J. (2016) A multiple length scale correlation operator for ocean data assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, 68, 29744.
- Penny, S.G., Akella, S., Alves, O., Bishop, C., Buehner, M., Chevalier, M., Counillon, F., Draper, C., Frolov, S., Fujii, Y., Karspeck, A., Kumar, A., Laloyaux, P., Mahfouf, J.F., Martin, M., Peña, M., de Rosnay, P., Subramanian, A., Tardif, R., Wang, Y. and Wu, X. (2017) *Coupled data assimilation for integrated Earth system analysis and prediction: goals, challenges and recommendations*. Geneva, Switzerland: World Meteorological Organization. Technical Report WWRP 2017-3.
- Price, J.F. (1981) Upper ocean response to a hurricane. *Journal of Physical Oceanography*, 11, 153–175.
- Purser, R.J., Wu, W.-S., Parrish, D.F. and Roberts, N.M. (2003) Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: spatially homogeneous and isotropic Gaussian covariances. *Monthly Weather Review*, 131, 1524–1535.
- Rabier, F. (2005) Overview of global data assimilation developments in numerical weather-prediction centres. *Quarterly Journal of the Royal Meteorological Society*, 131, 3215–3233.
- Ridley, J.K., Blockley, E.W., Keen, A.B., Rae, J.G.L., West, A.E. and Schroeder, D. (2018) The sea ice model component of HadGEM3-GC3.1. *Geoscientific Model Development*, 11, 713–723.
- RSMC-Tropical Cyclones New Delhi (2019) *Report on cyclonic disturbances over the north Indian Ocean during 2018*. https://rsmcnwdelhi.imd.gov.in/uploads/report/27/27_60dae9_rsmc-2018.pdf [Accessed 22nd June 2021].
- Schiller, A. and Ridgway, K.R. (2013) Seasonal mixed-layer dynamics in an eddy-resolving ocean circulation model. *Journal of Geophysical Research: Oceans*, 118, 3387–3405.
- Storkey, D., Blaker, A.T., Mathiot, P., Megann, A., Aksenov, Y., Blockley, E.W., Calvert, D., Graham, T., Hewitt, H.T., Hyder, P., Kuhlbrodt, T., Rae, J.G.L. and Sinha, B. (2018) UK Global Ocean GO6 and GO7: a traceable hierarchy of model resolutions. *Geoscientific Model Development*, 11, 3187–3213.
- Storto, A. and Andriopoulos, P. (2021) A new stochastic ocean physics package and its application to hybrid-covariance data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 147, 1691–1725.
- Walters, D., Baran, A.J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whittall, M., Williams, K. and Zerroukat, M. (2019) The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations. *Geoscientific Model Development*, 12, 1909–1963.
- Wang, X., Parrish, D., Kleist, D. and Whitaker, J. (2013) GSI 3DVar-based ensemble-variational hybrid data assimilation for NCEP Global Forecast System: single-resolution experiments. *Monthly Weather Review*, 141, 4098–4117.
- Waters, J., Lea, D.J., Martin, M.J., Mirouze, I., Weaver, A. and While, J. (2015) Implementing a variational data assimilation system in an operational 1/4 degree global ocean model. *Quarterly Journal of the Royal Meteorological Society*, 141, 333–349.
- Weaver, A. and Courtier, P. (2001) Correlation modelling on the sphere using a generalized diffusion equation. *Quarterly Journal of the Royal Meteorological Society*, 127, 1815–1846.
- Weaver, A.T., Chrust, M., Ménétrier, B., Piacentini, A., Tshimanga, J., Yang, Y., Gürol, S. and Zuo, H. (2018) *Using ensemble-estimated background error variances and correlation scales in the NEMOVAR system*. Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique. <https://cerfacs.fr/publication> [Accessed 23rd June 2021].
- Weaver, A.T., Deltel, C., Machu, E., Ricci, S. and Daget, N. (2005) A multivariate balance operator for variational ocean data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 131, 3605–3625.
- Weaver, A.T. and Mirouze, I. (2013) On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation. *Quarterly Journal of the Royal Meteorological Society*, 139, 242–260.
- Weaver, A.T., Tshimanga, J. and Piacentini, A. (2016) Correlation operators based on an implicitly formulated diffusion equation solved with the Chebyshev iteration. *Quarterly Journal of the Royal Meteorological Society*, 142, 455–471.
- Whitaker, J.S. and Hamill, T.M. (2012) Evaluating methods to account for system errors in ensemble data assimilation. *Monthly Weather Review*, 140, 3078–3089.
- Williams, K.D., Copsey, D., Blockley, E.W., Bodas-Salcedo, A., Calvert, D., Comer, R., Davis, P., Graham, T., Hewitt, H.T., Hill, R., Hyder, P., Ineson, S., Johns, T.C., Keen, A.B., Lee, R.W., Megann, A., Milton, S.F., Rae, J.G.L., Roberts, M.J., Scaife, A.A., Schiemann, R., Storkey, D., Thorpe, L., Watterson, I.G., Walters, D.N., West, A., Wood, R.A., Woollings, T. and Xavier, P.K. (2017) The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) Configurations. *Journal of Advances in Modeling Earth Systems*, 10, 357–380.

How to cite this article: Leung, T.Y., Lawless, A.S., Nichols, N.K., Lea, D.J. & Martin, M.J. (2022) The impact of hybrid oceanic data assimilation in a coupled model: A case study of a tropical cyclone. *Quarterly Journal of the Royal Meteorological Society*, 148(746), 2410–2430. Available from: <https://doi.org/10.1002/qj.4309>