University of Reading

Validating and improving the uncertainty assumptions for the assimilation of ocean colour derived chlorophyll a into a marine biogeochemistry model of the North-West European Shelf Seas

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Fowler, A. M. ORCID: https://orcid.org/0000-0003-3650-3948, Skákala, J. and Ford, D. (2023) Validating and improving the uncertainty assumptions for the assimilation of ocean colour derived chlorophyll a into a marine biogeochemistry model of the North-West European Shelf Seas. Quarterly Journal of the Royal Meteorological Society, 149 (750). pp. 300-324. ISSN 1477-870X doi: 10.1002/qj.4408 Available at https://centaur.reading.ac.uk/109040/

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.4408

Publisher: Royal Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law,

University of Reading

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

Received: 13 May 2022 | Revised: 14 November 2022 | Accepted: 15 November 2022 | Published on: 23 December 2022

DOI: 10.1002/qj.4408

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

Validating and improving the uncertainty assumptions for the assimilation of ocean-colour-derived chlorophyll *a* into a marine biogeochemistry model of the Northwest European Shelf Seas

Alison M. Fowler[1,2] | **Jozef Skákala**[2,3] | **David Ford**[4]

[1]Department of Meteorology, University of Reading, Reading, UK

[2]National Centre for Earth Observation, Reading, UK

[3]Plymouth Marine Laboratory, Plymouth, UK

[4]Met Office, Exeter, UK

Correspondence
A. M. Fowler, Department of Meteorology, University of Reading, Reading, UK.
Email: a.m.fowler@reading.ac.uk

Funding information
Natural Environment Research Council, Grant/Award Number: PR140015

Abstract

The correct specification of all sources of uncertainty is critical to the success of data assimilation (DA) in improving the realism and accuracy of forecasts and reanalyses. This work focuses on improving the uncertainty assumptions made during the assimilation of ocean-colour-derived chlorophyll *a* into an operational marine coupled physical–biogeochemical DA system, which produces daily biogeochemistry forecasts on the Northwest European Shelf Seas. Analysis of the observation–model misfits shows significant biases in chlorophyll *a*, which vary strongly with season. The behaviour of these misfits agrees well with previous studies and can be attributed to systematic errors within the coupled model. Diagnostic metrics, used frequently within numerical weather prediction, are applied to separate out the random component of the observation and model errors, allowing for the derivation of new error covariance matrices. These new error covariance matrices are then modified to account for the biases in the model that cannot be treated explicitly within the operational DA system. This has the effect of inflating both the error variances and the correlation length-scales. Experiments show that the new error covariances can result in significant improvements in the accuracy of the analysis and forecast. In particular, the new error covariance matrices reduce the bias in the spring phytoplankton bloom present when using the previous error covariances. Validation against independent glider observations in the North Sea also shows reductions in bias in chlorophyll *a* and oxygen that extend below the surface to the depth of the mixed layer. Accounting for the biases in the model in the error correlations can lead to much larger improvements than not accounting for them; however, there are also regions where large degradations are seen that may indicate model instabilities. This may be improved by estimating the bias separately for the different regions on the shelf.

KEYWORDS

data assimilation, model bias, Northwest European Shelf Seas, phytoplankton

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 Crown Copyright and The Authors. *Quarterly Journal of the Royal Meteorological Society* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society. This article is published with the permission of the Controller of HMSO and the King's Printer for Scotland.

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

1 | INTRODUCTION

Shelf seas are crucial to all life on our planet. They are known to be responsible for 90% of global fisheries, 20% of marine primary production, and 20% of atmospheric carbon dioxide uptake (Pauly *et al.*, 2002; Borges *et al.*, 2006; Jahnke, 2010). Here we focus on the operational analysis and forecasting of surface total chlorophyll *a* on the Northwest European (NWE) shelf. Chlorophyll *a* is a proxy for phytoplankton biomass, which forms the base of the marine food web and has important economic, environmental, and public safety consequences. Each year, phytoplankton blooms are a major ecosystem driver for the NWE shelf (Lutz *et al.*, 2007; Henson *et al.*, 2009). The blooms can occur throughout the year, but generally the most significant occur in February–May.

The regular assimilation of observations to initialise models of the coupled physics–biogeochemistry marine system has proven to be essential to the realism of biogeochemistry forecasts (e.g., Nerger and Gregg, 2007; Teruzzi *et al.*, 2014; Skákala *et al.*, 2018). An important source of information for constraining the models via data assimilation (DA) is available using satellite remote-sensing data (e.g., McClain, 2009; Aiken *et al.*, 2014; Groom *et al.*, 2019), which are capable of viewing the ocean with high temporal and spatial coverage. Ocean-colour-derived total chlorophyll *a*, in particular, has been demonstrated to be essential (e.g., Carmillet *et al.*, 2001; Hoteit *et al.*, 2005; Nerger and Gregg, 2007; Ciavatta *et al.*, 2011; 2016; Ford *et al.*, 2012; Ford and Barciela, 2017).

The assimilation of data at many operational centres is performed using a class of variational DA (e.g., Waters *et al.*, 2015). An important input into variational DA is the error statistics of the observations and the background (the prior estimate of the model state of interest made before the observations are given, usually provided by a previous forecast). The assumption that both the observations and background errors are unbiased and follow a Gaussian distribution means that their error statistics are fully described by the predefined error covariance matrices $\mathbf{R}_a \in \mathbb{R}^{p\times p}$ and $\mathbf{B}_a \in \mathbb{R}^{n\times n}$ respectively, where p is the number of observations and n is the size of the model state (number of model variables $\times$ number of grid points). The subscript a denotes that these are the matrices assumed during the assimilation of the data and may differ from the true statistics, as discussed later. The diagonal elements of these matrices represent the error variances, whilst the off-diagonal elements represent the error covariances in space and between different variables. Normalising the covariance of two variables by the variance of each variable gives the (nondimensional) correlations. The error correlations are often parameterised in terms of a correlation length-scale, given a predefined correlation function. The matrices $\mathbf{B}_a$ and $\mathbf{R}_a$ specify how closely the analysis of the model initial state should be constrained by the background and the observations. It is therefore crucial that $\mathbf{B}_a$ and $\mathbf{R}_a$ account correctly for all sources of uncertainty, and that the unbiased assumption is achieved, so that the weighting of the model and the observations is valid and the most accurate analysis is obtained.

Many different approaches have been proposed for estimating $\mathbf{B}_a$ and $\mathbf{R}_a$ (see Section 2). Here we will focus on the use of consistency diagnostics as proposed by Desroziers *et al.* (2005), which uses samples of output from an already operationally running marine biogeochemistry analysis system. The Desroziers *et al.* (2005) method has the benefit of representing all sources of uncertainty in the assimilation system implicitly. A disadvantage is that the estimate of $\mathbf{B}_a$ is restricted to the space of the observations; however, for large observation datasets, such as ocean-colour-derived total chlorophyll *a*, that observe a key model variable directly, and by carefully binning the sample, the technique can provide meaningful estimates with minimal sampling noise. An issue specific to this application that we must contend with is significant biases in the marine biogeochemistry model, which cannot be accounted for explicitly within the current setup of the variational assimilation system. We therefore propose a method for modifying the Desroziers *et al.* (2005) estimate to account for these biases.

The structure of this article is as follows. In Section 2 we provide a brief overview of how $\mathbf{B}_a$ and $\mathbf{R}_a$ may be estimated in practice, with a focus on the consistency diagnostics of observation–model misfits (Desroziers *et al.*, 2005). This method is then modified to account for biases within the estimate of $\mathbf{B}_a$. In Section 3 the 2018 case study is presented, for which we provide new estimates of $\mathbf{B}_a$ and $\mathbf{R}_a$ in Section 4. In Section 5 we apply these new estimates to demonstrate the impact on the forecast using the new estimates of $\mathbf{B}_a$ and $\mathbf{R}_a$, focusing on the impact of accounting for biases in $\mathbf{B}_a$. In Section 6 we investigate whether repeated estimation of $\mathbf{B}_a$ and $\mathbf{R}_a$ results in convergence of the estimates. Finally, in Section 7 we summarise our key findings and discuss the limitations and implications of this work.

2 | ESTIMATING UNCERTAINTIES IN DA

The importance of $\mathbf{B}_a$ and $\mathbf{R}_a$ can be appreciated by looking at the analytical form of the analysis, $\mathbf{x}_a$, given by the minimum of the variational cost function:

$$\mathbf{x}_a = (\mathbf{I} - \mathbf{K}_a\mathbf{H})\mathbf{x}_b + \mathbf{K}_a\mathbf{y}, \quad (1)$$

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

where $\mathbf{y} \in \mathbb{R}^p$ is a vector of the observations and $\mathbf{x}_b \in \mathbb{R}^n$ is the background state (see Kalnay, 2003 for a derivation). $\mathbf{K}_a \in \mathbb{R}^{n \times p}$ is given by $\mathbf{B}_a\mathbf{H}^T(\mathbf{H}\mathbf{B}_a\mathbf{H}^T + \mathbf{R}_a)^{-1}$, where $\mathbf{H} \in \mathbb{R}^{p \times n}$ is the observation operator, the mapping from the modelled variables to those observed. The observation operator can account for observations that have a different spatial distribution from the model grid on which $\mathbf{x}$ is defined, and can also allow for the observations to measure different variables from those modelled. If the observations are distributed in time over the assimilation window, then the observation operator may also include a dynamical model to account for this. To simplify the notation, we assume the observation operator to be linear; however, the theory can be modified for small nonlinearities in the observation operator as is typical in DA (Lorenc *et al.*, 2000). $\mathbf{x}_a$ provides the most probable state given the background and observations, assuming that the errors in the data are unbiased and Gaussian and the error covariances are correctly specified by $\mathbf{B}_a$ and $\mathbf{R}_a$ (Kalnay, 2003).

Within this section we give a brief overview of the various possibilities for estimating $\mathbf{B}_a$ and $\mathbf{R}_a$. We begin by defining the error and its statistics.

Let $\mathbf{z}$ be the estimate of the truth contained in the vector $\mathbf{z}_t$. The error in $\mathbf{z}$ can then be defined as $\boldsymbol{\epsilon}_z = \mathbf{z} - \mathbf{z}_t$.

The background estimate of the true state, $\mathbf{x}_t$, given by $\mathbf{x}_b$ comes from a forecast initiated from the previous analysis of $\mathbf{x}$. The errors in the background, $\boldsymbol{\epsilon}_b = \mathbf{x}_b - \mathbf{x}_t$, are therefore an amalgamation of the errors in the previous analysis propagated by the (potentially) erroneous model: for example, due to missing processes or boundary condition uncertainty, including atmospheric forcing uncertainty, as well as errors in the physical and biogeochemical models.

For the observations, $\mathbf{y}$, the true state we wish to estimate mapped to the observed variables is given by $\mathbf{y}_t = \mathbf{H}\mathbf{x}_t$. Therefore $\boldsymbol{\epsilon}_y = \mathbf{y} - \mathbf{H}\mathbf{x}_t$ incorporates not only instrument error but also forward model error (due to uncertainty in specifying the mapping $\mathbf{H}$) and representation errors (due to the difference in scales and processes represented by $\mathbf{y}$ and $\mathbf{x}$: Janjic *et al.*, 2017). Ocean colour data, in particular, are subject to uncertainty in the retrieval and postprocessing to the final gridded product. If the observations are distributed in time within an assimilation window and a dynamical model is included in $\mathbf{H}$, then the observation error should arguably also incorporate model error (Gejadze *et al.*, 2017; Howes *et al.*, 2017).

Due to the multifarious origins of the errors in both $\mathbf{y}$ and $\mathbf{x}_b$, it is difficult to estimate all components individually. Fortunately for DA we only need the combined error covariance matrices. However, these matrices may still vary in time and space.

The bias in $\mathbf{z}$ is given by the expectation of $\boldsymbol{\epsilon}_z$, $\mathbf{b}_z = E[\boldsymbol{\epsilon}_z]$. An equation for the error covariance of $\mathbf{z}$ is then given by

$$\mathbf{Z} = E[(\boldsymbol{\epsilon}_z - \mathbf{b}_z)(\boldsymbol{\epsilon}_z - \mathbf{b}_z)^T]. \quad (2)$$

To estimate $\mathbf{b}_z$ and $\mathbf{Z}$, an estimate of the truth is necessary. For this purpose, it could be assumed that in situ data have a negligible error compared with satellite data and the model and so can be used as a proxy for the truth (e.g., Brewin *et al.*, 2017). Due to the limited sampling of in situ data, it is difficult to obtain a large enough sample to obtain estimates of the observation and background-error variances, let alone the covariances. However, Brewin *et al.* (2017) used a total of 2,791 samples collected in the North Atlantic region analysed by high-performance liquid chromatography, spanning 1995–2014, to estimate ocean-colour-derived chlorophyll *a* biases and root-mean-square errors (RMSE), by parameterising the uncertainties as a function of optical water type (OWT) (Jackson *et al.*, 2017). These estimates of bias and RMSE as a function of OWT were then used to map uncertainties in ocean-colour-derived chlorophyll *a* on a per-pixel basis (i.e., at every data point from the ocean-colour-derived chlorophyll *a* product). This approach can, to a certain degree, overcome issues with the distribution of data used in the validation, and accounts for uncertainties varying with the conditions and the magnitude of the observed chlorophyll *a*. The assumption that in situ measurements of phytoplankton group chlorophyll *a* have negligible error is of course problematic, but in practice it can be difficult to quantify the uncertainties in in situ data meaningfully (Boss *et al.*, 2008; Brewin *et al.*, 2014). Additionally, for the estimation of satellite observation-error statistics, in situ data comparisons do not allow for an estimate of uncertainty due to the mismatch in scales between the satellite observation and model, as the in situ observations represent a different scale from either (Groom *et al.*, 2019). For estimates of surface chlorophyll *a*, this includes the differences in horizontal and temporal scales, as well as the depths represented by the in situ and satellite data (Sathyendranath *et al.*, 2019). To understand the consequences of the mismatch of scales on the estimate of the uncertainty in the ocean colour product, in situ methods capable of measuring the optical and biogeochemical properties of the water continuously would be needed.

Many studies have attempted to estimate the background-error covariances using model outputs. Examples include using an ensemble of perturbed model runs (e.g., Ciavatta *et al.*, 2018), using differences in lagged forecast (e.g., Ford and Barciela, 2017), and computing empirical orthogonal functions (EOFs) from a long time series (e.g., Cossarini *et al.*, 2019). These methods all assume that the variability in the model encapsulates its uncertainty fully. However, it is entirely possible that

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

sources of uncertainty are missed and model biases will not be captured.

Another approach to estimating the covariance matrices is based on consistency diagnostics of observation–background misfits, making use of the theory used to assimilate the data. This method makes no attempt to untangle the separate sources of uncertainty in $\mathbf{B}_a$ and $\mathbf{R}_a$, so should, by definition, account for everything.

Let the innovation be defined as $\mathbf{d} = \mathbf{y} - \mathbf{H}\mathbf{x}_b$. This can be seen to be equal to the difference of errors in the observations $\left(\boldsymbol{\epsilon}_y = \mathbf{y} - \mathbf{H}\mathbf{x}_t\right)$ and the background in observation space ($\mathbf{H}\boldsymbol{\epsilon}_b = \mathbf{H}\mathbf{x}_b - \mathbf{H}\mathbf{x}_t$), such that

$$\mathbf{d} = \boldsymbol{\epsilon}_y - \mathbf{H}\boldsymbol{\epsilon}_b. \tag{3}$$

An estimate of the differences in the biases in the observation and background ($\boldsymbol{b}_y$ and $\boldsymbol{b}_b$ respectively) is then given by the sample mean of a population of innovations:

$$\begin{aligned}\overline{\mathbf{d}} &= E[\mathbf{d}] \\ &= E[\boldsymbol{\epsilon}_y] - E[\mathbf{H}\boldsymbol{\epsilon}_b] \\ &= \boldsymbol{b}_y - \mathbf{H}\boldsymbol{b}_b \\ &\approx \frac{1}{N-1}\sum_{i=1}^{N}\mathbf{d}^{(i)},\end{aligned} \tag{4}$$

where N is the size of the sample.

An estimate for the sum of $\mathbf{R}$ and $\mathbf{HBH}^T$ (from which we can then attempt to separate out the two matrices) can similarly be given by the sample covariance of the innovations,

$$\begin{aligned}\mathbf{D} &= E\left[(\mathbf{d} - \overline{\mathbf{d}})(\mathbf{d} - \overline{\mathbf{d}})^T\right] \\ &= E\left[(\boldsymbol{\epsilon}_y - \boldsymbol{b}_y)(\boldsymbol{\epsilon}_y - \boldsymbol{b}_y)^T\right] + E\left[(\mathbf{H}(\boldsymbol{\epsilon}_b - \boldsymbol{b}_b))(\mathbf{H}(\boldsymbol{\epsilon}_b - \boldsymbol{b}_b))^T\right] \\ &= \mathbf{R} + \mathbf{HBH}^T \\ &\approx \frac{1}{N-1}\sum_{i=1}^{N}\left(\mathbf{d}^{(i)} - \overline{\mathbf{d}}\right)\left(\mathbf{d}^{(i)} - \overline{\mathbf{d}}\right)^T,\end{aligned} \tag{5}$$

assuming that the errors in the background and observations are uncorrelated. Here we neglect the "a" suffix on $\mathbf{R}$ and $\mathbf{B}$, as these are a sample estimate of the true error covariances rather than the estimation used within the assimilation.

If $\mathbf{HBH}^T$ is thought to be characterised well by a previous estimate of $\mathbf{HB}_a\mathbf{H}^T$, then this can be subtracted from Equation (5), leaving just $\mathbf{R}$, or vice versa. If this is not the case, then various possible solutions have been proposed to untangle $\mathbf{R}$ and $\mathbf{HBH}^T$ from the sample covariance of the innovations, which we now discuss.

The approach of Hollingsworth and Lönnberg (1986) is to assume that the observation errors are spatially uncorrelated, so that by fitting a function to the correlations of the innovations the variances can be separated out. In the review of Sitwell and Ménard (2020) this is referred to as an a priori method.

If significant biases are present, then ideally these would be corrected offline before the assimilation (as is attempted with the ocean-colour-derived chlorophyll *a* observations in our case, see Section 3) or corrected during the assimilation with methods such as weak-constraint 4DVar (e.g., Trémolet, 2007) and VarBC (Auligné *et al.*, 2007). However, if biases are present in the innovations (violating the DA theory) and are not subtracted when estimating the innovation covariances, then

$$\begin{aligned}\mathbf{D}_{\text{bias}} &= E\left[\mathbf{d}(\mathbf{d})^T\right] \\ &= \mathbf{D} + \overline{\mathbf{d}}(\overline{\mathbf{d}})^T.\end{aligned} \tag{6}$$

The bias present in the innovations will result in the estimate of the variances as well as the correlation length-scales increasing. Therefore, the Hollingsworth and Lönnberg (1986) approach will diagnose more significant error correlations in the background than if the biases had been removed. In addition to this, it will also diagnose a proportionally larger background-error variance. Assuming the biases originate from the background and not the observations, this inflation of the background-error covariances is consistent with the suggestion by Dee and Da Silva (1998), who propose that significant biases present in the background that are not corrected before or during the assimilation should be incorporated into the covariance matrices used within the assimilation. This will have the effect that the biased background is downweighted and the bias in the analysis estimated by Equation (1) will be reduced, although the analysis-error covariance will be increased. However, overall, inflating the background-error covariances in this way will result in the analysis with the smallest RMSE possible without correcting the biases present explicitly.

An alternative method to Hollingsworth and Lönnberg (1986), which makes less strong assumptions on the structure of the covariance matrices, is the method of Desroziers *et al.* (2005) (from now on referred to as DBCP05). DBCP05 proposed a method to estimate the covariance matrices directly from the output of the assimilation system. Hence these are referred to as a posteriori methods in the review of Sitwell and Ménard (2020).

Let us refer to $\mathbf{d}_a^o = \mathbf{y} - \mathbf{H}\mathbf{x}_a$ as the residual and $\mathbf{d}_b^a = \mathbf{H}\mathbf{x}_a - \mathbf{H}\mathbf{x}_b$ as the increment (in observation space). Using the equation for the analysis (Equation 1), we can rewrite these as $\mathbf{d}_a^o = (\mathbf{I} - \mathbf{H}\mathbf{K}_a)\mathbf{d}$ and $\mathbf{d}_b^a = \mathbf{H}\mathbf{K}_a\mathbf{d}$. DBCP05 showed that an estimate for the observation-error

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

covariance matrix can then be given by

$$\begin{aligned}\mathbf{R}_{\text{new}} &= E\left[\left(\mathbf{d}_{\text{a}}^{\text{o}} - \overline{\mathbf{d}_{\text{a}}^{\text{o}}}\right)\left(\mathbf{d} - \overline{\mathbf{d}}\right)^{\text{T}}\right] \\ &= (\mathbf{I} - \mathbf{H}\mathbf{K}_{\text{a}})\mathbf{D}, \end{aligned} \tag{7}$$

where $\mathbf{H}\mathbf{K}_{\text{a}} = \left(\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}}\right)_{\text{a}}\left(\left(\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}}\right)_{\text{a}} + \mathbf{R}_{\text{a}}\right)^{-1}$ is a function of the covariance matrices used within the assimilation, which may differ considerably from the true error covariances (without the *a* suffix) present in $\mathbf{D}$ (Equation 5).

In a similar way, an estimate for the background-error covariance matrix in observation space can be given by

$$\begin{aligned}\left(\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}}\right)_{\text{new}} &= E\left[\left(\mathbf{d}_{\text{b}}^{\text{a}} - \overline{\mathbf{d}_{\text{b}}^{\text{a}}}\right)\left(\mathbf{d} - \overline{\mathbf{d}}\right)^{\text{T}}\right] \\ &= \mathbf{H}\mathbf{K}_{\text{a}}\mathbf{D}. \end{aligned} \tag{8}$$

From Equations (7) and (8), it is clear that the DBCP05 diagnostics result in the innovation covariance, $\mathbf{D}$, being partitioned into observation and background (in observation space) components according to the identity minus the weights they are given in deriving the analysis in observation space (see Equation 1). $\mathbf{R}_{\text{new}}$ and $(\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}})_{\text{new}}$ can also be interpreted in terms of an adjustment to the matrices used within the assimilation:

$$\mathbf{R}_{\text{new}} = \mathbf{R}_{\text{a}}(\mathbf{D}_{\text{a}})^{-1}\mathbf{D}, \tag{9}$$

$$\left(\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}}\right)_{\text{new}} = \left(\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}}\right)_{\text{a}}(\mathbf{D}_{\text{a}})^{-1}\mathbf{D}, \tag{10}$$

where $\mathbf{D}_{\text{a}} = (\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}})_{\text{a}} + \mathbf{R}_{\text{a}}$.

If the covariances are given correctly, then $\mathbf{D}_{\text{a}} = \mathbf{D}$ and Equations (9) and (10) will simply return the matrices used within the assimilation. However, if the covariance matrices are misspecified, then Equations (9) and (10) will return different matrices from those used within the assimilation. These will not be the true error covariances. How they relate to the true statistics was explored theoretically by Waller *et al.* (2016), who showed a complex interaction between misspecifying the observation- and background-error variances and length-scales. However, the new estimates can provide guidance on how to adjust the covariance matrices currently used, and have proven beneficial in numerical weather prediction, in particular for elucidating observation-error correlations (e.g., Weston *et al.*, 2014; Bormann *et al.*, 2016; Campbell *et al.*, 2017; Simonin *et al.*, 2019).

The method of DBCP05 has also been applied by Mattern *et al.* (2018) to the tuning of background- and observation-error variances in a physical–biogeochemical model for the assimilation of physical and chlorophyll *a* observations. They found that tuning only the variances in this way resulted in improved prior and posterior model-observation fits, as well as a reduction in overfitting to the observations and a better balance in the weighting of physical and chlorophyll *a* observations. A limitation of that study, which is addressed here, is the assumption that the underlying spatial structure of $\mathbf{B}_{\text{a}}$ is correct.

The method of DBCP05 has also been used for marine biogeochemical monitoring by Cossarini *et al.* (2019). In that study, the DBCP05 diagnostic was used to estimate the observation-error variances for new biogeochemical-Argo float data. Performing this monthly allowed them to account for the dynamically varying representation error associated with the biogeochemical-Argo float data.

If Equations (7) and (8) are used simultaneously to update the full $(\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}})_{\text{a}}$ and $\mathbf{R}_{\text{a}}$ matrices and the data are unbiased, then performing the assimilation again with these new matrices should mean that the data assimilation system will appear consistent with the innovation statistics, as $(\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}})_{\text{new}} + \mathbf{R}_{\text{new}} = \mathbf{D}$ from Equations (7) and (8). Therefore, applying Equations (7) and (8) again will return the same estimates for $\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}}$ and $\mathbf{R}$. If not updating the full $\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}}$ and $\mathbf{R}$ matrices simultaneously, then the estimates will differ as the diagnostic is reapplied. In the study of Mattern *et al.* (2018), where the covariances are rescaled without altering the length-scales, the estimates are updated iteratively to convergence. In other studies it has been assumed that one of the matrices, usually $(\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}})_{\text{a}}$, is correct, so that the observation-error covariance matrix can be found by reapplying the estimate iteratively (Gauthier *et al.*, 2018). However, the covariances estimated will only converge on the true error covariance matrices if the error in $(\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}})_{\text{a}}$ is negligible (Bathmann, 2018; Gauthier *et al.*, 2018).

In Equations (7) and (8), the biases in the innovations, residuals, and increments are removed when estimating the error covariances, as this is consistent with the assumptions made in data assimilation theory. However, as discussed above, if significant biases are present in the data then this information should not be ignored during the assimilation. We could compute $\mathbf{R}$ and $\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}}$ from Equations (7) and (8) without removing the biases. This will attribute some of the innovation bias to the background and some to the observations, depending on the weights $\mathbf{I} - \mathbf{H}\mathbf{K}_{\text{a}}$ and $\mathbf{H}\mathbf{K}_{\text{a}}$. However, in our case, as the observations are independently bias-corrected using the estimates from Brewin *et al.* (2017), in the following we attribute the biases in the innovations solely to the prior. Therefore, in line with Dee and Da Silva (1998), we propose a modified estimate for $\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}}$ that incorporates the innovation biases fully:

$$(\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}})_{\text{new,bias}} = (\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}})_{\text{new}} + \overline{\mathbf{d}}(\overline{\mathbf{d}})^{\text{T}}. \tag{11}$$

The effect of including the biases in the estimate of $\mathbf{H}\mathbf{B}\mathbf{H}^{\text{T}}$ is to inflate the matrix so that during the

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

assimilation the weighting of the background is reduced, and therefore the analysis is less likely to be pulled towards a biased state. The bias in the analysis will therefore be reduced at the expense of increasing its error covariance. Increasing the background-error correlation length-scales by incorporating the biases will also have the effect of spreading information across larger regions, reducing spatial biases. Again this comes at the expense of no longer extracting as much small-scale information from the observations.

As $(\mathbf{HBH}^{\mathrm{T}})_{\mathrm{new,bias}} + \mathbf{R}_{\mathrm{new}}$ equals $\mathbf{D}_{\mathrm{bias}}$ rather than $\mathbf{D}$, applying the diagnostic a second time would give

$$\mathbf{R}_{\mathrm{new,II}} = \mathbf{R}_{\mathrm{new}}(\mathbf{D}_{\mathrm{bias}})^{-1}\mathbf{D}, \tag{12}$$

$$(\mathbf{HBH}^{\mathrm{T}})_{\mathrm{new,biasII}} = (\mathbf{HBH}^{\mathrm{T}})_{\mathrm{bias}}(\mathbf{D}_{\mathrm{bias}})^{-1}\mathbf{D}. \tag{13}$$

That is, the estimates of the covariance matrices will not converge in one step. By accounting for the biases in $(\mathbf{HBH}^{\mathrm{T}})_{\mathrm{new,bias}}$, we expect the biases to be reduced and so in the next iteration $\overline{\mathbf{d}}(\overline{\mathbf{d}})^{\mathrm{T}}$ will be reduced compared with the previous iteration. However, if the inflation is removed then the bias will return and so it is not desirable to apply Equation (13) iteratively. In Section 6, we will investigate the convergence of the diagnostics further for our application.

Equations (7), (8), or (11) can be estimated from samples of innovations, residuals, and increments. The advantages of this approach over in situ comparisons is that the sample size, N, is now the number of ocean colour data, which is orders of magnitude greater than the number of inhomogeneously placed in situ observations. Therefore, there are many more data to provide robust estimates of correlations as well as variances and biases for both the observations and the model equivalent. This could be used to allow for dependences of the uncertainties on time/location and so forth to be estimated. To increase the sample size for any given strata, assumptions of isotropy, ergodicity (perhaps within a given season), and homogeneity can be made. Careful consideration should therefore be given as to how to bin the data.

3 | 2018 CASE STUDY

An operational prediction system for the biogeochemistry of the NWE shelf is maintained by the Met Office (Edwards *et al.*, 2012; McEwan *et al.*, 2021). This system is based on the coupling of the ocean physics component, Nucleus for European Modelling of the Ocean (NEMO: Madec *et al.*, 2015; O'Dea *et al.*, 2017) and a biogeochemical component based on the European Regional Seas Ecosystem Model (ERSEM: Baretta *et al.*, 1995; Blackford, 1997; Butenschön *et al.*, 2016). The coupling is achieved using the Framework for Aquatic Biogeochemical Models (FABM) coupler (Bruggeman and Bolding, 2014). The coupling is one-way, with the physics affecting the biogeochemistry but no feedback from the simulated biogeochemistry to physics (Skákala *et al.*, 2022).

To compute the new error covariance matrices, we use output from a previously published run of the NEMO–ERSEM analysis system for 2018 (Skákala *et al.*, 2021). Details of this setup are given in the following subsections.

3.1 | The physical component: NEMO

NEMO is a primitive equation model adapted to regional and global ocean circulation problems. We use version 3.6, in which the prognostic variables are the three-dimensional velocity field, nonlinear sea-surface height, potential temperature, and practical salinity. In our experiments, NEMO has 7-km spatial resolution on the Atlantic Margin Model (AMM7) domain using a terrain-following z^*–σ coordinate system with 51 vertical levels (Siddorn and Furner, 2013; O'Dea *et al.*, 2017). The lateral boundary conditions for physical variables at the Atlantic boundary were taken from the outputs of the Met Office operational 1/12° North Atlantic model (Storkey *et al.*, 2010); the Baltic boundary values were derived from a reanalysis produced by the Danish Meteorological Institute for Copernicus Marine Environment Monitoring Service (CMEMS). Inputs of river discharge were based on data from Lenhart *et al.* (2010). The model was forced at the surface by atmospheric fluxes provided by an hourly and 31-km resolution realization (HRES) of the European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis 5 (ERA5) dataset (Hersbach *et al.*, 2020).

3.2 | The biogeochemical component: ERSEM

ERSEM (Baretta *et al.*, 1995; Butenschön *et al.*, 2016) is an ecosystem model for marine biogeochemistry, pelagic plankton, and benthic fauna. The model partitions the phytoplankton into four functional types (PFT): picophytoplankton (<2 μm), nanophytoplankton (2–20 μm), and microphytoplankton (>20 μm), with the latter containing diatoms (having silicate cell walls) and dinoflagellates. ERSEM uses variable stoichiometry for the simulated plankton groups (Baretta Bekker *et al.*, 1997; Geider *et al.*, 1997), representing each PFT biomass in terms of chlorophyll *a*, carbon, nitrogen, and phosphorus, with

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

diatoms also represented by silicon. ERSEM predators are composed of three zooplankton types (mesozooplankton, microzooplankton, and heterotrophic nanoflagellates), with organic material being decomposed by one functional type of heterotrophic bacteria (Butenschön *et al.*, 2016). The ERSEM inorganic component consists of nutrients (nitrate, phosphate, silicate, ammonium, and carbon) and dissolved oxygen. The carbonate system is also included in the model (Artioli *et al.*, 2012).

In our experiments, the Atlantic boundary values for nitrate, phosphate, silicate, and oxygen were taken from the World Ocean Atlas (Garcia *et al.*, 2013) and dissolved inorganic carbon from the Global Ocean Data Analysis Project (GLODAP) gridded data set (Key *et al.*, 2015; Lauvset *et al.*, 2016), while plankton and detritus variables were set to have zero fluxes at the Atlantic boundary.

3.3 | The assimilative component: NEMOVAR

The physical and biogeochemical data were assimilated on a daily basis into NEMO–ERSEM using NEMOVAR (Mogensen *et al.*, 2009; Waters *et al.*, 2015; King *et al.*, 2018), a Three-Dimensional Variational (3DVAR)–First Guess at Appropriate Time (FGAT) method. A weakly coupled data assimilation approach is taken, whereby physical and biogeochemical variables are assimilated separately. The innovations are calculated using the background from the coupled model; separate analyses are carried out for the physical and biogeochemical variables, with the resulting increments being added back into the relevant component of the coupled model.

Observations of total chlorophyll *a* were provided by European Space Agency's Ocean Colour–Climate Change Initiative (ESA OC-CCI) v4.2. The ESA OC-CCI global dataset is comprised of merged Medium Resolution Imaging Spectrometer (MERIS), *Aqua*–Moderate Resolution Imaging Spectroradiometer (MODIS), Sea-viewing Wide Field-of-view Sensor (SeaWiFS), and Visible Infrared Imaging Radiometer Suite (VIIRS) data, providing observations of water-leaving radiance in the visible domain, derived surface chlorophyll *a*, and inherent optical properties (Mélin *et al.*, 2017; Sathyendranath *et al.*, 2019). The data are provided daily on a 4 km × 4 km grid with gaps in cloud-contaminated and low-light pixels. The data were then superobbed by taking the median of any observations within a 7-km radius to match the resolution of the model and reduce the representation uncertainty. The assimilation of chlorophyll *a* was performed in log space, considering that the chlorophyll *a* concentration is approximately log-normally distributed (Campbell, 1995) and it is therefore natural to assume that the errors will also be log-normally distributed.

As well as ocean colour observations, observations of physical variables were also assimilated. These were sea-surface temperature (SST) data from the *GCOM-W1*/Advanced Microwave Scanning Radiometer 2 (AMSR-2), National Oceanic and Atmospheric Administration (NOAA)/Advanced Very High Resolution Radiometer (AVHRR), *MetOp*/AVHRR, *Meteosat Second Generation* (*MSG*)/Spinning Enhanced Visible and InfraRed Imager (SEVIRI), *Sentinel-3*/Sea and Land Surface Temperature Radiometer (SLSTR), and *Suomi–National Polar-orbiting Partnership* (NPP)/VIIRS satellite products, and temperature and salinity in situ data from the EN4 dataset (Good *et al.*, 2013), which includes profiles from Argo floats, fixed moored arrays, expendable bathythermographs (XBTs), conductivity, temperature, and depth sensors (CTDs), gliders, and marine mammals.

The background-error variances for the biogeochemical variables were derived in Skákala *et al.* (2018). These are based on a monthly climatology of log-transformed error variances obtained from the 100-member ensemble Kalman filter (EnKF) Proudman Oceanographic Laboratory Coastal Ocean Modelling System (POLCOMS)–ERSEM reanalysis of Ciavatta *et al.* (2018). These variances were then regularised and smoothed using the moving-averages algorithm and rescaled to the range 0.02–1.5 $\log_{10}(\mathrm{mg{\cdot}m^{-3}})$, so that the average ratio of background error to observation error was similar to that calculated in the region when assimilating OC-CCI data into a global coupled physical–biogeochemical assimilative system (Ford and Barciela, 2017).

In NEMOVAR, the background-error correlations are modelled using a diffusion operator (Weaver *et al.*, 2016). Successive applications of an implicit diffusion operator are used to simulate the matrix multiplication of an autoregressive correlation matrix. The method is easily able to account for complex boundary conditions and allows for geographically varying length-scales (Mirouze and Weaver, 2010) and multiple length-scales (Mirouze *et al.*, 2016). As the number of applications of the implicit diffusion operator tends to infinity, the modelled correlation function is Gaussian. In our application, we use 10 applications of the diffusion operator, which provides a good approximation to a Gaussian correlation function. For both the physical and biogeochemical variables, the horizontal correlation length-scales are specified a priori and are based on a combination of two different length-scales, a longer 111-km correlation scale (L_2) and a shorter length-scale (L_1) based on the first baroclinic Rossby radius of deformation, with a minimum value of 21 km set (King *et al.*, 2018). The correlation between two

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

points separated by Δx is given by

$$w \exp\left(-\frac{\Delta x^2}{L_1^2}\right) + (1-w)\exp\left(-\frac{\Delta x^2}{L_2^2}\right). \quad (14)$$

For the physical correlations, the weighting, w, between the two length-scales varies spatially (King *et al.*, 2018). For chlorophyll a, the weighting is set to 0.5.

For physics variables, NEMOVAR computes a 3D analysis using vertical correlation length-scales, which are a function of the mixed-layer depth in the background (King *et al.*, 2018). For chlorophyll a, a 2D surface analysis is computed, and these increments are applied equally throughout the mixed layer (Skákala *et al.*, 2018; 2020).

The values of the ocean-colour-derived chlorophyll a RMSE and bias are supplied by OC-CCI (Brewin *et al.*, 2017). These per pixel values are derived from in situ matchups that provide an estimate of the bias and RMSE as a function of OWT. Higher OWT values (indicative of turbid waters) tend to dominate around the coast throughout the year, but are also abundant in the more open oceans during the spring and summer. Lower OWTs (indicative of oligotrophic waters) tend to dominate in the open ocean from late summer. For each pixel of the ocean colour product, a daily weight is assigned to describe how well the pixel is characterised by each OWT (Jackson *et al.*, 2017). To enable enough samples, there is no seasonal variation in these statistics. The RMSE and bias values are converted to error standard deviations using the method of Ciavatta *et al.* (2016). To this estimate, a fixed value of 0.01 $[\log_{10}(\text{mg}\cdot\text{m}^{-3})]$ is added to account for representation uncertainty (Skákala *et al.*, 2018) not included in the OC-CCI uncertainties, whilst maintaining the average ratios of the background- to observation-error variances suggested by Ford and Barciela (2017). The estimate of the bias supplied by OC-CCI is also used to correct the mean value of the data before assimilation. The observation errors are assumed to be uncorrelated, so that $\mathbf{R}_a$ is diagonal.

4 | NEW COVARIANCE ESTIMATES

4.1 | The innovations

The operationally used error variances and correlations described above are referred to, hereafter, as the current $\mathbf{B}_a$ and $\mathbf{R}_a$. New estimates of $\mathbf{B}_a$ and $\mathbf{R}_a$ derived in this section are based primarily on the DBCP05 diagnostics. As discussed in Section 2, these are based on separating the innovation statistics out into the observation and background component. Figure 1 shows the mean and standard deviations of the innovation as a function of five bathymetry bins and month for the 2018 period. The bathymetry bins have been chosen so that they cover roughly equal numbers of grid points. It can be seen that the innovation mean and standard deviation have similar magnitude, especially in winter and late spring/early summer. The model appears to consistently underestimate the winter concentration of chlorophyll a and overestimates chlorophyll a during the spring bloom, compared with the satellite-derived estimates. The background biases (assuming the observations are unbiased) are therefore significant in a way not seen for the sea-surface temperature when similar analysis is performed (not shown).

It can be expected that the DBCP05 diagnostics will be sensitive to the choice of bins used to separate out the data. In choosing the bins, it is important that they describe the variability in the error statistics appropriately and also allow for a large enough sample within each bin. To compute the background-error statistics, we have partitioned a year's worth of innovation, residual, and increment statistics by month and by bathymetry (on-shelf (bottom depth < 250 m) and off-shelf (bottom depth > 250 m)); this partitioning is shown by the white contour in the lower panel of Figure 1. Using just two bins ensures a large sample size, which is especially crucial when estimating error correlations. To compute the observation-error statistics for $\log_{10}$ chlorophyll a derived from ocean colour, we have partitioned the same sample of innovations, residuals, and increments by month and the dominant (largest weight) OWT, as is done by Brewin *et al.* (2017). Partitioning the data differently to estimate $\mathbf{B}$ and $\mathbf{R}$ may aid the ability of the DBCP05 diagnostic to untangle the two different components of the total innovation covariances. Within these bins, we assume homogeneity, ergodicity, and isotropy over the given sample.

To ensure the new estimates of $\mathbf{B}_a$ and $\mathbf{R}_a$ are free from sampling noise, we impose a lower limit on the sample size (given by the number of superobbed ocean colour data within a given bin) when computing the error covariances. This is set at 100 for the error variances and 250 for the correlations. In practice, this lower limit is only ever met when computing the observation-error variances for very high and very low OWTs (see Section 4.3).

4.2 | Estimated background-error covariances

Within this section, we show the results from estimating new background-error covariances for surface $\log_{10}$ chlorophyll a. We do not adjust the vertical correlation, as ocean colour data do not provide information at depth.

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

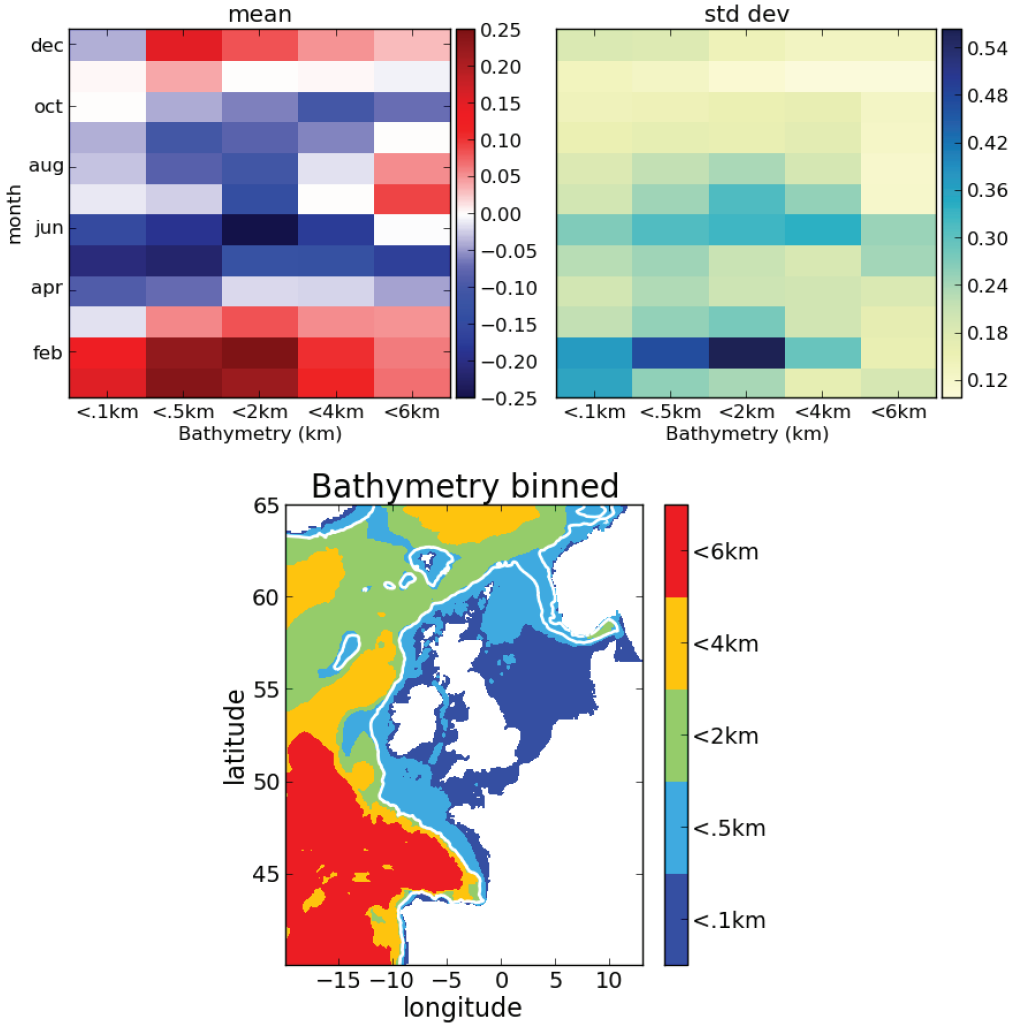

FIGURE 1 Mean of the innovations (top left) and standard deviation of the innovations (top right) in log chlorophyll a [$\log_{10}(\text{mg}\cdot\text{m}^{-3})$] as a function of bathymetry and month. The bathymetry bins (bottom) have been chosen so that they cover roughly equal numbers of grid points; the white contour marks the continental shelf defined by 250-m bathymetry [Colour figure can be viewed at wileyonlinelibrary.com]

Two estimates of both the variances and correlations are computed: one where biases are not included, as in Equation (8), and one in which the biases are included, as in Equation (11).

The NEMOVAR-ERSEM DA system has inputs of monthly varying background-error standard deviations for total chlorophyll a defined at each model grid point (Skákala *et al.*, 2018). We wish to keep some of the spatial variability in the current estimate of the standard deviations, but this is a higher resolution than can be achieved with two bathymetry bins. Therefore, we only update the magnitudes of the variances according to the bathymetry bins by performing the following steps.

1. First, we estimate the background-error standard deviations as a function of month and the two bathymetry bins, $\sigma^{\text{b}}_{\text{new}}(month, bathymetry)$, from a year's innovation and residual samples using the DBCP05 method (Equation 8). These have been estimated by removing all biases from the innovations and residuals. We also provide a separate estimate of the background-error standard deviations, which is derived including the absolute innovation biases:

$$\sigma^{\text{b}}_{\text{new,bias}}(month, bathymetry) = \sigma^{\text{b}}_{\text{new}}(month, bathymetry) + \|E[\mathbf{d}(month, bathymetry)]\|. \quad (15)$$

2. For each month and bathymetry bin, a perturbation to the currently used background-error standard deviations is found based on the mean value of the currently used background-error standard deviations for each bin:

$$\Delta\sigma^{\text{b}}(month, bathymetry) = \sigma^{b}_{\text{new}}(month, bathymetry) - \overline{\sigma^{b}_{\text{curr}}(month, bathymetry)}. \quad (16)$$

This is repeated for the $\sigma^{\text{b}}_{\text{new,bias}}$ estimate. Figure 2 shows how different the DBCP05 estimates of the

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

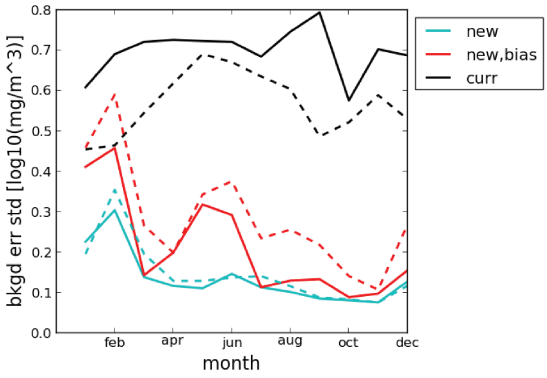

FIGURE 2 Comparison of new estimates of the background-error standard deviations for $\log_{10}$ chlorophyll [$\log_{10}$(mg·m^{-3})], shown in cyan (not accounting for innovation bias) and red (accounting for innovation bias), and the average currently used background-error standard deviations (black) as a function of month for the two bathymetry bins (on-shelf: solid, off-shelf: dashed) [Colour figure can be viewed at wileyonlinelibrary.com]

background-error standard deviations are from the currently used values as a function of month and bathymetry. Including the bias acts to inflate the error variances for the periods when the biases are largest (cf. Figure 1). In most seasons the DBCP05 estimate gives smaller values than those currently used, the main exception being for the off-shelf values in February when the innovation bias is included.

3. To provide a new estimate of monthly varying background-error standard deviations for total $\log_{10}$ chlorophyll a defined at each model grid point, the following is then performed:

$$\sigma^{\mathrm{b}}_{\mathrm{new}}(month, gridpoint) = \sigma^{\mathrm{b}}_{\mathrm{curr}}(month, gridpoint) + \Delta\sigma^{\mathrm{b}}(month, bathymetry). \quad (17)$$

This provides new estimates of the background-error standard deviations that have the same spatial variability as the previous values within a given bathymetry bin, but with updated magnitudes. To avoid a sudden jump between the bins, the perturbations added to $\sigma^{\mathrm{b}}_{\mathrm{new}}$ are linearly interpolated as a function of bathymetry in the range 125 m< *bathymetry* <3,125 m. A lower limit of 0.05 [$\log_{10}$(mg·m^{-3})] is imposed. Again this is repeated for the $\sigma^{\mathrm{b}}_{\mathrm{new,bias}}(month, bathymetry)$ estimate to give $\sigma^{\mathrm{b}}_{\mathrm{new,bias}}(month, gridpoint)$.

Examples of the output of this process for March–August are shown in Figure 3. We see that we maintain the spatial variability in the error standard deviations present in the currently used values on and off the shelf (marked by the black contour), but have changed the magnitudes.

Using the same DBCP05 output used to update the background-error standard deviations, we also update the background-error correlations. Figure 4 shows raw estimates of the correlations estimated for each month and the two bathymetry bins. Again, two separate estimates of the correlations are provided, for when the biases are (red) and are not (cyan) included. Including the bias acts to increase the error correlation length-scales for the periods when the biases are largest (cf. Figure 1). These estimates can be compared with the currently used correlations (black lines in Figure 4), which do not vary with bathymetry or month. The new estimates have a large amount of monthly variability, but in general the length-scales are longer than those currently used, especially when the biases are included.

The current correlations (black lines) are based on those used for the assimilation of SSTs. We performed the same analysis for SST data using data from the VIIRS instruments on board *Suomi–NPP*. This verified that the correlations used for SST are a good estimate, with the assumption of monthly invariant length-scales being a good approximation. The bias in the SST innovations is also much less significant. The difference between the SST and chlorophyll a correlation estimates is therefore interesting, as it had been assumed that it is the physics which drives the correlations in chlorophyll a. These results show that there are other processes responsible for the correlations and they are much more complex for chlorophyll a than for SST.

To use these new estimates of the correlations, Equation (14) has been fitted to the correlations using curve_fit from the Python scipy.optimize package, summarising the curves in terms of three parameters for each month: w, L_1, L_2 ($w_{\mathrm{bias}}, L_{1,\mathrm{bias}}, L_{2,\mathrm{bias}}$ when bias is included). These parameters are provided in Table 1. Recall that the currently used parameters are $w = 0.5$, $L_1 \approx 25$ km and $L_2 = 111$ km. There is a fair amount of monthly variability in the optimised parameters, but, comparing the yearly average, we see that when the innovation bias is included this has the effect of giving less weight to the smaller length-scale (w) and increasing the larger length-scale (L_2). The impact of including the innovation bias has negligible effect on the smaller length-scale (L_1). This is similar if we compare the optimised parameters for on the shelf ($B < 0.25$ km) and off the shelf ($B > 0.25$ km). Off the shelf, less weight is given to the smaller length-scale and the larger length-scale is increased compared with on the shelf. Again there is little change in the smaller length-scale.

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

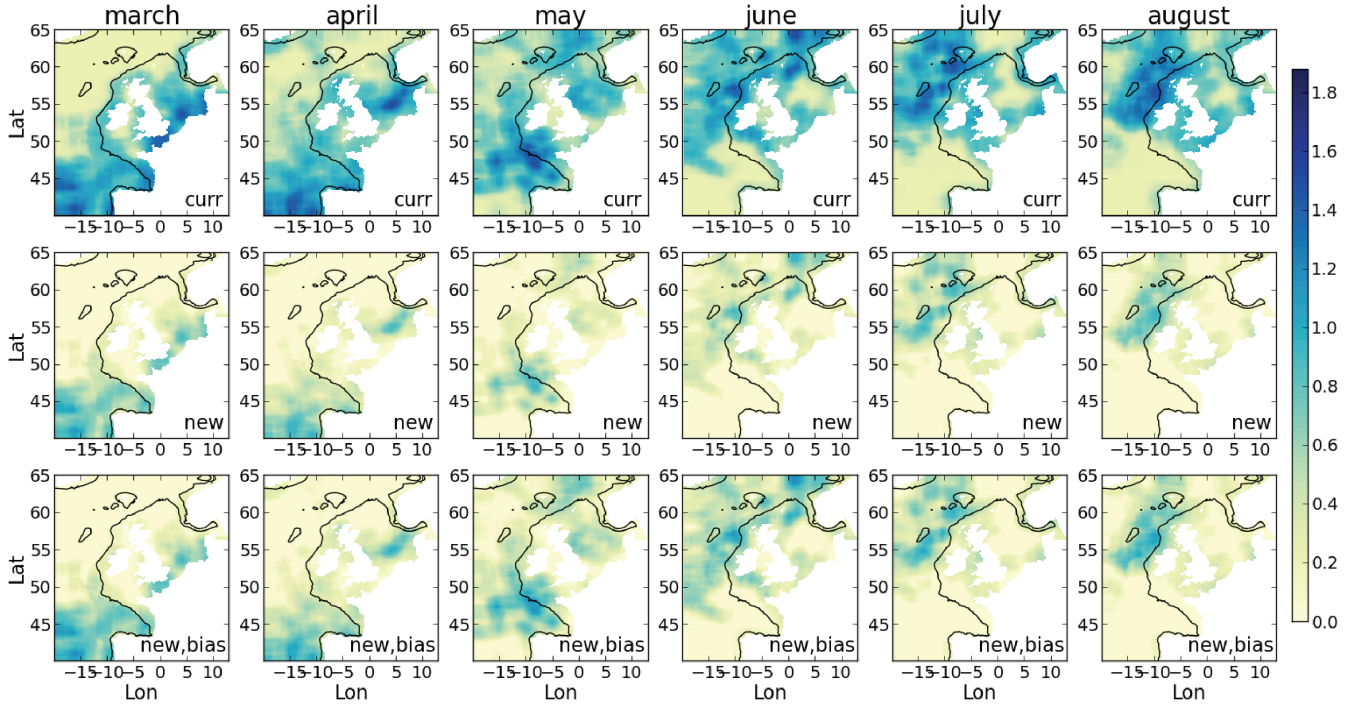

FIGURE 3 Comparison of new estimates of the background-error standard deviations for $\log_{10}$ chlorophyll a [$\log_{10}(\text{mg}\cdot\text{m}^{-3})$], shown in the middle row (not accounting for innovation bias) and bottom row (accounting for innovation bias), and the currently used background-error standard deviations in the top row as a function of latitude and longitude. From left to right: estimates for March–August. [Colour figure can be viewed at wileyonlinelibrary.com]

4.3 | Estimated ocean colour observation-error variances

Here, we use the DBCP05 diagnostic to produce new estimates of the ocean-colour error standard deviations that, by default, include the representation, instrument, and forward model uncertainty. The current estimates of observation-error variances, estimated from in situ data, are a constant function of OWT over time. The increased sample size due to comparing with the model instead of in situ observations means, again, that we are able to provide monthly estimates of the standard deviations. We do not update the bias estimates as we have assumed that the innovation bias is dominated by the model bias so have no additional information on this. We also do not provide estimates of the error correlations. This is because a good estimate of $\mathbf{B}_a$ is needed to estimate the correlations in $\mathbf{R}_a$. This will be addressed in Section 6 once we have a better estimate of the background-error correlations.

Estimates of the ocean-colour error standard deviations as a function of OWT are given in Figure 5 for each month. Values are not available for every OWT, due to the sample size being too small. These values can be compared with the currently used values provided by OC-CCI. As with the background errors, the currently used values seem generally to overestimate the ocean-colour error standard deviations. As discussed in Section 2, the OC-CCI estimates do not include the uncertainty due to the mismatch in scales between the ocean colour data and the modelled surface chlorophyll a, nor do they account for uncertainty in the observation operator. It is therefore interesting that the OC-CCI estimates appear to overestimate the ocean-colour error standard deviations. This could perhaps be explained by the mismatch in the scales represented by the ocean colour data and the in situ data used to provide the OC-CCI estimates, such that the OC-CCI estimate of the ocean-colour error standard deviation has the undesirable property of including the representation uncertainty of the in situ data. The OC-CCI estimates were also estimated for the ocean colour data before they were superobbed. If the errors in the data are uncorrelated, then we can expect the superobbing to reduce the error variances. This highlights an advantage of the DBCP05 diagnostic in that it provides an estimate of the observation-error variances for the actual observation assimilated, without the need to hypothesise the missed source of uncertainties or the effect of processing the observations before assimilation.

One important consideration is how the ratio of the background-error standard deviations to the observation-error standard deviations changes for the new estimates compared with the current values. From

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

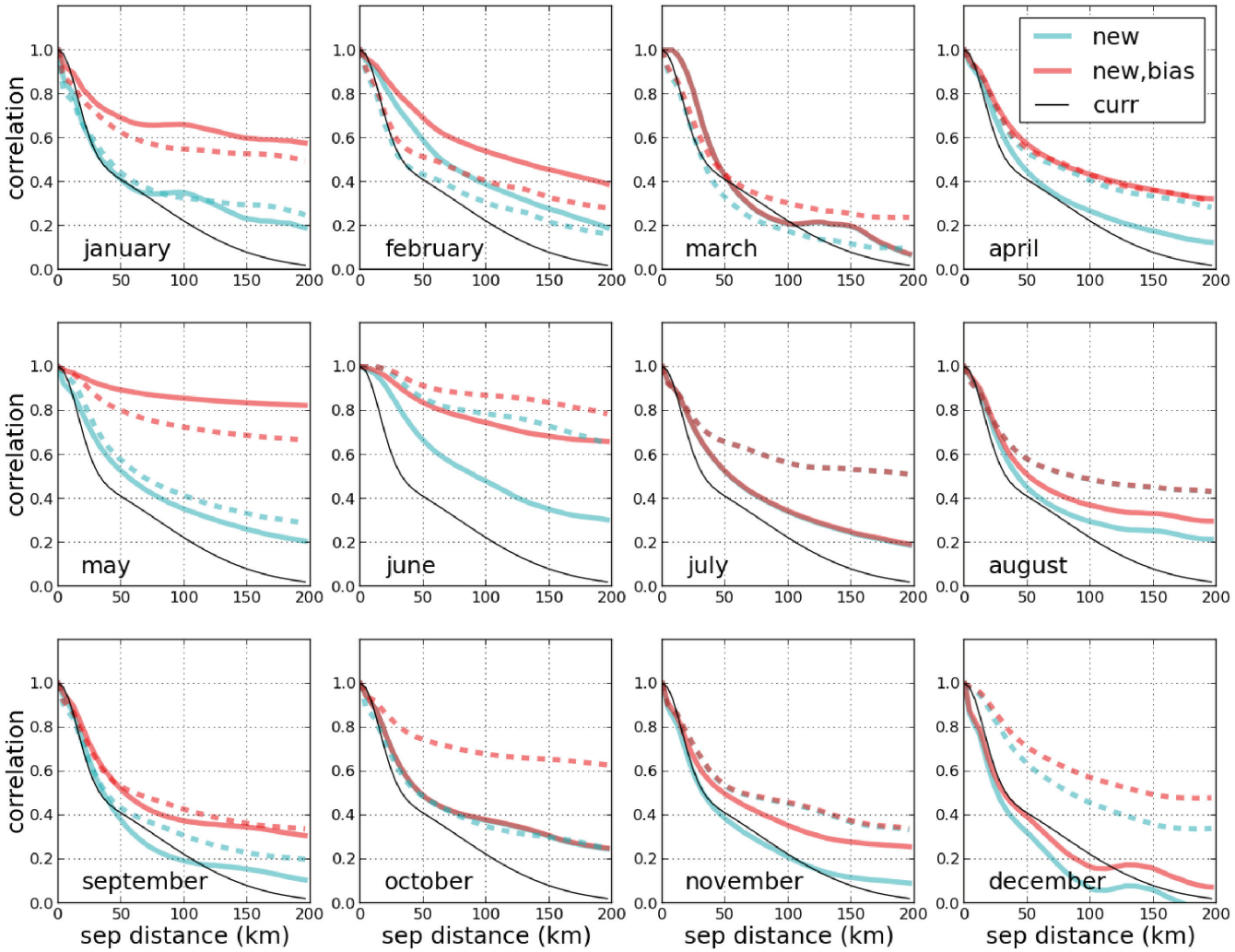

FIGURE 4 Comparison of raw output from DBCP05 diagnostics of background-error correlations in $\log_{10}$ total chlorophyll *a* for each month and bathymetry bin (on-shelf: solid, off-shelf: dashed), shown in cyan (not accounting for innovation bias) and red (accounting for innovation bias). The black lines are the currently used correlations, which do not vary with bathymetry or month. [Colour figure can be viewed at wileyonlinelibrary.com]

Figures 2 and 5, we see that in general the estimates of the background-error variances are reduced more than the observation-error variances, so that in general less weight will be given to the observations. However, this is very variable and for some months and locations the opposite is true. If the weighting given to the observations is not justified, this may result in initialisation shock, in which the analysis output from assimilation of the observations is not a consistent solution to the model equations and so a large error quickly manifests in the forecast as the model tries to regain balance. This will be investigated in Section 5, where we look at the impact of the new error statistics on forecast skill.

These estimates of the error standard deviations are converted to per-pixel values using daily maps of weighted per-pixel OWT. As each pixel is characterised by more than one OWT, we are still able to provide a value of the error standard deviations from each pixel despite the fact that for some OWTs we did not have a large enough sample to estimate the error standard deviation (cf. Figure 5). Example composites of the currently used error standard deviations (top rows) and the new estimates (bottom rows) are given in Figure 6. With the current estimates, it can be seen that there is a large variability in the spatial distribution of the standard deviations each month, but not in the magnitudes. In contrast, there is much more seasonal variability in both the magnitudes and spatial distribution in the new estimates.

This monthly variability in the magnitude of the error standard deviations could plausibly be due to changes in the optical properties, for example zenith angle or changes in the representation uncertainty. It could also be an artefact of the DBCP05 diagnostic effectively leaving an imprint of the background-error standard deviations.

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Optimised parameter values fitting Equation (14) to the estimated correlations in Figure 4

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Average
On shelf: $0< B <0.25$ km													
w	0.63	0.54	0.72	0.70	0.60	0.49	0.60	0.67	0.72	0.61	0.69	0.50	0.62
w_{bias}	0.34	0.48	0.72	0.58	0.15	0.27	0.60	0.61	0.60	0.61	0.63	0.61	0.52
L_1	24.5	41.9	41.4	37.0	36.3	49.8	37.2	34.9	32.9	33.5	27.9	16.0	35.3
$L_{1,\text{bias}}$	26.2	50.3	41.5	40.8	41.7	56.4	37.5	35.7	35.9	33.6	32.7	22.6	37.9
L_2	283	225	204	246	260	271	244	313	196	358	186	75.9	214
$L_{2,\text{bias}}$	683	425	208	499	1000	640	249	402	425	362	397	138	452
Off shelf: $0.25\text{ km} < B$													
w	0.63	0.56	0.80	0.55	0.55	0.16	0.40	0.49	0.66	0.60	0.49	0.58	0.54
w_{bias}	0.42	0.53	0.69	0.53	0.27	0.11	0.40	0.49	0.56	0.31	0.49	0.46	0.44
L_1	24.0	20.8	33.1	33.7	39.4	42.2	28.4	30.5	30.5	29.5	26.9	52.6	32.6
$L_{1,\text{bias}}$	25.9	23.4	34.2	34.2	42.7	45.8	28.4	30.5	31.8	31.7	27.0	52.9	34.0
L_2	290	187	281	313	301	388	488	511	303	308	294	567	352
$L_{2,\text{bias}}$	538	283	501	340	723	566	488	512	452	740	298	761	517

Note: The weights are given to two significant figures and length-scales are in metres to three significant figures.

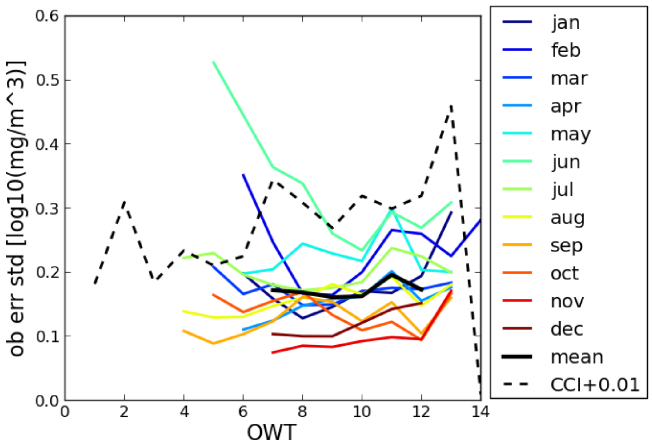

FIGURE 5 Comparison of new estimates of ocean-colour-derived $\log_{10}$ chlorophyll a error standard deviations [$\log_{10}(\text{mg}\cdot\text{m}^{-3})$] as a function of OWT for each month (coloured lines, see legend) and the currently used values provided by OC-CCI (black dashed line). [Colour figure can be viewed at wileyonlinelibrary.com]

5 | TRIALS WITH NEW UNCERTAINTY ESTIMATES

To understand the impact of the new estimates of the error statistics for the observations and background, a year's worth of analyses and six-day forecasts have been initiated daily throughout 2018, using the same model and assimilation setup used to compute the new covariances. Four different experiments have been performed (also summarised in Table 2).

- **Control**: original uncertainties currently used operationally.
- **New-nobias**: New **R** and **B** without accounting for biases (Equations 7 and 8).
- **New-biasstd**: New **R** and **B** accounting for bias in the background-error standard deviations only (Equations (7) and (11) for background-error standard deviations and Equation (8) for background-error correlations).
- **New-bias**: New **R** and **B**accounting for bias, (Equations 7 and 11).

To recap: the main differences between the control and new statistics are essentially a change in the ratios of the error variances, the temporal variability in observation-error variances (cf. Figure 5), and the temporal variability in the background-error correlations (cf. Figure 4).

5.1 | Impact on analysis

Within this section, we first look at the impact the new error covariance estimates have on the analysis, which is then used to initialise the forecasts.

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

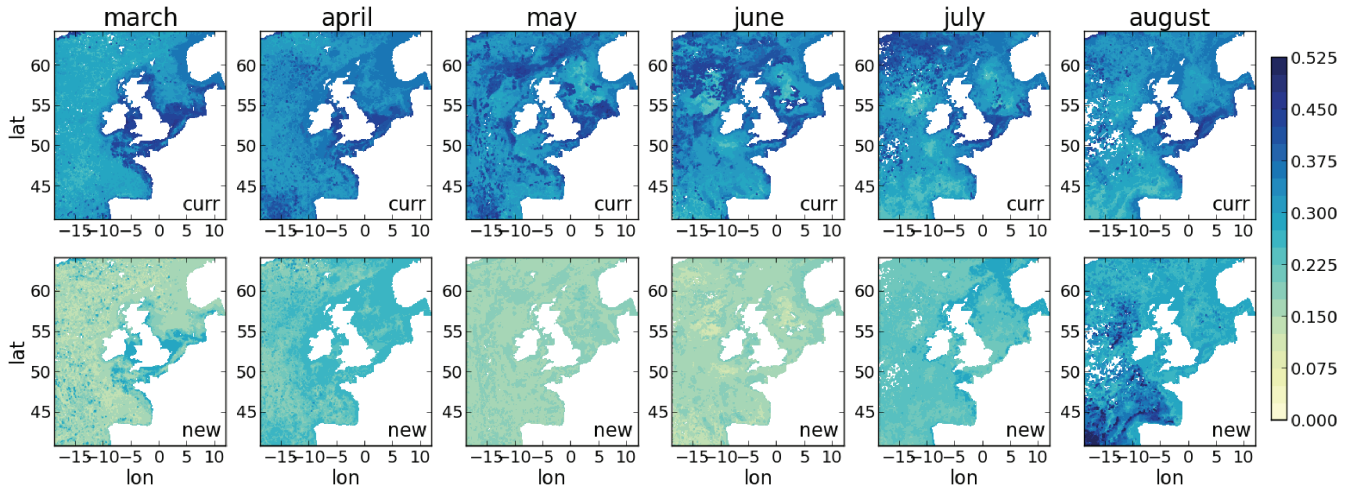

FIGURE 6 Comparison of monthly composites of the new estimates of ocean-colour-derived $\log_{10}$ chlorophyll *a* error standard deviations [$\log_{10}(\text{mg}\cdot\text{m}^{-3})$] (bottom row) with the currently used values (top row). Left to right: estimates for March–August. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Description of variances and correlations used in the trials with new uncertainty estimates (Section 5)

Experiment name	Ob. err. var.	Ob. err. corr.	Bkgd err. var.	Bkgd err. corr., see Equation (14)
Control	diag($\mathbf{R}_a$): Currently used, function of OWT.	None	diag($(\mathbf{HBH})^T_a$): Currently used, fn of month and grid point.	$w = 0.5$, L_1 based on Rossby radius, $L_2 = 111$ km.
New-nobias	diag($\mathbf{R}_{new}$): Estimated using DBCP05 diagnostic, Equation (7), fn of OWT and month.	None	diag($(\mathbf{HBH})^T_{new}$): Estimated using DBCP05 diagnostic, Equation (8), fn of month and grid point.	Monthly varying and shelf-dependent. w, L_1, and L_2 from Table 1.
New-biasstd	diag($\mathbf{R}_{new}$): Estimated using DBCP05 diagnostic, Equation (7), fn of OWT and month.	None	diag($(\mathbf{HBH})^T_{new,bias}$): Estimated using DBCP05 diagnostic, Equation (11), fn of month and grid point.	Monthly varying and shelf-dependent. w, L_1, and L_2 from Table 1.
New-bias	diag($\mathbf{R}_{new}$): Estimated using DBCP05 diagnostic, Equation (7), fn of OWT and month.	None	diag($(\mathbf{HBH})^T_{new,bias}$): Estimated using DBCP05 diagnostic, Equation (11), fn of month and grid point.	Monthly varying and shelf-dependent. w_{bias}, $L_{1,bias}$, and $L_{2,bias}$ from Table 1.

5.1.1 | Validation against ocean colour data

Figure 7 shows a time series for the average chlorophyll *a* of the different experiments over the whole model domain (left) and split into on- and off-shelf regions (middle and right respectively). These can be compared with the time series of the OC-CCI data (black line). A small negative winter bias in the analysis and a larger positive bias in spring/early summer and late summer can be seen on comparing the control experiment (red) with the observations. Using the new statistics when the biases are not included in the background-error correlations (green and yellow lines) can be seen to reduce this bias consistently. The timing of the spring bloom was already captured fairly accurately by the control; however, the magnitude was far too strong. The new statistics do not change the timing of the spring bloom across the region, but do help to reduce its magnitude, bringing it more in line with that observed.

The performance when the biases are/are not accounted for in the variances is very similar (green/yellow lines). The only difference between the two analyses is the magnitude of the background-error variances used. This suggests it is the seasonal variability in the observation-error variances and background-error correlations that is important, rather than the overall change in ratios of the background- to observation-error variances. This is consistent with the findings of Skákala *et al.* (2018), who found little sensitivity when tuning the ratio of the background- to observation-error variances.

The behaviour of the analysis time series is much more sensitive to accounting for the biases in the correlations (cyan lines). In many places the increase in correlation length-scales is seen to be beneficial. However, during the

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

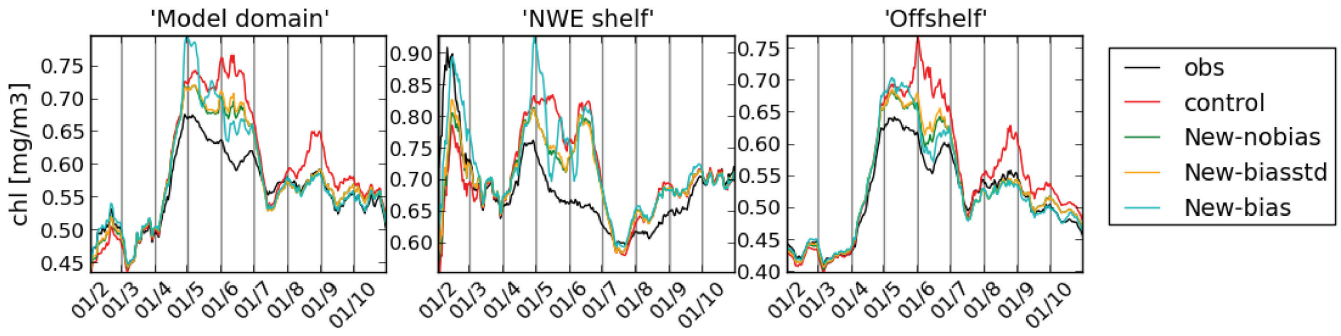

FIGURE 7 Time series of the analysis of median chlorophyll *a* [mg·m^{-3}] for the different experiments (see legend) compared with the OC-CCI observations (black line), for (left to right) the full model domain, on-shelf, and off-shelf. Smoothed over 15-day running averages. [Colour figure can be viewed at wileyonlinelibrary.com]

spring/summer bloom, the large oscillations in chlorophyll *a* on the shelf are unrealistic. This may be due to the larger length-scales in the **B** matrix spreading the information in the observations too far and impacting the stability of the model. This is supported by spatial maps of analysis increments, which indicate that the longer correlation length-scales appear, on occasion, to be spreading information from very localised features more widely than is appropriate.

Figure 8 separates out the analysis time series into different regions on the NWE shelf for April–July only (the months in which the worst performance from including the biases in the correlations was seen in Figure 7). The behaviour of chlorophyll *a* and the characteristics of the model bias are seen to vary across the different regions. We see that the unrealistic behaviour of the analysis when accounting for the biases in the correlations is mostly confined to the Norwegian Trench and North Sea regions, whereas accounting for the biases in the correlations is seen to be beneficial in the performance of the analysis in the NW approach, Irish Sea, and the English Channel. This suggests that, in order to gain the benefits of accounting for the biases in the correlations, the **B** matrix should be specified separately for the different regions to allow for spatial variability in the bias characteristics.

5.1.2 | Validation against independent glider data

So far we have validated the analysis against the OC-CCI observations that are also assimilated. Glider observations, although limited in number, provide an independent validation for the analyses of chlorophyll *a* and oxygen, as well as allowing us to see the impact below the surface. A number of gliders were deployed during 2018 as part of the Alternative Framework to Assess Marine Ecosystem Functioning in Shelf Seas (AlterECO) programme (http://projects.noc.ac.uk/altereco/). The AlterEco gliders that operated in the central North Sea between May and August 2018 provided data for temperature, salinity, chlorophyll *a* (derived from fluorescence), and oxygen concentrations (Skákala *et al.*, 2021).

Figure 9 shows the analysis validated against the chlorophyll *a* observations provided by the gliders. In the control simulation (Figure 9b), a large positive bias in chlorophyll *a* can be seen around the base of the mixed layer during the summer months. As discussed in Skákala *et al.* (2021), this may occur as a consequence of model dynamical adjustment to surface chlorophyll *a* assimilation, that is, resulting from the imbalance between the incoming irradiance and nutrient abundance beneath the mixed layer and the limited ability of ocean-colour data assimilation to correct chlorophyll *a* below the mixed layer. The improvement in the bias compared with the control for the new statistics is shown in Figure 9c–e. When the bias is not included in the background-error correlations (Figure 9c,e), we see a clear reduction in the positive bias throughout May and June but some increases in the bias afterwards. When the bias is included in the background-error correlations (Figure 9d), we see a similar effect to that in the surface chlorophyll *a*, that is, a mixture of large improvements along with large degradations. Again, this is indicative of the larger length-scales in the **B** matrix impacting the stability of the model. Recall that the North Sea, the location of the gliders, was the region most sensitive to the inclusion of biases in the background-error correlations (cf. Figure 8).

Figure 10 shows the analysis validated against the oxygen observations provided by the gliders. Accurate monitoring and forecasts of oxygen are paramount to identifying oxygen depletion events (i.e., hypoxia: Vaquer Sunyer and Duarte, 2008), which can have disastrous impacts on marine life. We see that the control (Figure 10b) has a positive bias at the majority of depths and times, increasing in magnitude from May onwards. Assimilating the surface

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

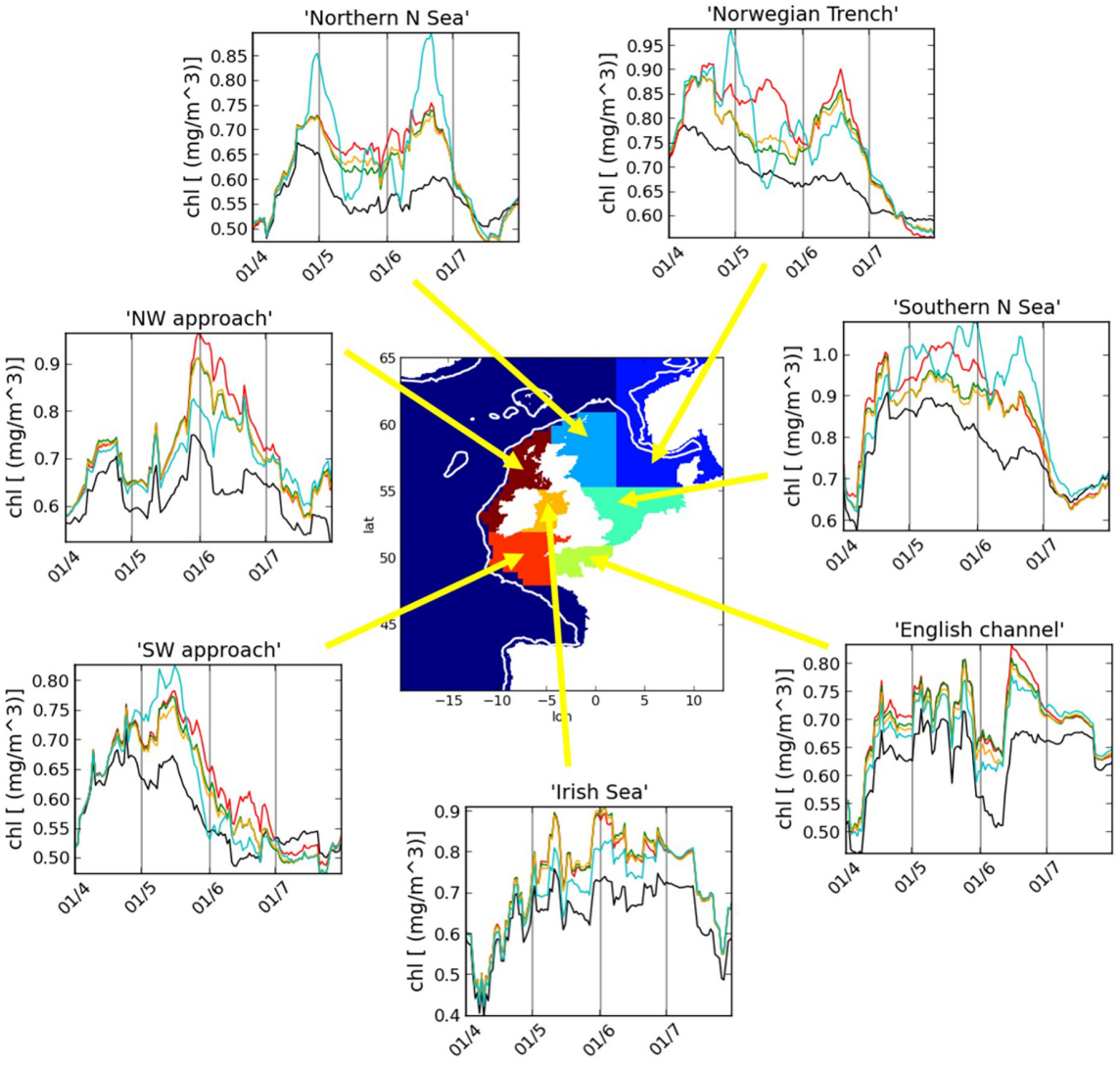

FIGURE 8 As in Figure 7, but separated out into different on-shelf regions from April–July. [Colour figure can be viewed at wileyonlinelibrary.com]

chlorophyll *a* observations with the new error covariances is seen to reduce the large positive bias from May onwards. Again, including the biases in the background-error correlations can lead to some degradations over the control (Figure 10e).

5.2 | Impact on forecast skill

Within this section we show how the changes to the analysis (mostly improvements) are propagated in the skill of the forecast. Figure 11 shows the bias versus root-median-square difference (RMSD) in six-day on-shelf forecasts of surface chlorophyll *a* validated against ocean-colour-derived surface chlorophyll *a* for February–September. As with the analysis time series, we again see that the behaviour of the forecast skill when the biases are/are not accounted for in the variances is very similar (green and yellow lines), but much more sensitivity is seen when accounting for the biases in the correlations. Note the different scales on each axis; the grey box indicates the same value ranges in each subplot.

We would expect that accounting for the biases in the background-error covariances would reduce the bias in analysis and forecast but increase the RMSD. This is generally true; however, again the improvement over the control is mixed and, in general, not accounting for the biases in the correlations gives more reliable improvements over the control. Similar conclusions can also be drawn from the forecast skill computed for the off-shelf region (not shown).

The forecast skill with lead time can provide a good indication of initialisation shock due to overfitting to the

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

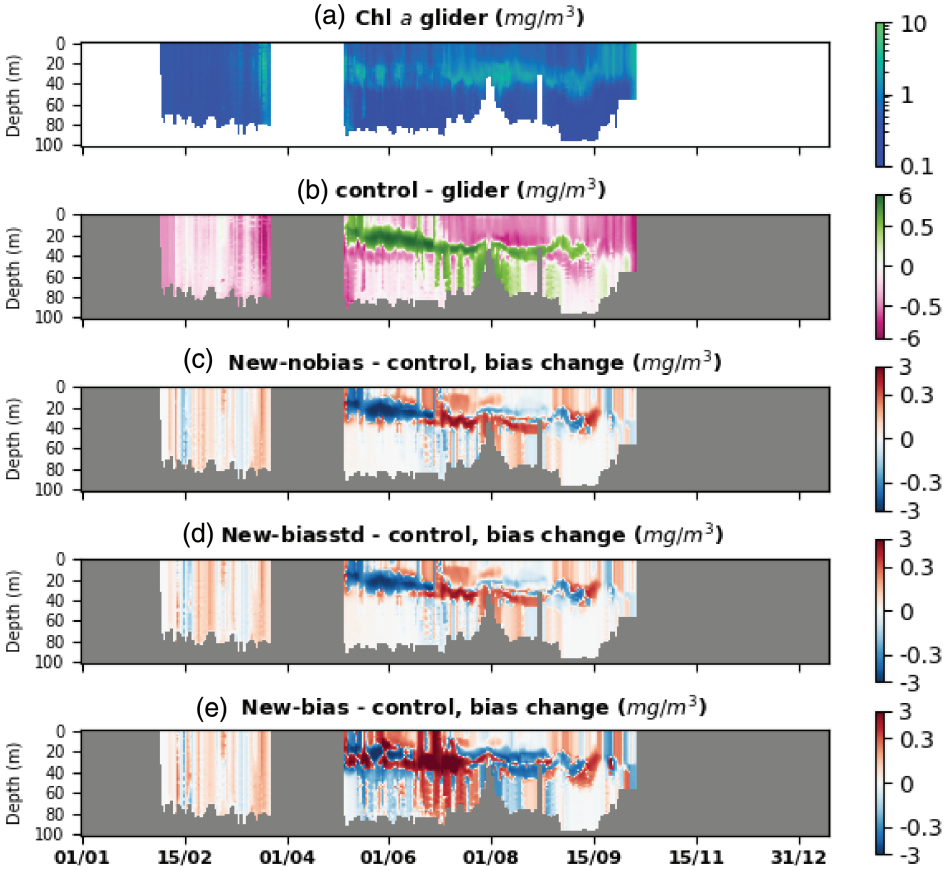

FIGURE 9 Analysis validated against unassimilated glider observations of chlorophyll *a* [mg·m^{-3}]. (a) Chlorophyll *a* observed by glider, (b) bias in control, (c–e) change in absolute bias for the three experiments (positive is an increase in bias, negative is a reduction in bias). [Colour figure can be viewed at wileyonlinelibrary.com]

observations and an imbalanced analysis. As the degradation in forecast skill with lead time is consistent across the different experiments, there is no obvious sign of an increased initialisation shock with the new error covariance statistics. Mostly the improvements at the initial time are maintained throughout the six-day forecast.

Another interesting feature seen in these forecast skill plots is the sensitivity to accounting for the biases in the background-error correlations for the July period. Figure 4 showed that there was little difference in the correlations in July, due to the bias being negligible. Therefore this sensitivity must be propagated from May and June, when the differences in correlations when accounting for biases were large.

The forecast skill was also validated against the glider observations (not shown). It was found that the improvements seen at the analysis time were again maintained in the six-day forecast.

6 | CONVERGENCE OF UNCERTAINTY ESTIMATES

Given the improvements in the analysis and forecast seen with the new estimates of **B** and **R**, obvious questions are as follows.

- Can we do any better if we run the DBCP05 diagnostics again?
- Do we expect the estimates of **B** and **R** to converge?

As discussed in Section 2, the presence of significant biases in the innovations, which, although reduced, still remain, disrupts the usual assumption of convergence of the DBCP05 diagnostics. The answers to these two questions are therefore not obvious.

Figure 12 shows the new standard deviation estimates for **B** for the first and second iteration of the DBCP05 diagnostics for the three experiments as a function of month for on the shelf only. The first iteration comprises the new estimates used in the three experiments (see Figure 2). For the second iteration, we provide values for when the biases were and were not included in the variance estimation using the output from the experiments described in the previous section. We see that, in the case in which the biases *were not* added to the standard deviations in the first iteration (New-nobias, left panel), the first and second iterations are very similar when the bias was again *not* included in the second iteration. Similarly, when the biases *were* added to standard deviations in the first iteration (New-biasstd and New-bias, middle and right panels), the first and second iterations are very similar when the bias *is* included in the second iteration. Therefore,

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

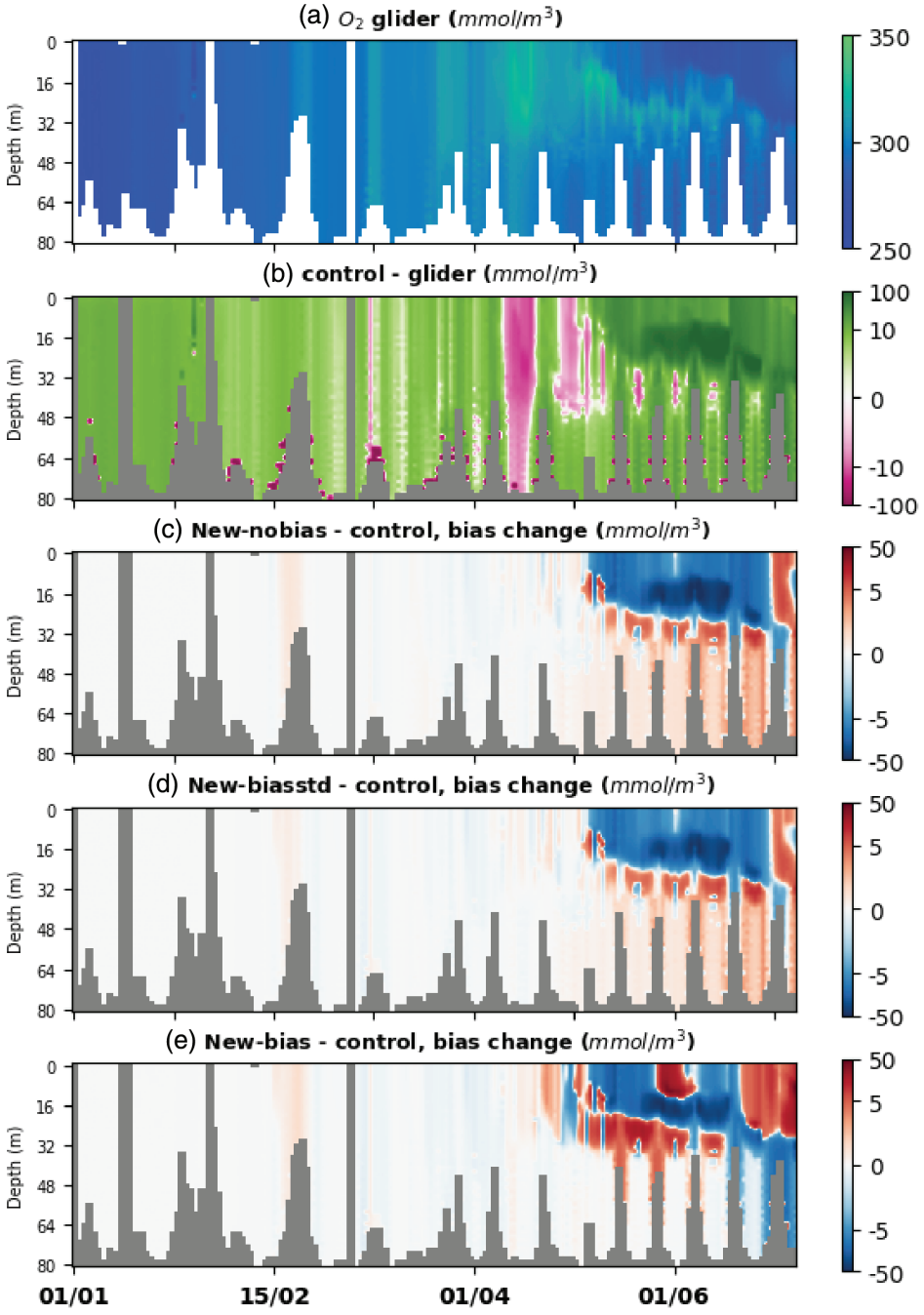

FIGURE 10 As in Figure 9 but for glider observations of oxygen [mmol·m^{-3}]. [Colour figure can be viewed at wileyonlinelibrary.com]

surprisingly, the error variances are close to convergence, especially when compared with the original values used (black line).

Figure 13 shows the new variance estimates for **R** as a function of OWT averaged over the year for the first and second iterations of the DBCP05 diagnostics for three experiments. Note that all experiments used the same **R**, so there is only one line for the first iteration. This time there is a clearer shift in the standard deviations estimated between the first and second iterations, although interestingly the second estimates are largely insensitive to the different **B** matrices used in the different experiments. The outputs from the second iteration indicate a greater weighting to the observations than was achieved using the estimates of **B** and **R** from the first iteration.

Figure 14 shows the new correlation estimates for **B** for the first and second iterations of the DBCP05 diagnostics for the three experiments for May, June, and July for on the shelf only. The first iterations are the values used in the three experiments (see Table 1). For the second iteration, we provide values for cases in which the biases were and were not included in the correlation estimate. Similarly to the error standard deviations, we see that, in the case when the biases *were not* added to the correlations in the

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

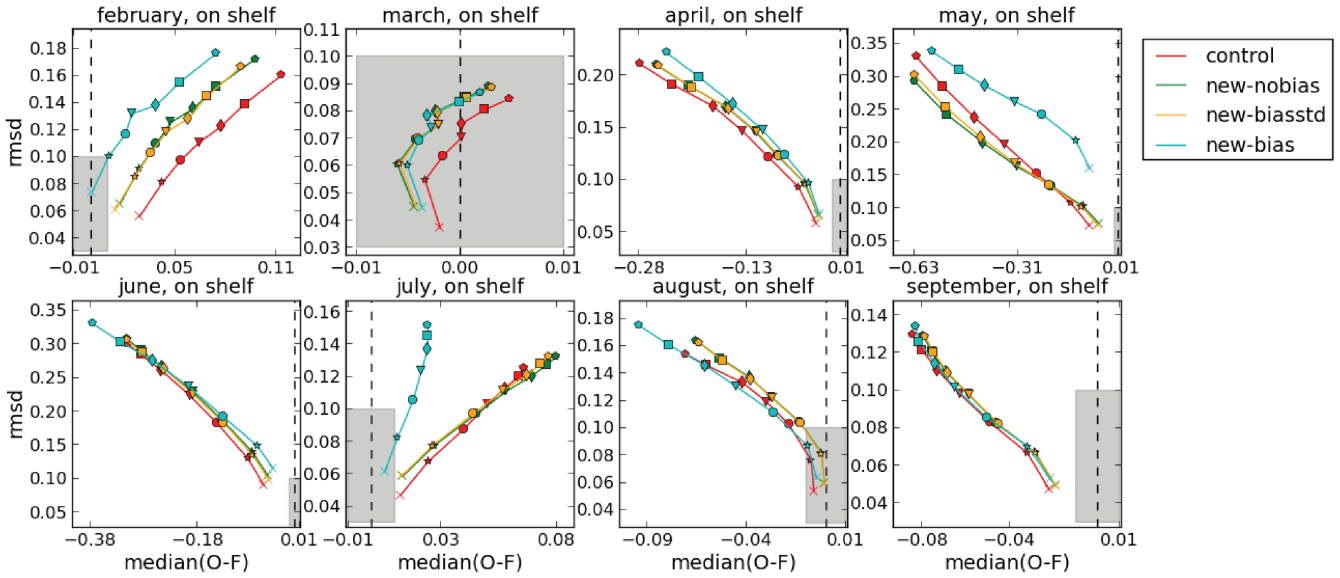

FIGURE 11 On-shelf bias versus RMSD of forecast error validated against ocean colour data. Each symbol represents a forecast of different length in integer days; the skill of the analysis is indicated by the cross and the skill of day 6 is indicated by the pentagon. Note the different scales on each axis; the vertical dashed line indicates zero bias and the grey box indicates the same value ranges in each subplot. [Colour figure can be viewed at wileyonlinelibrary.com]

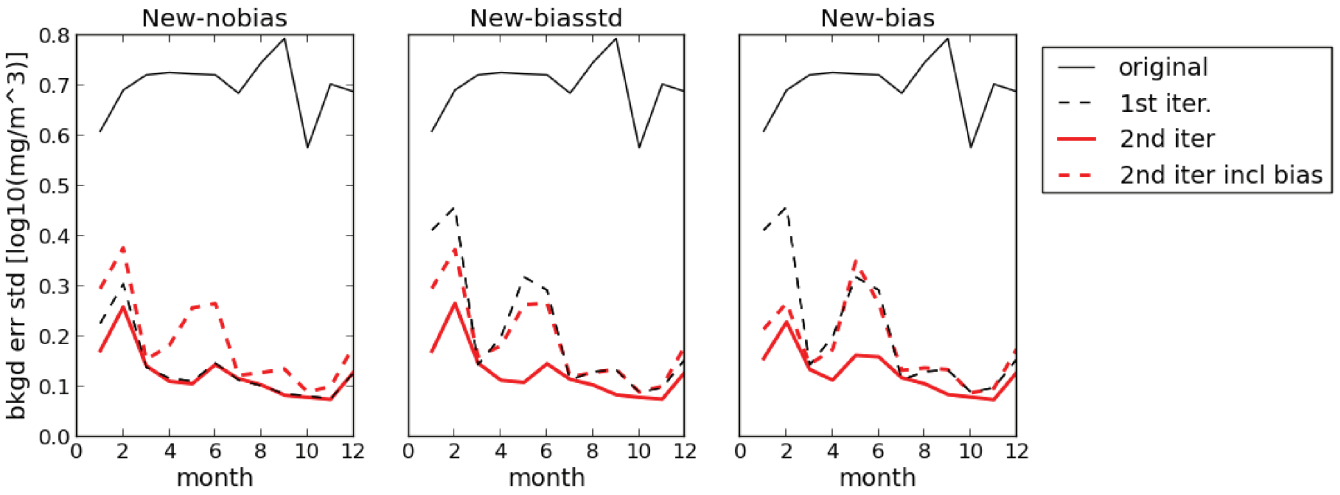

FIGURE 12 Estimates of the background-error standard deviations for $\log_{10}$ chlorophyll a [$\log_{10}$ (mg·m^{-3})] as a function of month for on the shelf. We compare the original values used (black line, same as the black line in Figure 2), estimates derived from the first iteration of the DBCP05 diagnostic (black dashed line, same as the red and cyan lines in Figure 2, depending on whether the bias was or was not included in the experiment), and estimates derived from the second iteration of the DBCP05 diagnostic with and without the bias included (red lines, solid when the bias is not included and dashed when the bias is included). [Colour figure can be viewed at wileyonlinelibrary.com]

first iteration (New-nobias and New-biasstd, top and middle panels), the first and second iterations are very similar when the bias again *is not* included in the second iteration (red solid line). Similarly, when the biases *were* added to correlations in the first iteration (New-bias, bottom row of panels), the first and second iterations are very similar when the bias *is* included in the second iteration (red dashed line). Therefore, again the correlation estimates are close to convergence, as long as the biases are treated consistently.

Also shown in Figure 14 are the new correlation estimates for **R** for the second iteration of the DBCP05 diagnostics for the three experiments. Note that the original and first iterations both assumed no correlations were present in **R**. In each case, we see that correlations in the ocean colour data are significant (correlations

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

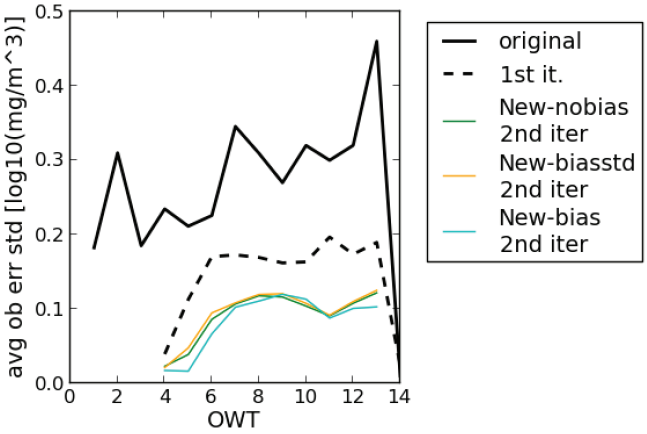

FIGURE 13 Estimates of the observation-error standard deviations averaged over the year as a function of OWT for $\log_{10}$ chlorophyll *a* [$\log_{10}$ mg·m^{-3}] on the shelf. We compare the original values used (black line, same as the black dashed line in Figure 5), estimates derived from the first iteration of the DBCP05 diagnostic (black dashed line, same as black line in Figure 5), and estimates derived from the second iteration of the DBCP05 diagnostic for the three different experiments. [Colour figure can be viewed at wileyonlinelibrary.com]

greater than 0.2 are seen at tens of km separation distances). The inclusion of these error correlations can be expected to allow for the use of much finer scale information in the observations (Rainwater *et al.*, 2015; Fowler *et al.*, 2018). Therefore, benefits from including correlations when assimilating ocean colour data could be expected, although this is not currently easy to implement. However, for May, in particular, it is clear that the estimate of the correlations in the observations depends upon the **B** matrix used within the assimilation, with longer length-scales diagnosed when the length-scales in **B** used in the assimilation are longer. This makes it hard to have confidence in the results, as **R** should be independent of **B**. The sensitivity of the estimates of observation-error correlations to length-scales in the **B** matrix using the DBCP05 diagnostics are discussed extensively in Waller *et al.* (2016).

7 | SUMMARY AND DISCUSSION

The aim of this work was to derive new estimates of **B** and **R** for the assimilation of ocean-colour-derived chlorophyll *a* into a marine biogeochemistry model. Using the method of DBCP05 allowed us to derive new estimates of the error statistics as a function of month as well as bathymetry and optical water type (OWT) for the background and observations respectively. Within these bins, we made strong assumptions of ergodicity, isotropy, and homogeneity. The new covariances derived from these consistency diagnostics differed from those currently used operationally in the following ways.

1. The observation-error variances varied as a function of month as well as OWT, rather than just OWT,

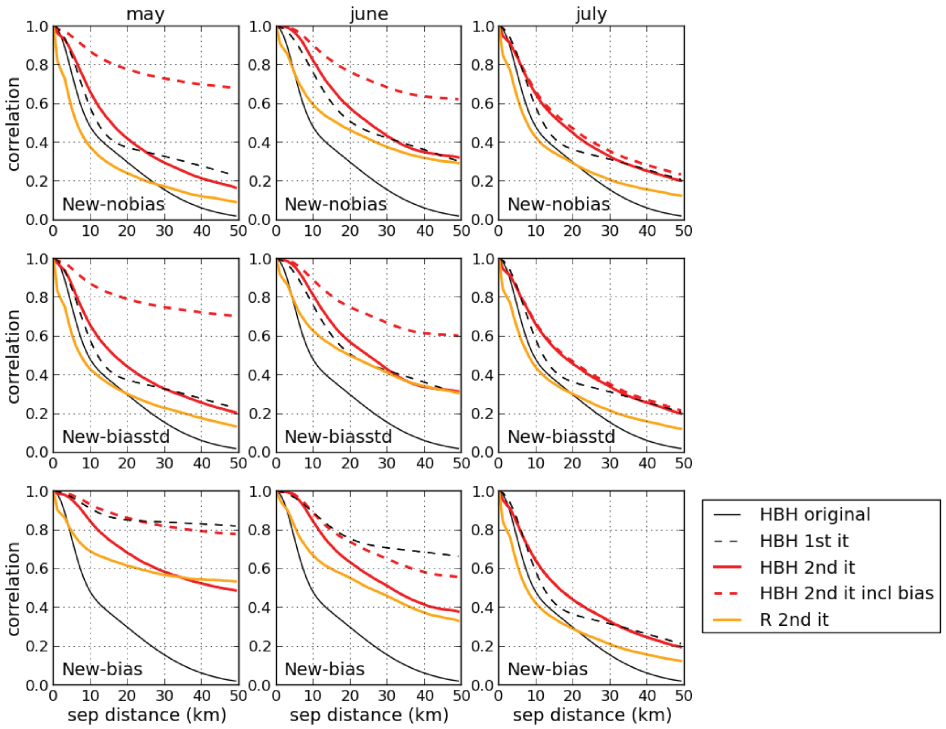

FIGURE 14 Estimates of the background-error correlations for $\log_{10}$ chlorophyll *a* on the shelf, May–July. We compare the original values used (black line, same in all panels and same as the black line in Figure 4), estimates derived from the first iteration of the DBCP05 diagnostic (black dashed line here the same as the purple solid and dashed lines in Figure 4, depending on whether the bias was or was not included in the experiment), and estimates derived from the second iteration of the DBCP05 diagnostic for the three different experiments (red lines, solid when the bias is not included and dashed when the bias is included). Also plotted in yellow are the estimates of observation-error correlations, which were originally and in the first iteration assumed to be uncorrelated. [Colour figure can be viewed at wileyonlinelibrary.com]

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and included the effect of representation uncertainty implicitly.
2. The background-error variances were also updated and no longer assumed that observation- to background-error variances satisfied a particular ratio.
3. The background-error correlations were no longer based on those used for the assimilation of SSTs, allowing for monthly varying estimates of the correlation length-scales on and off the shelf.
4. The significant bias in the model was accounted for by including the bias within the background-error covariance.

Including the biases in the background-error covariance resulted in the error variances and correlation length-scales being inflated, so that the analysis was less constrained by the biased model. The increase in correlation length-scales also implies that less small-scale information is extracted from the observations, allowing them to correct biases over a larger region. The alternative approach of removing the biases in the innovations before they are assimilated (e.g., Lea *et al.*, 2008) was not investigated, due to the difficulty in estimating the temporal and spatial variability of the biases accurately.

We tested the impact of the new error covariance matrices within the framework of the operational CMEMS system for the NWE shelf. We found that in general, when the covariances do not account for the biases, large improvements in the skill of the analysis and forecast can be gained by updating the background-error covariances and observation-error variances. Specifically, we showed the potential for a reduction in bias of area-averaged chlorophyll *a* compared with the satellite data in most regions for most months.

The results are apparently insensitive to whether the background-error variances were inflated to include the model bias. This suggests that it is the seasonal variability in the covariances that is important. More sensitivity was achieved by increasing the correlation length-scales to account for model bias. However, the results were mixed, giving large improvements as well as degradations in the analysis. This may be due to the different nature of the bias in the different regions of the shelf, so that the treatment of model bias should not be applied in the same way across the whole shelf. The assumption of isotropy of the errors on the shelf may also not be justified. Future work could relax this assumption by estimating the correlations separately in the latitudinal and longitudinal directions.

No attempts were made to update the vertical correlations in **B**, as the application of the DBCP05 diagnostic within this article is limited to the space of the OC-CCI dataset (i.e., surface total chlorophyll *a*). However, validation against independent glider observations of chlorophyll *a* and oxygen showed that the impact on analysis skill seen at the surface largely followed through to the impact at lower depths, although the impact is more subtle. Again, the impact on variables below the surface was much more mixed when the biases were included in **B**. This suggests that the larger length-scales may be causing imbalances within the model. However, the improvements seen at the analysis time were found to carry through to the six-day forecast, with no evidence of initialisation shock with the new error covariance matrices. This suggests that the new error covariance matrices were more consistent with the data and the improvements at the analysis time were not simply due to overfitting the observations.

Repeated application of the DBCP05 diagnostic could be expected to provide better estimates of **B** and **R**, although this is complicated by the presence of significant (although reduced after the first iteration) biases in the innovations. We found that applying the diagnostic a second time resulted in little change in the estimate of the background-error variances and correlations, as long as the biases in the innovations were treated consistently. A second iteration of the diagnostic, however, resulted in a large reduction to the estimate of the observation-error variances. This could be due to the fact we have neglected the observation-error correlations, which may be significant for the OC-CCI dataset. At present it is not possible to account for observation-error correlations within NEMOVAR and so inflation of the observation-error variances is common practice (e.g., Hilton *et al.*, 2009), in parallel to thinning and spatially averaging (superobbing) the data. Inclusion of observation-error correlations in the future would allow for finer scale information to be extracted from the dataset, which is not possible by simply inflating the error variances (Rainwater *et al.*, 2015; Fowler *et al.*, 2018).

A remaining question is how transferrable these estimates are to other years and datasets. Cossarini *et al.* (2019) found that online estimation of the error variances of biogeochemical-Argo float data was necessary to account for the flow-dependent nature of the representation error. However, online estimation of both **B** and **R** would be prohibitively costly. The suitability of these estimated error covariances for other years and datasets should therefore be evaluated. However, given the improvements shown in the analysis and forecast, we recommend the new error covariances be trialled for operational use in their current form. The new estimates of OC-CCI uncertainties will also be of interest to the wider ocean colour community.

Further improvements could also be expected in combining the monthly varying **B** defined in this work with a dynamically evolving ensemble representation of the background-error covariances, commonly known as

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

hybrid DA (Clayton *et al.*, 2013). Hybrid DA aims to overcome both the issues with a predefined **B** (Skákala *et al.*, 2020) and the limitations of a small ensemble representation of **B**. This is clearly more expensive than just using the monthly varying **B** defined here, as an ensemble system must also be developed and maintained. The benefit also relies on the accuracy of both the predefined and ensemble representations of **B** (Bowler *et al.*, 2017). Therefore, these new estimates of **B** will also be beneficial to these future developments.

AUTHOR CONTRIBUTIONS

Alison M. Fowler: conceptualization; data curation; formal analysis; investigation; methodology; resources; software; validation; visualization; writing – original draft; writing – review and editing. **Jozef Skákala:** conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization; writing – review and editing. **David Ford:** data curation; formal analysis; investigation; methodology; resources; software; validation; writing – review and editing.

ACKNOWLEDGEMENTS

The authors thank Stefano Ciavatta (PML) for his guidance in the conception and initial stages of this work and Matt Martin (Met Office) for his insightful feedback on the final article. We also thank the two anonymous reviewers, whose thoughtful comments greatly improved the article. This study was funded as part of NERC's support of the National Centre for Earth Observation, contract number PR140015.

DATA AVAILABILITY STATEMENT

The ocean colour data were provided by the European Space Agency Climate Initiative "Ocean Color" (https://esa-oceancolour-cci.org/). The glider data used for validation are publicly available on https://www.bodc.ac.uk/data/published_data_library/catalogue/. The model was forced by the atmospheric ERA5 product of The European Centre for Medium-Range Weather Forecasts (ECMWF, https://www.ecmwf.int/). The river forcing data used by the model were prepared by Sonja van Leeuwen and Helen Powley as part of UK Shelf Seas Biogeochemistry programme (contract no. NE/K001876/1) of the NERC and the Department for Environment Food and Rural Affairs (DEFRA). We acknowledge use of the MONSooN system, a collaborative facility supplied under the Joint Weather and Climate Research Programme, a strategic partnership between the Met Office and the NERC. The different outputs for the experiments are stored on the MONSooN storage facility MASS and can be obtained upon request.

ORCID

Alison M. Fowler https://orcid.org/0000-0003-3650-3948

REFERENCES

Aiken, J., Alvain, S., Barlow, R., Bouman, H., Bracher, A., Brewin, R.J.W., Bricaud, A., Brown, C.W., Ciotti, A.M., Clementson, L., Craig, S.E., Devred, E., Hardman Mountford, N., Hirata, T., Hu, C., Kostadinov, T.S., Lavender, S., Loisel, H., Moore, T.S., Morales, J., Moulin, C., Mouw, C.B., Nair, A., Raitsos, D., Roesler, C., Sathyendranath, S., Shutler, J.D., Sosik, H.M., Soto, I., Stuart, V., Subramaniam, A. and Uitz, J. (2014) *Phytoplankton functional types from space* Vol. 15. Dartmouth, NS: International Ocean Colour Coordinating Group (IOCCG). (Reports and Monographs of the International Ocean Colour Coordinating Group (ed. S. Sathyendranath)).

Artioli, Y., Blackford, J.C., Butenschön, M., Holt, J.T., Wakelin, S.L., Thomas, H., Borges, A.V. and Allen, J.I. (2012) The carbonate system in the North Sea: sensitivity and model validation. *Journal of Marine Systems*, 102-104, 1–13.

Auligné, T., McNally, A.P. and Dee, D.P. (2007) Adaptive bias correction for satellite data in a numerical weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 133, 631–642.

Baretta, J., Ebenhoh, W. and Ruardij, P. (1995) The European Regional Seas Ecosystem Model, a complex marine ecosystem model. *Netherlands Journal of Sea Research*, 33, 233–246.

Baretta Bekker, J., Baretta, J. and Ebenhöh, W. (1997) Microbial dynamics in the marine ecosystem model ERSEM II with decoupled carbon assimilation and nutrient uptake. *Journal of Sea Research*, 38, 195–211. (European Regional Seas Ecosystem Model II).

Bathmann, K. (2018) Justification for estimating observation-error covariances with the Desroziers diagnostic. *Quarterly Journal of the Royal Meteorological Society*, 144, 1965–1974.

Blackford, J. (1997) An analysis of benthic biological dynamics in a North Sea Ecosystem Model. *Journal of Sea Research*, 38, 213–230. (European Regional Seas Ecosystem Model II).

Borges, A., Schiettecatte, L.S., Abril, G., Delille, B. and Gazeau, F. (2006) Carbon dioxide in European coastal waters. *Estuarine, Coastal and Shelf Science*, 70, 375–387. (Trace gases in the European coastal zone).

Bormann, N., Bonavita, M., Dragani, R., Eresmaa, R., Matricardi, M. and McNally, A. (2016) Enhancing the impact of IASI observations through an updated observation-error covariance matrix. *Quarterly Journal of the Royal Meteorological Society*, 142, 1767–1780.

Boss, E., Swift, D., Taylor, L., Brickley, P., Zaneveld, R., Riser, S., Perry, M.J. and Strutton, P.G. (2008) Observations of pigment and particle distributions in the western North Atlantic from an autonomous float and ocean color satellite. *Limnology and Oceanography*, 53, 2112–2122.

Bowler, N.E., Clayton, A.M., Jardak, M., Jermey, P.M., Lorenc, A.C., Wlasak, M.A., Barker, D.M., Inverarity, G.W. and Swinbank, R. (2017) The effect of improved ensemble covariances on hybrid variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 143, 785–797.

Brewin, R.J.W., Ciavatta, S., Sathyendranath, S., Jackson, T., Tilstone, G., Curran, K., Alrs, R.L., Cummings, D., Brotas, V., Organelli,

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

E., Dall'Olmo, G. and Raltsos, D.E. (2017) Uncertainties in ocean color estimates of Chlorophyll for Phytoplankton Groups. *Frontiers in Marine Science*, 4, 104.

Brewin, R.J.W., Sathyendranath, S., Lange, P.K. and Tilstone, G. (2014) Comparison of two methods to derive the size structure of natural populations of phytoplankton. *Deep Sea Research Part I: Oceanographic Research Papers*, 85, 72–79.

Bruggeman, J. and Bolding, K. (2014) A general framework for aquatic biogeochemical models. *Environmental Modelling & Software*, 61, 249–265.

Butenschön, M., Clark, J., Aldridge, J.N., Allen, J.I., Artioli, Y., Blackford, J., Bruggeman, J., Cazenave, P., Ciavatta, S., Kay, S., Lessin, G., van Leeuwen, S., van der Molen, J., de Mora, L., Polimene, L., Sailley, S., Stephens, N. and Torres, R. (2016) ERSEM 15.06: a generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels. *Geoscientific Model Development*, 9, 1293–1339.

Campbell, J.W. (1995) The lognormal distribution as a model for bio optical variability in the sea. *Journal of Geophysical Research: Oceans*, 100, 13237–13254.

Campbell, W.F., Satterfield, E.A., Ruston, B. and Baker, N.L. (2017) Accounting for correlated observation error in a dual formulation 4D variational data assimilation system. *Monthly Weather Review*, 145, 1019–1032.

Carmillet, V., Brankart, J.M., Brasseur, P., Drange, H., Evensen, G. and Verron, J. (2001) A singular evolutive extended Kalman filter to assimilate ocean color data in a coupled physical biochemical model of the North Atlantic ocean. *Ocean Modelling*, 3, 167–192.

Ciavatta, S., Brewin, R.J.W., Skákala, J., Polimene, L., de Mora, L., Artioli, Y. and Allen, J.I. (2018) Assimilation of ocean color plankton functional types to improve marine ecosystem simulations. *Journal of Geophysical Research: Oceans*, 123, 834–854.

Ciavatta, S., Kay, S., Saux Picart, S., Butenschön, M. and Allen, J.I. (2016) Decadal reanalysis of biogeochemical indicators and fluxes in the Northwest European shelf sea ecosystem. *Journal of Geophysical Research: Oceans*, 121, 1824–1845.

Ciavatta, S., Torres, R., Saux Picart, S. and Allen, J.I. (2011) Can ocean color assimilation improve biogeochemical hindcasts in shelf seas?. *Journal of Geophysical Research: Oceans*, 116, C12043.

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.

Cossarini, G., Mariotti, L., Feudale, L., Mignot, A., Salon, S., Taillandier, V., Teruzzi, A. and D'Ortenzio, F. (2019) Towards operational 3D Var assimilation of chlorophyll Biogeochemical Argo float data into a biogeochemical model of the Mediterranean Sea. *Ocean Modelling*, 133, 112–128.

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

Desroziers, G., Berre, L., Chapnik, B. and Poli, P. (2005) Diagnosis of observation, background and analysis error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131, 3385–3396.

Edwards, K.P., Barciela, R. and Butenschön, M. (2012) Validation of the NEMO ERSEM operational ecosystem model for the North-west European continental shelf. *Ocean Science*, 8, 983–1000.

Ford, D. and Barciela, R. (2017) Global marine biogeochemical reanalyses assimilating two different sets of merged ocean colour products. *Remote Sensing of Environment*, 203, 40–54. (Earth observation of essential climate variables).

Ford, D.A., Edwards, K.P., Lea, D., Barciela, R.M., Martin, M.J. and Demaria, J. (2012) Assimilating globcolour ocean colour data into a pre operational physical biogeochemical model. *Ocean Science*, 8, 751–771.

Fowler, A., Dance, S. and Waller, J. (2018) On the interaction of observation and a priori error correlations. *Quarterly Journal of the Royal Meteorological Society*, 144, 48–62.

Garcia, H.E., Locarnini, R.A., Boyer, T.P., Antonov, J.I., Baranova, O.K., Zweng, M.M., Reagan, J.R. and Johnson, D.R. (2013). World ocean atlas 2013. Volume 4: Dissolved inorganic nutrients (phosphate, nitrate, silicate). In S. Levitus and A. Mishonov (Eds.), *NOAA Atlas NESDIS 76*. Silver Spring, MD: National Oceanographic Data Center.

Gauthier, P., Du, P., Heilliette, S. and Garand, L. (2018) Convergence issues in the estimation of interchannel correlated observation errors in infrared radiance data. *Monthly Weather Review*, 146, 3227–3239.

Geider, R., MacIntyre, H. and Kana, T. (1997) Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a: carbon ratio to light, nutrient limitation and temperature. *Marine Ecology Progress Series*, 148, 187–200.

Gejadze, I., Oubanas, H. and Shutyaev, V. (2017) Implicit treatment of model error using inflated observation-error covariance. *Quarterly Journal of the Royal Meteorological Society*, 143, 2496–2508.

Good, S.A., Martin, M.J. and Rayner, N.A. (2013) En4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of Geophysical Research: Oceans*, 118, 6704–6716.

Groom, S., Sathyendranath, S., Ban, Y., Bernard, S., Brewin, R., Brotas, V., Brockmann, C., Chauhan, P., Choi, J.k., Chuprin, A., Ciavatta, S., Cipollini, P., Donlon, C., Franz, B., He, X., Hirata, T., Jackson, T., Kampel, M., Krasemann, H., Lavender, S., Pardo Martinez, S., Mélin, F., Platt, T., Santoleri, R., Skakala, J., Schaeffer, B., Smith, M., Steinmetz, F., Valente, A. and Wang, M. (2019) Satellite ocean colour: current status and future perspective. *Frontiers in Marine Science*, 6, 485.

Henson, S.A., Dunne, J.P. and Sarmiento, J.L. (2009) Decadal variability in North Atlantic phytoplankton blooms. *Journal of Geophysical Research: Oceans*, 114, C04013.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.-N. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049.

Hilton, F., Collard, A., Guidard, V., Randriamampianina, R. and Schwaerz, M. (2009). Assimilation of IASI radiances at European NWP centres. In: Proceedings of Workshop on the assimilation of IASI data in NWP, ECMWF, 6–8 May 2009, Reading, UK.

Hollingsworth, A. and Lönnberg, P. (1986) The statistical structure of short range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus A*, 38A, 111–136.

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Hoteit, I., Triantafyllou, G. and Petihakis, G. (2005) Efficient data assimilation into a complex, 3 D physical biogeochemical model using partially local Kalman filters. *Annales Geophysicae*, 23, 3171–3185.

Howes, K.E., Fowler, A.M. and Lawless, A.S. (2017) Accounting for model error in strong constraint 4D var data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 143, 1227–1240.

Jackson, T., Sathyendranath, S. and Mélin, F. (2017) An improved optical classification scheme for the ocean colour essential climate variable and its applications. *Remote Sensing of Environment*, 203, 152–161.

Jahnke, R.A. (2010) *Global Synthesis, Carbon and Nutrient Fluxes in Continental Margins*, (pp. 597–615). Berlin: Springer Verlag.

Janjic, T., Bormann, N., Bocquet, M., Carton, J.A., Cohn, S.E., Dance, S.L., Losa, S.N., Nichols, N.K., Potthast, R., Waller, J.A. and Weston, P. (2017) On the representation error in data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 144, 1257–1278.

Kalnay, E. (2003) *Atmospheric Modeling, Data Assimilation and Predictability*. New York: Cambridge University Press.

Key, R.M., Olsen, A., van Heuven, S., Lauvset, S.K., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F.F. and Suzuki, T. (2015) *Global ocean data analysis project, version 2 (GLODAPv2)*. Oak Ridge National Laboratory, Carbon Dioxide Information Analysis Center. Earth System Science Data. https://doi.org/10.5194/essd-2015-42

King, R.R., While, J., Martin, M.J., Lea, D.J., Lemieux Dudon, B., Waters, J. and O'Dea, E. (2018) Improving the initialisation of the Met Office operational shelf-seas model. *Ocean Modelling*, 130, 1–14.

Lauvset, S.K., Key, R.M., Olsen, A., van Heuven, S., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F.F., Suzuki, T. and Watelet, S. (2016) A new global interior ocean mapped climatology: the $1^\circ \times 1^\circ$ GLODAP version 2. *Earth System Science Data*, 8, 325–340.

Lea, D.J., Drecourt, J.P., Haines, K. and Martin, M.J. (2008) Ocean altimeter assimilation with observational and model bias correction. *Quarterly Journal of the Royal Meteorological Society*, 134, 1761–1774.

Lenhart, H.-J., Mills, D.K., Baretta-Bekker, H., van Leeuwen, S.M., van der Molen, J., Baretta, J.W. Blaas, M., Desmit, X., Kühn, W., Lacroix, G., Los, H.J., Ménesguen, A., Neves, R., Proctor, R., Ruardij, P., Skogen, M.D., Vanhoutte-Brunier, A., Villars, M.T., and Wakelin, S.L. (2010) Predicting the consequences of nutrient reduction on the eutrophication status of the North Sea. *Journal of Marine Systems*, 81, 148–170.

Lorenc, A.C., Ballard, S.P., Bell, R.S., Ingleby, N.B., Andrews, P.L.F., Barker, D.M., Bray, J.R., Clayton, A.M., Dalby, T., Li, D., Payne, T.J. and Saunders, F.W. (2000) The Met. Office global three dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 126, 2991–3012.

Lutz, M.J., Caldeira, K., Dunbar, R.B. and Behrenfeld, M.J. (2007) Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean. *Journal of Geophysical Research: Oceans*, 112, C10011.

Madec, G., Bourdallé-Badie, R., Chanut, J., Clementi, E., Coward, A., Ethé, C., Iovino, D., Lea, D., Lévy, C., Lovato, T., Martin, N., Masson, S., Mocavero, S., Rousset, C., Storkey, D., Müeller, S., Nurser, G., Bell, M., Samson, G., Mathiot, P., Mele, F. and Moulin, A. (2015). NEMO ocean engine, version 3.6 stable. Technical report, ISPL. Note du Pôle de modélisation de l'institut pierre Simon Laplace No. 27. [Accessed 17th January 2019].

Mattern, J.P., Edwards, C.A. and Moore, A.M. (2018) Improving variational data assimilation through background and observation error adjustments. *Monthly Weather Review*, 146, 485–501. https://doi.org/10.1175/MWR-D-17-0263.1

McClain, C.R. (2009) A decade of satellite ocean color observations. *Annual Review of Marine Science*, 1, 19–42.

McEwan, R., Kay, S. and Ford, D. (2021). CMEMS-NWS-QUID-004-002. (Copernicus Marine Environment Monitoring Service—QUID).

Mélin, F., Vantrepotte, V., Chuprin, A., Grant, M., Jackson, T. and Sathyendranath, S. (2017) Assessing the fitness for purpose of satellite multi mission ocean color climate data records: a protocol applied to OC CCI chlorophyll a data. *Remote Sensing of Environment*, 203, 139–151. (Earth observation of essential climate variables).

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.

Mirouze, I. and Weaver, A.T. (2010) Representation of correlation functions in variational assimilation using an implicit diffusion operator. *Quarterly Journal of the Royal Meteorological Society*, 136, 1421–1443.

Mogensen, K., Balmaseda, M., Weaver, A., Martin, M. and Vidard, A. (2009) NEMOVAR: a variational data assimilation system for the NEMO ocean model. *ECMWF Newsletter*, 120, 17–22.

Nerger, L. and Gregg, W.W. (2007) Assimilation of SeaWiFS data into a global ocean biogeochemical model using a local SEIK filter. *Journal of Marine Systems*, 68, 237–254.

O'Dea, E., Furner, R., Wakelin, S., Siddorn, J., While, J., Sykes, P., King, R., Holt, J. and Hewitt, H. (2017) The CO5 configuration of the 7 km Atlantic Margin Model: large scale biases and sensitivity to forcing, physics options and vertical resolution. *Geoscientific Model Development*, 10, 2947–2969.

Pauly, D., Christensen, V., Guénette, S., Pitcher, T.J., Sumaila, U.R., Walters, C.J., Watson, R. and Zeller, D. (2002) Towards sustainability in world fisheries. *Nature*, 418, 689–695.

Rainwater, S., Bishop, C.H. and Campbell, W.F. (2015) The benefits of correlated observation errors for small scales. *Quarterly Journal of the Royal Meteorological Society*, 141, 3439–3445.

Sathyendranath, S., Brewin, R.J., Brockmann, C., Brotas, V., Calton, B., Chuprin, A., Cipollini, P., Couto, A.B., Dingle, J., Doerffer, R., Donlon, C., Dowell, M., Farman, A., Grant, M., Groom, S., Horseman, A., Jackson, T., Krasemann, H., Lavender, S., Martinez Vicente, V., Mazeran, C., Mélin, F., Moore, T.S., Müller, D., Regner, P., Roy, S., Steele, C.J., Steinmetz, F., Swinton, J., Taberner, M., Thompson, A., Valente, A., Z uhlke, M., Brando, V.E., Feng, H., Feldman, G., Franz, B.A., Frouin, R., Gould, R.W., Hooker, S.B., Kahru, M., Kratzer, S., Mitchell, B.G., Muller Karger, F.E., Sosik, H.M., Voss, K.J., Werdell, J. and Platt, T. (2019) An ocean-colour time series for use in climate studies: the experience of the ocean-colour climate change initiative (OC-CCI). *Sensors*, 19, 4285.

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Siddorn, J. and Furner, R. (2013) An analytical stretching function that combines the best attributes of geopotential and terrain following vertical coordinates. *Ocean Modelling*, 66, 1–13.

Simonin, D., Waller, J.A., Ballard, S.P., Dance, S.L. and Nichols, N.K. (2019) A pragmatic strategy for implementing spatially correlated observation errors in an operational system: an application to Doppler radial winds. *Quarterly Journal of the Royal Meteorological Society*, 145, 2772–2790.

Sitwell, M. and Ménard, R. (2020) Framework for the comparison of a priori and a posteriori error variance estimation and tuning schemes. *Quarterly Journal of the Royal Meteorological Society*, 146, 2547–2575.

Skákala, J., Bruggeman, J., Brewin, R.J.W., Ford, D. and Ciavatta, S. (2020) Improved representation of underwater light field and its impact on ecosystem dynamics: a study in the North Sea. *JGR Oceans*, 125, e2020JC016122.

Skákala, J., Bruggeman, J., Ford, D., Wakelin, S., Akpinar, A., Hull, T., Kaiser, J., Loveday, B.R., O'Dea, E., Williams, C.A. and Ciavatta, S. (2022) The impact of ocean biogeochemistry on physics and its consequences for modelling shelf seas. *Ocean Modelling*, 172, 101976.

Skákala, J., Ford, D., Brewin, R.J.W., McEwan, R., Kay, S., Taylor, B., de Mora, L. and Ciavatta, S. (2018) The assimilation of phytoplankton functional types for operational forecasting in the Northwest European shelf. *JGR Oceans*, 123, 5230–5247.

Skákala, J., Ford, D., Bruggeman, J., Hull, T., Kaiser, J., King, R.R., Loveday, B., Palmer, M.R., Smyth, T., Williams, C.A.J. and Ciavatta, S. (2021) Towards a multi platform assimilative system for North Sea biogeochemistry. *Journal of Geophysical Research: Oceans*, 126, e2020JC016649.

Storkey, D., Blockley, E.W., Furner, R., Guiavarc'h, C., Lea, D., Martin, M.J., Barciela, R.M., Hines, A., Hyder, P. and Siddorn, J.R. (2010) Forecasting the ocean state using NEMO: the new FOAM system. *Journal of Operational Oceanography*, 3, 3–15.

Teruzzi, A., Dobricic, S., Solidoro, C. and Cossarini, G. (2014) A 3-D variational assimilation scheme in coupled transport biogeochemical models: Forecast of Mediterranean biogeochemical properties. *Journal of Geophysical Research: Oceans*, 119, 200–217.

Trémolet, Y. (2007) Model error estimation in 4D var. *Quarterly Journal of the Royal Meteorological Society*, 133, 1267–1280.

Vaquer Sunyer, R. and Duarte, C.M. (2008) Thresholds of hypoxia for marine biodiversity. *Proceedings of the National Academy of Sciences*, 105, 15452–15457.

Waller, J.A., Dance, S.L. and Nichols, N.K. (2016) Theoretical insight into diagnosing observation-error correlations using observation minus background and observation minus analysis statistics. *Quarterly Journal of the Royal Meteorological Society*, 142, 418–431.

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.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.

Weston, P.P., Bell, W. and Eyre, J.R. (2014) Accounting for correlated error in the assimilation of high resolution sounder data. *Quarterly Journal of the Royal Meteorological Society*, 140, 2420–2429.

How to cite this article: Fowler, A.M., Skákala, J. & Ford, D. (2023) Validating and improving the uncertainty assumptions for the assimilation of ocean-colour-derived chlorophyll *a* into a marine biogeochemistry model of the Northwest European Shelf Seas. *Quarterly Journal of the Royal Meteorological Society*, 149(750), 300–324. Available from: https://doi.org/10.1002/qj.4408

1477870x, 2023, 750, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4408 by University of Reading, Wiley Online Library on [08/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License