**General Response:**

We would like to thank all three reviewers for their encouraging comments and thorough review of our revised manuscript. We have addressed all the comments and concerns below and added additional analysis and figure details where needed. To address a specific request of Dr Ciavatta, we have added an additional co-author: Dr Jenny Lovell.

**Emmanuel Boss**

Dear authors,

10  I am happy with the revisions you made, though I do think you could have pushed the state of the science more by assimilating more wavelengths, in particular in the red to constrain TSM. The wavelength you assimilated is the one that varies least with change in Chl (it is the hinge point). I also do not agree that NPQ may not affect the glider data. I will want to see the backscattering channel to be convinced of that (that is that both Fchl and bbp are homogeneous near the surface).

*We have included the backscatter coefficient in Figure 1 below, but are reticent to include it in the final MS, as it is becoming "very busy" and distracting. As can be seen, the raw backscatter coefficient is a little spikey, but is generally homogeneous in the upper 100m. We cannot find evidence of NPQ in the profiles contained in Figure 1. If the editor wishes we can include Figure 1 in*
20  *the MS, but our preference is to exclude it and make reference to the fact there is no evidence of NPQ in the text as supported by the homogeneous backscatter profiles. To the effect, we have added this sentence to page 14:*

"There is no quenching evident in this or other profiles, as the observed fluorescence and
25  backscatter coefficients (not shown) are constant in the upper 50 m."
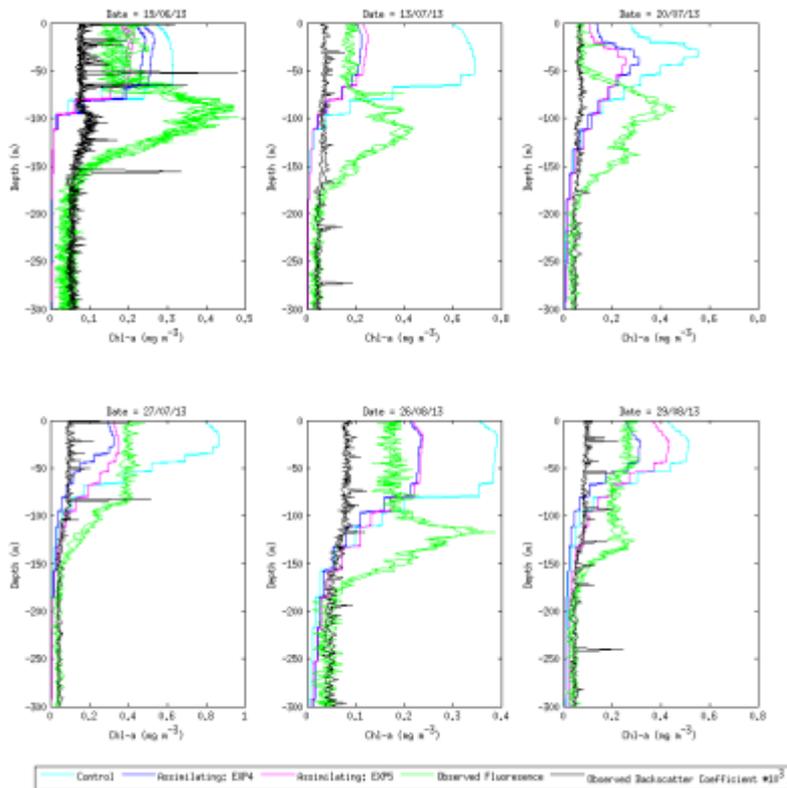
Figure 1: This is the same figure as used in Figure 8 in the MS, with the addition of the backscatter coefficient.

5    In any case, all the best and good luck,
Emmanuel

Please fix the following:
Your Dutkiewicz citation is incomplete.
10
*Now corrected in the references section.*

Your Fujii reference is not the appropriate one to use. It was published in Biogeosciences after it appeared in the BGD (http://www.biogeosciences.net/4/817/2007/bg-4-817-2007.pdf)
15
*Apologies, we have corrected this.*

I don't think you need to include reviewers number in your acknowledgment.

20    *We have removed the reviewer numbers from the Acknowledgement section.*

**David Ford:**

Minor comments

P10 L44-46: "EXP1-5 shows the progressive reduction in difference
in kind error associated with each experimental configuration". Semantics, but I'm not sure this is
exactly true as stated: if I interpret correctly, the difference in kind error (ED) is set the same in
EXP1-2 (artificially so in EXP2), and is set to zero in EXP3-5.

*Your interpretation is correct. We have removed this sentence from the MS as it is repetitious given
the statement in the previous paragraph of the MS.*

P11 L1-2: Please clarify, either in section 2.5.2 or the caption to Table 1, whether Etot is being
expressed as a fractional or an absolute error. I assume it's fractional, so if 20% has been used for EA
for EXP5, and Etot is 0.2, does this mean ER is zero? I assume then this is the case for all
experiments, has this been assumed due to the super-obbing?

*Extra detail has been added to section 2.5.2 -*
*We have altered the following sentence to read:*
*"The total observation error used on the diagonal of $R$ is then given by the sum of error types: $E_{tot}$ =
$E_R + E_D + E_A$ and is expressed as a standard deviation in log-space and can therefore be thought of as
a fraction of the observed quantity."*

*We have added the following text to the description of $E_R$:*
*"In the case of the MODIS observations, we set $E_R$ to be the standard deviation of all the 1 km log-
scaled observations that fall within a 4 km model grid cell."*

*We have also updated the caption and column headings of Table 1.*

P12 L14-15: "we can conclude that the model represents accurately
the general distribution of Chl-a throughout the region" - this should
perhaps be clarified as "OC3M Chl-a", since a distinction's being drawn
between the OC3M and surface Chl-a distributions.

*Done, thank you for pointing this out.*

S3.2: Thank you for including more validation against in situ observations for the different runs.
However, it's not made clear in the manuscript why EXP2-3 haven't been included in Fig. 7-11, and
EXP1 in Fig. 8,10,11. I do appreciate that including all six runs would potentially make Fig. 8-11
unclear (Fig. 7 less so, but I'll leave that to the editor to decide), and there's not enough data to
compare EXP1 (and EXP3?) to nutrients, so I'm happy for these to remain excluded unless the editor
disagrees. However, please at least comment in the text on the relative performance of EXP2-3
against in situ data com-pared with EXP1 and EXP4. For instance, P15 L11 you attribute part of the
success to "the inclusion of the TSS constituents in the state vector". This implies that EXP4
outperforms EXP3 - is this the case, that would be useful to know? Furthermore, EXP3 appears to
have been stopped early, were there issues with this configuration?

*We have added addition experiments to Figures 7, 9, 10 and 11 and updated the captions
accordingly. And in-text discussion of the results accordingly.*

*We have added a column to table 4 detailing the mean RMSD for EXP2.*

We have added text to the bottom of page 13:

*"In EXP3, the magnitude of the diagonal elements of **R ($E_{tot}$)** are reduced from 0.8 to 0.4, yielding a very minor improvement in forecast statistics over EXP2, and was discontinued after 10 cycles. Additional variables were added to the assimilation state vector for EXP4, while keeping the observation error equal to EXP3. The additional assimilation state variables improved the forecast statistics and decreased the mean absolute forecast innovation when compared with EXP2 and EXP3."*

Fig. 16 and Fig. 17 could potentially be combined into a single 5-panel figure for easier comparison, but that's just personal preference, I'm happy either way.

*We have left Figure 16 and 17 as is. Though we are happy to change this should we receive further advice from the editorial team.*

*All other minor typos and grammar mentioned have been corrected. We have altered a sentence in the last para of the discussion to read:*

*"In the absence of an optical model, Ford and Barciela (2015) suggest further exploration of the assimilation of empirical statistical products such as Chl-a and Phytoplankton Functional Types (PFTs)"*

**Stefano Ciavatta:**

**The authors addressed all my comments in a constructive way. Though they did not accept some recommendations (e.g. to include TSS and nutrients in the control vector of EXP1, or include more optically active variables in EXP2-4), they justified soundly their methodological choices and pointed out assumptions and limits of their work.**
**In particular, I am glad and grateful that the authors added EXP5, which demonstrated the feasibility and advantages of direct assimilation of reflectance data (R551).**
**Furthermore, the authors reorganized and polished the manuscript, such that the text flows much better in the current version.**
**Overall, this work represents a major step in biogeochemical data assimilation and, potentially, in our capability to simulate and predict crucial dynamics in marine ecosystems.**
**For the above reasons, I think that this revised manuscript is suitable for publication after a minor revision.**
**The most relevant comments are listed below; others were included in the pdf file.**

**1) I kindly insist that the author replace the current Appendix 1 with an appendix that lists the most important equations and parameterisation of the bio-optical model. This would increase the readability of the manuscript, as well as well as the reproducibility of the work. I do appreciate that the bio-optical model was published in a previous paper, but so was the biogeochemical model which is lengthy described in the current appendix;**

*We have included an additional appendix (Appendix B) that outlines the bio-optical model used to calculate the IOPs.*

2) The authors should be more objective in judging the results of the simulation shown in figure 4 (e.g. the overestimated OC3M south of Papua). Furthermore, in figure 4, the simulated OC3M matches better the NASA OC3M rather than the ANN OC3M, which has been assimilated. Please comment in the manuscript. See more comments in the pdf;

*We've added the following text to page 12 addressing these comments:*

*"The model predicts a large Chl-a plume south of Papua New Guinea, that is not present in the remote sensing observations. In the reference simulation, the model over predicts inshore sediment resuspension, which causes high simulated OC3M values in shallow coastal regions"*

*and*

*" The reflectances derived from the ANN method have been shown to have substantially lower errors due to the improved atmospheric correction near the coast for case 2 waters (Schroeder et al., 2007), and we have therefore used these observations in the assimilation experiments below."*

*The in-text PDF comments have been addressed below.*

3) Figure 5: The non-linear relationship between the two simulated variables is an obvious result, because you have included in your model a non-linear relationship between the two variables. It would be more relevant a plot of in situ chlorophyll versus satellite OC3M, or, at least, simulated chlorophyll versus satellite OC3M. Weak or non-linear relationships in these plots would support better your point. Please add such plot to the manuscript;

*The key point we want to make here is that in optically complex waters, there is a an inherent non-linear relationship between Chl-a and OC3M. To further highlight this, we have taken your advice and included an extra panel that shows a comparison between in-situ Chl-a and MODIS OC3M. All plots are log-scaled, and show that for the inshore water types of the GBR, OC3M overestimates the in-situ Chl-a by upto an order of magnitude as low concentrations. We have added the following text to the MS to highlight this point:*

*"A scatter plot of in situ Chl-a plotted against MODIS OC3M observations (Figure 5, lower panel) further highlights the magnitude of the OC3M errors in complex coastal waters. While it is obvious that coastal water types are outside of the design specification of the OC3M algorithm, MODIS OC3M contains errors upwards of 400% at low Chl-a."*

4) Figure 6. The top graph does not provide a clear overall ranking of the performance of the experiments. It is not easy to assess which exp has the "best" mean forecast innovation (i.e. ~0 most of the analyses). The authors should plot the absolute values of the mean forecast innovations, or provide in a table the temporal average of the absolute values of the mean forecast innovations for each experiment;

*Done, the lower panel of Figure 6 presents a time series of the Mean Absolute Forecast Innovation (error).*

5) Figure 11, top plot: you can reduce the y-scale to 10 mg m-3;

*Done*

6) In the conclusion, section 4.2, I do agree with the authors that assimilation of reflectance has the benefit to use data with lower error, because data do not rely on empirical models. However, they should recognize that modelling (and assimilating) optical properties might increase the complexity and/or uncertainty of the ecosystem model. To a certain extent, the problem is moved from satellite retrievals into the structure of the marine model. One needs to add equations and parameterization that are rough approximation and highly uncertain. For example: which experimental phytoplankton absorption coefficient should I use for the "small phytoplankton" represented in the model? What are the implications of representing CDOM as a linear regression of salinity? Is this more accurate than retrieving CDOM absorption from satellite data?
Therefore, the authors should mention the above issues in the conclusions and list the most relevant bio-optical equations and parameterizations in the appendix.

Comments from the PDF:
Page 1: Should this on be 2016a, rather than b, since it's the first paper you cite?

*Correct, this has been fixed.*

Page 3: it's quite a repetition of the previous sentence

*Removed*

Page 4: line 3: please explain why this is relevant (e.g., ", ehich provided data for skill assessment")

Page 4 line 5: please mention why cloud free is relevan (e.g. increased spatial coverage of assimilated data)

*This para now reads:*
*"The assimilation system was tested between the 25$^{th}$ May 2013 to 22$^{nd}$ September 2013. This period was chosen as it coincides with a field program by the Great Barrier Reef Marine Park Authority (GBRMPA) inshore Marine Monitoring Program (MMP) and an autonomous glider deployment (Integrated Marine Observing System, IMOS), these in situ observations are used for an independent assessment of the assimilation system. Additionally, this period also has a large number of cloud free days, which increases the amount of observational data available for assimilation. Details of the observations, model and assimilation system are given below. "*

Page 7: To my knowledge, "univariate assimilation" is commonly used to mean you are updating just the model variable you are assimilating; in multivariate assimilation, you can still assimilate just one variable, but you are correcting >1 model variable.

This is different than assimilating data of 1 or >1 variables, to which you refer here.

*We have changed the text to:*

*"We acknowledge that a 36 member ensemble is small, however as demonstrated in the results section this small ensemble performs adequately for the assimilation of a univariate observation type. The ensemble size will need to be increased to assimilate multivariate observations."*

5    Page 11 – line 10: Please report the values you applied in a table in this manuscript, to guarantee reproducibility of your work.

*Done, this text now reads:*
*Where $a0$, $a1$, … $a4$ are a set of empirically determined coefficients (e.g. NOMAD version 2,*
10   [http://seabass.gsfc.nasa.gov/wiki/article.cgi?article=NOMAD)](http://seabass.gsfc.nasa.gov/wiki/article.cgi?article=NOMAD) *with the respective values of: a0 =*
*0.2424, a1 = -2.7423, a2 = 1.8017, a3 = 0.0015 and a4 = -1.2280*
*([http://oceancolor.gsfc.nasa.gov/cms/atbd/chlor_a)](http://oceancolor.gsfc.nasa.gov/cms/atbd/chlor_a).*

Page 12: please mention that the reference simulation overestimated badly the OC3M values
15   observed south of Papua New Guinea

Page 12: In figure 4, simulated OC3M is much closer to the NASA product than to the ANN product (by-eye; please provide quantitative evidence if I am wrong). Please state clearly this evidence in the text, as well as the reason why you still preferred to assimilate ANN.

The following sentences have been added in response to the two points raised above:

*"The model predicts a large Chl-a plume south of Papua New Guinea, that is not present in the*
25   *remote sensing observations. In the reference simulation, the model over predicts inshore sediment*
*resuspension, which causes high simulated OC3M values in shallow coastal regions."*

*"The reflectances derived from the ANN method have been shown to have substantially lower errors*
*due to the improved atmospheric correction near the coast for case 2 waters (Schroeder et al., 2007),*
30   *and we have therefore used these observations in the assimilation experiments below."*

Page 13: These sentences look to be contrasting, please clarify.

*We have altered these sentences to read:*
35   *"Of the experiments that assimilated OC3M, the configuration used for EXP4 (Figure 6, blue line)*
*gave the lowest mean absolute forecast innovation. While the EXP5 configuration (noting that this*
*experiment assimilates R551, not OC3M) beat EXP4 with a lower mean absolute innovation, but*
*these innovations are for R551, not OC3M and therefore cannot be directly compared. To undertake*
*an assessment of the relative performance between EXP4 and EXP5, an assessment against*
40   *independent (non-assimilated) observations is given in Sections 3.2.2 – 3.2.5"*

Page 14: The reader cannot see this DCM on that date
*We point out that this feature can be seen from the 7/7/13 – 14/7/13 –*
*"The configuration used in EXP5 captures a weak DCM that persists for 7 days from the 7/7/13 -*
45   *14/7/13 and also maintains the lowest mixed layer Chl-a for the feature captured on from the 7/7/13*
*– 13/7/13 (Figure 7: bottom panel). Assimilating OC3M still over-predicts the mixed layer Chl-a."*

Page 14: Not really. I see that EXP5 has a lower RMSD, at least in the range 0-60 meters. Please clarify.

*We see that within the 0-200m range, the RMSD for EXP5 is slightly less than EXP4. Our text originally read:*
*The assimilation of R551 (EXP5) results in a minor improvement in RMSD when compared with the assimilation of OC3M (EXP4)*

*To clarify, we have changed it to:*
*"In the upper 60 m, the assimilation of R551 (EXP5) results in a minor improvement in RMSD when compared with the assimilation of OC3M (EXP4)."*

10     With my compliments for this nice piece of work,
       kind regards,
       Stefano Ciavatta

# Use of remote-sensing reflectance to constrain a data assimilating marine biogeochemical model of the Great Barrier Reef.

Emlyn M. Jones[1], Mark E. Baird[1], Mathieu Mongin[1], John Parslow[1], Jenny Skerratt[1], Jenny Lovell[1], Nugzar Margvelashvili[1], Richard J. Matear[1], Karen Wild-Allen[1], Barbara Robson[2], Farhan Rizwi[1], Peter Oke[1], Edward King[1], Thomas Schroeder[3], Andy Steven[3] and John Taylor[4]

[1] CSIRO Oceans and Atmosphere, Hobart, Australia, 7000

[2] CSIRO Land and Water, Canberra, Australia, 2601

[3] CSIRO Oceans and Atmosphere, Brisbane, Australia, 4102

[4] CSIRO Data61, Canberra, Australia, 2601

*Correspondence to*: Emlyn M. Jones (emlyn.jones@csiro.au)

Key words: Great Barrier Reef, Data Assimilation, Remote Sensing, observed OC3M, Reflectances, Biogeochemical (BGC) Model, Chlorophyll

**Abstract:** Skillful marine biogeochemical (BGC) models are required to understand a range of coastal and global phenomena such as changes in nitrogen and carbon cycles. The refinement of BGC models through the assimilation of variables calculated from observed in-water inherent optical properties (IOPs), such as phytoplankton absorption, is problematic. Empirically-derived relationships between IOPs and variables such as Chlorophyll-a concentration (Chl-a), Total Suspended Solids (TSS) and Colored Dissolved Organic Matter (CDOM) have been shown to have errors that can exceed 100% of the observed quantity. These errors are greatest in shallow coastal regions, such as the Great Barrier Reef (GBR), due to the additional signal from bottom reflectance. Rather than assimilate quantities calculated using IOP algorithms, this study demonstrates the advantages of assimilating quantities calculated directly from the less error-prone satellite remote-sensing reflectance (RSR). To assimilate the observed RSR, we use an in-water optical model to produce an equivalent simulated RSR, and calculate the mis-match between the observed and simulated quantities to constrain the BGC model with a Deterministic Ensemble Kalman Filter (DEnKF). Following theThe traditional assumption that simulated surface Chl-a is equivalent to the remotely-sensed OC3M estimate of Chl-a resulted in a forecast error of approximately 75%. We show this error can be halved by instead using simulated RSR to constrain the model via the assimilation system. When the analysis and forecast fields from the RSR-based assimilation system are compared with the non-assimilating model, a comparison against independent *in situ* observations of Chl-a, TSS, and dissolved inorganic nutrients ($NO_3$, $NH_4$ and DIP) showed that errors are reduced by up to 90%. In all cases, the assimilation system improves the simulation compared to the non-assimilating model. Our approach allows for the incorporation of vast quantities of

remote-sensing observations that have in the past been discarded due to shallow water and/or artefacts introduced by terrestrially-derived TSS and CDOM, or the lack of a calibrated regional IOP algorithm.

**1 Introduction:**

Aquatic biogeochemical (~~biogeochemical (BGC)~~BGC) models have been used to understand a range of coastal and global phenomena such as ocean acidification (Mongin et al., 2016), nutrient pollution (Skerratt et al., 2013) and carbon cycles, and are central to our predictions of global climate (Sarmiento and Gruber, 2006). At the coastal / regional scale,

10 non-linear ~~biogeochemical~~BGC processes driven by planktonic interactions, as well as non-linear circulation features such as mesoscale eddies, limit the timescale over which ~~biogeochemical~~BGC properties are deterministically predictable (Baird, 2010). For the purposes of prediction, it is therefore necessary to assimilate observations to correct for model errors and non-linear processes.

*In situ* observations of phytoplankton pigments and macro-nutrients are sparse in space and time due to the prohibitive expense of collecting them. Optical sensors on gliders and floats provide high resolution *in situ* observations that are used to estimate pigment and nutrient concentrations. Nonetheless, these observations in large parts of the ocean remain sparse.

20 The most spatially comprehensive dataset available for BGC assimilation is from Ocean Color (OC) remote sensing. The assimilation of remotely-sensed data into marine ~~biogeochemical~~BGC models has been problematic due to differences between the variables represented in models and the variables that are routinely observed (~~Baird et al., 2016b~~Baird et al., 2016a), typically referred to as "difference in kind" errors. Satellites

25 measure the top of atmosphere radiance, not chlorophyll-a concentration (Chl-a; or other modelled variables) directly. To attain estimates of Chl-a or other variables, the spectrally-resolved top of atmosphere radiance is converted into atmospherically-corrected remote-sensing reflectances (RSRs), which are then related to Chl-a via empirical statistical relationships derived from *in situ* observations. Figure 1 graphically demonstrates the key

30 steps in the OC processing chain, and points at which models and remotely-sensed and *in situ* observations can be compared.

Early studies investigating the benefits of assimilating OC products, predominantly SeaWiFS-derived Chl-a, into BGC models include those of Carmillet et al., (2001) and Natvik and

35 ~~Evenson~~ Evensen (2003), with a comprehensive review of algorithms used and observations assimilated detailed in Gregg (2008). Considerable effort has been invested in Data Assimilation (DA) algorithm development, with ensemble and variational approaches being the most common. A thorough review of these approaches in a statistical sense is presented in Dowd et al., (2014). There are now examples of operational and pre-operational global

40 systems that routinely assimilate Chl-a products (Ford et al., 2012). Additionally, there has been further experimentation with assimilating alternative remotely-sensed Apparent Optical Properties (AOPs) such as the vertical attenuation coefficient at 443 nm, $Kd_{443}$ (Ciavatta et al., 2014) and Inherent Optical Properties (IOPs) such as phytoplankton absorption ($a_{ph}$), as described in Shulman et al.,. (2013).

It is well known that OC algorithms, such as observed OC3M (Moderate Resolution Imaging Spectroradiometer, MODIS, three-band Chl-a algorithm) that are optimized for global applications suffer from errors due to a variety of optically-active constituents in coastal and shelf waters (Odermatt et al., 2012). Furthermore, it has also been noted that even globally, there is a non-uniform distribution of error, and substantial bias, in the remotely-sensed OC3M-derived Chl-a. Satellite-derived ~~ocean color~~OC ~~products~~ algorithms such as OC3M are a function of observed RSRs. A substantial effort is invested in empirical studies that convert RSRs to ~~biogeochemical~~BGC quantities such as Chl-a, Total Suspended Solid concentration (TSS), phytoplankton functional types (PFTs), and Colored Dissolved Organic Matter (CDOM) (Odermatt et al., 2012). Each of these empirical relationships have differing error magnitudes stemming not only from a difference in kind, but also representation errors. In optically deep regions (e.g. offshore waters) not influenced by sediment resuspension and terrestrial runoff, typical errors for OC-derived Chl-a (e.g. OC3M) are less than 40%, and as low as 5-20% for IOPs and AOPs. However, in optically-complex coastal areas where there is river discharge, sediment resuspension and ~~optically shallow water~~a surface expression of benthic reflectance, errors can exceed 300% (e.g. Schroeder et al., 2016; Qin et al., 2007).

Numerical weather prediction (NWP) avoids difference in kind errors through using the model to simulate directly-observed quantities, such as temperature brightness, in preference to deriving other quantities such as temperature and humidity profiles from the brightness measurements. ~~Thus NWP systems commonly assimilate brightness temperature in preference to satellite-derived temperature and humidity~~ (Derber and Wu, 1998; Dee et al., 2011). The goal of this study is to apply this approach to marine ~~biogeochemical~~BGC modelling, assimilating RSRs rather than quantities calculated using empirical-statistical relationships. We assess ~~this~~ the approach of assimilating RSRs through a comparison against withheld *in situ* observations of a range of BGC quantities such as Chl-a, Total Suspended Solids (TSS) and dissolved nutrients.

The Great Barrier Reef (GBR) region, located along the northeast coastline of Australia, is used to demonstrate the assimilation~~s~~ of RSRs in optically-complex and shallow waters (Blondeau-Patissier et al., 2009). The GBR is characterized by fringing reefs along the continental slope that create a semi-connected inshore lagoon that spans over 3000 km of coastline (Figure 2). The GBR ecosystem, described as one of the seven natural wonders of the world, is under increasing pressure from local and global anthropogenic stressors (De'ath et al., 2012). Decreasing water clarity due to nutrient and sediment pollution is considered a serious threat to the GBR ecosystem (Thompson et al., 2014), with major concerns ~~being~~ including the impact of ~~lower~~ reduced benthic light levels on coral and seagrass communities (Collier et al., 2012; ~~Baird et al., 2016a~~Baird et al., 2016b) and the impacts of invasive species (e.g. Morello et al., 2014).

The paper is structured in the following manner, in Section 2 (Methods) the data for assimilation and skill assessment is presented, along with a description of the model and assimilation methods used. Section 3 contains the results from the control (non-assimilating) run of the model, and subsequent data assimilation experiments. In Section 4 we discuss the approach used and implications of the findings more generally. And we conclude with major findings in Section 5.

## 2 Methods:

The assimilation system was tested between the 25th May 2013 to 22nd September 2013. This period was chosen as it coincides with a field program by the Great Barrier Reef Marine Park Authority (GBRMPA) inshore Marine Monitoring Program (MMP) and an autonomous glider deployment (Integrated Marine Observing System, IMOS), that are used for assessment of the assimilation system. Additionally, this period also has a large number of cloud free days, which increases the amount of remotely-sensed observational data available for assimilation. Details of the observations, model and assimilation system are given below.

### 2.1 In-situ observations

All *in situ* observations have been withheld from the assimilation system for validation purposes, and are primarily obtained from two different programs.

The Integrated Marine Observing System (IMOS)IMOS has deployed fluorometers on moorings at Yongala and North Stradbroke National Reference Stations (NRS) (Figure 2). This study uses the monthly observations of dissolved inorganic nutrients ($NO_3$, $NH_4$ and DIP) at the NRS sites (Lynch et al., 2014). Additionally, glider data obtained from the IMOS-operated Australian National Facility for Ocean Gliders (ANFOG) provide cross-shelf profiles sections of water column properties, including temperature, salinity, and chlorophyll fluorescence. These locations are shown in Figure 2.

The GBMRMPA MMP (Figure 2, yellow circles) contains samples 13 sites in inshore regions and in the GBR lagoon, and samples nutrients and Chl-a extractions 3 times a year (Thompson et al., 2011; Rolfe and Gregg, 2015); no MMP bottle sampling occurred during the simulation period. The Australian Institute of Marine Sciences (AIMS) deployed The moorings moorings were deployed at the GBRMPA MMP sites from 2009 to 2014 and included a Sea-Bird water quality monitors (WQM) that measure chlorophyll fluorescence and turbidity (NTU), which are used for assessment purposes in this study. A comparison of the 2011-2014 control run simulation against the GBRMPA MMP observations, and other observations, is available in a 180 page skill assessment report at: https://research.csiro.au/ereefs/models.

### 2.2 MODIS observations

The daily observations of RSR are obtained from MODIS-Aqua and an atmospheric correction developed for the region is applied (Schroeder et al., 2007). The atmospheric correction applied an Artificial Neural Network (ANN) approach trained by a radiative transfer model to invert the top of atmosphere (TOA) signal measured by MODIS-Aqua. The ANN algorithm was adapted to an approach previously-developed for the MEdium Resolution Imaging Spectrometer (MERIS) sensor but on the basis of a different learning algorithm (Schroeder et al., 2007). Algorithm performance is described in detail in Goyens et al. (2013) and King et al. (2014).

SeaDAS-provided Level-2 flags were used to quality control the observed RSR and to exclude erroneous and out-of-range pixels. We filtered the data for land and severe sun glint

affected pixels, cloud contamination including cloud shadows, and rejected pixels with observing and solar zenith angles above 52° and 70°, respectively. Super-observations (Cummings et al, 2005; Oke et al., 2008b), of RSR and OC3M (see Section 2.5.2, eq. 21) are generated for the assimilation system by taking the mean and variance of all the 1 km resolution observations that fall within a 4 km model grid cell. The mean value of the super-observation is then assimilated, while the variance is used as an estimate of the representation error in the observation error covariance matrix. More detail on this process in given in Section 2.5.2.

**2.3 The eReefs modelling system**

We used the eReefs coupled hydrodynamic, sediment and ~~biogeochemical~~BGC modelling system (Schiller et al., 2014). The hydrodynamic model is a fully three-dimensional (3D) finite-difference baroclinic model based on the three dimensional equations of momentum, continuity and conservation of heat and salt, employing the hydrostatic and Boussinesq assumptions (Herzfeld 2006; Herzfeld and Gillibrand 2015). The sediment transport model adds a multilayer sediment bed to the hydrodynamic model grid and simulates sinking, deposition and resuspension of multiple size-classes of suspended sediment (Margvelashvili 2009; Margvelashvili et al., 2016). The complex ~~biogeochemical~~BGC model simulates optical, nutrient, plankton, benthic organisms (seagrass macroalgae and coral), detritus, chemical and sediment dynamics across the whole GBR region, spanning estuarine systems to oligotrophic offshore reefs (Figure 3, ~~Baird et al., 2016b~~Baird et al., 2016a). A, an expanded description of the BGC model is given in Appendix A, with a brief description of the optical model in Appendix B. Briefly, the ~~biogeochemical~~BGC model considers four groups of microalgae (small and large phytoplankton, *Trichodesmium* and microphytobenthos), two zooplankton groups, three macrophytes types (seagrass types corresponding to *Zostera* and *Halophila*, macroalgae) and coral communities. Photosynthetic growth is determined by concentrations of dissolved nutrients (nitrogen and phosphorous) and photosynthetically active radiation. Microalgae contain two pigments (chlorophyll-a and an accessory pigment), and have variable carbon:pigment ratios determined using a photo-adaptation model (described in Baird et al., 2013). Overall, the model contains 23 optically-active constituents (~~Baird et al., 2016b~~Baird et al., 2016a; and Appendix A).

The model is forced with freshwater inputs at 21 rivers along the GBR and the Fly River in southwest Papua New Guinea. River flows input into the model are obtained from the DERM (Department of Environment and Resource Management) gauging network. Statistical flow/load relationships are used to account for nutrient and sediment inputs from rivers into the model (statistical relationships between river flow and nutrient concentrations; (Furnas 2003). Nutrient concentrations flowing in from the ocean boundaries were obtained from the CSIRO Atlas of Regional Seas (CARS) 2009 climatology (Ridgway et al., 2002).

**2.4 Calculation of remote-sensing reflectance from ~~biogeochemical~~BGC state**

The model contains 23 optically-active constituents (~~Baird et al., 2016b~~Baird et al., 2016a; and Appendix A and B). To calculate the RSR at the surface, we need to consider the light returning from multiple depths, and from the bottom. Rather than using a computationally-

expensive radiative transfer model, we approximate RSR based on an optical-depth weighted scheme (~~Baird et al., 2016b~~Baird et al., 2016a), alternative methods are given in Fuji~~j~~i et al., (2007) and Dutkiewicz et al., (2015). The ratio of the backscattering coefficient to the sum of backscattering ($b_b$) and absorption ($a$) coefficients for the whole water column at wavelength, $\lambda$, is:

$$\mathbf{u}_\lambda = \sum \frac{\mathbf{w}_{\lambda,z'} \mathbf{b}_{b,\lambda,z'}}{\mathbf{a}_{\lambda,z'} + \mathbf{b}_{b,\lambda,z'}} \qquad \text{Eq 1}$$

where $\mathbf{w}_{\lambda,z'}$ is a weighting for each layer representing the component of the RSR due to the absorption and scattering at depth $\mathbf{z'}$.
The weighting fraction is given by:

$$\mathbf{w}_{\lambda,z} = \frac{1}{z_1 - z_0} \int_{z_0}^{z_1} \exp(-2\mathbf{K}_{\lambda,z'}) \mathbf{dz'} \qquad \text{Eq 2}$$

where $\mathbf{K}_\lambda$ is the vertical attenuation coefficient at wavelength $\lambda$, $z_0$ and $z_1$ are the top and bottom depths of the layer and the factor of 2 accounts for the path length of both downwelling and upwelling light. The vertical attenuation coefficient is calculated from the sum of the absorption and scattering properties of each of the optically-active constituents, and the zenith angle (for each of these relationships, and more information, see ~~Baird et al., 2016b~~Baird et al., 2016a).

The integral of $\mathbf{w}_{\lambda,z'}$ to infinite depth is 1. In areas where light reaches the bottom, the integral of $\mathbf{w}_{\lambda,z'}$ to the bottom is less than one, and benthic reflectance is included in the sum as an extra term with a weighting of $1 - \sum \mathbf{w}_{\lambda,z'}$.

The sub-surface RSR, $\mathbf{r_{rs}}$, is given by:

$$\mathbf{r_{rs}} = \mathbf{g_0} \mathbf{u}_\lambda + \mathbf{g_1} \mathbf{u}_\lambda^2 \qquad \text{Eq 3}$$

where $\mathbf{g_0}$ = 0.0895 and $\mathbf{g_1}$ = 0.1247 are coefficients for the nadir-view in oceanic waters that vary with wavelength and other optical properties (Morel et al., 2002), but can be approximated as constants (Lee et al., 2002). The constants result in a change of units from the unitless $\mathbf{u}$ to a per unit of solid angle, sr$^{-1}$, quantity, $\mathbf{r_{rs}}$.

The above-surface RSR is given by (Lee et al., 2002):

$$\mathbf{R}_{rs,\lambda} = \frac{0.52 \, \mathbf{r}_{rs,\lambda}}{1 - 1.7 \mathbf{r_{rs}}} \qquad \text{Eq 4}$$

Thus, the above-surface RSR is calculated from the inherent optical properties of the optically-active constituents in the ~~biogeochemical~~BGC model.

## 2.5 Data Assimilation System

The Data Assimilation (DA) algorithm used in this study is the Deterministic Ensemble Kalman Filter (DEnKF; Sakov and Oke, 2008). The full ~~biogeochemical~~BGC state variable list contains over 130 2D and 3D variables. Including all of these variables in the assimilation

system is impractical due to memory constraints, but we also acknowledge that for many variables the observations will be uninformative, and therefore not good candidates to include in the assimilation state vector. We therefore limit the variables that are updated within the system to a select subset that ~~are~~ is detailed in Section 2.5.2.

### 2.5.1 Data Assimilation Algorithm

The DEnKF is based on the Kalman filter analysis equation, of which various flavors have had general success in state estimation in other marine BGC data assimilation problems (e.g. Hu et al., 2012, Ciavatta et al., 2016). The derivation of the DEnKF is given in Sakov and Oke (2008) and is a modification to the traditional Kalman filter equation:

$$\mathbf{x^a} = \mathbf{x^f} + \mathbf{K}(\mathbf{Y} - \mathbf{Hx^f}) \qquad \text{Eq 5}$$

Where **x** is the model state, **Y** is the vector of observations, **H** is the observation operator, and the superscripts of $^a$ and $^f$ denote the analysis and forecast fields respectively. In this study, we only use a subset of the state variables in the DEnKF update~~;~~ those variables not included in the state vectors denoted in Table 1 are not altered by the state update. The Forecast Innovations (**FI**) are defined by:

$$\mathbf{FI} = (\mathbf{Y} - \mathbf{Hx^f}) \qquad \text{Eq 6}$$

The Kalman gain matrix, *K*, is given by:

$$\mathbf{K} = \mathbf{lP^f H^T}(\mathbf{HP^f H^T} + \mathbf{R})^{-1} \qquad \text{Eq 7}$$

Where **l** is the localization operator, applied in the form of covariance localization (Sakov and Bertino, 2011), with an isotropic localization radius of 60 km, and **R** is the observation error covariance matrix. Off diagonal elements of **R** are set to 0, while the diagonal elements of **R** are the sum of the representation error, difference in kind error and analytical measurement error, termed **E_{tot}**. These error terms are discussed in detail~~s~~ in Section 2.5.2 (additionally see Schaeffer et al., 2016 for a further discussion of these error sources relating to BGC variables and methods to estimate them from glider data).

The background error covariance matrix, **P**, is given by:

$$\mathbf{P^f} = \frac{1}{m-1}\sum_{i=1}^{m}(\mathbf{X_i^f} - \mathbf{x^f})(\mathbf{X_i^f} - \mathbf{x^f})^T = \frac{1}{m-1}\mathbf{AA^T} \qquad \text{Eq 8}$$

where **m** is the ensemble size, and **i** denotes the **i**th member of the ensemble. ~~Given that we are using a flavor of the ensemble Kalman filter, t~~The background error covariance is approximated by a 36 member dynamic ensemble where~~by~~ $\mathbf{X_i^f}$ is the ith ensemble member, and **x** is the ensemble mean. We acknowledge that a 36 member ensemble is small, however as demonstrated in the results section this small ensemble performs adequately for ~~univariate~~ the assimilation of a univariate observation type. The ensemble size will need to be increased to assimilate multivariate observations. To avoid negative values and normalize the state, we log-transform the state before forming the state vector. The

application of the log-transform to biological variables is discussed in Parslow et al., (2013), and is commonly used in other BGC assimilation schemes (e.g. Ciavatta et al., 2014). The background error covariance matrix is never computed; rather a series of anomaly fields are constructed and denoted by **A**. We then construct the Kalman gain matrix in the observation subspace as per Sakov and Oke (2008):

$$\mathbf{A} = [\mathbf{A_1} \dots \mathbf{A_m}] \qquad \text{Eq 9}$$

where the **i**th anomaly field is given by:

$$\mathbf{A_i^f} = \mathbf{X_i^f} - \mathbf{x^f} \qquad \text{Eq 10}$$

where **x** is the the ensemble mean and is updated via Eq 5, each anomaly field is updated by:

$$\mathbf{A_i^a} = \mathbf{A_i^f} - \frac{1}{2}\mathbf{KHA^f} \qquad \text{Eq 11}$$

The full analyzed ensemble is then given by:

$$\mathbf{X_i^a} = \mathbf{A_i^a} + [\mathbf{x^a}, \dots,] \qquad \text{Eq 12}$$

The assimilation system iterates through time using a five-day forecast duration. The assimilation system is cycled by calculating the analysis fields at time t, using the forecast from the previous cycle, **X**$^f$(t), and observations **Y**(t) at time t, using all observations that fall within a window of t +/- 3 hrs. The numerical model is initialized using the analysis fields **X**$^a$(t) and the next five day forecast is made. This forecast at t+5 days, **X**$^f$(t+5), is then used in the next assimilation cycle.

The DEnKF requires the ensemble to be perturbed in such a way that it captures the main source of error. These perturbations are introduced in a way that captures our prior understanding of the dominant errors. In this system we expect that errors will stem from uncertainty in the Initial Conditions (ICs) as per most assimilation system. Additional sources of error can stem from uncertainty associated with BGC process parameters, which has been discussed at length in Parslow et al. (2013), and river boundary conditions (BCs). In this study we have ignored the effects of uncertainty that propagate from the model physics (hydrodynamics), short wave radiation forcing and open ocean boundary fluxes of BGC tracers, but do acknowledge that they contribute to the overall uncertainty in the predicted system state.

In the context of this study, we have introduced perturbations to the ensemble by sampling initial conditions randomly from a four year run (where **T** = 4 years) of the BGC model:

$$\mathbf{X(t = 0)_i} \sim \mathbf{Uniform}(\mathbf{X(t = 0)}, \mathbf{X(t = T)}) \quad \mathbf{for\ i = 2 \dots m} \qquad \text{Eq 13}$$

Where $\mathbf{X}$(t=0)$_i$ is the initial condition for the model state for the **i**th member sampled for a uniform distribution with no replacement~~.~~, ~~Where~~ and $\mathbf{X}$(t=0)$_{i=1}$ is the ensemble mean. Sensitivity experiments have shown the model is sensitive to perturbations in the quadratic zooplankton mortality rate for ~~large~~ small (**m**$_{Q,ZS}$) and ~~small~~ large ~~zooplankton~~ (**m**$_{Q,ZL}$) zooplankton with units of d$^{-1}$ (mg N m$^{-3}$)$^{-1}$. These are considered system parameters, and are as such uncertain. To this end we have perturbed the ensemble, by sampling space and time invariant parameters from:

$$\mathbf{m_{q,ZL,i}} \sim \mathbf{LN(0.012, 1)} \quad \mathbf{for\ i = 2 \ldots m} \qquad \text{Eq 14}$$
$$\mathbf{m_{q,ZS,i}} \sim \mathbf{LN(0.007, 1)} \quad \mathbf{for\ i = 2 \ldots m} \qquad \text{Eq 15}$$

Where **LN** is a log-normal distribution, and respective means of 0.012 and 0.007 are typical values as used in the control run (~~Baird et al., 2016b~~Baird et al., 2016a), and are given relatively broad standard deviations of 1 for both parameters. The river nutrient and sediment loads were altered by a time invariant scaling factor ($\mathbf{\theta}$) to all rivers:

$$\mathbf{\theta_{NO3,i}} \sim \mathbf{N(1, 0.3)} \quad \mathbf{for\ i = 2 \ldots m} \qquad \text{Eq 16}$$
$$\mathbf{\theta_{NH4,i}} \sim \mathbf{N(1, 0.3)} \quad \mathbf{for\ i = 2 \ldots m} \qquad \text{Eq 17}$$
$$\mathbf{\theta_{DIP,i}} \sim \mathbf{N(1, 0.3)} \quad \mathbf{for\ i = 2 \ldots m} \qquad \text{Eq 18}$$
$$\mathbf{\theta_{FineSed,i}} \sim \mathbf{N(1, 0.3)} \quad \mathbf{for\ i = 2 \ldots m} \qquad \text{Eq 19}$$

Where **N** is a normal distribution truncated at 0. Each ensemble member has their load (**Q**$_i$) scaled according to:

$$\mathbf{Q_i = \theta_i Q_{control}} \qquad \text{Eq 20}$$

where $\mathbf{Q_{control}}$ is the load entering the control run.

### 2.5.2 Assimilation system experiments and configuration

The assimilation system was assessed using five assimilation system configurations (Table 1) using a subset of the full model state in the assimilation state vector, and corresponding diagonal elements (**E**$_{tot}$; see below) of the observation error covariance matrix (**R**). The assimilation experiments were conducted on Raijin, a super-computer hosted at the National Computational Infrastructure (NCI; http://nci.org.au/). The assimilation cycle progressed via a two-step process. The first step generated the forecast fields by integrating the ensemble forwards on 36 nodes of the system. Each node has a dual 8 core processor, with 16 Gb of RAM. This step takes approximately 90 minutes to simulate 5 model days. The second step generates the analysis fields and updates model scripts to allow for the next forecast cycle. The analysis step is undertaken on a single high-memory node and requires 800 Gb of RAM and takes 50 minutes using 16 cores.

We only allow the observations to update the variables contained in the assimilation state vector. The analyzed assimilation state vector is then inserted into the full model state vector.

There are three sources of observational error ($E_{tot}$) that must be accounted for when relating remotely-sensed observations to ~~the~~ a modelled state variable:

1. Representation error ($E_{RF}$) – errors that arise due to the approximation that the modelled tracer quantities are an average over a whole model cell. This can be thought of as unresolved spatial variability. In the case of the MODIS observations, we set $E_R$ to be the standard deviation of all the 1 km log-scaled observations that fall within a 4 km model grid cell.

2. Difference in kind error ($E_D$) – these errors arise when the variables that are being modelled and included in the assimilation state vector differ from the observations. For example, many studies have included surface Chl-a (or some optical depth weighted average) in the assimilation state vector, and assume there is a direct relationship with Chl-a estimated using OC and the OC3M algorithm (or other quantities). The OC3M algorithm is known to have typical errors of between 30-70% in blue water domains, and errors that exceed 300% in optically complex (or optically shallow) waters (Qin et al., 2007).

3. Analytical/sensor/processing error ($E_A$) – ~~Depending~~ depending on the observational platform in use, these errors can be small (e.g. ~~ARGO~~ Argo floats), or in the case of remote -sensing products, even with the ANN atmospheric correction, as large as 15-20% (see Schroeder et al., 2007).

The total observation error used on the diagonal of **R** is then given by the sum of error types: $E_{tot} = E_R + E_D + E_A$ and is expressed as a standard deviation in log-space and can therefore be thought of as a fraction of the observed quantity.

The sum of these error types can be large and forms the diagonal element of the observation error covariance matrix (**R**). The larger $E_{tot}$, the lower the impact of the observations in the assimilation system. If we can remove the difference in kind error ($E_D$), then we only have representation error and analytical/sensor/processing error. Given that Level 3 ~~ocean color~~ OC remote sensing products rely on empirical/statistical relationships, $E_D$ dominates, therefore if we can minimize $E_D$ (or remove it entirely), then the information content of the observations increases. Conversely, if we don't have a large enough observation error, we run the risk of overfitting the observations and generating unrealistically large increments for the unobserved model state variables.

Experiment 1 (EXP1) was designed to test the assimilation system under the assumption that modelled surface total Chl-a (the sum of small and large Phytoplankton Chl-a and *Trichodesmium* Chl-a) was equivalent to the Chl-a from MODIS-observed OC3M. This experiment is analogous to those of Natvik and ~~Evenson~~ Evensen (2003), Gregg (2008) and Ford et al., (2012). This is a reasonable assumption in offshore waters, however, OC3M is known to be unreliable in coastal waters where sediments (e.g. TSS), bottom reflectance and CDOM cause artificially high OC3M values. We expect that EXP1 will contain difference in kind errors that stem from the assumption that surface Chl-a is equivalent to observed OC3M. The observation error prescribed for this experiment is considered to be the lower bound of OC3M errors as reported in Qin et al., (2007).

In Experiments 2-4 (EXP2-4) we assume that the simulated OC3M (i.e. calculated from simulated RSR) is equivalent to the observed OC3M (i.e. calculated from observed RSR) and

it is used as an input into the observation operator. The simulated OC3M (Eq 21) contains the signature from simulated TSS, CDOM and bottom reflectance as per ~~Section~~ Sections 2.3, 2.4 and Appendix A. Thus for EXP2–4, the observed and simulated OC3M contains no ~~the same error characteristics~~ difference in-kind errors (although we leave $E_D$ high during EXP2 for experimental purposes). In other words, the configuration used for EXP2 is the same as EXP1, except the effect of difference in kind error has been removed by using simulated OC3M, in place of total surface Chl-a. In EXP3, we reduce the observation error to account for the reduction in $E_D$, and in EXP4 we add additional variables to the assimilation state vector.

In EXP5, we assimilate observed RSR at 551 nm (R551) using the simulated RSR at the equivalent wavelength. ~~The progressively lowering of the $E_{tot}$, inserted on the diagonal element of R, through EXP1-5 shows the progressive reduction in difference in kind error associated with each experimental configuration (Table 1).~~ Typically errors in RSR as reported in Schroeder et al., 2007 are in the order of 10-20%, and we have used ~~a figure of 20%~~ 0.2 for this experiment.

In all experiments we have generated a set of super-observations (Cummings et al, 2005; Oke et al., 2008) by spatially averaging the 1 km observations, onto the 4 km model grid. Prior to super-obing in Experiments 1-4 (EXP1-4), the observed atmospherically-corrected RSRs (Schroeder et al., 2007) at 443, 488 and 551 nm were transformed into a single observation using the OC3M algorithm:

$$\mathbf{OC3M} = \mathbf{10}^{(a0 + a1 \cdot B + a2 \cdot B^2 + a3 \cdot B^3 + a4 \cdot B^4)} \qquad \text{Eq 21}$$

Where **a0**, **a1**, … **a4** are a set of empirically determined coefficients (e.g. NOMAD version 2, http://seabass.gsfc.nasa.gov/wiki/article.cgi?article=NOMAD) with the respective values of: **a0** = 0.2424, **a1** = -2.7423, **a2** = 1.8017, **a3** = 0.0015 and **a4** = -1.2280 (http://oceancolor.gsfc.nasa.gov/cms/atbd/chlor_a).~~and~~ **B** is:

$$\mathbf{B} = \mathbf{log}_{10}\left(\frac{R_{rs,\lambda 1}}{R_{rs,\lambda 2}}\right) \qquad \text{Eq 22}$$

$R_{rs,\lambda 1}$ and $R_{rs,\lambda 2}$ are determined by the absolute magnitude (whichever is greater) of the remote-sensing reflectance, and $R_{rs,\lambda 1}$ is either the band centered on 443 nm or 488 nm and $R_{rs,\lambda 2}$ is the band centered on 551 nm. We apply the OC3M algorithm (Eq 21) to both the observed and simulated RSRs.

The assimilation system preserves the stoichiometry of the small and large phytoplankton (referred to as PhyS_Chl-a and PhyL_Chl-a respectively) as follows. In the ~~biogeochemical~~ BGC model, each phytoplankton cell (small, large, benthic or *Trichodesmium*) is represented by a quantity of structural material, B, and reserves of nitrogen, $R_N$, reserves of phosphorus, $R_P$, reserves of energy, $R_I$, and an intracellular chlorophyll-a concentration, $c_i$. Our intention in the assimilation is to change the number of cells, as quantified by B, not the physiological status of the cell, as represented by $R_N$, $R_P$, $R_I$, and $c_i$. Since the reserves are quantified as the total of these reserves across the entire population, each of the reserves is changed by the same proportion as the biomass. ~~Thus,~~

for example, the nitrogen reserve of an individual cell, $R_N / B$, is unchanged. Once the analyzed quantity of $c_i$ is determined (e.g. PhyL_Chl-a and PhyS_Chl-a), the quantities of $R_N$, $R_P$, $R_I$ and B are updated such that the respective ratios prior to assimilation are preserved.

**3 Results:**

3.1 **Control Run**

The modelling system has been designed to represent the spatially-resolved water quality dynamics (phytoplankton, nutrients, turbidity and oxygen) of the GBR World Heritage Area for informed management. A number of indicators have been used to assess the skill of the model, including RMS errors, Pearson's correlation coefficients and, Wilmott's skill indicators (Wilmott et al., 1985; https://research.csiro.au/ereefs/models).

The simulated state variable concentrations resemble both the regional climatology for offshore-reef, lagoon-reef and near-shore zones and water quality observations under contrasting seasons/loads and flood events [not shown here, but detailed at https://research.csiro.au/ereefs/models, with optical (Baird et al., 2016bBaird et al., 2016a) and carbon chemistry (Mongin et al., 2016) skill assessment published elsewhere]. As mentioned above, the model simulates RSR. It is therefore possible to incorporate these RSRs into standard, well-recognized remote-sensing products. Figure 4 presents a snapshot of simulated surface Chl-a and the simulated OC3M, as well as remotely-sensed products (regional ANN-observed OC3M and NASA-observed OC3M). The images are during the dry season conditions with little cloud contamination along the inshore region.

The two panels on the right side of Figure 4 represent the simulated (top) and NASA remotely-sensed (bottom) OC3M estimate of Chl-a. Both combine individual RSRs into proxies for Chl-a using the OC3M algorithm (Eq 21). OC3M poorly represents surface Chl-a close to the coast where sediment resuspension and CDOM absorption dominate. By comparing the two panels, we can conclude that the model represents accurately the general distribution of OC3M Chl-a throughout the region, with high values along the coast and above each reef systems, and low concentrations offshore. The model predicts a large Chl-a plume south of Papua New Guinea, that is not present in the remote sensing observations. In the reference simulation, the model over predicts inshore sediment resuspension, which causes high simulated OC3M values in shallow coastal regions. The simulated surface Chl-a inside the coastal band is lower than in the remotely-sensed OC3M observation. The two panels on the left of Figure 4 represent simulated surface Chl-a (top) and OC3M based on the regionally-optimized RSR (bottom; ANN-observed OC3M). Regardless of the remote-sensing products used, there are clear differences between the simulated Chl-a and the simulated OC3M, and these exceed the differences between those of the ANN-observed and NASA-observed Chl-a. The reflectances derived from the ANN method have been shown to have substantially lower errors due to the improved atmospheric correction near the coast for case 2 waters (Schroeder et al., 2007), and we have therefore used these observations in the assimilation experiments below.

Surface Chl-a and OC3M are typically assumed to be equivalent. However, many studies have shown that there are regional biases and substantial overestimation in optically-

complex coastal waters (Qin et al., 2007). This is demonstrated in Figure 5, where we plot simulated surface Chl-a against simulated OC3M for deep water (Figure 5, top-left) and the whole domain (Figure 5, top-right) for the 26th May 2013. It is immediately obvious that these variables are not equivalent and the error structure is non-linear, especially when coastal regions are included. ~~Figure 5 shows that for the GBR region, there is a substantial risk of OC3M over-estimating the *in situ* Chl-a even in optically simple deep water regions of the domain.~~ A scatter plot of *in situ* Chl-a plotted against MODIS OC3M observations (Figure 5, lower panel) further highlights the magnitude of the OC3M errors in complex coastal waters. While it is obvious that coastal water types are outside of the design specification of the OC3M algorithm, MODIS OC3M contains errors upwards of 400% at low Chl-a. Figure 5 shows that for the GBR region, there is a substantial risk of OC3M over-estimating the *in situ* Chl-a even in optically simple deep water regions of the domain.

### 3.2 Assimilation system configuration experiments

To choose the best configuration for the assimilation of RSR into a coastal ~~biogeochemical~~BGC model, five experiments were undertaken using a variety of state variables in the state vector (**X**, Table 1), and by altering the diagonal elements of the observation error covariance matrix (**R**, Eq 7).

### 3.2.1 Forecast Innovations

The forecast innovations (Eq 6) for the five assimilation system configuration experiments are shown in Figure 6:

- EXP 1 (~~green~~ black lines, Figure 6): Assumes that total surface Chl-a is equivalent to observed OC3M and is used to calculate the forecast innovations, which are then used to update small and large phytoplankton Chl-a. An 80% error in the ANN-observed OC3M observation is prescribed on the diagonal elements of **R**.
- EXP 2 (red line, Figure 6): Simulated OC3M (calculated from simulated RSR at 443, 488 and 551 nm, as described by ~~Baird et al., 2016b~~Baird et al., 2016a) is used to calculate the forecast innovations. The state variable and observation errors are the same as EXP1.
- EXP 3 (~~black~~ green line, Figure 6): The same configuration as EXP 2, with a reduced observation error imposed on the diagonal of the observation error covariance matrix.
- EXP 4 (blue line, Figure 6): Additional variables are included in the assimilation state vector, which now comprises of small and large phytoplankton, ammonia, nitrate and total suspended solids concentrations. The observation error is the same as EXP3.
- EXP5 (dashed magenta line, Figure 6): The assimilation vector as EXP4, but using R551 observations, with a lower observation error of 0.2. Note that these innovations relate to the difference in simulated and observed R551, not OC3M as in EXP1-4.

The forecast innovation (Eq 6) statistics for Experiments 1-5 (Figure 6) provide an insight into the assimilation system performance. An optimal assimilation system should result in mean forecast innovations (mismatches between observations and the model) of close to 0, and low mean absolute innovations. The assimilation system where the observation operator assumed that there was a direct relationship between simulated surface Chl-a and ANN-observed OC3M (EXP1, Figure 6 green black line), performed very poorly and was discontinued after 9 cycles; the model at times became numerically stiff, requiring the adaptive $4^{th}-5^{th}$ order Ordinary Differential Equation (ODE) integrator to take progressively smaller steps. The innovation statistics for EXP1 suggested the model was constantly over-predicting Chl-a with the mean absolute innovation exceeding 0.7 more than 50% of the time. Calculating the forecast innovations with simulated OC3M and ANN-observed OC3M, rather than simulated surface Chl-a and observed ANN-observed OC3M, improved innovation statistics dramatically (EXP2-4, Figure 6). In EXP3, the magnitude of the diagonal elements of **R ($E_{tot}$)** are reduced from 0.8 to 0.4, yielding a very minor improvement in forecast statistics over EXP2, and was discontinued after 10 cycles. Additional variables were added to the assimilation state vector for EXP4, while keeping the observation error equal to EXP3. The additional assimilation state variables improved the forecast statistics and decreased the mean absolute forecast innovation when compared with EXP2 and EXP3.

Of the experiments that assimilated OC3M, tThe configuration used for EXP4 (Figure 6, blue line) gave the best performance when assimilating OC3Mlowest mean absolute forecast innovation. While the EXP5 configuration (noting that this experiment assimilates R551, not OC3M) gave best performance in terms of a forecast statisticbeat EXP4 with a lower mean absolute innovation, but these innovations are for R551, not OC3M and therefore cannot be directly compared. An assessment against independent (non-assimilated) *in situ* observations is given in Sections 3.2.2 – 3.2.5. to assess the relative performance between the non-assimilating control run and EXP1 (glider only), EXP2, EXP3 (glider only) EXP4 and EXP5., a further examination against independent (non-assimilated) observations is given in Sections 3.2.2 – 3.2.5.

3.2.2 Independent assimilation system assessment: Glider

The control and assimilating runs were compared with to with-held ocean glider data that was deployed on 26th May 2013, and recovered on the $4^{th}$ of August 2013. The glider track largely followed the shelf break and headed in a south-easterly direction. To make a comparison between glider observations and the model, we take a sub sample of glider observations centered at the time of model output, with a time window of two hours. For each glider observation that falls within this time period, we find a corresponding 3-D3D cell from the model and extract the equivalent model solution. No 3-D interpolation is performed as the non-interpolated solution will yield insight into the unresolved sub grid scale variability.

A persistent feature of the observed Chl-a when interpolated onto the model time-space grid (Figure 7, top panel) is the relatively low values in the upper 100 m of water column. For much of the record, there is a persistent weak deep chlorophyll maximum (DCM) that is centered at between 80 and 120 m depth. Rarely do concentrations in the upper 80 m exceed 0.5 mg/m³. When the control run is examined, Chl-a in the top 80 m regularly

exceeds 1 mg/m³, and DCMs when they exist, are located between 30 and 60 m deep. A detailed analysis of the control run demonstrates the model is able to reliably produce DCMs (contained in the ~~180 page~~ skill assessment report at: https://research.csiro.au/ereefs/models), however, in this particular location in time and space, the model does not generate one consistent with the observations. Using the EXP1 configuration, the assimilation system does not improve the solution when compared with the control run. The EXP2 configuration marginally improves the solution by reducing the mixed layer Chl-a. However, using the EXP4 and EXP5 configurations, the assimilation system substantially improves on the control when compared to the glider observations. The difference between the EXP2 and EXP4 configuration is the reduction in observation error and the inclusion of additional variables in the assimilation state vector. Regardless of the assimilation configuration used, the ~~The~~ assimilation system cannot place the DCM in the correct location because the remote sensing observations provide no information about such a deep feature. The remote sensing observations do remove the bias in the upper 80 m with concentrations in the assimilating run ranging between 0.1 - 0.3 mg/m³ (Figure 7). The configuration used in EXP5 captures ~~the~~ a weak DCM that ~~is predicted on the~~ persists for 7 days from the 7/7/13 - 14/7/13 and also maintains the lowest mixed layer Chl-a for the feature captured on from the 7/7/13 – 13/7/13 (Figure 7: bottom panel). Assimilating OC3M still over-predicts the mixed layer Chl-a.

A comparison between non-interpolated ~~individual~~ profiles from the glider and equivalent sampling of the model shows substantial unresolved sub-grid scale variability (or instrument noise). Figure 8 (top left) shows that in the upper 70 m the observed Chl-a as measured by the fluorometer ranges between 0.14 and 0.3 mg/m³, with a mean value of 0.18 mg/m³. There is no quenching evident in this or other profiles, as the observed fluorescence and backscatter coefficients (not shown) are~~is~~ constant in the upper 50 m. Within each of the subplots, the glider profiles collected over the 2 h window may sample multiple model grid cells. However, due to the relatively slow horizontal glider speed, these glider profiles fall within two adjacent model cells. In most cases, the control and assimilating runs gave indistinguishable solutions between adjacent cells with the exception of Figure 8, top left. In all cases the assimilation of observations of both OC3M and R551 have reduced the error in the simulated Chl-a when assessed against independent glider data. The noise in the fluorometer observations appears to range between 0.04 and 0.08 mg m⁻³.

The Chl-a RMSD for each layer of the model, and aggregated in time for all glider observations, is shown in Figure 9. Above 80 m, there is a substantial reduction in RMSD between the control (cyan) and assimilating runs (EXP2-red, EXP4-blue and EXP5-magenta). EXP3 behaved similarly to EXP2. The RMSD from EXP1 (dashed black) is marginally higher than the RMSD in the control run. The assimilating runs (EXP4 and EXP5) have an RMSD of between 0.10 and 0.17 mg m⁻³ compared with 0.30 to 0.41 mg m⁻³ in the control run. In the upper 60 m, t~~T~~he assimilation of R551 (EXP5) results in a minor improvement in RMSD when compared with the assimilation of OC3M (EXP4). Below 80 m, the RMSD profile is similar between the control and assimilating runs. However, the assimilation of R551 results in the lowest RMSD below 80 m. The largest impact of assimilating remote sensing observations is in the top 80 m of the water column.

3.2.3 Independent assimilation system assessment: GBRMPA MMP Chl-a and TSS

The AIMS GBRMPA MMP moorings were deployed at the 13 GBRMPA MMP sites in the shallow inshore regions of the Great Barrier Reef Lagoon (Figure 1). The control run typically had RMSDs of between 0.4 and 0.6 mg/m³ for *in situ* Chl-a (Figure 10; upper panel). In all cases the EXP2 configuration performed worse than the control run. In most cases the assimilation of OC3M (EXP4)EXP4 configuration reduced the RMSD of *in situ* Chl-a by 0 – 10 %, when compared with the control run. The assimilation of R551 (EXP5) was more variable. Some sites improved, whereas others were degraded by the assimilation of this observation type when compared with the control run. In only one case did the assimilation of R551 beat OC3M for Chl-a at the GBRMPA MMP sites.

The TSS RMSD varies widely across all the GBRMPA MMP sites (Figure 1110; lower panel), driven by the strong variation in magnitude of the spring – neap tidal forcing. The combination of the initialization of the ensemble and perturbed forcing caused the ensemble mean for EXP2 to have a substantially reduced RMSD when compared with the control run. , and Tthe inclusion of the TSS constituents in the state vector in the assimilating model further reduced the RMSD, and has generated realistic time varying correlations between the observed OC3M/R551 and inshore TSS. These cross-correlations allow for the correction of simulated TSS from OC3M/R551 observations. The TSS RMSD at all sites for the assimilating run is less than 5 to 20% that of the control run. The performance of EXP4 and EXP5 were indistinguishable, but the inclusion of TSS in the assimilation state vector was beneficial.

3.2.4 Independent assimilation system assessment: Nutrients

Within the GBR region, there are two IMOS NRS sites (Yongala and North Stradbroke, NS, Figure 1). The dissolved inorganic nutrients of $NO_3$, $NH_4$ and DIP are taken monthly. At Yongala, water samples are taken at the surface (0 m), 10 m, 20 m and bottom (26 m). At NS, samples are taken at the surface (0 m) and 10 m. It should be noted that there are only 3 to 4 samples per depth at each site during the 4 month simulation period. At Yongala, typical control run RMSDs for $NO_3$ range from 5 to 12 mg/m³. Assimilating OC3M/R551 halved these errors using the EXP4/5 configurations. The improvement at NS is evident with improvements found in the upper 10 m of the water column, with EXP4 and EXP5 giving indistinguishable results. The EXP2 configuration yielded higher RMSDs than EXP4 and EXP5.

With the exception of the surface samples at Yongala, the assimilation system improved the prediction of $NH_4$ at all depths for each site. Most notably was the 70 to 90% reduction in RMSD at the deeper locations at Yongala. EXP4 and EXP5 out performed EXP2 in all cases for $NH_4$. There were marginal improvements to DIP, which displayed a 0 to 30% reduction in RMSD across all sites. EXP4 resulted in lower DIP errors at NS, but EXP5 resulted in lower DIP errors at Yongala. EXP2 had the lowest DIP RMSDs.

3.2.5 Summary of assimilation system experiments

The direct assimilation of RSR (EXP5: R551), or a function of RSR, *f(RSR),* (EXP4: OC3M) improved the model solution when compared with the control run against three independent non-assimilated in-situ observational datasets. In most cases the EXP4 and

EXP5 configurations outperformed the EXP2 configuration, which was to be expected due to the additional variables included in the assimilation state vector. The subtle differences between the two approaches will be discussed in Section 4. Based on these findings, the preferred assimilation system for state estimation on the GBR- is that used in EXP4. A further examination of the EXP4 forecast error statistics and examples forecast, analysis and increment field is given Section 3.3.

### 3.3 EXP4: Assimilation system forecast errors.

The assimilation system was run with a 5-day forecast cycle. Using the forecast at $t+5$ days and comparing the temporal mean (across all cycles) of the Root Mean Square Difference (RMSD) and Percentage Error against observations provides insight into the value of the assimilation system, when compared with a non-assimilating system. By comparing the forecast fields against yet to be assimilated observations, we are providing a semi-independent estimate of forecast skill. Additionally, by comparing the forecast against the persisted analysis field from the previous analysis cycle (e.g. the analysis field from t-5 days), it can be determined if the dynamic model is adding skill to the forecast.

A comparison of simulated OC3M and observed OC3M for the non-assimilating control run gives a domain wide median error (range) of 0.32 (0.27 – 0.48) mg m$^{-3}$ (Figure 12). This is approximately equivalent to a domain-wide median percentage error (range) of 100% (80% - 130%) (Figure 12). The data assimilation system reduces the forecast errors and percentage errors to a median value (range) of 0.23 (0.20 – 0.30) mg m$^{-3}$ and 55% (43% - 63%) respectively. The analysis errors are again reduced when observations are assimilated, with median and percentage errors (range) of 0.19 (0.14 – 0.23) mg m$^{-3}$ and 39% (37% - 42%) respectively. When the analysis field from the previous assimilation cycle is persisted forward, the errors (and percentage errors) slightly exceed that of the forecast field with values of 0.26 (0.21 – 0.29) mg m$^{-3}$ and 52% (44% - 65%) respectively. However, it is not expected that these error statistics are spatially uniform given the large percentage of area that is dominated by deep oceanic waters. To understand the spatial variability of the forecast error statistics, the whole domain is divided into three regions representing shallow coastal waters (depth < 30 m), lagoon and shelf waters (30 m < depth < 500 m), and deep oceanic waters (depth > 500 m).

In shallow coastal areas (Figure 12, 2$^{nd}$ column), the non-assimilating control run has a median error (range) of 1.35 (1.1 – 2.45) mg m$^{-3}$, which corresponds to a percentage error (range) of 130 (105 – 180) %. The distribution of control run errors in the coastal zone is positively skewed, with the mean value of the distribution sitting some way from the median. The assimilation system marginally reduces the median forecast error when compared with the control run, though most notably, it reduces the median percentage error and associated variability. The forecast also beats persistence in this region. There is a marked improvement for lagoon and shelf waters with the assimilation system reducing the median error from 0.34 to 0.25 mg m$^{-3}$, which corresponds to a reduction in percentage error from 96% to 48%. In the oceanic regions of the domain, the assimilation system reduces the error from 0.16 to 0.10 mg m$^{-3}$, corresponding to a percentage error reduction from 91% to 45%. In all cases the forecast fields beat persistence. A summary of the results can be found in Table 2 and Table 3.

### 3.3.1 EXP4: Forecast, Increment and Analysis Fields

The sum of the surface *Trichodesmium* Chl-a, Small Phytoplankton (PhyS) Chl-a and Large Phytoplankton (PhyL) Chl-a biomass differs substantially from the simulated OC3M as shown in Figure 5. The assimilation system updates all of the state variables included in the assimilation state vector. Simulated OC3M is a diagnostic variable that is a function of all the optically-active dynamic state variables as described in Section 2. To demonstrate the impact on the dynamic variables of PhyS and PhyL, results from the forecast step, and the assimilation update, are presented in Figure 13, Figure 14 and Figure 15. Cycle 22 was chosen to demonstrate the spatial impact of the system as it is representative of the last 10 cycles and was relatively cloud free.

The simulated OC3M forecast field for cycle 22 (12th September 2013) displays elevated OC3M in the shallow near shore environment throughout the whole of the GBR region and southern shelf of Papua New Guinea (PNG) (Figure 13). Additional features are elevated OC3M in the vicinity of the central and southern fringing reefs, and a plume originating from the eastern region of PNG. Offshore oceanic waters generally have low OC3M of 0.2 mg m$^{-3}$ or less. There is some evidence of mesoscale blooms in the northern and southern sections of the domain. Observed OC3M is overlaid on Figure 13 (left). Where there is a difference in colour, the simulated OC3M differs from the observed OC3M.

The forecast surface layer fields for PhyS and PhyL appear substantially different to the simulated OC3M field (Figure 13). The differences near the coast are where TSS and CDOM are known to cause artefacts in OC3M. While there are patchy blooms of small phytoplankton at various locations within the domain, rarely does the PhyS Chl-a exceed 0.5 mg m$^{-3}$. The exception to this is in the inner central coastal region of the GBR and in the vicinity of the Fly River plume on the south coast of PNG. Similarly, the PhyL Chl-a remains very low for large areas of the domain, however in regions with additional nutrient supply (e.g. in upwelling regions, mesoscale eddies and some river mouths) blooms do occur.

When the observed OC3M is assimilated, increment fields are calculated using Eq 5 and are presented in Figure 14 for simulated OC3M, PhyS Chl-a and PhyL Chl-a. The innovations are overlaid on the increment field for OC3M to give an indication of how well we are fitting the observations. In areas where the model is over-predicting OC3M, the increments will be negative. In areas where the model is under-predicting OC3M, the increments will be positive. The increments and innovations here are presented as a fractional change with respect to the background (forecast) field.

For this particular analysis cycle, it appears that the model is under-estimating inshore OC3M by up to 10-30% and over-estimating OC3M by upwards of 50% offshore (Figure 14, left). By using the background ensemble correlation structure, the increments applied to PhyS biomass increase its concentration in the inner lagoon by up to 20%, and substantially increase the PhyS biomass offshore of the central outer reefs by more than 50% (Figure 14, centre). It should be noted that the increments being applied to the background fields contain meso and sub-meso scale information. Significantly, features such as upwelling filaments, eddies and plumes are maintained through the assimilation procedure, demonstrating that they are allowed to dynamically evolve in the assimilation system. The

increments applied to PhyL biomass (Figure 14, right) differ substantially to those of PhyS biomass (Figure 14, centre). For large areas of the domain, the assimilation system decreases the PhyL biomass by up to 50%, whereas there are some areas that it increases. These areas correspond to regions where a bloom may be occurring (there is a small

5  westward shift in the major bloom off the Papua New Guinea coast). The increment applied to the central region of the domain, offshore of the outer reefs, is linear and coherent and likely a result of shifting a dynamic feature such as an upwelling-induced bloom to better match observations.

10  When the increments contained in Figure 14 are applied to the forecast fields, the resulting analysis field for simulated OC3M better fits the observed OC3M, with a substantially reduced error inshore and in the vicinity of the outer reefs (Figure 15). The difference between simulated and observed OC3M is small in the deeper offshore regions and shallow sections of the lagoon. The greatest error in OC3M occurs in the central lagoon and the

15  outer reefs where spatial variability is highest. The corresponding analysis fields for small phytoplankton and large phytoplankton are contained in Figure 15. There are elevated concentrations of small phytoplankton biomass in the near shore region near river mouths and the outer fringing reefs.  The large phytoplankton biomass is concentrated in the region of Broad Sound, the Fly River plume, and the Papua New Guinea upwelling.  Each of these

20  features was predicted by the forecast, as little biomass is added or subtracted by the assimilation update. However, there is substantial removal of large phytoplankton biomass from the northern and central offshore regions. This leaves very little large phytoplankton biomass present in substantial areas of the domain during this particular analysis cycle.

**4 Discussion**

25  In the optically-complex waters of the GBR, the use of an optics model to calculate a simulated RSR and subsequent simulated OC3M to constrain the BGC model substantially reduces the errors in biogeochemicalBGC state. This has been achieved by explicitly assimilating like-for-like variables of simulated R551 or simulated OC3M. The data assimilation is constrained by the mis-match between simulated and observed R551 and

30  OC3M. A summary of the root mean square differences (RMSDs) is contained in Table 4. Our approach of simulating the observation is the opposite to the conversion of observed RSRs into modelled variables, e.g. the assimilation of phytoplankton functional types, TSS, CDOM and Chl-a. We therefore advocate for the use of a "forward" model to predict optical properties, rather than rely on the "inversion" to back calculate the model state from

35  reflectances. The conversion of RSRs into derived variables (e.g. Chl-a, PFTs) have associated errors that are as large as 300% in some locations, and can be biased by up to 70% (Qin et al., 2007), whereas the forwards approach has errors of approximately 20% (Baird et al 2016b2016a). The errors stemming from the "inversion" technique are at times difficult to characterise in data assimilation systems. Using a forwards model avoids these

40  errors and therefore the dominant source of error stems from model error, rather than observation error. (via difference in kind errors). The approach of Ciavatta et al., (2014) contained similar results where they found the assimilation of $Kd_{443}$, to be superior to remote-sensed Chl-a.

A significant source of error in algorithms such as OC3M is that they produce a single value for each horizontal pixel, generally considered to be representative of the first optical depth of the water column. If this is to be compared to a single value in a ~~biogeochemical~~BGC model, then it must be assumed that the water column is well-mixed to the optical depth,
5    and that there is an equal optical depth of each of the wavebands used in the algorithm. Both of these conditions are rarely met in coastal waters. Matching RSR requires no assumptions about the structure of the water column, or of the vertical distribution of the optically-active constituents, because both observed and modelled quantities are two dimensional fields.

EXP1 performed poorly for a number of reasons. It is likely that even an 80% observation error was insufficient to adequately account for positive biases in the OC3M observations. These biases, present even in offshore waters, lead to positive innovations that result in adding phytoplankton biomass in the increments. These large increments lead to
15    persistently high biomass, that draw available nutrients down to very low levels. The only way to account for this form of observation error is to inflate the "difference in kind error" to a large value. Given that the non-assimilating control run of the model had forecast errors that range between 70%-100% (region dependent), running an assimilation system with an observation error larger than the error present in the non-assimilating model, does
20    not make sense.

The similarity in RMSDs between EXP4 and EXP5 (Table 4) suggests there is little difference between using simulated R551 and simulated OC3M. Univariate observations were assimilated in each case, and the information content of the observations is likely going to
25    be similar in deep water, hence the similar results for the Ocean Glider. However, due to the simulated OC3M containing information from 443 and 488 nm, which are important in shallow regions due the absorption of these wavelengths from sediments and CDOM, there is additional information in the OC3M observation -that is not present in the 551 nm RSR. Ultimately the greatest information content will come from the simultaneous assimilation of
30    multivariate observations obtained from the multi-spectral OC sensors and including the longer wavelengths.

4.1 True color visualization of EXP4
In order to visualize the impact of the assimilation system on the prediction of water clarity,
35    we compare the observed true color (Figure 16) with the simulated true color of the control run (Figure 17, top left) and the EXP4 assimilating run (Figure 17, top right). Simulated true colour images are generated from RSR at the red, green and blue wavelengths calculated using the optical model and the three dimensional fields of the model-predicted 23 optically-active constituents.

The observed true colour image on the 12 Sep 2013 shows brown / yellow features associated with high suspended sediment concentrations. As these concentrations become more diluted, and mixed with phytoplankton, the water appears more greenish blue. Offshore reefs, with clear water above white substrates, appear as light blue features, with
45    the intensity depending on the reef depth. Qualitatively, the control run (Figure 17, top left) does a reasonably good job of reproducing the observed true colour. The quantification of this mismatch can be done on individual color bands (not shown, ~~Baird et al., 2016b~~Baird et

al., 2016a). Qualitatively, the control run does not have enough suspended solids in the surface water in the mouth of Broad Sound (22.2°S, 149.5°E), and has too high phytoplankton concentrations offshore, especially in a feature centred at 23°S, 151.5°E. The assimilated run (EXP4), while not that different to the control run, corrects some of these errors.

To approximately quantify impacts of the assimilation of water clarity, it is possible to consider the colour of the added (and subtracted) constituents in the assimilation procedure. To avoid confusion with the phrases 'falsely-coloured' or 'negative', which have distinct meanings in visualisation science, but to still provide a phrase for true colour error, we use the term "off-colour", and distinguish between off-colour that requires correction through addition (Figure 1417, bottom left) and subtraction (Figure 1417, bottom right). The assimilation procedure added yellow colours (suspended sediment) within Broad Sound and green colours (phytoplankton) in the mouth. Offshore the assimilation removed green, particularly, as noted above at 23°S, 151.5°E.  By removing green it made the water more blue (Figure 1417, top right).

### 4.2 Like-for-like assimilation
The general methodology presented in this study is similar to that used in numerical weather prediction (NWP). More than two decades ago, t̶T̶he NWP community moved away from assimilating satellite-derived temperature and humidity profiles more than two decades ago, in favour of assimilating radiances (or temperature brightness), this approach is detailed in Derber and Wu (1998). By assimilating satellite-derived radiance data, the NWP community avoided any reliance on empirical statistical relationships used to predict the temperature and humidity profiles. Similarly, the approach taken here is to avoid the use of an empirical/statistical inverse model, and use a physics-based forward model to predict RSR centred at the MODIS bandwidths. We then post-process these simulated RSRs into a simulated OC3M. The simulated OC3M is directly comparable to the observed OC3M with both containing quantitatively similar sources of error derived from bottom reflectance, and turbid coastal waters. By avoiding the use of an inverse empirical statistical model, we are presenting a BGC DA approach that had been adopted by the NWP community decades ago (Derber and Wu, 1998; Dee et al., 2015).

An additional benefit of avoiding the use of IOP / AOP based empirical/statistical products, is that assimilation of RSRs can take advantage of non-ocean-color specific missions such as Himawari 8. The spectral resolution of simulated RSR can be altered to simulate reflectances at the Himawari 8 true colour bands, . If the Himawari 8 data can be assimilated it will providinge a step change in the data available for areas such as the GBR, due to the high spatial resolution (nominally 500 m) and temporal resolution (every 30 minutes). This data density far exceeds that available from orbiting satellites, and will provide coverage similar to the products being assimilated in NWP systems.

### 4.3 Multi-band assimilation
If multiple reflectance bands are to be simultaneously assimilated, it will be necessary to account for cross-correlation between observation errors. For example, the RSR at 443 nm is strongly correlated with the RSR at 488 nm. Therefore, the observations of adjacent bands are no longer independent and it is likely we need to reconsider that the assumption that

the off diagonal elements of the observation error covariance matrix (**R**) are 0~~needs to be reconsidered~~. Using simulated and observed OC3M eliminates the possibility of cross-correlation and contains information derived from multiple bands. The OC3M algorithm can be considered a band-ratio function, f(RSR), that transforms multi-variate observations into univariate observations. It is likely that there are other functional forms that could combine information from multiple bands into a single non-correlated observation.

The ~~underlying~~ configuration of the data assimilation presented in this study requires the dominant error sub-space to be spanned by the ensemble. Pragmatic choices have been made to allow the system to run on the available compute resources. To this end we have perturbed 2 sensitive model parameters, and river loads of nutrients and sediments. The distributions that have been sampled to perturb the zooplankton mortality rates and $\theta_i$ along with their respective shape parameters could be considered a subjective choice. There is substantial scope to recast the problem with a Bayesian Hierarchical Modelling (BHM) framework (as in Parslow et al., 2013; and Dowd et al., 2014), whereby the prior distribution are assigned to uncertain parameters, and a thorough meta-analysis of the literature could be used to construct informative distributions. The observations could then be used to construct not only a posterior over the state, but a full joint posterior over the state and parameters. Furthermore, we have not allowed uncertainty in the physics to propagate into the BGC solution. We recognise this is a shortcoming of the study, however, given the computational constraints, we are not in a position to expand the ensemble to include physics perturbations (which would require an ensemble that is up to an order of magnitude larger). As more computing power becomes available, ensemble sizes could be increased, stochastic parameterisations introduced (Garnier et al., 2016), and DA methods with less parametric assumptions (e.g. Parslow et al., 2013), could be adopted.

There have been two recent discussion papers that detail the pathway towards operationalising BGC forecasting systems (Gehlen et al., 2015; and Ford and Barciela, 2015), analogous to the current NWP and hydrodynamic prediction system that routinely run at numerous operational centres. It has been acknowledged that satellite remote sensing will play a key role in such systems, there appears to be two possible pathways to achieve this vision, dependent upon on whether bio-optics are taken into account. In the absence of an optical model, Ford and Barciela (2015) suggest further exploration of the assimilation of empirical statistical products such as Chl-a and Phytoplankton Functional Types (PFTs). If simple bio-optical properties are taken into account, an alternative approach is the assimilation of diffuse attenuation coefficient(s) (e.g. Ciavatta et al., 2014) or phytoplankton absorption (e.g. Shulman et al., 2013). For complex coastal regions that are dominated by case 2 waters, an explicit spectrally-resolved in-water optics model opens the possibility of directly assimilating RSRs and avoids the costly requirement of calibrating an empirical IOP algorithm that is regionally-specific. Whilst the results from this study have shown to be valuable in the GBR region, further work needs to be undertaken to demonstrate the broad applicability of this approach. Nonetheless, we would advocate a third approach should be considered – the assimilation of RSR.

**5 Conclusion**

In this study we have used a spectrally-resolved optical model coupled to a BGC model to simulate the remote-sensing reflectances (RSR) centred at the MODIS OC bands. A series of assimilation system configuration experiments were undertaken to test the assimilation system performance. When the simulated OC3M (EXP4) and remote-sensing reflectances (EXP5) were assimilated into the model, the forecast errors in Chl-a fell from 100% to 55% when compared to the non-assimilating model. By using a function of the remote-sensing reflectances (OC3M), information from multiple bands are included in a univariate observation and the forecast error is halved compared to simply assuming the OC3M is directly related to the model prediction of surface total Chl-a. A comparison against in-situ observations of $NO_3$, $NH_4$, DIP and TSS shows the assimilating model (EXP4) reduces the MAPE from 90% to less than 20% at most stations. By using a forward model that includes a majority of error sources present in the observed OC3M, we have shown that the assimilation of remotely-sensed products in optically complex case 2 waters can be achieved, and adds substantial predictive skill when compared to the non-assimilating model. Furthermore, this approach can be generalized to non OC specific missions by assimilating the remote-sensing reflectances directly (e.g. EXP5), liberating a vast quantity of data that cannot be used in traditional BGC assimilation systems.

**Acknowledgements**

We would firstly like to acknowledge the thorough and insightful comments provided by the three reviewers: Prof Emmanuel Boss, Dr Stefano Ciavatta and David Ford. The scope and clarity of this manuscript have improved greatly due to their input. The model simulations were developed as part of the eReefs project, a public-private collaboration between Australia's leading operational and scientific research agencies, government, and corporate Australia (research.csiro.au/ereefs). Atmospherically-corrected MODIS products were sourced from NASA and the Integrated Marine Observing System (IMOS) - IMOS is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS) and the Super Science Initiative. Simulations and the processing of the remote sensing data was undertaken with the assistance of resources from the National Computational Infrastructure (NCI), which is supported by the Australian Government via NCRIS. We thank the many colleagues involved in developing the eReefs model, particularly Mike Herzfeld, John Andrewartha, Philip Gillibrand, Cedric Robillot and Richard Brinkman. Mark Baird was additionally funded by the CSIRO Wealth from Oceans Flagship, the Gas Industry Social & Environmental Research Alliance (GISERA), CSS FSP and the CSIRO Coastal Carbon Cluster. River flow data and information used to specify river nutrient and sediment loads were provided by the Queensland Government.

**References**

Ariathurai, R. & and Krone, R. B. 1976. Finite element model for cohesive sediment transport. Journal of the Hydraulics Division, ASCE, 104, HY2, 323–328.

Baird, M. E. (2010) Limits to prediction in a size-resolved pelagic ecosystem model. J. Plankton Res. 32, 1131-1146.

Baird, M. E., P. J. Ralph, K. Wild-Allen, F. Rizwi and A. D. L. Steven (2013). A dynamic model of the cellular carbon to chlorophyll ratio applied to a batch culture and a continental shelf ecosystem. Limnol. Oceanogr. 58, 1215-1226.

Baird, M. E., N. Cherukuru, E. Jones, N. Margvelashvili, M. Mongin, K. Oubelkheir, P. J. Ralph, F. Rizwi, B. J. Robson, T. Schroeder, J. Skerratt, A. D. L. Steven and K. A. Wild-Allen (2016a) Remote-sensing reflectance and true colour produced by a coupled hydrodynamic, optical, sediment, biogeochemical model of the Great Barrier Reef, Australia: comparison with satellite data. Env. Model. Software 78: 79-96.

Baird, M. E., M. P. Adams, R. C. Babcock, K. Oubelkheir, M. Mongin, K. A. Wild-Allen, J. Skerratt, B. J. Robson, K. Petrou, P. J. Ralph, K. R. O'Brien, A. B. Carter, J. C. and Jarvis, M. A. Rasheed (2016b) A biophysical representation of seagrass growth for application in a complex shallow-water biogeochemical model Ecological Modelling 325: 13-27.

Blondeau-Patissier, D., Brando, V. E., Oubelkheir, K., Dekker, A. G., Clementson, L. A., and Daniel, P. (2009). Bio-optical variability of the absorption and scattering properties of the Queensland inshore and reef waters, Australia. *Journal of Geophysical Research: Oceans*, *114*(C5).

Brando VE, Schroeder T, King E, and Dyce P, 2015. Reef Rescue Marine Monitoring Program: Using remote sensing for GBR-wide water quality. Final Report for 2012/13 Activities. CSIRO Report to the Great Barrier Reef Marine Park Authority, 213 pp.

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*(3), 167-192.

Ciavatta, S., Torres, R., Martinez-Vicente, V., Smyth, T., Dall'Olmo, G., Polimene, L. and Allen, J.I., 2014. Assimilation of remotely-sensed optical properties to improve marine biogeochemistry modelling. *Progress in Oceanography*, *127*, 74-95.

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

Collier, C. J., Waycott, M., and& McKenzie, L. J. (2012). Light thresholds derived from seagrass loss in the coastal zone of the northern Great Barrier Reef, Australia. *Ecological Indicators*, *23*, 211-219.

Cummings, J.A., 2005. Operational multivariate ocean data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *131*(613), 3583-3604.

De'ath, G., Fabricius, K. E., Sweatman, H., and Puotinen, M. (2012). The 27–year decline of coral cover on the Great Barrier Reef and its causes. *Proceedings of the National Academy of Sciences*, *109*(44), 17995-17999.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., and Bechtold, P. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, *137*(656), 553-597.

Derber, J.C. and Wu, W.S., 1998. The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Monthly Weather Review*, *126*(8), 2287-2299.

Dowd, M., Jones, E. and Parslow, J., 2014. A statistical overview and perspectives on data assimilation for marine biogeochemical models. *Environmetrics*, *25*(4), 203-213.

Dutkiewicz, S., Hickman, A. E., Jahn, O., Gregg, W. W., Mouw, C. B., and& Follows, M. J. (2015). Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model. *Biogeosciences, 12, 4447-4481*

Egbert, G.D. and Erofeeva, S.Y., 2002. Efficient inverse modeling of barotropic ocean tides. *Journal of Atmospheric and Oceanic Technology*, 19(2), 183-204.

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*(5), 751-771.

Ford, D., and& Barciela, R., 2015. Forecasting Research Technical Report No: 609 - Marine biogeochemical data assimilation–literature review and scoping report. 46 pages

Fujii, M., Boss, E., and& Chai, F. (2007). The value of adding optics to ecosystem models: a case study. *Biogeosciences*, *4*, 817-835.

Furnas, M. 2003. Catchments and corals: terrestrial runoff to the Great Barrier Reef. Australian Institute of Marine Science, Queensland., 334 p.

Garnier, F., Brankart, J.M., Brasseur, P. and Cosme, E., 2016. Stochastic parameterizations of biogeochemical uncertainties in a 1/4° NEMO/PISCES model for probabilistic comparisons with ocean color data. *Journal of Marine Systems*, *155*, 59-72.

5    Gaspari, G. and Cohn, S.E., 1999. Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, *125*(554), 723-757.

Gehlen, M., Barciela, R., Bertino, L., Brasseur, P., Butenschön, M., Chai, F., Crise, A., Drillet, Y., Ford, D., Lavoie, D. and Lehodey, P., 2015. Building the capacity for forecasting marine
10   biogeochemistry and ecosystems: recent advances and future developments. *Journal of Operational Oceanography*, *8*(sup1), s168-s187.

Gillibrand, P.A. and Herzfeld, M., 2016. A mass-conserving advection scheme for offline simulation of scalar transport in coastal ocean models. *Ocean Modelling*, *101*, 1-16.

Goyens, C., Jamet, C., and Schroeder, T. (2013). Evaluation of four atmospheric correction algorithms for MODIS-Aqua images over contrasted coastal waters. *Remote Sensing of Environment*, *131*, 63-75.

Gregg, W.W., 2008. Assimilation of SeaWiFS ocean chlorophyll data into a three-dimensional global ocean model. *Journal of Marine Systems*, *69*(3), 205-225.

Herzfeld, M. 2006, An alternative coordinate system for solving finite difference ocean
25   models., *Ocean Modelling*, 14, 174-196.

Herzfeld, M. 2015. Methods for freshwater riverine input into regional ocean models. *Ocean Modelling,* 90**,** 1-15.

30   Herzfeld, M., and P. A. Gillibrand (2015). "Active open boundary forcing using dual relaxation time-scales in downscaled ocean models." *Ocean Modelling* 89 (2015): 71-83.

Hu, J., Fennel, K., Mattern, J.P. and Wilkin, J., 2012. Data assimilation with a local Ensemble Kalman Filter applied to a three-dimensional biological model of the Middle Atlantic
35   Bight. *Journal of Marine Systems*, *94*, 145-156.

King, E. A., Schroeder, T., Brando, V. E., and Suber, K., 2014. A pre-operational System for Satellite Monitoring of the Great Barrier Reef Marine Water Quality. Tech. rep., CSIRO Report, pp56.

Lee, Z., Carder, K. L., and& Arnone, R. A. (2002). Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Applied optics*, *41*(27), 5755-5772.

Lynch, T. P., Morello, E. B., Evans, K., Richardson, A. J., Rochester, W., Steinberg, C. R., and Sherrington, R. (2014). IMOS National Reference Stations: a continental-wide physical, chemical and biological coastal observing system. *PloS one*, *9*(12), e113652.

Madsen, O.S., 1994. Spectral wave-current bottom boundary layer flows, in Coastal Engineering 1994 Proceedings, 24th International Conference Coastal Engineering Research Council/ASCE, 384-398.

Margvelashvili, N. (2009). Stretched Eulerian coordinate model of coastal sediment transport. *Computers & Geosciences*, *35*(6), 1167-1176.

Margvelashvili, N., Saint-Cast, F., and Condie, S., 2008. Numerical modelling of the suspended sediment transport in Torres Strait. *Continental Shelf Research* 28, 2241-2256.

Margvelashvili, N. Y., Herzfeld, M., Rizwi, F., Mongin, M., Baird, M. E., Jones, E., ... & Schroeder, T. (2016). Emulator-assisted data assimilation in complex models. *Ocean Dynamics*, *66*(9), 1109-1124.

Mongin, M., M. E. Baird, B. Tilbrook, R. J. Matear, A. Lenton, M. Herzfeld, K. A. Wild-Allen, J. Skerratt, N. Margvelashvili, B. J. Robson, C. M. Duarte, M. S. M. Gustafsson, P. J. Ralph, and A. D. L. Steven (2016). The exposure of the Great Barrier Reef to ocean acidification. *Nature Communications* 7, 10732.

Moore, T. S., J. W. Campbell and M. D. Dowell (2009). A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. *Remote Sensing of Environment 113*(11), 2424-2430.

Morel, A., Antoine, D., and Gentili, B., 2002. Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function. *Applied Optics* 41(30), 6289-6306.

Morello, E. B., Plagányi, É. E., Babcock, R. C., Sweatman, H., Hillary, R., and Punt, A. E. (2014). Model to manage and reduce crown-of-thorns starfish outbreaks. *Marine Ecology Progress Series*, *512*, 167-183.

Natvik, L.J. and Evensen, G., 2003. Assimilation of ocean colour data into a biochemical model of the North Atlantic: Part 1. Data assimilation experiments. *Journal of Marine Systems*, *40*, 127-153.

5   Odermatt, D., Gitelson, A., Brando, V.E. and Schaepman, M., 2012. Review of constituent retrieval in optically deep and complex waters from satellite imagery. *Remote sensing of environment*, 118, 116-126.

Oke, P.R., Brassington, G.B., Griffin, D.A. and Schiller, A., 2008a. The Bluelink ocean data
10  assimilation system (BODAS). *Ocean Modelling*, 21(1), 46-70.

Oke, P.R. and Sakov, P., 2008b. Representation error of oceanic observations for data assimilation. *Journal of Atmospheric and Oceanic Technology*, *25*(6), 1004-1017.

15  Parslow, J. S., Cressie, N., Campbell, E., Jones. E. M., and Murray, L. 2013. Bayesian learning and predictability in a stochastic nonlinear dynamical model. *Ecological Applications* 23:4, 679-698

Qin, Y., Brando, V.E., Dekker, A.G. and Blondeau-Patissier, D., 2007. Validity of SeaDAS water
20  constituents retrieval algorithms in Australian tropical coastal waters. *Geophysical Research Letters*, *34*(21).

Ridgway, K. R., Dunn, J. R., and Wilkin, J. L. (2002). Ocean interpolation by four-dimensional weighted least squares-application to the waters around Australasia. *Journal of atmospheric and oceanic technology, 19*(9), 1357-1375.

Ridgway K.R., J.R. Dunn, and J.L. Wilkin 2002 Ocean interpolation by four-dimensional least squares  Application to the waters around Australia. J. Atmos. Ocean. Tech., 19, 1357-1375.

30  Robson, B.J., Baird, M. and Wild-Allen, K., 2013. A physiological model for the marine cyanobacteria, *Trichodesmium*. In MODSIM2013, 20th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, ISBN ( 978-0).

35  Rolfe, J. and Gregg, D., 2015. Factors affecting adoption of improved management practices in the pastoral industry in Great Barrier Reef catchments. *Journal of Environmental Management*, 157, 182-193.

Sakov, P. and Oke, P.R., 2008. A deterministic formulation of the ensemble Kalman filter: an
40  alternative to ensemble square root filters. *Tellus A*, *60*(2), 361-371.

Sakov, P. and Bertino, L., 2011. Relation between two common localisation methods for the EnKF. *Computational Geosciences*, *15*(2), 225-237.

Sarmiento, J. L and N. Gruber (2006) Ocean Biogeochemical Dynamics. Princeton University Press. ISBN: 9780691017075. 528

Schiller, A., Herzfeld, M., Brinkman, R., and Stuart, G. (2014). Monitoring, predicting, and managing one of the seven natural wonders of the world. *Bulletin of the American Meteorological Society*, *95*(1), 23-30.

Schroeder, T., Behnert, I., Schaale, M., Fischer, J., and Doerffer, R. (2007). Atmospheric correction algorithm for MERIS above case-2 waters. *International Journal of Remote Sensing*, *28*(7), 1469-1486.

Schroeder, T., Devlin, M.J., Brando, V.E., Dekker, A.G., Brodie, J.E., Clementson, L.A., and McKinna, L., (2012). Inter-annual variability of wet season freshwater plume extent into the Great Barrier Reef lagoon based on satellite coastal ocean colour observations. *Marine Pollution Bulletin*, 65, 210-223.

Schroeder, T., Lovell J., King E., Clementson L, and Scott R, (2016), *IMOS Ocean Colour Validation Report 2015-16*.

Schaeffer, A., Roughan, M., Jones, E. M. and White, D. (2016) Physical and biogeochemical spatial scales of variability in the East Australian Current separation from shelf glider measurements. *Biogeosciences* 13(6): 1967-1975.

Shulman, I., Frolov, S., Anderson, S., Penta, B., Gould, R., Sakalaukus, P. and Ladner, S., 2013. Impact of bio-optical data assimilation on short-term coupled physical, bio-optical model predictions. *Journal of Geophysical Research: Oceans*, *118*(4), 2215-2230.

Skerratt, J., Wild-Allen, K., Rizwi, F., Whitehead, J. and Coughanowr, C., 2013. Use of a high resolution 3D fully coupled hydrodynamic, sediment and biogeochemical model to understand estuarine nutrient dynamics under various water quality scenarios. *Ocean & coastal management*, 83, 52-66.

Thompson, A., Costello, P., Davidson, J., Logan, M., Schaffelke, B., Uthicke, S. and Takahashi, M., 2011. Reef rescue marine monitoring program. Report of AIMS Activities–Inshore coral reef monitoring, p.128.

Thompson, A., Schroeder, T., Brando, V.E., Schaffelke, B., 2014. Coral community responses to declining water quality: Whitsunday Islands, Great Barrier Reef, Australia. Coral Reefs 33, 923-938.

5    Willmott, C.J., Ackleson SG, Davis RE, Feddema JJ, Klink KM,Legates DR, ODonnell J, Rowe CM. 1985. Statistics for the evaluation of model performance. *Journal of Geophysical Research* 90: 8995-9005.

Formatted: Font: Italic

**Appendix A: Detailed description of the eReefs modelling system**

The eReefs modelling system is a suite of coupled hydrodynamic, sediment, optical and ~~biogeochemical~~BGC models specifically tailored to the Great Barrier Reef. The
5   hydrodynamic model is a ~~three-dimensional~~3D, finite-difference, baroclinic model based on the ~~three-dimensional~~3D equations of momentum, continuity and conservation of heat and salt, employing the hydrostatic and Boussinesq assumptions (Herzfeld et al., 2006, Schiller et al., 2015). The equations of motion are discretized on a finite-difference stencil corresponding to the Arakawa C grid. In the vertical z-coordinate scheme, there are 47 fixed
10  z-levels. The atmospheric forcing products (wind, pressure, rain and heat fluxes) are supplied by Bureau of Meteorology (BOM) reanalysis products. A tidal signal was superimposed on the low-frequency sea level oscillation provided by BRAN2.3 (Oke et al., 2008a) on the regional grid open boundary. This tidal signal was introduced via a local flux adjustment. The OTIS tidal model (Egbert and Erofeeva, 2002) was used to generate the
15  tidal signal from amplitude and phase information for 8 constituents. The local grid open boundary was forced with temperature, salinity and velocity (with local flux adjustment) derived from the regional grid. A mass conserving flux-based advection scheme is used to transport sediment and ~~biogeochemical~~BGC tracers.

20  The sediment transport model adds a multilayer sediment bed to the hydrodynamic model grid and simulates sinking, deposition and resuspension of ~~multiply~~multiple size-classes of suspended sediment (Margvelashvili et al., 2008). The model solves advection-diffusion equations of the mass conservation of suspended and bottom sediments and is particularly suitable for representing fine sediment dynamics, including resuspension and transport of
25  ~~biogeochemical~~BGC particles. The model is initialised with the observed distribution of gravel, sand and mud in the seabed of the shelf region. Sediment particles settle on the seabed due to gravity and resuspend into the water column whenever the bottom shear stress, exerted by waves and currents, exceeds the critical shear stress of erosion. The resuspension and deposition fluxes are parameterised with the Ariathurai and Krone (1976)
30  formula. The bottom friction under combined waves and currents is estimated through the nonlinear bottom boundary layer model (Madsen, 1994).

Sediments in benthic layers undergo vertical mixing due to bioturbation, represented by local diffusion. The corresponding diffusion coefficient scales with the sediment depth so
35  that the bioturbation ceases to operate beneath the biologically active layer. The resistance of sediments to resuspension also varies with the sediment depth to reflect the consolidated nature of deep sediments. The numerical grid for sediment variables in the water column coincides with the numerical grid for the hydrodynamic model. Within the bottom sediments, the model utilises a time-varying sediment-thickness-adapted grid,
40  where the thickness of sediment layers varies with time to accommodate the deposited sediment. Horizontal resolution within sediments follows the resolution of the water column grid.

The ~~biogeochemical~~BGC model is organised into 3 zones: pelagic, epibenthic and sediment.
45  The epibenthic zone overlaps with the lowest pelagic layer and the top sediment layer, sharing the same dissolved and suspended particulate material fields. Dissolved and

particulate ~~biogeochemical~~BGC tracers are advected and diffused throughout the model domain. Additionally, ~~biogeochemical~~BGC particulate substances sink and are resuspended in the same way as sediment particles. ~~Biogeochemical~~BGC processes are organized into pelagic processes of phytoplankton and zooplankton growth and mortality, remineralisation of particulate and organic material, and fluxes of dissolved oxygen, nitrogen, phosphorus and carbon (including nitrogen fixation, phosphorus adsorption and desorption, surface gas exchanges, respiration and ~~phorosythensis~~photosynthesis, and fluxes to and from biotic pools); epibenthic processes of growth and mortality of macroalgae, seagrass and corals, and sediment based processes of phytoplankton mortality, microphytobenthos growth, detrital remineralisation and fluxes of dissolved substances (~~Fig.2~~Figure 3).

The ~~biogeochemical~~BGC model includes four groups of microalgae (small and large phytoplankton, *Trichodesmium* and microphytobenthos) and three macrophytes types (seagrass types corresponding to *Zostera* and *Halophila*, macroalgae and coral communities). Photosynthetic growth is determined by concentrations of dissolved nutrients (nitrogen and phosphate) and photosynthetically active radiation. Autotrophs take up dissolved ammonium, nitrate, phosphate and inorganic carbon, and in the case of *Trichodesmium*, fix atmospheric nitrogen (Robson et al., 2014). Microalgae incorporate carbon (C), nitrogen (N) and phosphorus (P) at the Redfield ratio (106C:16N:1P, Redfield 1963) while macrophytes do so at the Atkinson ratio (550C:30N:1P, Atkinson1983). Microalgae contain two pigments (chlorophyll-*a* and an accessory pigment), and have variable carbon:pigment ratios determined using a photoadaptation model (described in Baird et al., 2013).

Micro-zooplankton graze on small phytoplankton and meso-zooplankton graze on large phytoplankton and microzooplankton, at rates determined by particle encounter rates and maximum ingestion rates. Of the grazed material that is not incorporated into zooplankton biomass, half is released as dissolved and particulate carbon, nitrogen and phosphate, with the remainder forming detritus. Additional detritus accumulates by mortality. Detritus and dissolved organic substances are remineralised into inorganic carbon, nitrogen and phosphate with labile detritus transformed most rapidly (days), refractory detritus slower (months) and dissolved organic material transformed over the longest timescales (years). The production (by photosynthesis) and consumption (by respiration and remineralisation) of dissolved oxygen is also included in the model and depending on prevailing concentrations, facilitates or inhibits the oxidation of ammonia to nitrate and its subsequent denitrification (in the sediment) to di-nitrogen gas which is then lost from the system. Full details of equations used in the ~~biogeochemical~~BGC model are given by Baird et al.~~,~~ (~~2016b~~2016a) and details of parameter values and implementation for the Great Barrier Reef are given by Herzfeld et al. 2016

The model is forced using flow and concentrations of dissolved and particulate constituents from 21 rivers along the Queensland coast (north to south: Normanby, Daintree, Barron, combined Mulgrave+Russell, Johnstone, Tully, Herbert, Haughton, Burdekin, Don, O'Connell, Pioneer, Fitzroy, Burnett, Mary, Calliope, Boyne, Caboolture, Pine, combined Brisbane+Bremer, and combined Logan+Albert) and the Fly River in Papua New Guinea (Herzfeld~~ et al.~~, 2015). To determine river concentrations, sediment and nutrient observations were statistically evaluated over 10 years (Furnas 2003). Separate analysis was

undertaken for wet- (the Fly, and the northern most 6 rivers in Queensland) and dry-(remainder) catchment rivers. Volume-averaged wet season export coefficients based on this observed dataset were derived for wet- and dry-catchment river types, and mean flow-weighted concentrations determined. These constant concentrations are multiplied by higher frequency (daily) observed discharge data to calculate the flux of constituents at the river mouths.

The eReefs BGC and sediment model has 3 open ocean boundaries. Nutrient concentrations flowing in from the boundaries were obtained from the CSIRO Atlas of Regional Seas (CARS) 2009 climatology (Ridgway et al., 2003) and empirical nutrient-temperature relationships. The initial conditions are specified by a generalised empirical relationship and scaled nutrient profiles on the model density profile specifying top and bottom water column values from the CARS ocean atlas. Surface $NO_3$ is usually low (< 3 mg $m^{-3}$). In deeper waters nutrient concentrations increase from 0 to 1500 m depth and then remain constant down to the ocean floor (4000 m depth, 500 mg $m^{-3}$). The initial conditions for most other tracers were not spatially resolved, since observations for the outer reef and Coral Sea are limited temporally and spatially.

**Appendix B: Calculation of Inherent Optical Properties (IOPs).**

Formatted: Font: Bold, English (United States)

The optical model considers the processes of absorption and scattering by pure seawater, coloured dissolved organic matter (CDOM), non-algal particulates (NAP) and phytoplankton cells, as well as benthic reflectance. Here we describe the calculation of the inherent optical properties (IOPs), such as total phytoplankton absorption at a specific wavelength (those required for the assimilation), that are calculated from the model state variables (e.g. phytoplankton chlorophyll biomass) and model parameters (e.g. cell radius). A full description is given in Baird et al. (2016a), including benthic reflection calculations that are not given here and a description for how the optical model is used to calculate photosynthetic processes.

*B.1 Phytoplankton absorption.*
The absorption-cross section ($\alpha$) of a spherical cell of radius ($r$) pigment-specific absorption coefficient ($\gamma$), and homogeneous intracellular pigment concentration ($c_i$), calculated using geometric optics:

$$\alpha = \pi r r^2 \left( \frac{1 - 2(1 - (1 + 2\gamma c_i r)e^{-2\gamma c_i r})}{(2\gamma c_i r)^2} \right)$$

where $\pi r^2$ is the projected area of a sphere.

The use of an absorption cross-section of an individual cell has two significant advantages. Firstly, the same model parameters used here to calculated absorption in the water column are used to determine photosynthesis in the biogeochemical model, including the effect of packaging of pigments within cells. Secondly, the dynamic chlorophyll concentration modelled in the biogeochemical model can be explicitly included in the calculation of phytoplankton absorption (Baird et al., 2013). The absorption of a population of $n$ cell m$^{-3}$ is given by $n\alpha$ m$^{-1}$.

Formatted: Font: Italic
Formatted: Font: Italic

*B.2 Coloured Dissolved Organic Matter (CDOM) absorption.*
The absorption of CDOM at 443 nm, $a_{CDOM,443}$, is determined from a relationship with salinity in the region (Schroeder et al., 2012):

$$a_{CDOM,443} = -0.0332\, S + 1.2336$$

where $S$ is the salinity. In order to avoid unrealistic extrapolation, the salinity used in this relationship is the minimum of the model salinity and 36. In some cases coastal salinities exceed 36 due to evaporation. The absorption due to CDOM at other wavelengths is calculated using a CDOM spectral slope for the region (Blondeau-Patissier et al., 2009):

$$a_{CDOM} = a_{CDOM,443}\, exp(-S_{CDOM}(\lambda - 443))$$

where $S_{CDOM}$ = 0.012 nm$^{-1}$ is an approximate spectral slope for CDOM, with observations in the GBR ranging from 0.01 to 0.02 nm$^{-1}$ for significant concentrations of CDOM.

*B.3 Absorption due to non-algal particulate material.*

In the model, optically-significant non-algal particulates (NAPs) include mineral particulates and detritus, with NAP absorption given by:

$$a_{NAP} = c_1 \, NAP$$

Where $c_1$ is spectrally-resolved coefficients determined from in-situ observations in Gladstone Harbour at times when absorption was dominated by particles.

*B.4 Total absorption*

The total absorption, *a*, is given by:

$$a = a_w + a_{NAP} + a_{CDOM} + \sum_{x=1}^{N} n_x \alpha_x$$

Where $a_w$ is pure seawater absorption and *N* is the number of phytoplankton classes.

*B.5 Scattering.*

The total scattering coefficient is given by:

$$b = b_w + c_3 NAP + b_{phy} \sum_{x=1}^{N} n_x c_{i,x} V_x$$

where *NAP* is the concentration of non-algal particulates, $b_W$ is the scattering coefficient due to pure seawater, $c_3$ is the NAP-specific scattering coefficient and $b_{phy}$ is the chlorophyll-specific scattering coefficient with the water column chlorophyll concentration of each classes is given by $n_x c_{i,x} V_x$ (where $c_i$ is the intracellular chlorophyll concentration and *V* is the cell volume). Similarly to absorption of NAPs, spectrally-resolved values for $c_3$ were based on field observations in Gladstone Harbour at times when absorption was dominated by particles.

**Tables:**

Table 1: The subset of state variables included in the state vector and corresponding sum of the analytical ($E_A$) and difference in kind errors ($E_D$) expressed as standard deviations in log-space used for the five assimilation system configurations, note that $E_{tot} = E_A + E_D + E_R$, as per section 2.5.2, and $E_{tot}$ is used on the diagonal elements of the observation error covariance matrix (R). The bold variables in the state vector are used in the input to the observation operator. It should be noted than the state variables are transformed by taking the natural logarithm of the variables. The observation error is then applied to the log-transformed state vector.

| | Assimilation State Vector (**X**) | $E_A + E_D$ |
|---|---|---|
| EXP1 | Ln(**Surface Total Chl-a**, PhyS Chl-a, PhyL Chl-a) | 0.8 |
| EXP2 | Ln(**Simulated OC3M**, PhyS Chl-a, PhyL Chl-a) | 0.8 |
| EXP3 | Ln(**Simulated OC3M**, PhyS Chl-a, PhyL Chl-a) | 0.4 |
| EXP4 | Ln(**Simulated OC3M**, PhyS Chl-a, PhyL Chl-a, NO3, NH4, TSS) | 0.4 |
| EXP5 | Ln(**Simulated R551**, PhyS Chl-a, PhyL Chl-a, NO3, NH4, TSS) | 0.2 |

Table 2: Forecast error statistics for OC3M (mg Chl-a m$^{-3}$) ~~in EXP4~~ by region (inshore, lagoon, offshore) for the Control (C) and EXP4 , Forecast (F), Analysis (A) and Persistence (P) fields.

| Region | Whole of Domain | | | | Coastal | | | | Lagoon and Shelf | | | | Oceanic | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Field | C | F | A | P | C | F | A | P | C | F | A | P | C | F | A | P |
| Median | 0.32 | 0.23 | 0.19 | 0.26 | 1.35 | 1.29 | 1.12 | 1.46 | 0.34 | 0.25 | 0.19 | 0.25 | 0.16 | 0.1 | 0.06 | 0.11 |
| Mean | 0.37 | 0.24 | 0.2 | 0.27 | 1.92 | 1.37 | 1.25 | 1.46 | 0.38 | 0.24 | 0.2 | 0.25 | 0.16 | 0.1 | 0.06 | 0.11 |
| 25% Quartile | 0.27 | 0.2 | 0.14 | 0.21 | 1.1 | 0.95 | 0.92 | 0.94 | 0.29 | 0.16 | 0.14 | 0.21 | 0.14 | 0.08 | 0.04 | 0.7 |
| 75% Quartile | 0.48 | 0.3 | 0.23 | 0.29 | 2.45 | 1.87 | 1.64 | 1.08 | 0.43 | 0.29 | 0.22 | 0.29 | 0.19 | 0.13 | 0.08 | 0.12 |

Table 3: Forecast Percentage (%) errors for OC3M ~~in EXP4~~ by region (inshore, lagoon, offshore) for the Control (C) and EXP4 , Forecast (F), Analysis (A) and Persistence (P) fields.

| Region | Whole of Domain | | | | Coastal | | | | Lagoon and Shelf | | | | Oceanic | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Field | C | F | A | P | C | F | A | P | C | F | A | P | C | F | A | P |
| Median | 100 | 53 | 39 | 53 | 130 | 95 | 90 | 105 | 95 | 47 | 38 | 52 | 93 | 48 | 31 | 48 |
| Mean | 107 | 54 | 38 | 55 | 180 | 97 | 90 | 107 | 102 | 51 | 37 | 56 | 96 | 50 | 31 | 51 |
| 25% Quartile | 81 | 41 | 35 | 42 | 105 | 105 | 95 | 74 | 75 | 43 | 36 | 41 | 71 | 44 | 35 | 39 |
| 75% Quartile | 131 | 61 | 41 | 62 | 181 | 85 | 81 | 145 | 118 | 62 | 42 | 62 | 126 | 55 | 30 | 62 |

Table 4: Root Means Square Differences (RMSDs) for the withheld observations and the control run (control), and assimilation Experiment 2 (EXP2), Experiment 4 (EXP4) and Experiment 5 (EXP5). The lowest RMSD is given in bold.

| Observed Variable | Control | EXP2 | EXP4 | EXP5 |
|---|---|---|---|---|
| Chl-a (mg Chl-a m$^{-3}$) -GBRMPA MMP | 0.5327 | 0.7270 | **0.4524** | 0.4957 |
| Chl-a (mg Chl-a m$^{-3}$) – IMOS Glider | 0.2064 | 0.1828 | 0.1344 | **0.1232** |
| TSS (mg solids m$^{-3}$) -GBRMPA MMP | 0.0081 | 0.0027 | **0.0017** | 0.0018 |
| NO3 (mg N m$^{-3}$) - IMOS NRS | 6.5126 | 3.8284 | **1.1361** | 1.6508 |
| NH4 (mg N m$^{-3}$) - IMOS NRS | 4.2795 | 2.3052 | 1.6384 | **1.6180** |
| DIP (mg P m$^{-3}$) - IMOS NRS | 2.1579 | **1.2574** | 1.3785 | 1.5387 |

**Figures:**
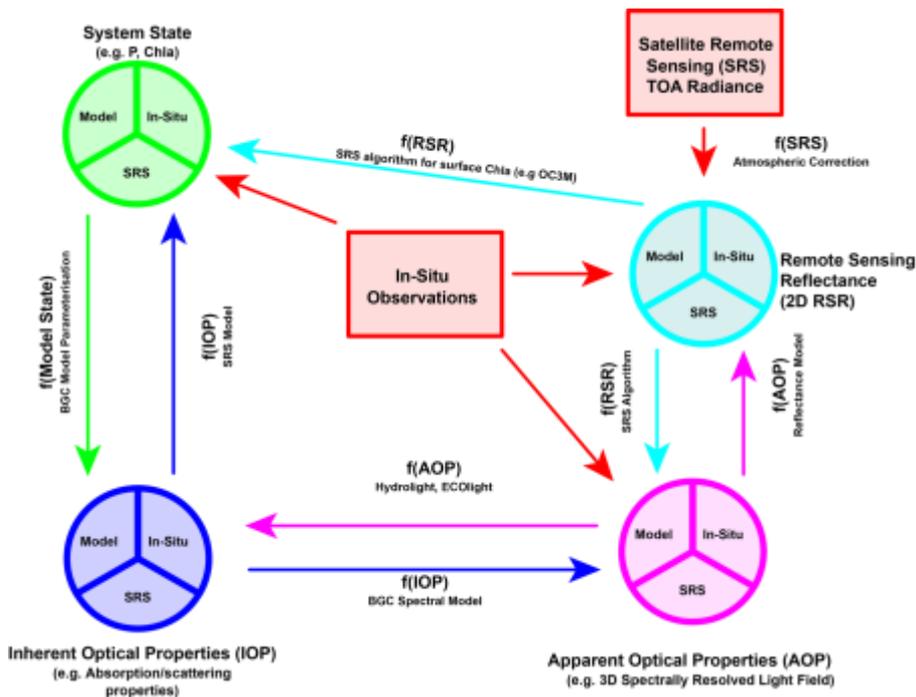
Figure 1: A diagram denoting the steps used to relate and transform various optical properties with the system state and observations. The green circle represents the in-water system state as predicted by a marine BGC model. The blue circle represents the Inherent Optical Properties (IOPs) of the system state. The magenta circle represents the depth resolved Apparent Optical Properties (AOPs). The cyan circle represents the 2D remote-sensing reflectance (RSR). The two red boxes represent either *in -situ* or satellite remote-sensing observations. Each circle is partitioned into three segments where each segments represents the possibility to compare like for like variables.

Formatted: Font: Italic
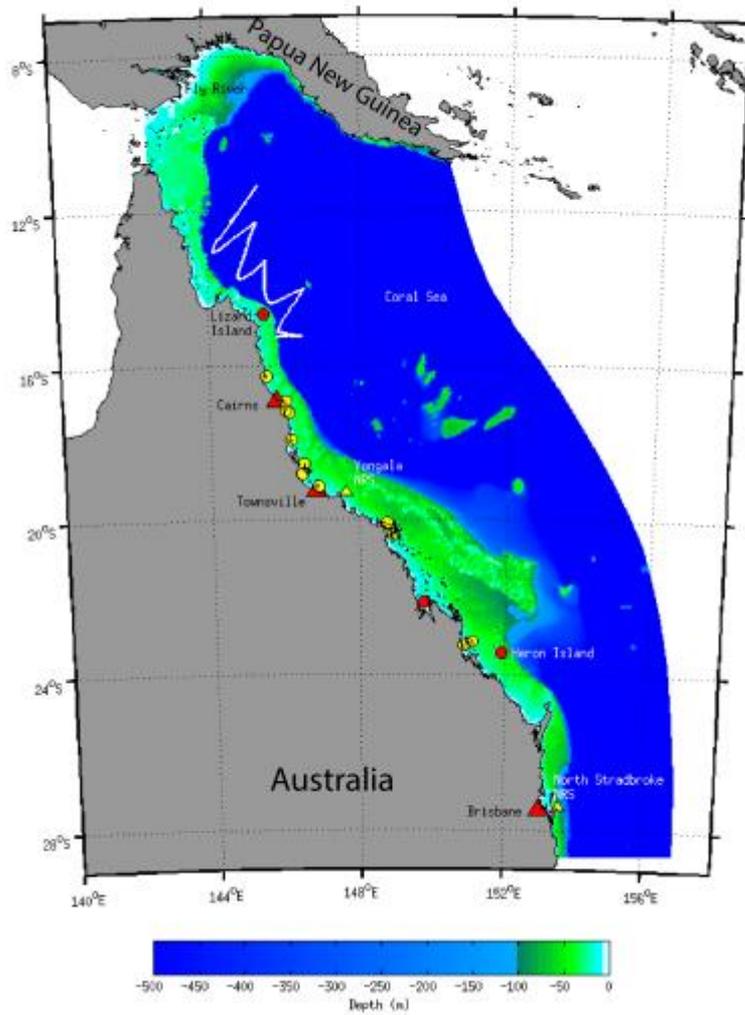
**Figure 2: A map of the Great Barrier Reef region, with the color bar denoting the water depth, markers denot~~inge~~ the population centres (red triangles), IMOS NRS sites (yellow triangles), GBRMPA MMP Water Quality Meters (WQMs; yellow circles) and points of interest referred to in the text (red circles), with the glider track (white line adjacent to Lizard Island). The *in situ* sampling locations and glider observations are used to assess the data assimilation system performance.**
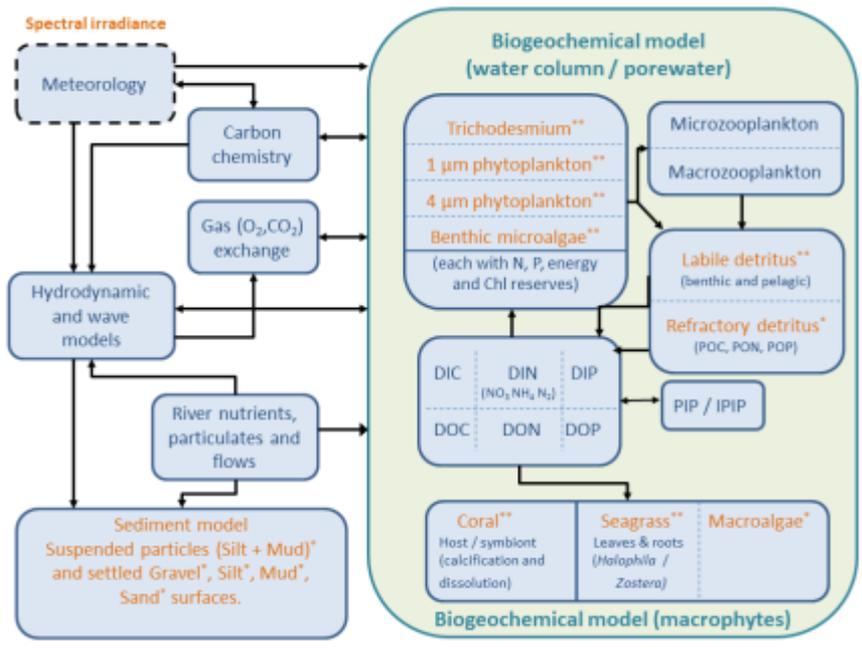
**Figure 3: The eReefs modelling system with optically-active components identified with beige colouring and an asterisk, with the number of asterisks denoting the number of different optically-active elements with this component. Thus, each of the four microalgae have two pigment types, one absorbing like divinyl chlorophyll-a, and the other like photosynthetic carotenoids; there are two seagrass types, corals have both skeletons and zooxanthellae; three types of detritus absorb and scatter, and the sediment model contains a suspended fraction and four (mixed) sediment compositions. Additionally, pure seawater both absorbs and scatters light.**
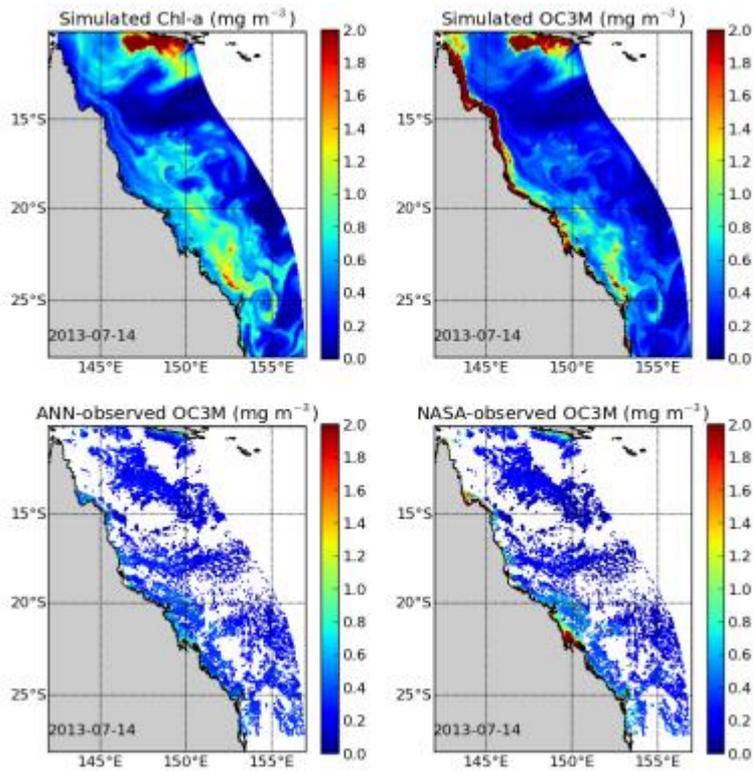
**Figure 4: Simulated surface Chl-a (mg m$^{-3}$) of the non-assimilating control run (top left) and the simulated OC3M (top right) derived from the simulated remote-sensing reflectance for the 14/7/2013. The observed OC3M with ANN-derived observed remote-sensing reflectance (bottom left) and NASA-derived observed remote-sensing reflectance (bottom right).**
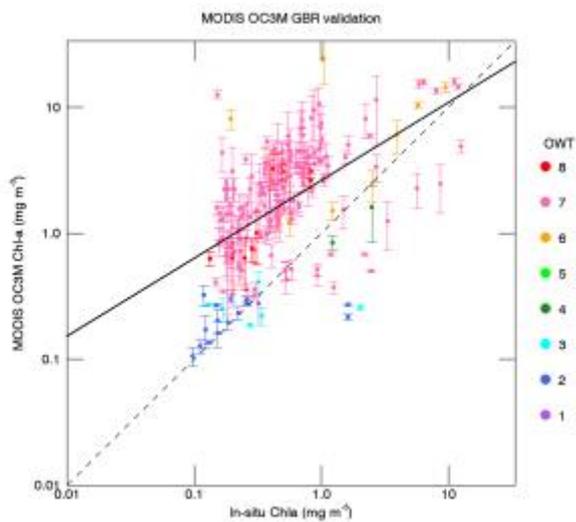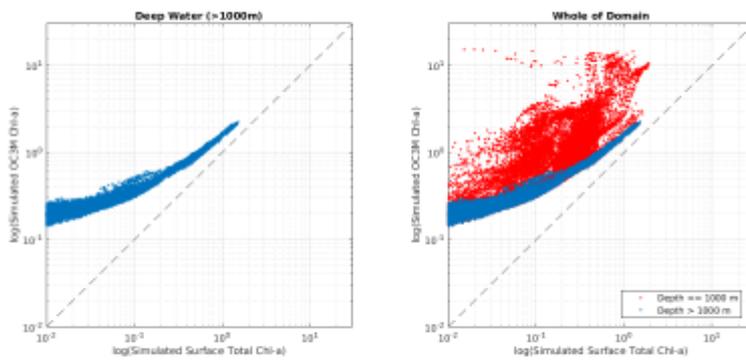
Figure 5: A log$_{10}$-scaled scatter plot of simulated surface total Chl-a and simulated OC3M for deep water regions (top-left) and whole of domain (top-right). The lower panel contains a scatter plot of MODIS OC3M against *in situ* Chl-a data obtained from the IMOS Bio-Optical Database which is publicly available through The Australian Ocean Data Network (AODN) portal (https://portal.aodn.org.au/ ). Matchups were included for MODIS overpasses within +/- 24 hours of the in-situ sampling. Points are coloured according to Optical Water Type [Moore et al 2009], ranging from low-chlorophyll open water (types 1-2) to coastal, CDOM and sediment-dominated waters (types 6–8).
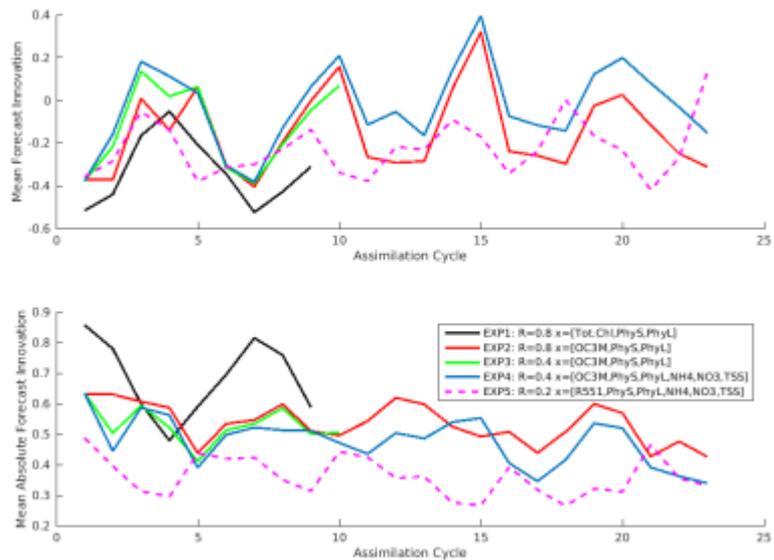
**Figure 6: Comparison of mean innovation statistics for EXP 1-5. EXP 5 should be analysed with caution as these innovations relate to R551, not OC3M. An innovation of 0 indicates perfect agreement between model and observations. The top panel plots the mean innovation for each assimilation cycle. The lower panel plots the mean absolute innovation against assimilation cycle. The colours correspond to: black (EXP1), red (EXP2), green (EXP3), blue (EXP4) and magenta (EXP5). The variables in the legend correspond to the observation error (R) and the assimilation state vector (X).**

.
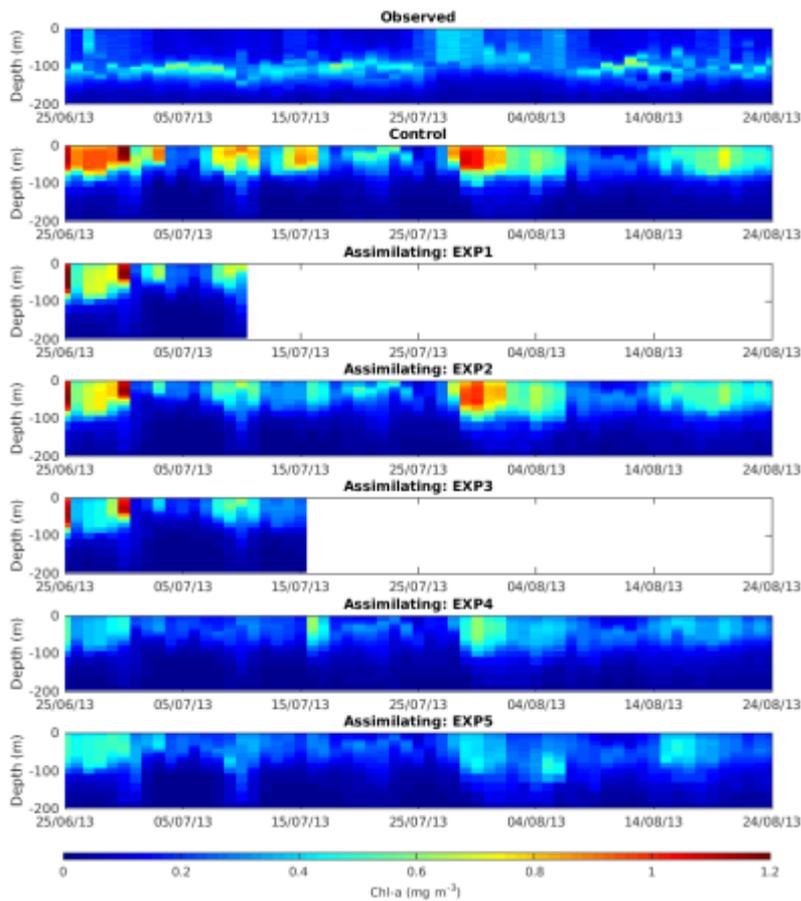
**Figure 7:** Comparison against independent IMOS glider observations, with the observed section of Chl-a derived from the onboard fluorescence sensor. The comparison is undertaken in model space, whereby all glider data that falls within a 1 hour window either side of when there is a 3D model output available is extracted, and equivalent simulated Chl-a are extracted from the model. The glider and model data is then spatially aggregated and interpolated onto the vertical model grid. The resulting section for the glider (top), and the simulated glider sections from the non-assimilating model (control) and ~~assimilating model (EXP1, EXP4 and EXP5).~~the five assimilation experiments.
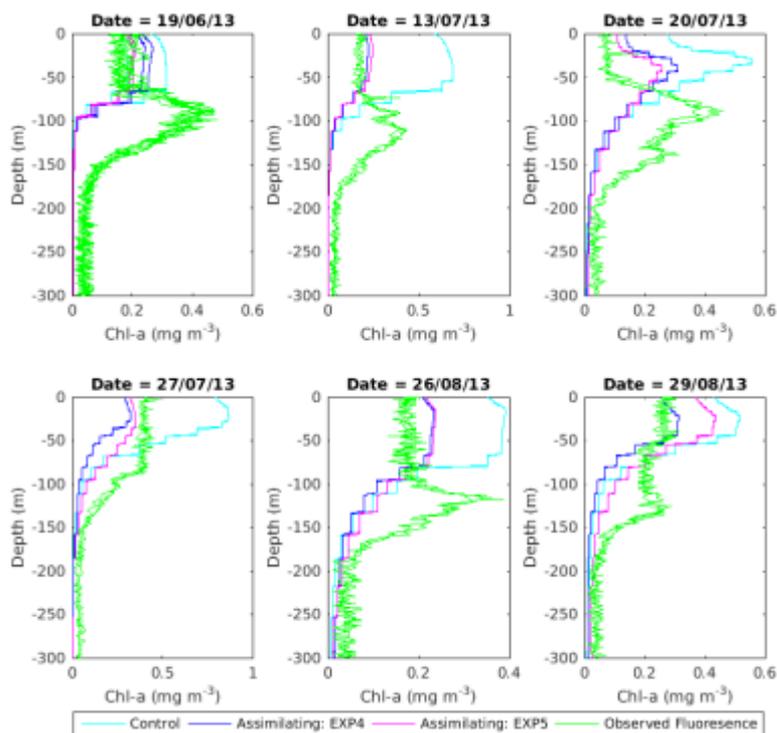
**Figure 8: Comparison of depth-resolved Chl-a against six individual glider profiles using the method described in the caption of Figure 7. Glider observations are green, the non-assimilating control run profiles are cyan, and the assimilating run profiles for EXP4 are blue and EXP5 are magenta.**
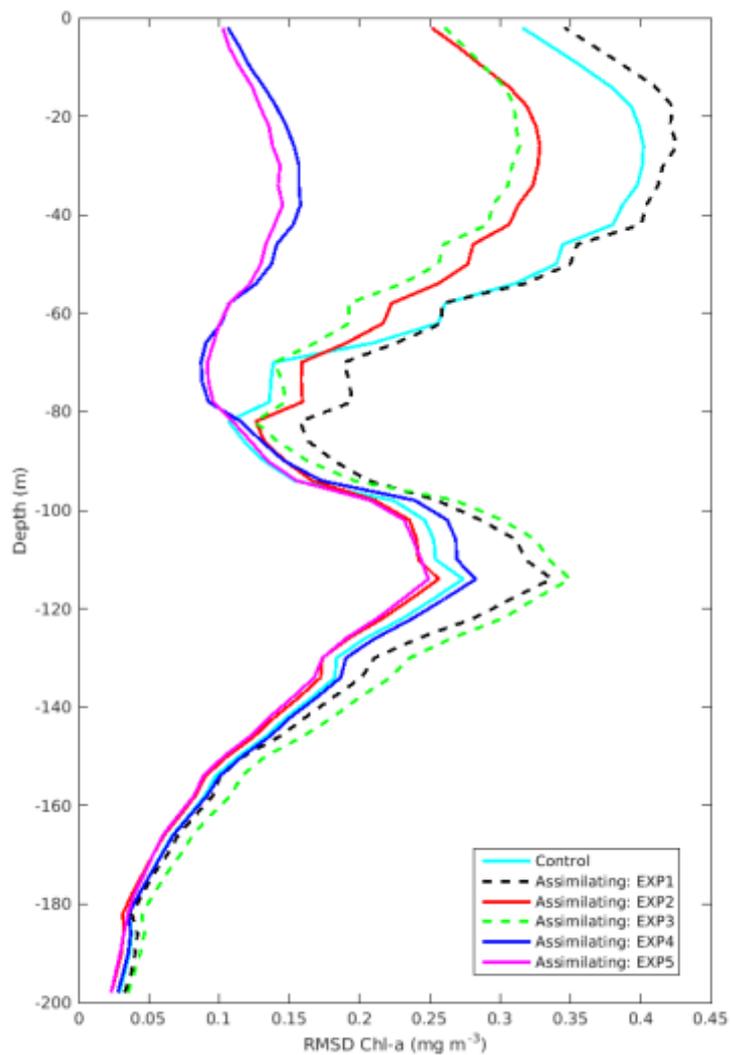
**Figure 9: A profile of the temporal mean Chl-a RMSD between the glider observations presented in Figure 7 and the non-assimilating control (cyan), EXP1 (black dashed, note that the mean RMSD is calculated using a short time period), EXP2 (red), EXP3 (green dashed, note that the mean RMSD is calculated using a shorter time period), EXP4 (blue) and EXP5 (magenta).**

Formatted: Font: 11 pt, Complex Script Font: 11 pt
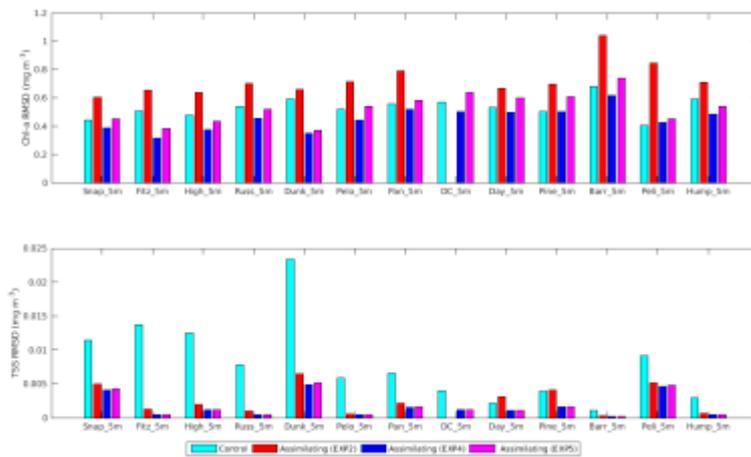
**Figure 10: A comparison of Chl-a and TSS RMSDs between the in-situ GBRMPA MMP moorings for the non-assimilating (cyan) and assimilating runs of EXP2 (red), EXP4 (blue) and EXP5 (magenta), the GBRMPA MMP sites are denoted by the yellow circles in Figure 1.**

Formatted: Font: 11 pt, Complex Script Font: 11 pt
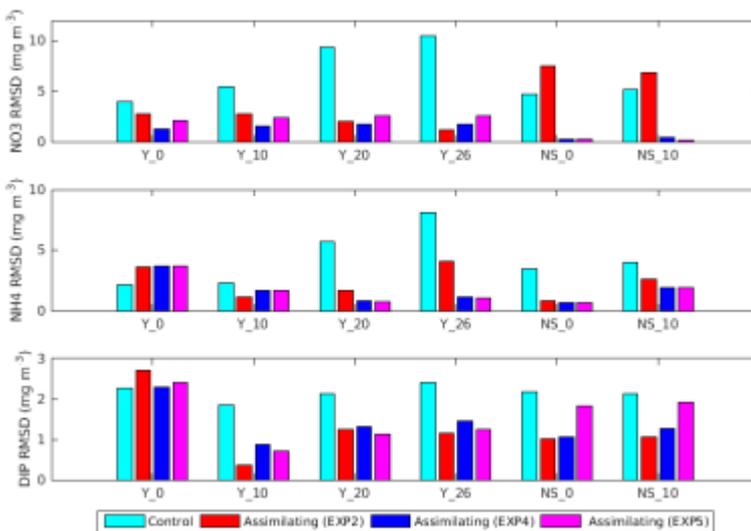
**Figure 11: A comparison of RMSD of simulated nutrients with in-situ bottle samples for the non-assimilating (cyan) and assimilating run of EXP2 (red), EXP4 (blue) and EXP5 (magenta). Observations are obtained at the Queensland IMOS site (yellow triangles in Figure 1) at Yongala (Y) and North Stradbroke (NS) Island. Y_0 are the Yongala surface samples, while Y_26 are the samples taken from 26 m depth.**
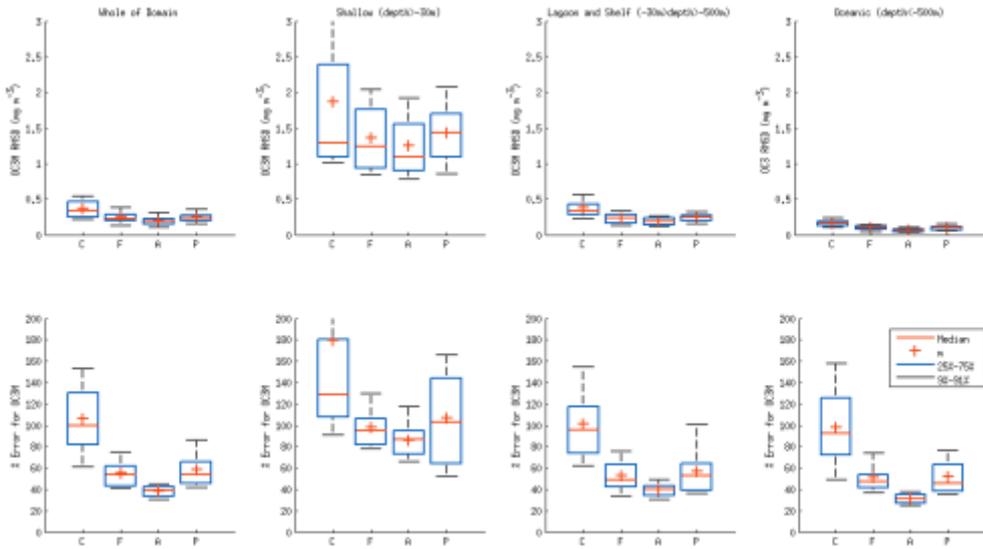
**Figure 12: Box and whisker plots of RMSD (top row) and MAPE (bottom row) of the mis-match between simulated OC3M and ANN-corrected observed OC3M. Each panel contains the control run (C) and EXP4 showing forecast (F), persistence (P) and analysis (A). Presented are statistics for the whole domain (left column), and regions as defined by bottom depth depth (three rightmost columns), for which the mean and range of values are presented in Tables 2 and 3.**
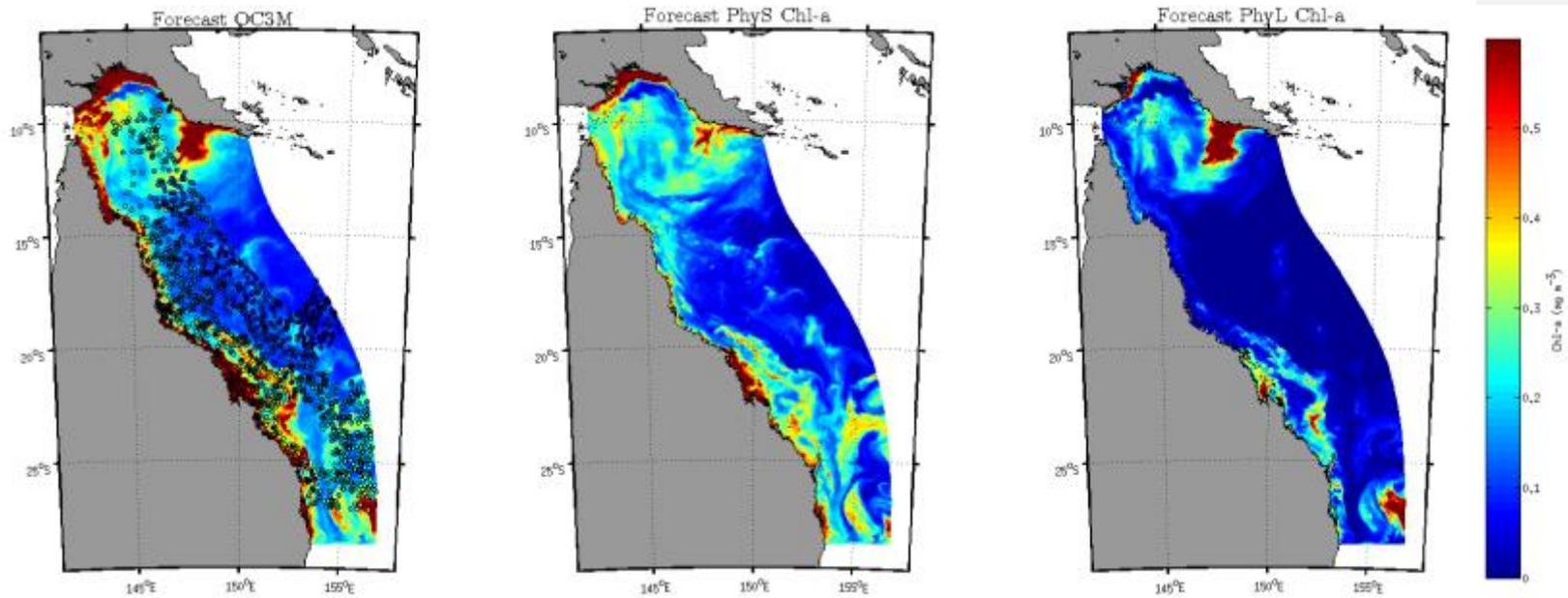
**Figure 13: Simulated ~~-~~(forecast) OC3M (left) for cycle 22 (12th September 2013)** of EXP4 **with** ~~observatns~~observations **overlaid, surface Small Phytoplankton (PhyS) Chl-a (centre) and surface Large Phytoplankton (PhyL) Chl-a (right).**
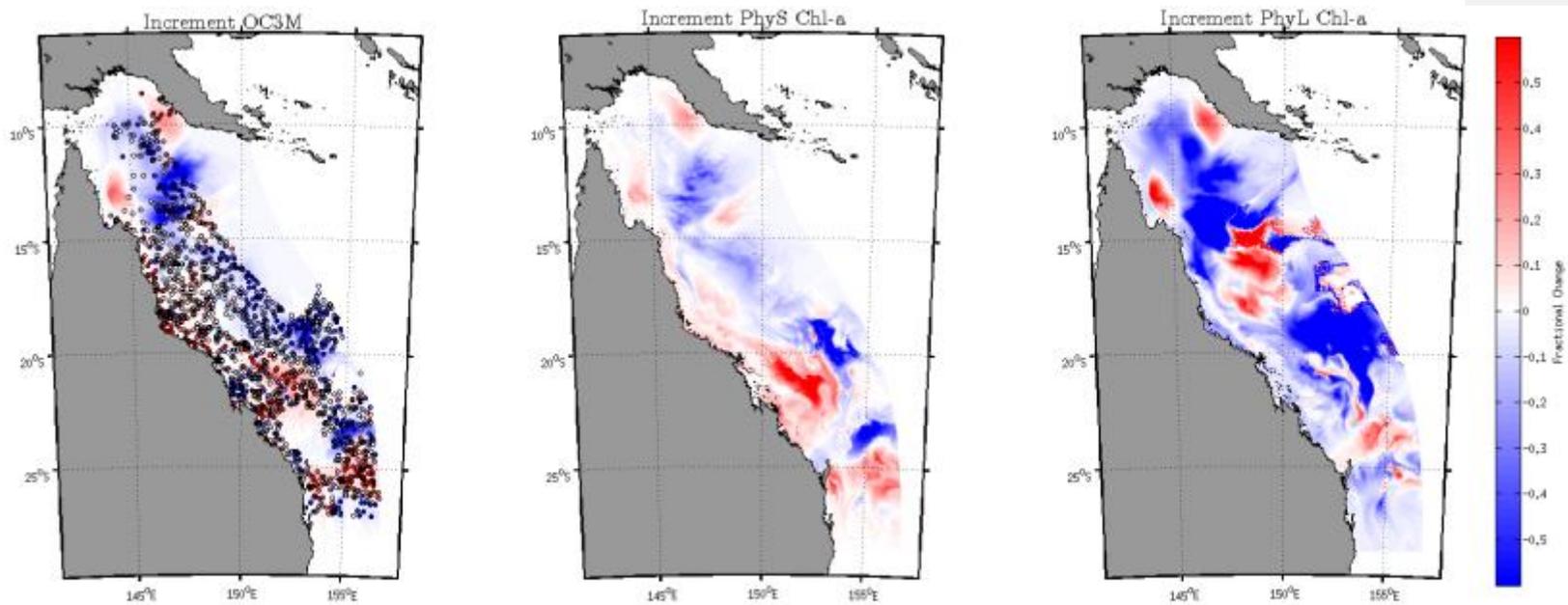
**Figure 14: Increments that are added to the forecast fields generated by the assimilation system (EXP4 configuration) for simulated OC3M with innovations overlaid (left), and the prognostic variables of surface small phytoplankton Chl-a (centre) and surface large phytoplankton Chl-a (right) for cycle 22.**
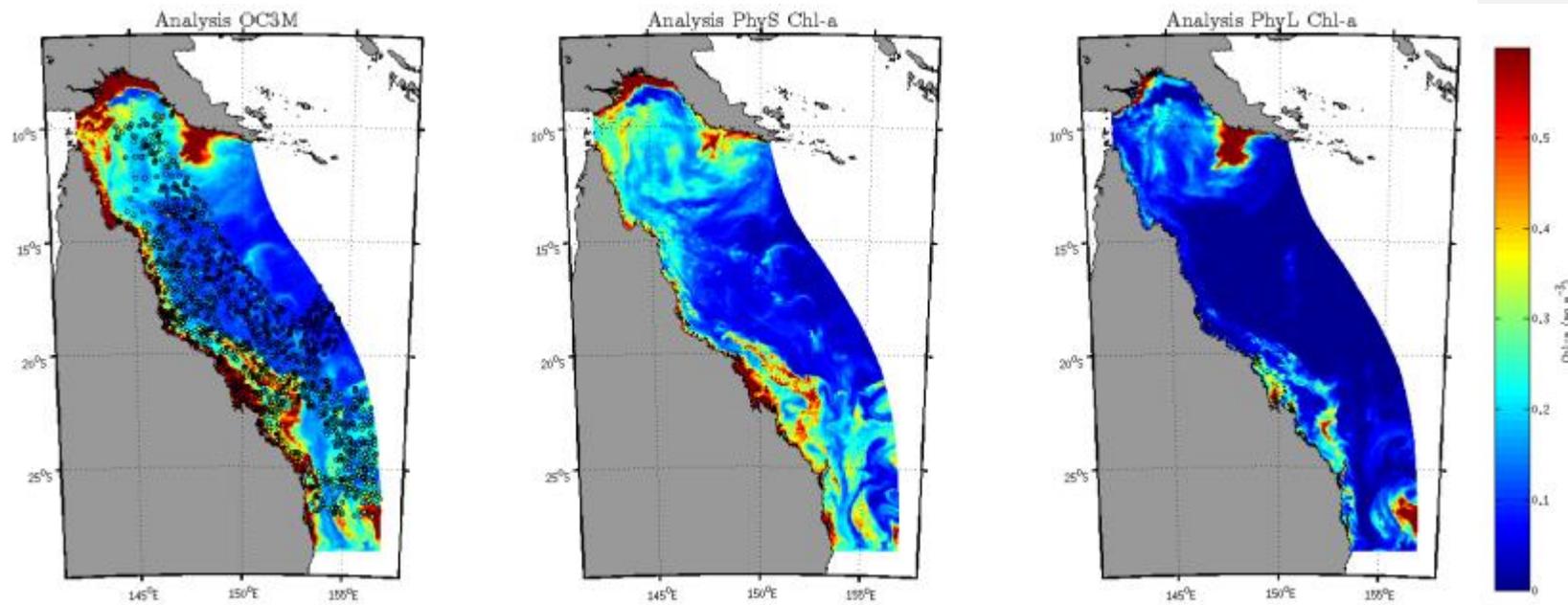
Figure 15: The resulting analysis fields for simulated OC3M and withheld ANN OC3M observations for EXP4 (cycle 22) overlaid (left), and the analysis fields for the prognostic variables of small phytoplankton (centre) and large phytoplankton (right) for the 12th September, 2013.
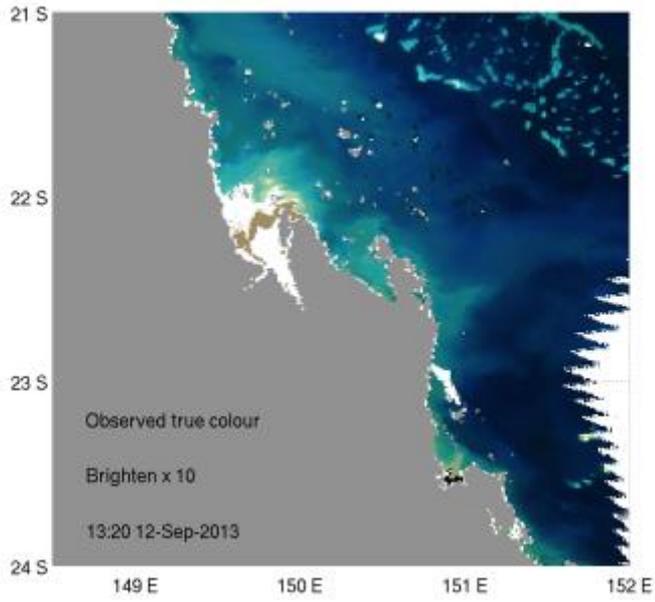
**Figure 16: Observed true-colour image on the 12 Sep 2013 obtained from 1 km resolution, atmospherically-corrected ANN remote-sensing reflectance. The RGB wavelengths used were 667, 551 and 488 nm and processed using the MODIS true colour algorithm (Gumley et al., 2010; ~~Baird et al., 2016~~Baird et al., 2016a). The white pixels are clouds, grey is land.**
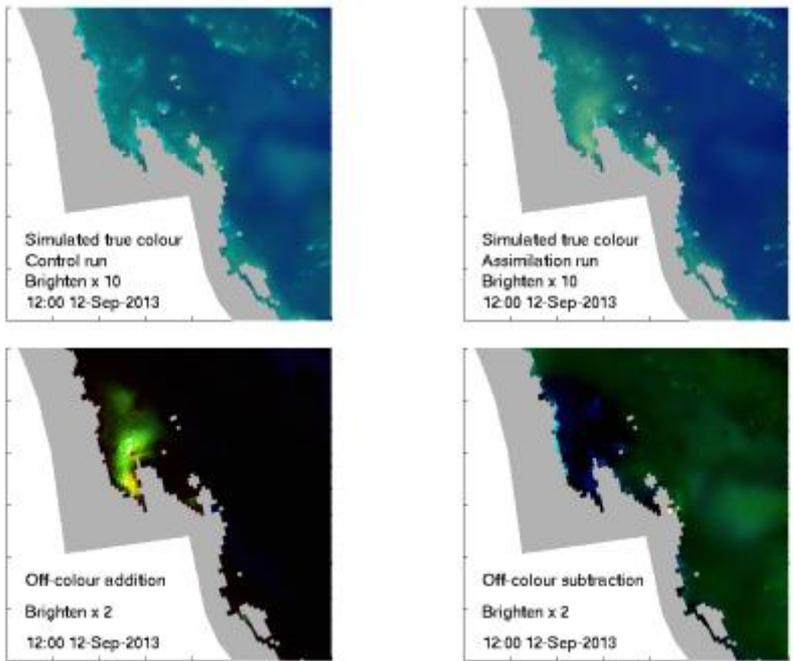
Figure 17: The simulated true-colour image on the 12 Sep 2013 of the control run (top left) and EXP4 assimilation run (top right). The difference between the remote-sensing reflectance in the control and assimilated runs was used to quantify the colour (referred to as off-colour) added (i.e. greater surface expression, bottom left) and subtracted (i.e. less surface expression, bottom right) due to the updating of optically-active constituents in the assimilation run (see Fig. 13 16 for more details). Note that the off-colour images have a smaller brightening factor as the MODIS true colour stretch saturates the features that are of most interest. Simulated true colour images are not falsely-coloured, thus do not require a colour map, nor are they 2D as they have a depth of field, being based on reflectance from multiple depths and the bottom (Baird et al., 2016aBaird et al., 2016a). Thus simulated true colour can be considered a photograph of the optical state of the different model runs, and, like observed true colour, a powerful and intuitive visualisation tool for water clarity in biogeochemicalBGC models.