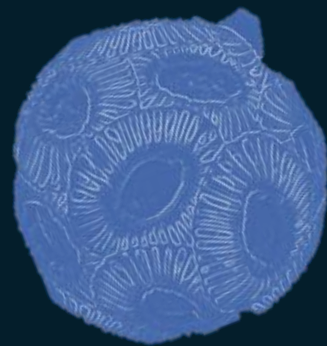
Evaluation Of Ocean Colour Approaches For Estimating Particulate Inorganic Carbon In The Ocean

Christina Eunjin Kong, Shubha Sathyendranath, Gemma Kulk, & Thomas Jackson

Introduction

Particulate Inorganic Carbon (PIC) plays an important role in the global ocean carbon cycle. Most PIC is in the form of calcium carbonate which is produced by various marine organisms though the process of calcification. Of the calcifying marine organisms, coccolithophores, which can be observed from space, especially during bloom conditions, are estimated to be the major producers of PIC in the open ocean [1,2].

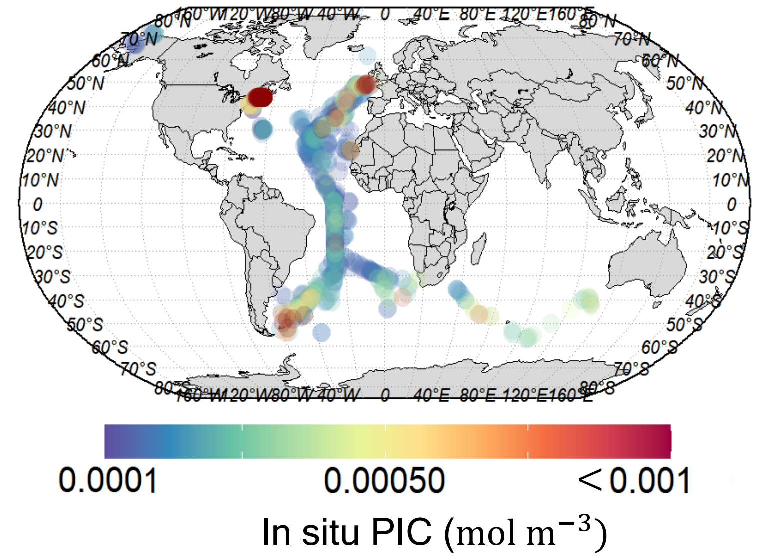


coccolithophores (*Emiliania huxleyi*)

High spatiotemporal resolution of PIC data can be obtained using in situ PIC measurements together with remotely sensed data. A considerable effort has been made to improve satellite-derived PIC products [3]. In this study, we evaluated two satellite approaches to estimate global PIC from the Ocean Colour Climate Change Initiative (OC-CCI) version 5 data (1997-2020). The accuracy of satellite-derived PIC products were evaluated using a large collection of near-surface in situ data (0 to 10 m depth) obtained from multiple databases (Figure 1).

Data

(a)



(b)

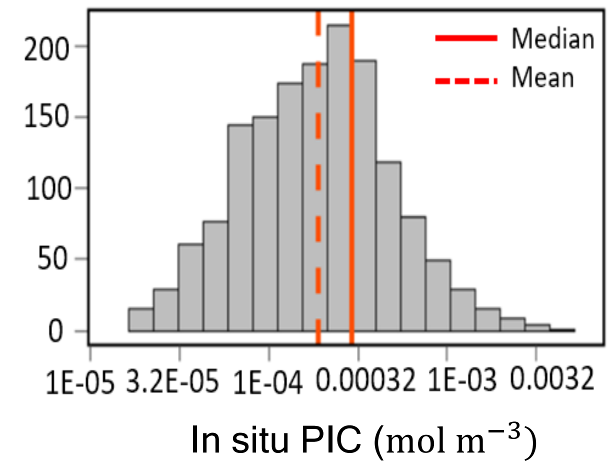


Figure 1. (a) Location of near surface (0 to 10 m) in situ PIC data used in this study (N: 1674). These represent multiple field measurements obtained from the SeaBASS, BCO-DMO and PANGAEA databases (please see QR code for more information). (b) Distribution of near surface in situ PIC concentration.

Method

Candidate PIC approaches for the OC-CCI products:

I. Colour-index-derived PIC (Mitchell et al., 2017)

- Mitchell et al. (2017) showed a relatively good performance for this algorithm compared with the merged global PIC algorithm (Gordon et al. 2001; Balch et al. 2005) currently adopted by NASA's Ocean Biological Processing Group.
- Since OC-CCI provide Rrs values at MERIS bands, whereas the Mitchell et al. algorithm is designed for MODIS bands, a band shift was performed to obtain Rrs(547) and Rrs(667), to implement Mitchell et al. (2017) algorithm with OC-CCI data.

II. Random-Forest-derived PIC (This study)

- We also explored the potential of random forest method for deriving PIC for satellite data. Details of the random forest method are provided in Figure 2.
- The candidate algorithm was evaluated the following statistical metrics: bias (φ), root-mean square deviation (ψ), central root mean square deviation(Δ), and coefficient of determination (r^2).

References to satellite-derived PIC algorithms

Random Forest
- **Mitchell et al. (2017)**

ML
- Sadeghi et al. (2012)
- Moor et al.(2012)
- Balch et al. (2005)
- Smyth et al. (2004)
- Gordon et al. (2001)
- Kondrik et al. (2018)

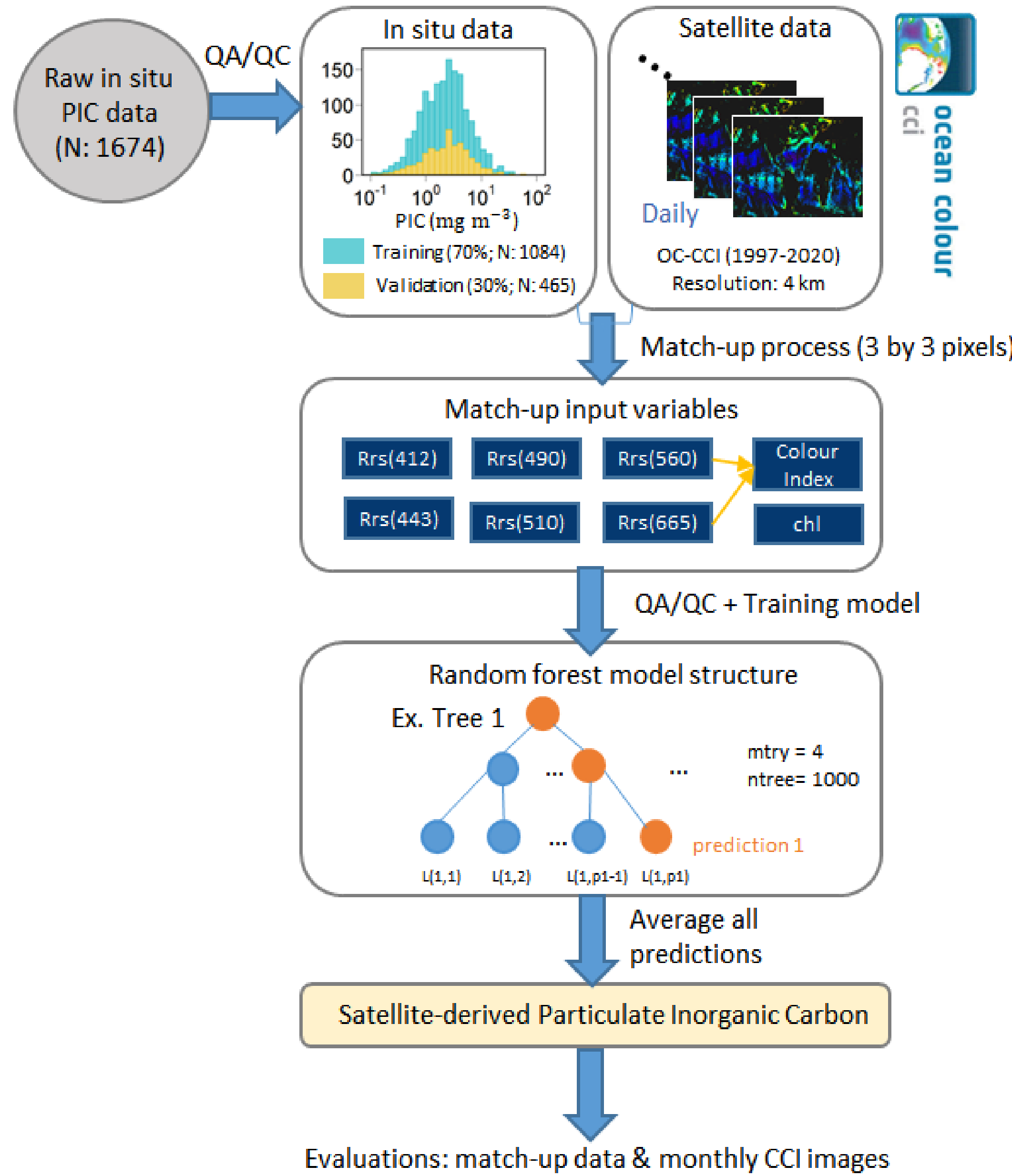


Figure 2. Schematic diagram illustrating the random forest algorithm development and evaluation method to estimate the PIC concentration from the OC-CCI products (1997-2020).

Results

- All OC-CCI input variables showed positive importance (Figure 3). The Rrs(560) and colour-index were two key drivers of random-forest-derived PIC values, followed by other Rrs values and chl concentration.

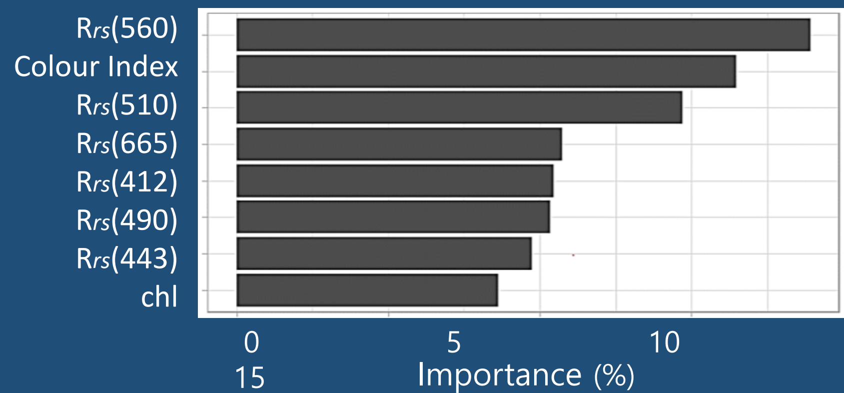


Figure 3. Importance (%) of input variables used in the random forest model.

- The results showed that the random-forest approach, which incorporated extra variables in addition to the colour index, performed well, with low uncertainties in the retrieved PIC (Figure 4).

- Both approaches showed high concentrations of PIC very close to the shore and in the Southern Ocean (Figure 5). In general, the difference between the two algorithms was positive in the tropics and the southern hemisphere, and negative in the temperate northern latitudes.

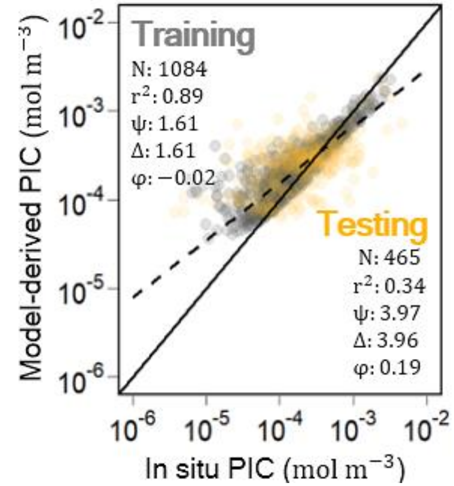


Figure 4. Scatter plot and statistics detailing the performance of random-forest-derived PIC versus in situ PIC match-up for training (grey) and testing data (yellow). The solid line is the 1:1 line, and the dashed line is the best fit for the linear regression of training data.

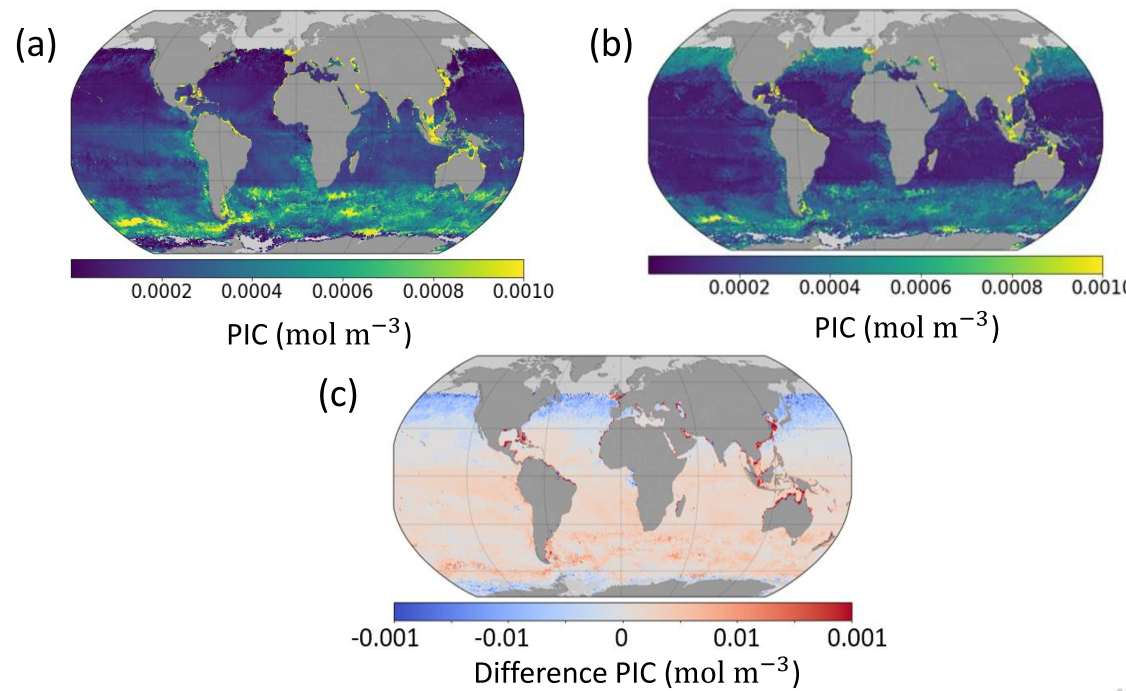


Figure 5. Spatial maps of (a) colour-index-derived PIC, and (b) random-forest-derived PIC using a monthly OC-CCIv5 data (Dec. 2020). (c) Absolute difference between colour-index-derived PIC and random-forest –derived PIC values of the same image.

Summary

- Mitchell et al. (2017) showed that a colour-index algorithm for retrieval of PIC performed better than the other algorithms that they tested.
- In this work, we explored a random forest method incorporating spectral Rrs values and chl concentration in addition to the colour-index values. The method was implemented using OC-CCI data.
- Comparison of both approaches with in situ data from AMT cruises in the Atlantic Ocean (Figure 6), showed that the random forest model captures the trend of in situ PIC well. The new result showed some improvement over the original colour-index-derived PIC algorithm.

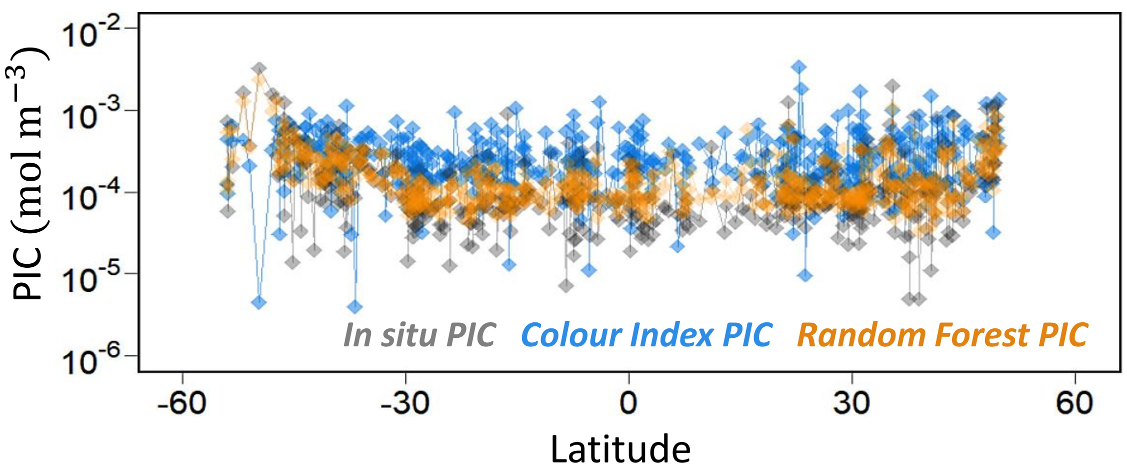


Figure 6. Scatter plot of in situ PIC, colour -ndex-derived PIC, and random-forest -derived PIC match-up data versus latitude of mean AMT cruise tracks.

Knowledge gaps & priorities for next steps

(a) We need more high quality in situ PIC measurements to improve spatial and temporal coverage, especially in the optically-complex coastal waters and high-latitude oceans for further algorithm development.

(b) Many previous algorithms have shown that use of near-infrared wavebands enhances the performances of PIC algorithms. Perhaps future versions of OC-CCI would include near-infrared wavebands to enable their use in PIC algorithms.

(c) The accuracy of random forest PIC model can be further improved by parameter tuning (i.e. tree depth, number of trees) and by adding more input variables (i.e. near-infrared Rrs(λ)) with high importance.

(d) Statistical methods such as random forest method presented here would benefit from a theoretical underpinning to add confidence in their use in the context of climate change.

Data

References

Acknowledgements

Funding