

Estimating total alkalinity for coastal ocean acidification monitoring at regional to continental scales in Australian coastal waters

Kimberlee Baldry^{1,2,*}, Nick Hardman-Mountford², Jim Greenwood²

5 ¹ University of Western Australia, Crawley, WA 6009, Australia

² CSIRO Oceans & Atmosphere, Floreat, WA 6913, Australia

* Current Affiliation: Red Sea Research Center, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

Correspondence to: Kimberlee Baldry (baldry.kimberlee@gmail.com, kimberlee.baldry@kaust.edu.sa)

10

15

20

25

30

35

Abstract.

Owing to a lack of resources, tools, and knowledge, the natural variability and distribution of Total Alkalinity (TA) has been poorly characterised in coastal waters globally, yet variability is known to be high in coastal regions due to the complex interactions of oceanographic, biotic, and terrestrially-influenced processes. This is a particularly challenging task for the vast Australian coastline, however, it is also this vastness that demands attention in the face of ocean acidification (OA). Australian coastal waters have high biodiversity and endemism, and are home to large areas of coral reef, including the Great Barrier Reef, the largest coral reef system in the world. Ocean acidification threatens calcifying marine organisms by hindering calcification rates, weakening the structural integrity of coral reefs and other ecosystems. Tracking the progression of OA in different coastal regions requires accurate knowledge of the variability in TA. Thus, estimation methods that can capture this variability at synoptic scales are needed. Multiple linear regression is a promising approach in this regard. Here, we compare a range of both simple and multiple linear regression models to the estimation of coastal TA from a range of variables, including salinity, temperature, chlorophyll-a concentration and nitrate concentration. We find that regionally parameterised models capture local variability better than continental or open ocean parameterised models, however NRS clusters that model TA to a comparable robustness of regional models were identified. The strongest contribution to model improvement came through incorporating temperature as an input variable as well as salinity. Further improvements were achieved through the incorporation of either nitrate or chlorophyll-a, with the combination of temperature, salinity, and chlorophyll-a constituting the minimum model in most cases. although the increases in robustness due to these improvements is regionally variable. These results provide an approach that can be applied to satellite Earth observation and autonomous in situ platforms to improve synoptic scale estimation of TA in coastal waters.

25

30

35

40

1 Introduction

Ocean acidification (OA), the reduction in oceanic pH caused by the oceans' uptake of anthropogenic carbon dioxide (CO₂) emissions, is a global phenomenon predicted to impact entire marine ecosystems, from microbial primary producers to top predators (Mostofa et al., 2016). Calcifying marine organisms generally show evidence of stress under ocean acidification scenarios, although the effects can vary widely, are species specific, and may depend on physiological traits (Comeau et al. 2017; Edmunds et al. 2016; Azevedo et al., 2015; Jokeil, 2016). Calcifying organisms potentially affected include corals, calcifying algae, molluscs, foraminifera, echinoderms, crustaceans, and bryozoans. In a high CO₂ world, elevated ocean CO₂ concentrations may stimulate photosynthesis (Gattuso et al., 2014) but may also decrease nitrogen fixation, leading to an overall decline in primary productivity (Hong et al. 2017), with uncertain but most-likely detrimental impacts on trophic interactions (Nagelkerken and Connell, 2015). Furthermore, the collapse of habitat builders such as corals, coralline algae and molluscs would have destructive impacts on entire marine ecosystems. Mesocosm and laboratory experiments have shown the majority of calcifying corals tested experience large declines in calcification and growth under OA scenarios (Gattuso et al., 2014).

Australia's coastline is over 36,000 km long, spanning ~33 degrees of latitude, from the tropics to the Southern Ocean. This coastline comprises unique and diverse marine ecosystems with high levels of endemism, of which the most famous is the Great Barrier Reef, the largest coral reef system on the planet and described as one of the seven natural wonders of the world (Mongin et al., 2016). The World Heritage-listed Ningaloo Reef system and remote reef systems of the Kimberley and Pilbara coasts in Western Australia are other examples of Australia's vulnerable coral habitats. Elsewhere, sponges, bryozoans, molluscs and crustaceans contribute to a significant presence of vulnerable calcifying fauna, including some commercially significant species of abalone and scallop. Tracking and predicting the rate of progression of OA in these systems, to inform local management actions, requires the development of robust, and cost-effective methods of monitoring the marine carbonate system at synoptic scales.

Understanding and quantifying distributions of total alkalinity (TA), the proton deficit of seawater relative to neutrality, is an indication of how much carbon dioxide seawater can hold. Waters with higher TA are less prone to rapid change in ocean pH, as they have a higher proton deficit to "consume" protons generated from CO₂ uptake, potentially offering refuge for marine biodiversity in the face of OA. Thus, TA is fundamental to understanding the rate of OA and oceanic uptake of CO₂. Salinity is a conservative tracer within a water mass, meaning that it only experiences changes due to mixing of different water masses or through the addition or removal of freshwater. This leads to a linear relationship between salinity and TA in a region where convective mixing occurs between two waters with differing TA signals (Cai et al., 2010; Jiang et al., 2014; Lee et al., 2006; Millero et al., 1998). This relationship has been exploited to predict alkalinity at the global scale from historical databases of ocean salinity (Lee et al., 2006; Millero et al., 1998). While this works well for open ocean regions, alkalinity in coastal regions can be more variable due to dynamic freshwater end-points in the TA-salinity relationship, the mixing of multiple oceanic end members, and the contribution of various processes that are non-conservative with salinity (e.g. dissolved organic inputs, organic matter respiration, and biological processes such as calcification and organic matter production) (Fig. 1). Thus, with the additional consideration

of non-conservative variability and convective mixing processes, it is clear that these relationships may not be robust in coastal regions (Bostock et al., 2013; Cai et al., 2010).

5 To address this limitation, additional proxy variables have been incorporated into predictive alkalinity models to account for processes that affect TA conservatively, as well as non-conservatively (Jiang et al., 2014). For example, seawater temperature has been used to help account for mixing between water masses (Jiang et al., 2014; Lee et al., 2006) and associated nutrient changes. Temperature also always has a seasonal component, so can be used to capture seasonal variations. To account for primary production, chlorophyll-a (CHL) (Hales et al., 2012), dissolved oxygen or nitrate (N) have also featured in such models (Bostock et al., 2013; Brewer and Goldman, 1976; Lee et al., 2008). However, there is still uncertainty about which proxies are the most robust to use, with the choice often being influenced by availability and location. In the literature, CHL is rarely used in linear regression (LR) models for TA, but rather N is included as a third explanatory variable after salinity (S) and temperature (T). This is not only because N directly affects TA, whereas CHL has indirect effects, but also that several other processes that influence TA variability can co-vary with N, making it a useful proxy (Hieronymus and Walin, 2013). Nonetheless, an important factor in considering such proxies is their amenity to broad-scale measurement. While N can be measured *in situ* using UV absorption sensors deployable on autonomous platforms (Johnson and Coletti, 2002), it cannot be measured directly from satellite instruments as it lacks an electromagnetic signature (Sarangi, 2011). There are, however, well developed remote sensing algorithms for retrieval of oceanic surface CHL concentration from satellites, suggesting a possible advantage for the use of CHL as an explanatory variable for TA, indicating changes due to primary productivity.

There has been little investigation of the distribution of TA in Australian coastal waters due to sparse availability of measurements. Global (Lee et al., 2006; Millero et al., 1998) and regional (Lenton et al., 2015) algorithms have been applied to the oceans surrounding Australia but little progress has been made in investigating variability in TA and its drivers in Australia's coastal waters. In this paper we analyse a seven year time series of observations from nine National Reference Stations (NRS) around Australia in order to quantify the suitability of a range of conservative and non-conservative regression models to predict TA in Australian coastal waters at regional (within the locality of the NRS) and synoptic (algorithms that combine at least 2 NRS) scales. From this we draw conclusions on the regional dependence of modelling TA in coastal waters and construct a visual guide for the modelling of TA at the nine NRS at synoptic scales. An important consideration of this work is the ability to increase spatial coverage using remote sensing techniques, thus we focus on proxy variables that can be measured from satellite Earth observation or autonomous *in situ* technologies such as gliders.

2 Data and Methods

2.1 National Reference Station (NRS) Data

5 As part of Australia's Integrated Marine Observing System (IMOS), time series data for nutrient and carbon variables are routinely measured at nine National Reference Stations (NRS) located around the Australian coastline (Lynch et al. 2014, Fig. 2). Measurements are made, processed, and quality controlled using consistent methods across all stations, as described by Morello, Galibert et al. (2014), and are available from the Australian Ocean Data View (AODV) portal (<https://imos.aodn.org.au>). For this study, time series of temperature (T),
10 salinity (S), total alkalinity (TA), chlorophyll-a (CHL) concentration and nitrate (N) concentration were used. These measurements were made at monthly-to-quarterly frequency as some NRS were established later than others and were sampled at different frequencies. One outlier, measured at the Maria Island NRS on the 14/02/2013 at 10m, was removed (S=23.15). The durations for each NRS are presented in Table S1.

15 Sampling and measurement protocols were undertaken according to the National Reference Stations Biogeochemical Operations Handbook (2016) and the Pre-Run Check and Field Sampling CTD Procedural Guide (2014), available for download from the IMOS website (http://imos.org.au/moorings_documentation.html). Triplicate samples of TA, S, N, and CHL were collected from Niskin bottles at 10m-20m depth intervals. TA samples were poisoned with mercury chloride solution
20 upon collection. All samples were then taken back to the laboratory for analysis. TA was determined by an automated open cell potentiometric titration using 0.1 M HCl as the titrant. The data for S used in this work was collected from bottle salinity data measured by a Guildline Autosal 8400B salinometer using conductivity ratios. CHL data was collected from filtered phytoplankton biomass, analysed using HPLC, and was a summed measurement of chlorophyll-a and divinyl chlorophyll-a. The nearest CHL measurement to a maximum
25 constraint of 10m was taken, as limited measurements were taken within the water column at each sampling point. Finally, N was measured using a Lachat 8000 flow injection analyser with detection limit of 0.1 μM . The resulting data were quality controlled (QC) and flagged according to the National Reference Stations Biogeochemical Operations handbook (2016). Only results flagged 1 or 2 were used.

30 Profiling SeaBird 19+ SEACAT Conductivity, Temperature, and Depth (CTD) instruments were used for continuous measurements of T, which were then binned into 1m depth intervals. Measurements were quality controlled according to the Australian National Moorings Network (ANMN) Standardised Profiling CTD Data Processing Procedures Appendix 4, using the SBE Data Processing-Win32 software and the IMOS MATLAB toolbox, before they were uploaded to the portal (procedures and toolbox available at
35 <https://github.com/aodn/imos-toolbox>). In this study, only TA measurements from the upper 60 m of the water column were used, and only CTD measurements sampled within one hour and 1 m vertically of bottle sampling were used in our analyses.

2.2 Linear regression (LR) analysis and model development

LR is well recognised as a useful predictive tool for spatial extrapolation, particularly in comparison to neural networks which are proven to have less predictive power in extrapolation (Lefèvre et al., 2008). Given the goal of enabling predictions of TA in areas of sparse *in situ* measurements, we restricted the range of input variables to those available with broad coverage from satellite Earth observation, namely T, S, and CHL. Additionally, a fourth BM that included nitrate (N) rather than CHL was included for comparison, which can be measured *in situ* using autonomous sensors. This variable choice accommodates the conservative three end-member mixing model presented in Fig. 1, in addition to testing for variability due to primary production and other non-conservative coastal processes.

General and regional models for the prediction of TA were constructed from LR analysis using the four base models (BM) shown below and the `lm()` function in R. General models refer to those derived from a combined dataset collected from all nine NRS. Regional models refer to those derived from data collected from singular NRS. In total, 40 models were derived (the 4 base models applied to 1 general coastal model and 9 regional models).

$$BM1: TA = aS + d$$

$$BM2: TA = aS + bT + d$$

$$BM3: TA = aS + bT + c\log[CHL] + d$$

$$BM4: TA = aS + bT + c\log[N] + d$$

where T is water temperature, S is salinity, CHL is chlorophyll-a concentration, N is nitrate concentration, and a-d are constants calculated via LR. A log transformation was applied to CHL to account for its well-described, log-normal distribution in the ocean (Campbell et al., 1998) and to satisfy the normality assumption of LR analysis. The same transformation is applied to N as it was strongly right skewed.

Some of the regional NRS data sets had small numbers of observations (n) for some variables, which is not ideal for LR (Table S2-S5), particularly the Ningaloo, Darwin, and Esperance NRS. For BM4, the number of observations used in LR analysis was significantly reduced at Yongala and Kangaroo Island NRS, with only four NRS possessing a robust number of observations ($n \geq 30 * [\text{number of explanatory variables}]$). The results of these models are still presented although they should be considered to be less robust than those for stations with higher n. For the combined data set, Shapiro-Wilk normality tests rejected the hypothesis that S, T, $\log[N]$, and $\log[CHL]$ were each individually normally distributed. It is rare for such data to resemble a normal distribution closely and it was concluded that the symmetrical distributions of S, T $\log[CHL]$, and $\log[N]$ were acceptable to proceed with LR analysis.

All residuals showed evidence of being normally distributed, appearing trendless, and homoscedastic, as should be seen for LR. Some quantile-quantile (Q-Q) plots (not presented) showed evidence of outliers, however the decision was made not to remove these apparent outliers due to the small size of some data sets.

2.3 Open ocean model (Lee *et al.* 2006)

In order to compare the performance of the models tested with an open ocean ‘base’ model, TA was reconstructed from an implementation of the model of Lee *et al.* (2006) using observed S and T measurements collected at the nine NRS. The open ocean model is a quadratic model and has one dynamic geographical boundary through Australian coastal waters, which varies seasonally with T. Like BM2, the number of observations able to be modelled by the open ocean model was restricted by temperature. The numbers of observations used for the open ocean model are the same for those used in the regional modelling of BM2 (Table S3).

2.4 Statistical analysis

Four statistical measures and one test were utilised in order to compare models, assess their robustness and determine the minimum model, which is the model that minimises information loss from the observations.

1. **Residual standard error (RSE)** was calculated as a measure of the error in a model, when compared to observations. By multiplying by the appropriate standard z-value, 1.96 from the standard normal distribution, we obtain an approximation of the 95% confidence error (CE) associated with the model. These estimates are not reliable for models with $n < 30$, which will have a larger CE in accordance with the central limit theorem.

2. **Mean absolute error (MAE)** was calculated as the mean of absolute residuals from each respective model. This is an important indication of how well a model captures the anomalous variations in the data which are more likely to be influenced by non-conservative processes, whilst keeping the measure comparable over data sets with different n . Outliers can obscure these values, so a visual assessment of residuals was conducted to assess if extreme residuals were characteristic anomalies (ie. occurred in groups or due to scatter), or single outliers. Only one outlier was found at the Esperance NRS, which was excluded in the calculation of MAE for models developed from Esperance NRS data.

3. Bootstrapped **Kolmogorov–Smirnov (KS) tests** were employed in order to test the hypothesis that reconstructed alkalinity values are drawn from the same distribution as observations. These were tested at a 5% significance level. As both data sets in the KS tests came from the same environment (same sample of water) the test had to be bootstrapped (Kleijnen, 1999). P-values are shown in Supplementary Table 5.

4. The **Akaike Information Criterion (AIC)** measures the relative quality of statistical models and is particularly useful when models with different numbers of variables are being compared. In calculating AIC there is a trade-off between the goodness-of-fit and the complexity of the model, adding an extra level of analysis compared to RSE. The minimum model, the model that minimises information loss, can then be determined as the model with the lowest AIC value. Using AIC values, the **relative probability of minimising information loss (RPMIL)** for each model can also be determined which normalises differences in AIC according to the number of observations. This allows a more intuitive and robust method for comparing models, by determining probabilities that another model is actually the minimum model given infinite data points.

2.5 Analysis of regional dependence for estimating TA

To assess the regional dependence of the distribution of TA in the Australian coastal zone two analyses were performed.

1. **2-D Multi-dimensional scaling (MDS)** was employed to investigate the regional differences in ocean variables. A 2-D MDS plot was produced using the `cmdscale()` function in R, on a the subset of the entire NRS data set that contained no missing values for TA, S, T and CHL. N was not included in this analysis as it was shown to have a smaller variability (Table 1).

2. A network was constructed from the results of K-S tests performed (as described in Section 2.4) between models from each sites. For each base model, a matrix was constructed as below, with the number 1 indicating that observations at one NRS were significantly similar to reconstructed TA based on a regression trained from the data of a different NRS.

$$M_{BMk} = \begin{bmatrix} x_{11} & \cdots & \cdots & x_{19} \\ \vdots & \ddots & x_{ij} & \vdots \\ x_{91} & \cdots & x_{9i} & \cdots & x_{99} \end{bmatrix}$$

where $x_{ij} = 1$ when observed values at NRS i are statistically similar to reconstructed values from base model k , trained using data from NRS j . Finally, uni-directional links (ie. where $x_{ij} \neq x_{ji}$) were ignored as to only obtain results in which both NRS are modelled to the same standard by one base model, as a cross-validation technique for cluster identification.

3. Results

3.1 Open ocean model (Lee *et al.* 2006)

Figure 3 shows the differences between modelled and *in situ* TA observations using the open ocean model at the nine NRS. All nine NRS showed RSE less than 14 $\mu\text{mol kg}^{-1}$. The model performed particularly well at the Kangaroo Island NRS, predicting TA with an average difference of - 0.70 $\mu\text{mol kg}^{-1}$ (i.e. lower than *in situ* observations) with a residual standard error (RSE) of 5.40 $\mu\text{mol kg}^{-1}$. However, the model underperformed significantly at the Darwin and Yongala NRS, while also overestimating at the remaining six NRS. At the Darwin NRS, on average the model predicted TA to be 20.28 $\mu\text{mol kg}^{-1}$ lower than observed values while at the Yongala NRS on average the model predicted TA of 14.65 $\mu\text{mol kg}^{-1}$ above observed values.

3.2 Kolgorov-Smirnov (KS) tests

Table 2 shows results of KS tests between respective models and observed TA with a 95% confidence level. Results indicate that the open ocean model produces statistically similar results to observed values at the Kangaroo Island and Ningaloo Island NRS. The statistical distribution of TA was only successfully modelled for all NRS using regionally developed algorithms that include N or CHL, T and S, and not by general models for all Australian waters. Nonetheless, regional models that only use S were also able to significantly reproduce the statistical distribution of TA, with the exception of the North Stradbroke Island, Maria Island, and Yongala NRS. At a regional level, observations from the Maria Island and North Stradbroke Island NRS were successfully modelled with BM2, BM3, and BM4, but the Yongala NRS was only successfully modelled with BM3 and BM4. All NRS that were successfully modelled regionally by BM1, were also successfully modelled regionally by all other base models.

3.3 95% Confidence Errors (CE)

95% CE are shown in Fig. 4. The combined general model showed a marked decrease in error over BM1-BM2, and comparable errors over BM2-BM4. Regionally, most NRS exhibited similar errors over the four base models, with the exceptions being Darwin and Ningaloo. The Darwin NRS showed particularly high errors over the 4 base models. Lowest errors were given by BM3 (Darwin, Esperance, Ningaloo, North Stradbroke Island, Port Hacking Bay, Yongala) or BM4 (Kangaroo Island, Maria Island, general coastal). Overall, errors were highest for the Darwin and Yongala regional models, and the general coastal models, with 95% CE $>10 \mu\text{mol kg}^{-1}$. All other models had 95% CE $< 10 \mu\text{mol kg}^{-1}$ for BM2-BM4.

3.4 AIC

AIC values are displayed in Fig. 5, and associated relative probabilities of minimising information loss (RPMIL) are presented in Table S6. AIC values are clearly higher for BM1 in all cases. AIC values indicate that BM3 is clearly the minimum model at 5 NRS, with the exception of the Esperance, Kangaroo Island, Ningaloo and Yongala NRS for which BM4 is the minimum model. Little difference in AIC is seen between BM2 and the minimum model at the Ningaloo, Esperance and Rottneest Island NRS, although when translated to RPMIL it is clear that probabilistically, the minimum model is almost certainly ($>90\%$) that indicated by AIC values. AIC values indicate that the minimum model for the Yongala NRS and the Maria Island NRS was not the best in

terms of RSE and MAE. Upon further analysis, at the Maria Island NRS BM3 removed two anomalous TA residuals as there was not CHL available to them. Thus BM3, the minimum model, is not the best model at this site and BM4 is a better choice.

5 **3.5 Regional dependence and clustering**

Mean and standard deviations for the studied ocean variables are presented in Table 1. A network map representing the results for cluster identification, using p-values from K-S tests as an indication of connectivity via different BM is presented in Fig. 6. A dominant cluster (C1) between three NRS was identified, containing the Rottneest Island, Esperance and Port Hacking Bay NRS, which are linked by BM1, BM2 and BM3. A second cluster (C2) linked by BM2 and BM3 also was identified, which includes all members from C1, and the North Stradbroke Island NRS. This aligns well with other K-S results, which show that at the North Stradbroke Island NRS TA cannot be modelled by S alone. A final cluster (C3) was identified, connected by BM4, which includes the Esperance, Maria Island and Ningaloo NRS. The 2-D MDS plot (Fig. 7) shows a clear regional gradient, when considering selected ocean variables, and it is evident that the members of C1 and C2 are geometrically close and have similar seawater characteristics. However, the members of C3 are relatively geometrically distant as indicated by median the geometric positions of each NRS as displayed on the 2-D MDS plot. Additionally, in C1 and C2 clusters members have differing minimum models, meaning that error could be introduced through the use of BM3 at NRS, which have a regional recommended model of BM4 (Esperance NRS). Regression parameters for each cluster are presented alongside the results from other models in the supplementary material.

20

3.6 A special case: The Yongala NRS

The Yongala NRS displayed a dominant inter-annual trend within the residuals of all regional BM. Upon further investigation it was revealed that this trend was a reflection of an inter-annual trend in TA, which co-varied closely with S (Fig. 8). It was also found that anomalous values were obscuring statistical tests, displaying high statistical leverage on models, which alters fit. This means that on average, BM3 and BM4 actually increased the error of a large portion of residuals further indicated by increases in MAE (Table S2-5).

30

35

4 Discussion

This paper presents a comparison of different linear regression (LR) models for the prediction of alkalinity in Australian coastal waters. In other regions, a simple linear TA-S dependence has often been assumed when estimating TA for use in calculations of other carbonate parameters (Bates et al., 2006; Hales et al., 2012; Lee et al., 2008; Majkut et al., 2014). In Australian waters, LR has also been utilised at a continental scale for the prediction of TA in carbon studies (Lenton et al., 2015; Takahashi, T., et al. 2014; Lenton et al., 2012), but not yet at a regional scale within coastal waters. Despite the wide application of such regression approaches in estimating TA, little investigation has been undertaken on the sensitivity of TA estimates to different input variables in this region. This is surprising given the wider range of processes that can influence TA in coastal waters beyond a simple water-mass mixing model, such as variable inputs of nutrients and dissolved organic material, and their influence on primary production. Here we provide recommendations for the modelling of TA for carbon studies at 9 NRS (Table 3), and have shown that the inclusion of not only salinity (S) but also temperature (T), and either chlorophyll (CHL) or nitrate (N) concentration in these models can significantly improve their performance. In addition to exploring regional relationships, we explore the possibility of estimating TA robustly at synoptic scales, which include the locality of multiple NRS. We find that the Esperance, Rottnest Island, North Stradbroke Island and Port Hacking Bay NRS can be combined for robust monitoring at a synoptic scale. However, regular cross-validation should be performed to identify anomalous events, or deviations from trends, as these NRS are on the opposite side of the continent and are influenced by different long-term climate patterns. This knowledge can be employed to construct cost-effective strategies for the monitoring of TA and considered for application to the monitoring other carbon variables in the Australian coastal zone.

A major finding relates to the use of globally-parameterised open ocean algorithms for modelling TA. It has been shown that such algorithms often fail in coastal waters due to the strong influence that coastal processes have on the distribution of TA (Bostock et al., 2013; Cai et al., 2010). Our results confirmed that such open ocean models are not necessarily optimal for predicting TA in coastal waters and their use can result in large systematic errors in some regions (Fig. 3). Nonetheless, the use of the open ocean model at the Kangaroo Island NRS appears to be consistent with regional parameterisations, and is further supported by KS tests. KS tests also suggest that the open ocean algorithm performs reasonably at Ningaloo, but still, a systematic error can be seen in Fig. 3. This result is due largely to the low number of observations obtained at the Ningaloo NRS and the large amount of scatter in observations, which reduces the sensitivity of the result.

It was found from AIC values that using S alone as a predictor for TA does not give the most informative results, and that the addition of T to the model substantially increases the information of the model at regional scales as well as at synoptic scales (Fig. 5). This is found at all NRS locations, and between the general and regional models. This conclusion is not as strongly reflected in model errors (Fig. 4) due to the substantial decreases in observation numbers (n) seen between the four models but is reflected in MAE values, and such AIC values are important to consider within model comparisons. Thus we recommend that as a minimum, T be included in regression models for the estimation of TA in Australian coastal waters. Further, AIC values indicated that the addition of a third variable increased the information of the model. As such, BM3 was the

minimum model for estimating TA in Australian coastal at 5 of the NRS; elsewhere BM4 was the minimum model. Based on RPMIL values there are no situations in which two models were comparably as likely to be the minimum model (Table S7), however the minimum model for the Yongala and the Maria Island NRS are not the best models when RSE and MAE are further considered.

5

The regional dependence of models for modelling TA in Australian coastal waters is evident throughout the results. The 2-D MDS plot shows a tendency for observations from the same NRS to cluster together (Fig. 7). This indicates that the distribution of selected ocean variables of a NRS is dependent on its location, and cannot be explained by variability in any of the ocean variables studied, and require regionally derived algorithms. KS tests showed that regionally-modelled TA values, constructed from BM4 and BM3, were statistically similar to NRS observations, at all locations, however modelled observations at a continental scale were not statistically similar to observations at all individual NRS locations for any of the BM. This shows that modelling TA at a continental scale is not the most robust method, and further illustrates how the addition of the third variable increases the confidence in successfully modelling TA regionally, if the dependence of TA on predictor variables is unknown.

15

When taking a closer look at the 2-D MDS plot, there is evidence of connectivity between individual NRS, which could lead to clustering. To assess the significance of this connectivity, a network was constructed from the 9 NRS as outlined in Section 2.6 (Fig. 6). Three clusters were identified, which were all successfully modelled from regressions developed from data collected at other NRS, which were members of the same cluster. Combining the results of the MDS plot and this cluster analysis, there is some doubt as to whether C3 is a viable model, as points on the MDS are geometrically distant meaning that sites do not display the same variability. Thus we recommend that C3 not be employed for the synoptic scale estimation of TA. By taking into account all statistical results we have proposed a simplistic guide to help users of the NRS data set, or of data collected in close proximity to NRS sites to understand which method of estimating TA is most robust (Table 3). However, knowledge on robustness alone is not enough to make informed decisions on the employment of environmental surrogates (Lindemayer et al. 2016), and users may wish to consider robustness trade offs for benefits such as cost reduction.

20

25

The results of Table 3 can be attributed to the differences in the oceanographic and benthic processes of the seven characteristic regions. The Yongala NRS and the Darwin NRS experience different distributions of the ocean variables studied compared to other NRS, according to the 2-D MDS plot. Darwin has a distinct geographical setting, located in the Beagle Gulf, between an island and the mainland. This location experiences highly variable freshwater inputs from land-masses, as well as seasonal perturbations from surrounding coral reefs. The Darwin site, exhibits a lower salinity, compared to other NRS, which supports the hypothesis that it is influenced by freshwater. This variability is best captured using BM3, indicating that non-conservative TA anomalies have an annual seasonal dependence, like those seen in phytoplankton variability. The Yongala NRS also experiences large variability, and residuals show that there is a large, inter-annual seasonal component that cannot be explained fully by any on the models (Fig. 8). The cause of this variability seems to be driven by salinity, leading to the conclusion that increased freshwater input is the cause of this variability. When TA is

40

further plotted with salinity in a time series, this is obviously the case, and we see that this variability is reduced by a large factor using a regional BM2. This region is prone to flooding, and river flow data presented in Lough et al. (2015) demonstrates that during the study period (2009-2016), the region was transitioning from a period of high flooding frequency to low flooding frequency, consistent with the trends shown in Fig. 8. Further, the maximum anomalies seen in the residuals corresponds to the worst flooding event, at the start of 2011 where three-quarters of the councils in the Australian state of Queensland were declared disaster zones. Here we only present the impacts of the flooding on employing pre-determined models in Yongala, however, understanding the impacts of this phenomenon on the carbon cycle would be an interesting consideration for future work.

The Ningaloo NRS also experiences large variability, as it is located over a coral reef. The remoteness of this NRS has resulted in relatively low levels of sampling effort (n), particularly CHL and N observations (Table 1). BM2 appears to capture most variability, and the further inclusion of nitrate to the model does appear to reduce error related to anomalies as inferred by RSE and MSE results, but more study of this area is required before an understanding of the mechanisms behind this can be reached. The Maria Island and Kangaroo Island NRS show evidence of similar characterisation, due to their latitude and influence from the Southern Ocean, but are still separated to a degree on the 2-D MDS plot. This is not surprising as Maria Island has a higher influence of oceanic waters, as it is located closer to the shelf edge, and is influenced by variability in subsurface currents where as Kangaroo Island is driven by seasonal currents (Lynch et al. 2014). There is most likely a high influence of upwelling at the Maria Island NRS, producing high levels of primary production, as indicated by higher nitrate and chlorophyll average distributions (Table 1). Although its water mass characteristics are very similar, the North Stradbroke Island NRS is not a member of the dominant cluster C1, and shows different mechanisms of variability. There is a large coastal bay in the vicinity of this NRS, which could be driving the dependence for the minimum model of TA on CHL, as waters from this region would exhibit higher CHL, however this is speculative and a more detailed study of biogeochemistry is required to confirm this hypothesis. Finally, the estimation of TA at the synoptic scale using C1 and C2 is possible as the three locations have a shared influence of boundary currents that flow from North to South, promoting upwelling coastal in these three regions (Lynch et al. 2014; Jones et al. 2015). This means that the source waters are comparable, as latitude is the primary driver of biogeochemical distribution in the open ocean (Lee et al. 2006), and that coastal processes have a comparable effect at the three locations. In this context, it makes sense that Maria Island NRS is excluded from C1 and C2, as here multiple currents converge and flow into the Tasman Sea, making the region unique (Jones et al. 2015).

Similarity deduced from average distributions of biogeochemical and physical ocean variables (Table 1) and latitude are not enough to predict the similarities observed between TA relationships at different NRS in coastal Australian waters. For example, Kangaroo Island is on the same latitudinal gradient, and in close proximity to the Esperance NRS, yet has no connectivity to it through any of the base models. The results of this work show why TA relationships, in addition to average distributions, must be studied to characterise regions with similar TA, as although their average distribution may exhibit similar characteristics, their anomalous TA observations that have largely been influenced by coastal processes may exhibit different driving characteristics. Consequently, interpolation between NRS cannot be recommended until further data is collected to increase

spatial resolution. Shipboard measurements offer an interesting resolution to this problem . Additionally low cost, autonomous carbon flux chambers may soon provide supplementary or alternative sampling to surface DIC measurements (Bastviken et al. 2015), allowing the entire carbonate system to be determined at significantly reduced cost.

5

The results of this study highlight a large limitation to broad-scale predictions of the progression of ocean acidification in vulnerable coastal regions, namely the paucity of high quality TA observations available for development of suitable algorithms. Australia has benefited from the establishment of national reference stations as part of an Integrated Marine Observing System (IMOS) that takes consistent time series observations around the coast. Nonetheless, TA observations have only been collected since 2009, so the temporal range of this data is minimal. Spatially, the data is also limited to only nine locations around a 36,000 km long coastline. For the Ningaloo, Darwin, Kangaroo Island, and Esperance stations, the number of available observations was particularly low, resulting in models for these locations being statistically less robust. The Ningaloo and Esperance NRS were removed in 2015 due to budget constraints, removing the opportunity for extending these relationships in the future (note also that this leaves only one reference station monitoring the western third of Australia's coastal environment). For many parts of the world, even this level of observation is not currently achievable, increasing the challenges of monitoring the progress and impacts of ocean acidification over coming decades.

10

15

20

A promising opportunity lies in the application of regional relationships to satellite Earth observation data, a direction that so far has been little investigated. Recent advances in Earth observation mean that salinity, temperature, and chlorophyll-a are able to be remotely sensed using a range of passive (visible spectrum radiometry) and active (microwave and radar) sensors on orbital satellites (Land et al., 2015). This opens up avenues for exploitation of LR models developed from in situ data to enable synoptic-scale monitoring of TA variability and other carbonate system parameters. While such approaches have been successfully trialled for open oceans (Lee et al., 2006; Millero et al., 1998), less effort has been invested on its application at the coastal scale. The success of this application will depend largely on the resolution of the satellite data that the algorithm is applied to, the accuracy of the algorithm itself and the ability to quantify associated errors, increasing the need for high quality *in situ* measurements. Satellite observations are vulnerable to inaccuracies in coastal waters due to factors including cloud cover, the presence of coloured dissolved organic matter (CDOM) and suspended sediments, the presence of both marine and terrestrial aerosols, land adjacency effects, and the electromagnetic complexity of coastal signals (in both optical and radio wave spectrum) (Schalles, 2006; Land et al., 2015). Future planned sensors with higher spatial and spectral resolution may help reduce these current limitations.

25

30

35

Our analysis has been conducted to additionally explore the possibilities of applying remote sensing platforms for the monitoring of TA. We find that this is a viable pathway, in which further study can be done, however first sampling needs to be performed on an increased spatial scale, so that algorithms can be interpolated accurately. The technology does not currently exist to remotely-sense nitrate from satellites, so BM4 is not useful when considering algorithms that can be applied to Earth observation. Nonetheless, BM4 can be utilised with data from autonomous platforms equipped with nitrate sensors, such as gliders and biogeochemical

40

profiling floats, thus was regarded important to include in this work. These autonomous platforms provide a cost effective solution for the monitoring of TA over long time periods (with intermittent model validation), and an avenue to which the spatial distribution of TA in Australian coastal waters can be resolved.

5 The chemistry of the ocean is dynamic and varies between seasons and years, as well as through direct uptake of anthropogenic CO₂ emissions, and the influence of changing water temperature and salinity from climate forcing. Empirically-parameterised algorithms for TA may therefore require regular retuning to remain robust through time. The presence of ocean acidification will change TA through increasing carbonate dissolution over time (Cross et al., 2013), a process which cannot be estimated from any of the proxy variables explored in this
10 paper. This might change the required algorithm inputs significantly and increase uncertainties in algorithms over time. As such, on-going *in situ* monitoring for alkalinity and other carbonate system parameters will continue to be required to support synoptic scale approaches to monitoring the progression of ocean acidification.

15

20

25

5

5 Conclusion

In addressing the two main applications of the results of this paper, we have defined two different minimum sets of variables for the prediction of TA in coastal waters: S, T, and log[CHL] for applications to satellite Earth observations, and S, T, and log[N] for *in situ* applications. We have shown that a uni-parameter model is not the best model for predicting TA from ocean observations. The use of T as a predictor will improve the model significantly and the addition of a third predictor offers further improvement. We find that the influence of biological responses on the distribution of TA can be significant at some locations in Australian coastal waters, and must be considered when estimating TA. Finally, we offer recommendations for the development of robust algorithms within the locality of the 9 NRS, and present cluster models of NRS that can be used to estimate TA at a synoptic scale. These recommendations have been made by considering the results from a number of statistical parameters to assess model robustness. With this information and the models presented in this paper, more informed decisions can be made about modelling TA in Australian and other coastal waters, assisting efforts to track the progress of ocean acidification.

20

25

30

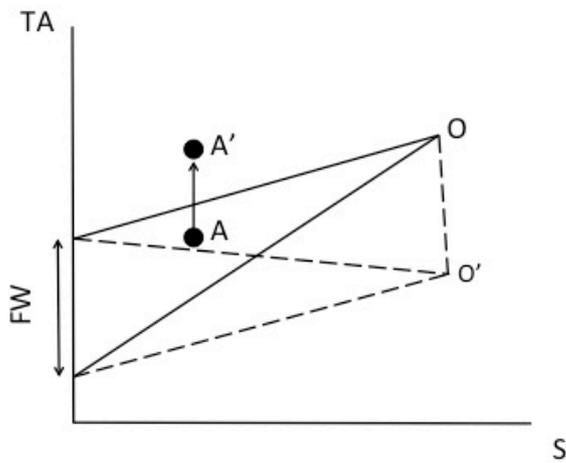
35

References

- Bastviken, D., Sundgren, I., Natchimuthu, S., Reyier, H. and Gålfalk, M.: Cost-efficient approaches to measure carbon dioxide (CO₂) fluxes and concentrations in terrestrial and aquatic environments using mini loggers. *Biogeosciences*, 12, 3849-3859, 2015.
- 5 Bates, N. R., Pequignet, A. C., and Sabine, C. L.: Ocean carbon cycling in the Indian Ocean: 1. Spatiotemporal variability of inorganic carbon and air-sea CO₂ gas exchange, *Glob. Biogeochem. Cycle*, 20, GB3020, 2006.
- Bostock, H. C., Fletcher, S. E. M., and Williams, M. J. M.: Estimating carbonate parameters from hydrographic data for the intermediate and deep waters of the Southern Hemisphere oceans, *Biogeosciences*, 10, 6199-6213, 2013.
- 10 Brewer, P. G., and Goldman, J. C.: Alkalinity changes generated by phytoplankton growth1, *Limnol. Oceanogr.*, 21, 108-117, 1976.
- Cai, W. J., Hu, X. P., Huang, W. J., Jiang, L. Q., Wang, Y. C., Peng, T. H., and Zhang, X.: Alkalinity distribution in the western North Atlantic Ocean margins, *J. Geophys. Res. Oceans*, 115, C08014, 2010.
- 15 Campbell, D., Hurry, V., Clarke, A. K., Gustafsson, P., and Öquist, G.: Chlorophyll fluorescence analysis of cyanobacterial photosynthesis and acclimation, *Microb. Mol. Biol. Rev.*, 62, 667-683, 1998.
- Comeau S., Tambutté E., Carpenter R.C., Edmunds P.J., Evensen N.R., Allemand D., Ferrier-Pagès C., Tambutté S, Venn A.A.: Coral calcifying fluid pH is modulated by seawater carbonate chemistry not solely seawater pH, *InProc. R. Soc. B.*, 284(1847), 20161669, 2017.
- 20 Cross, J. N., Mathis, J. T., Bates, N. R., and Byrne, R. H.: Conservative and non-conservative variations of total alkalinity on the southeastern Bering Sea shelf, *Mar. Chem.*, 154, 100-112, 2013.
- Edmunds, P.J., Comeau, S., Lantz, C., Andersson, A., Briggs, C., Cohen, A., Gattuso, J.P., Grady, J.M., Gross, K., Johnson, M. and Muller, E.B.: Integrating the effects of ocean acidification across functional scales on tropical coral reefs, *BioScience*, 66(5), 350-362, 2016.
- 25 Gattuso, J. P., Hansson, L., and Gazeau, F.: Ocean acidification and its consequences, *Ocean in the Earth System*, 189-253, 2014.
- Hales, B., Strutton, P. G., Saraceno, M., Letelier, R., Takahashi, T., Feely, R., Sabine, C., and Chavez, F.: Satellite-based prediction of pCO₂ in coastal waters of the eastern North Pacific, *Prog. Oceanogr.*, 103, 1-15, 2012.
- 30 Hieronymus, J. and Walin, G.: Unravelling the land source: an investigation of the processes contributing to the oceanic input of DIC and alkalinity, *Tellus Ser. B*, 65, 19683, 2013.
- Hong, H., Shen, R., Zhang, F., Wen, Z., Chang, S., Lin, W., Kranz, S.A., Luo, Y.W., Kao, S.J., Morel, F.M. and Shi, D.: The complex effects of ocean acidification on the prominent N₂-fixing cyanobacterium *Trichodesmium*, *Science*, 356(6337), 527-531, 2017.
- 35 Jiang, Z. P., Tyrrell, T., Hydes, D. J., Dai, M. H., and Hartman, S. E.: Variability of alkalinity and the alkalinity-salinity relationship in the tropical and subtropical surface ocean, *Glob. Biogeochem. Cycle*, 28, 729-742, 2014.
- Johnson, K.S. and Coletti, L.J.: In situ ultraviolet spectrophotometry for high resolution and long-term monitoring of nitrate, bromide and bisulfide in the ocean. *Deep-Sea Res. Pt I*, 49(7), 1291-1305, 2002.
- Jones, E.M., Doblin, M.A., Matear, R. and King, E.: Assessing and evaluating the ocean-colour footprint of a regional observing system, *J. Marine Syst.*, 143, pp.49-61, 2015.
- Kleijnen, J.P.C: Validation of models: statistical techniques and data availability, *Wint. Simul. C. Proc.*, 1, 647-654, 1999.
- 45 Land, P. E., Shutler, J. D., Findlay, H. S., Girard-Arduin, F., Sabia, R., Reul, N., Piolle, J.-F., Chapron, B., Quilfen, Y., and Salisbury, J.: Salinity from space unlocks satellite-based assessment of ocean acidification, *Environ. Sci. Technol.*, 49, 1987-1994, 2015.
- Lee, H.-W., Lee, K., and Lee, B.-Y.: Prediction of surface ocean pCO₂ from observations of salinity, temperature and nitrate: The empirical model perspective, *Ocean Sci.*, 43, 195-208, 2008.
- 50 Lee, K., Tong, L. T., Millero, F. J., Sabine, C. L., Dickson, A. G., Goyet, C., Park, G. H., Wanninkhof, R., Feely, R. A., and Key, R. M.: Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans, *Geophys. Res. Lett.*, 33, 5, 2006.
- Lefèvre, N., Guillot, A., Beaumont, L., and Danguy, T.: Variability of fCO₂ in the Eastern Tropical Atlantic from a moored buoy, *J. Geophys. Res. Oceans*, 113, C01015, 2008.
- 55 Lenton, A., Metzl, N., Takahashi, T., Kuchinke, M., Matear, R. J., Roy, T., Sutherland, S. C., Sweeney, C., and Tilbrook, B.: The observed evolution of oceanic pCO₂ and its drivers over the last two decades, *Global Biogeochem. Cycles*, 26, Gb2021, 2012.
- Lenton, A., Tilbrook, B., Matear, R., Sasse, T., and Nojiri, Y.: Historical reconstruction of ocean acidification in the Australian region, *Biogeosciences*, 13, 1753-1765, 2016.

- Lough, J.M., Lewis, S.E. and Cantin, N.E.: Freshwater impacts in the central Great Barrier Reef: 1648–2011. *Coral Reefs*, 34(3), 739-751, 2015.
- Lynch, T.P., Morello, E.B., Evans, K., Richardson, A.J., Rochester, W., Steinberg, C.R., Roughan, M., Thompson, P., Middleton, J.F., Feng, M. and Sherrington, R.: IMOS National Reference Stations: a continental-wide physical, chemical and biological coastal observing system, *PloS one*, 9(12), e113652, 2014.
- Majkut, J. D., Carter, B. R., Frolicher, T. L., Dufour, C. O., Rodgers, K. B., and Sarmiento, J. L.: An observing system simulation for Southern Ocean carbon dioxide uptake, *Philos. Trans. R. Soc. A-Math. Phys. Eng. Sci.*, 372, 17, 2014.
- Millero, F. J., Lee, K., and Roche, M.: Distribution of alkalinity in the surface waters of the major oceans, *Mar. Chem.*, 60, 111-130, 1998.
- Mongin, M., Baird, M.E., Tilbrook, B., Matear, R.J., Lenton, A., Herzfeld, M., Wild-Allen, K., Skerratt, J., Margvelashvili, N., Robson, B.J. and Duarte, C.M.: The exposure of the Great Barrier Reef to ocean acidification, *Nat. Commun.*, 7, 10732, 2016.
- Morello, E.B., Galibert, G., Smith, D., Ridgway, K.R., Howell, B., Slawinski, D., Timms, G.P., Evans, K. and Lynch, T.P.: Quality control (QC) procedures for Australia’s National Reference Station’s sensor data—comparing semi-autonomous systems to an expert oceanographer, *Methods in Oceanography*, 9, 17-33, 2014.
- Mostofa, K. M. G., Liu, C.-Q., Zhai, W., Minella, M., Vione, D., Gao, K., Minakata, D., Arakaki, T., Yoshioka, T., Hayakawa, K., Konohira, E., Tanoue, E., Akhand, A., Chanda, A., Wang, B., and Sakugawa, H.: Reviews and Syntheses: Ocean acidification and its potential impacts on marine ecosystems, *Biogeosciences*, 13, 1767-1786, 2016.
- Nagelkerken, I., and Connell, S. D.: Global alteration of ocean ecosystem functioning due to increasing human CO₂ emissions, *Proc. Natl. Acad. Sci. U.S.A.*, 112, 13272-13277, 2015.
- Oke, P.R. and Sakov, P.: Assessing the footprint of a regional ocean observing system. *J. Marine Syst.*, 105, 30-51, 2012.
- Sarangi, R.K.: Remote-Sensing-Based Estimation of Surface Nitrate and Its Variability in the Southern Peninsular Indian Waters, *Int. J. Oceanogr.*, 2011, 172731, 2011.
- Schalles, J. F.: Optical remote sensing techniques to estimate phytoplankton chlorophyll a concentrations in coastal, in: *Remote sensing of aquatic coastal ecosystem processes*, 9, Richardson, L.L., and LeDrew, E.F., Springer, 27-79, 2006.
- Shadwick, E. H., Thomas, H., Azetsu-Scott, K., Greenan, B. J., Head, E., and Horne, E.: Seasonal variability of dissolved inorganic carbon and surface water pCO₂ in the Scotian Shelf region of the Northwestern Atlantic, *Mar. Chem.*, 124, 23-37, 2011.
- Takahashi, T., Sutherland, S. C., Chipman, D. W., Goddard, J. G., Ho, C., Newberger, T., Sweeney, C., and Munro, D. R.: Climatological distributions of pH, pCO₂, total CO₂, alkalinity, and CaCO₃ saturation in the global surface ocean, and temporal changes at selected locations, *Mar. Chem.*, 164, 95–125, 2014.

Figures



5 Figure 1: A depiction of a two end-member mixing model that contains an open ocean end-member (O) and a
variable fresh-water end member (FW). Solid lines indicate different conservative mixing lines. Point A lies in the
conservative mixing region, however an arrow indicates how it can be perturbed away from conservative mixing
10 predictions by a non-conservative change in TA to Point A'. Thus, there are three distinct modes of variability; FW
variability, conservative mixing, and non-conservative changes in TA. The mixing of two oceanic end members in
coastal regions it also very likely, further complicating the problem, and extending the region of variability, indicated
by the addition of O' in the model, and associated mixing lines (dashed lines).

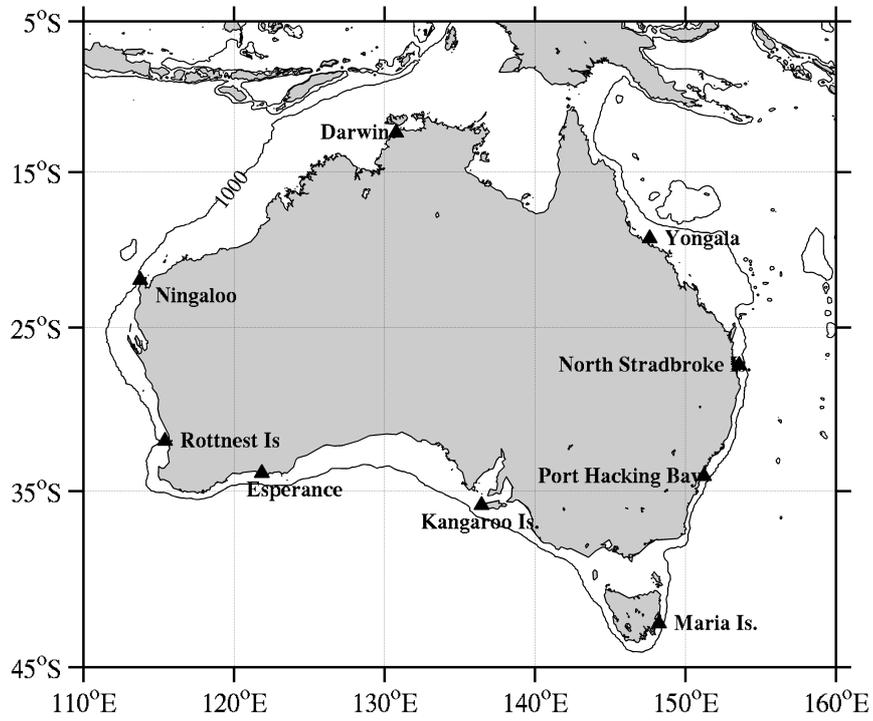
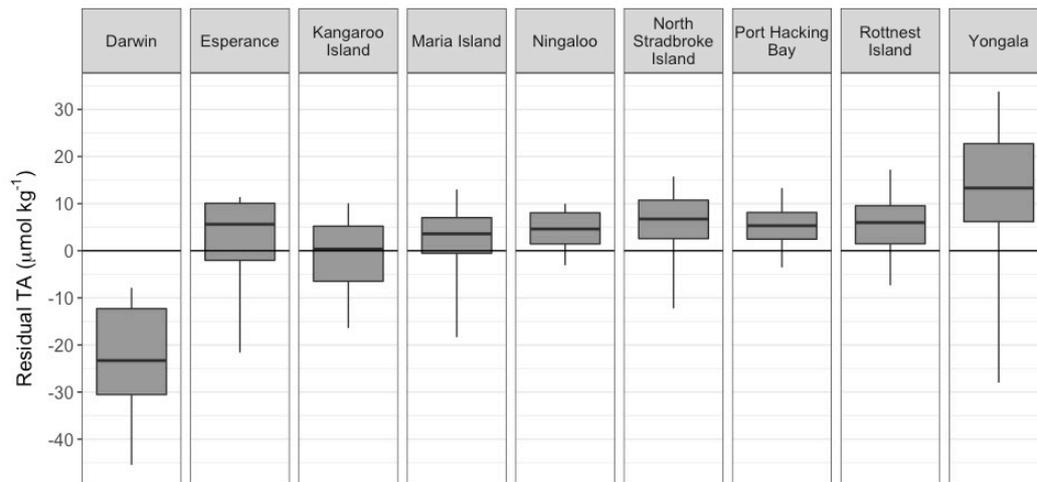


Figure 2: Map displaying the positions of nine national reference stations (NRS)



5

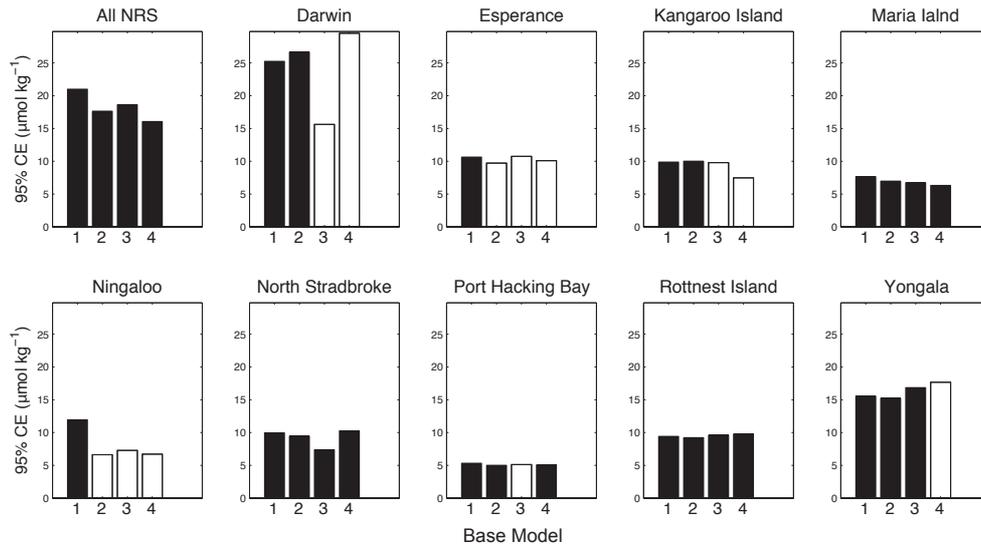
Figure 3: Plots of the Residual TA measurements for the nine NRS for The open ocean model. The mean residual is shown as a black line the height of the box corresponds to one standard deviation, and the extremities of the whiskers show maxima and minima.

10

15

20

25



5

Figure 4: 95% Confidence Errors for each of the four models tested at the coastal level and regional level for the nine NRS. Hollow bars indicate results obtained from algorithms developed from a low number of observations.

10

15

20

25

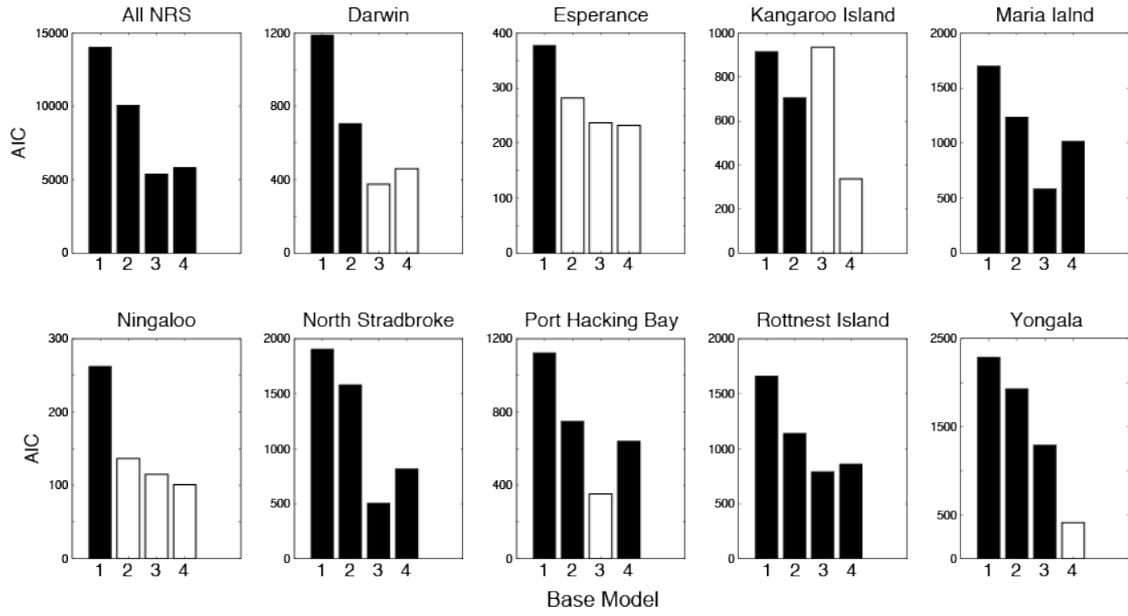


Figure 5: AIC values for each of the four models tested at the coastal level and regional level for the nine NRS. Hollow bars indicate results obtained from algorithms developed from a low number of observations.

5

10

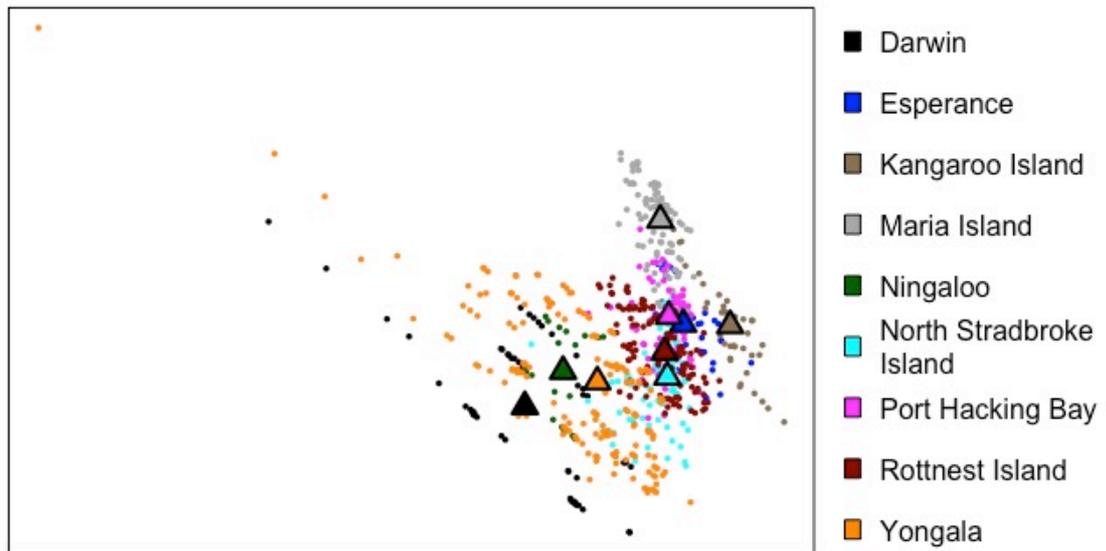


Figure 6: 2-D MDS plot showing observations (circles) that report TA, S, T and CHL measurements at one time stamp. The median geometric position of each NRS is overlaid on the graph (triangles).

5

10

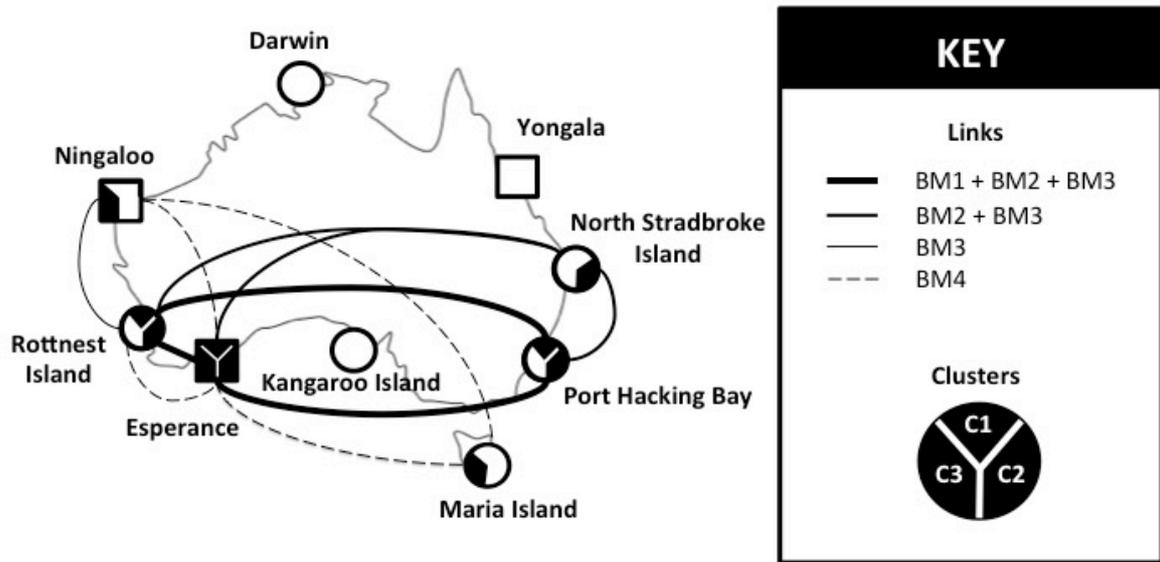


Figure 7: Results from K-S tests, as described in Section 2.5. Links symbolise that the TA distributions at a particular NRS can be modelled by regressions trained from connected NRS, to significantly similar distributions. The shape of the node represents the minimum model of each NRS; circles indicate BM3 is the minimum model, while squared indicate BM4 is the minimum model.

5

10

15

20

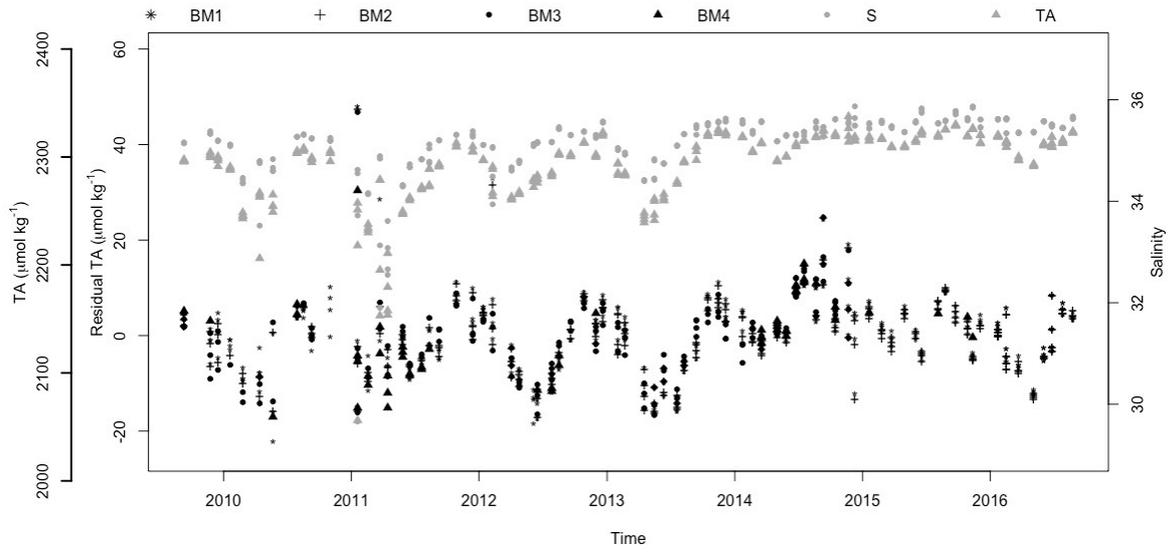


Figure 8: Time series for salinity, TA and residual errors for BM2-BM4 at the Yongala NRS.

5

10

15

20

25

Tables

Table 1: Latitude and mean distributions of parameters at each NRS. Means are presented for each variable with associated standard deviations

NRS	Latitude	TA	S	T	CHL	N
Darwin	-12.4	2265 (40)	34.07(0.68)	28.78(2.53)	0.688(0.687)	0.590(0.348)
Esperance	-33.9333	2337 (11)	35.62(0.14)	18.26(1.69)	0.267(0.058)	0.378(0.155)
Kangaroo Island	-35.8322	2355(11)	35.84(0.19)	16.78(1.57)	1.083(1.561)	0.324(0.197)
Maria Island	-42.5967	2324 (7)	35.31(0.15)	14.49(2.02)	2.369(1.316)	0.666(0.492)
Ningaloo	-21.99	2281(8)	34.80(0.13)	25.77(2.41)	0.514(0.449)	0.380(0.141)
North Stradbroke Island	-27.345	2324(12)	35.48(0.21)	22.56(2.45)	2.917(2.432)	0.263(0.313)
Port Hacking Bay	-34.1192	2326(9)	35.47(0.13)	18.92(2.02)	2.439(2.271)	0.688(0.408)
Rottneest Island	-32	2327(14)	35.49(0.22)	20.73(1.50)	0.514(0.328)	0.306(0.125)
Yongala	-19.3085	2296(32)	35.19(0.60)	25.86(2.37)	0.308(0.206)	0.248(0.145)

5

10

15

20

25

Table 2: Results for the K-S tests for differences in distribution.

✓ = statistically similar to observations

✗ = statistically different to observations

5 * algorithms developed with low numbers of observations employed (n < 30xnumber of explanatory variables)

NRS	Base Model 1		Base Model 2		Base Model 3		Base Model 4		Lee et al. (2006)
	Regional	General	Regional	General	Regional	General	Regional	General	
Darwin	✓	✗	✓	✗	✓*	✗	✓*	✗	✗
Esperance	✓	✗	✓*	✓	✓*	✓	✓*	✓	✗
Kangaroo Island	✓	✗	✓	✗	✓*	✗	✓*	✗	✓
Maria Island	✗	✗	✓	✓	✓	✓	✓	✓	✗
Ningaloo	✓	✗	✓*	✓	✓*	✓	✓*	✓	✓
North Stradbroke Island	✗	✗	✓	✗	✓	✗	✓	✓	✗
Port Hacking Bay	✓	✗	✓	✓	✓*	✓	✓	✓	✗
Rottneest Island	✓	✓	✓	✗	✓	✗	✓	✓	✗
Yongala	✗	✗	✗	✗	✓	✗	✓*	✗	✗

10

15

20

25

30

35

Table 3: Author recommendations for the modelling of TA in the locality of the nine NRS. Recommendations are based on a critical analysis of models using a number of statistical results, as reasoned in the text. Results are presented alongside minimum model predictions (from AIC) and the maximum effect on MAE, when BM2 is instead employed.

NRS	Scale	Recommended BM	Max effect on MAE using BM2	Reasoning
Darwin	Regional	BM3	4.08	Minimum model with agreeing RSE and MAE
Esperance	Regional	BM4	-0.23	Minimum model with agreeing RSE and MAE
Kangaroo Island	Regional	BM4	0.80	Minimum model with agreeing RSE and MAE
Maria Island	Regional	BM4	0.32	RSE and MAE contradict minimum model and indicate that BM4 is the best model
Ningaloo	Regional	BM4	0.06	Minimum model with agreeing RSE and MAE
North Stradbroke Island	Regional	BM3	0.27	Minimum model with agreeing RSE and MAE
Port Hacking Bay	Regional	BM3	0.04	Minimum model with agreeing RSE and MAE
Rottnest Island	Regional	BM3	-0.02	RSE and MAE contradict minimum model and indicate that BM4 is the best model, although the use of BM2-4 is comparable.
Yongala	Regional	BM2	NA	RSE and MAE contradict minimum model and indicate that BM2 is the best model.
C1	Synoptic	BM2	NA	RSE and MAE contradict minimum model and indicate that BM2 is the best model. This makes sense as the three members have different minimum models but are all successfully modeled by BM2 at a regional scale.
C2	Synoptic	BM2	NA	The Esperance NRS displays a different minimum model, so it is not advised to use this cluster at with BM3.