Determination of the acidification state of Canadian Pacific coastal waters using empirical relationships with hydrographic data

by

Alejandra Lara Espinosa
B.Sc., Universidad Autónoma de Baja California, 2007

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the School of Earth and Ocean Sciences

Supervisory Committee

Determination of the acidification state of Canadian Pacific coastal waters using empirical relationships with hydrographic data

by

Alejandra Lara Espinosa
B.Sc., Universidad Autónoma de Baja California, 2007

Supervisory Committee

Dr. Debby Ianson
(School of Earth and Ocean Sciences; Fisheries and Oceans Canada)
Co-Supervisor

Dr. Roberta C. Hamme
(School of Earth and Ocean Sciences)
Co-Supervisor

Dr. Adam H. Monahan
(School of Earth and Ocean Sciences)
Departmental Member

Dr. Richard E. Thomson
(School of Earth and Ocean Sciences; Fisheries and Oceans Canada)
Departmental Member

Abstract

Supervisory Committee
Dr. Debby Ianson
(School of Earth and Ocean Sciences; Fisheries and Oceans Canada)
Co-Supervisor
Dr. Roberta C. Hamme
(School of Earth and Ocean Sciences)
Co-Supervisor
Dr. Adam H. Monahan
(School of Earth and Ocean Sciences)
Departmental Member
Dr. Richard E. Thomson
(School of Earth and Ocean Sciences; Fisheries and Oceans Canada)
Departmental Member

Despite recent interest in understanding long-term trends in ocean acidity, natural variations of carbon chemistry on short timescales are still poorly understood. Unfortunately, historical observations of the oceanic CO_2 system are relatively few in number. Such data are particularly scarce along the highly productive Canadian Pacific coast. However, hydrographic data such as temperature, salinity, oxygen and nutrients have been collected regularly in this region. I developed a fully cross-validated statistical model to predict the aragonite saturation state (Ω_{arag}), a biologically relevant measure of the carbonate system. Different sensitivity tests were performed to assess the robustness of the statistical modelling skill to different model structures. In particular, this study found that *in situ* temperature and O_2 used together were strong predictors of Ω_{arag}. The carbon data used to build this statistical model came from five hydrographic surveys along the Pacific coast of Canada (in July 1998, August 2004, late May 2007, February 2010 and early August 2010) that contain direct measurements of CO_2 system parameters. Only data from a depth range of 0-750 m were used, as data from below 750 m showed biases due to calcium carbonate dissolution. Although processes such as solar warming and gas exchange occur in the surface and could possibly introduce biases in the model, I show that these surface data can be included. The ability of the statistical models to compute robust estimates of Ω_{arag} was assessed by exploring the generalizability of the model through cross-validation procedures using different partitions of the data. By predicting $\ln\Omega_{arag}$ rather than Ω_{arag} directly, I obtained a strong and robust predictive relationship. This MLR model form yielded a high value in the squared correlation coefficient between predicted and observed values (0.96) and a low percentage in erroneous prediction of undersaturated conditions (3.1%). This relationship was found to be insensitive to changes in spatial domain or interannual variability in the data. These results suggest that the model can be used to estimate the distribution of Ω_{arag} along the outer west coast of Canada when basic hydrographic data on temperature and O_2 are available. Predictions of Ω_{arag} from historical observations (1980-2009) in this region reveal that the saturation horizon (Ω_{arag}=1) tended to be more stable in winter and spring and highly variable and occasionally shallow in summer and fall during and following the upwelling season. Undersaturation with respect to aragonite was more likely to occur at shallower depths over the shelf relative to adjacent offshore waters likely as a result of upwelling. The Ω_{arag} saturation horizon tended to be more variable in depth on the shelf compared to offshore waters. The saturation horizon tended to occur at deeper depths over the Queen Charlotte Sound (QCS) shelf and be more stable with respect to the west coast of Vancouver island (WCVI). Thus, the WCVI may experience adverse effects of ocean acidification more acutely than QCS. The use of this approach may provide insight into natural variability and the key controls of Ω_{arag} in future studies at a low cost. However, this predictive model cannot hind-cast data to evaluate the presence of the anthropogenic signal.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I cannot begin this list of acknowledgements without first expressing my most sincere gratitude to my supervisor Dr. Debby Ianson, who has been nothing but supportive since I started working with her. Debby has been a great example of commitment and love for Science and I feel blessed and privileged to have learnt so much from her.

Dr. Adam Monahan and Dr. Roberta Hamme also deserve special thanks for the patience, dedication and care they provided me with. I feel very lucky to have been supervised by them.

Thanks are also due to Dr. Rick Thomson. He has shared his time and experience with me generously and I am truly grateful for that.

Thanks to the amazing friends I have made during this special chapter of my life: Nastasja, Christina, Leandro, Walter, Valeria, Alma, Martha, Maéva, Thomas, Todd, Vicki, Fabian and Norma.

Special thanks to my dear friends Jenni and Manuel. Their friendship was a special motivation during harsh times and a great fulfillment during happy times.

I will always be thankful for the friendship of my old friends (some of them life-long friends): Andrea, Mónica, Tihui, Ikerne, Mauricio and Abraham.

Special thanks are also due to my family: my two siblings Victor and Alma, my Mom, my Dad and my step Mom. Their love and support has overcome all the miles and borders between us. I love them from the deepest part of my heart.

Thanks to my amazing in-laws who have made me feel loved and part of our special family since the moment we met.

Last but not least, I want to thank my wonderful husband Mathieu for his unconditional love and his unique capacity to make me feel like the most loved person.

Dedication

To my dear grandparents: Aurora, Justino, Estela, Lauro and Armando Ceballos (my extra grandfather). For the incredible legacy they have left, for the inspiration they represent among each of the members of our family and for the guidance they give me in every new step of my life.

To my amazing parents, Aurora and Rodolfo, for whom I feel a profound admiration not only for all the unconditional and endless love they have always given to me but for the exceptional persons they are. They mean the world to me.

Chapter 1
Introduction

Oceanic uptake of anthropogenic carbon dioxide (CO_2) is altering the seawater chemistry of the world's ocean [Caldeira and Wickett, 2005; Orr *et al.*, 2005]. The reaction of CO_2 with seawater reduces the availability of carbonate ions. Marine biological calcification requires adequate availability of carbonate ions, which is often represented using the calcium carbonate ($CaCO_3$) saturation state of seawater Ω_{arag} [Fabry *et al.*, 2008]. In the Northeastern Pacific the Ω_{arag} saturation horizon, the horizontal boundary between non-corrosive (saturated with respect to Ω_{arag}) surface ocean and corrosive (i.e. undersaturated with respect to Ω_{arag}) deep ocean, is particularly shallow [Feely *et al.*, 2004] and it has been shown that undersaturated waters with respect to aragonite can be transported onto the continental shelf and into the surface during upwelling [Feely *et al.*, 2008].

There are relatively few historical observations of the oceanic CO_2 system. In chapter 2 of the present study, different approaches to predict Ω_{arag} from observations of commonly sampled hydrographic chemical quantities were assessed. Robustness of all model forms was investigated. A multiple linear regression (MLR) model was developed to determine aragonite saturation state (Ω_{arag}). In chapter 3, the statistical model was applied to historical records available for the Canadian west coast. Seasonal cycles of Ω_{arag} were reconstructed. Variations caused by seasonal changes and regional differences between shelf and offshore waters and between adjacent waters to the west coast of Vancouver Island and to the Queen Charlotte Sound region were described. Each of these chapters is a manuscript in preparation, intended to be submitted to the Journal of Geophysical Research - Oceans and Geophysical Research Letters, respectively.

Chapter 2

A method for determining the aragonite saturation state for Canadian Pacific coastal waters using empirical relationships with hydrographic data

2.1 Introduction

A wide range of direct and indirect measurements confirm that the atmospheric concentration of carbon dioxide (CO_2) has increased globally over the last 250 years [IPCC, 2007] as a result of the industrial and land use activities of humankind. Oceans appear to have absorbed nearly 25% of the total anthropogenic CO_2 emissions [Levine *et al.*, 2011]. A number of studies [e.g. Feely *et al.*, 2012] have predicted that this input of CO_2 has reduced the availability of CO_3^{2-} ions and it has increased the acidity of the seawater. This phenomenon, which is commonly called “ocean acidification”, could affect some of the most fundamental biological processes in the ocean and substantially alter the fundamental structure of pelagic and benthic ecosystems [Kleypas *et al.*, 2006].

The aragonite saturation state, Ω_{arag}, depends on the availability of carbonate ions $[CO_3^{2-}]$ through:

$$\Omega_{arag}=[Ca^{2+}][CO_3^{2-}]/K'_{sp\text{-}arag} \qquad (2.1.1)$$

where $[Ca^{2+}]$ is the concentration of calcium and $K'_{sp\text{-}arag}$ is the stoichiometric solubility product of aragonite. Ω_{arag} is the primary control of seawater calcification rate [Fabry *et al.*, 2008], which is of importance for calcifying organisms. In thermodynamic equilibrium, $\Omega_{arag}>1$ indicates that mineral precipitation (as aragonite) is favoured. For values $\Omega_{arag}<1$, seawater is corrosive to

$CaCO_3$ and, in the absence of biologically-mediated protective mechanisms, dissolution will begin [Fabry *et al.*, 2008]. Marine organisms have been observed to reduce their calcification rates in undersaturated waters [Feely *et al.*, 2004; Hoegh-Guldberg *et al.*, 2007; Doney *et al.*, 2009].

Carbon cycling over continental shelves is complex and highly variable over a range of space and time scales. The coastal ocean exchanges large amounts of matter and energy with the open ocean across continental slopes and constitutes the most biogeochemically active area in the biosphere [Borges *et al.*, 2005]. Until relatively recently, direct [e.g. Feely *et al.*, 2008] and indirect [e.g. Juranek *et al.*, 2009; Kim *et al.*, 2010; Nam *et al.*, 2011; Alin *et al.*, 2012] measurements of carbonate system parameters in these regions have been scarce and consequently, natural variations in ocean carbon chemistry on short time scales (i.e. seasonal, interannual) are still poorly understood. This lack of data hinders attempts to evaluate not only the current status of ocean acidification but future trends and likely impact on marine biota [Wei *et al.*, 2009]. Developing the ability to characterise the natural range in Ω_{arag} on short time scales would be an important contribution in our understanding of the ecological responses of carbonate-secreting organisms (such as coccolithophores, foraminifera, pteropods, mollusks and crustacean), which are already affected by the increased acidity of ocean waters [e.g. Fabry *et al.*, 2008].

On the North American Pacific continental shelf, corrosive conditions with respect to aragonite have already been reported [e.g. Feely *et al.*, 2008]. These conditions resulted from upwelling of intermediate depth water that is rich in CO_2. The northern section of this shelf (e.g. the west coast of Vancouver Island) has the highest primary production of the continental margin of western North America and hosts a lucrative fishery [Ware and Thomson, 2005].

In the absence of extensive observations, modelling approaches represent an important methodological tool to understand the natural variations of ocean carbon chemistry. In fact, knowledge of the current state of ocean acidification is largely the result of empirical model calculations [e.g. Lee *et al.*, 2006; Caldeira and Wickett, 2005] using other quantities such as temperature (T) and salinity (S), for which extensive observations are available.

In the Canadian Pacific continental shelf region, several process studies and semi-regular hydrographic surveys have been carried out over the past decades. As a result, nearly twelve thousand observations that contain hydrographic properties of seawater (i.e. T, S, oxygen (O_2) and nutrients) have been generated and are available through the Institute of Ocean Sciences (IOS) data archive. However, a relatively small number of coastal data containing carbon system parameters exist to assess natural variations in Ω_{arag} (at the surface and through the water column). My purpose in this study is to explore these limited data and their potential to predict the carbonate parameters such as Ω_{arag} using statistical relationships. Specifically, I investigate the ability of statistical models to compute realistic estimates of Ω_{arag} from existing non-carbon data over the continental slope and shelf of the west coast of Canada. It is also my purpose to investigate the robustness of this statistical relationship across different water masses with different chemical signatures present.

My findings will be presented as follows: In section 2.2, the data used to build the statistical relationship are described. In section 2.3, the factors and processes that were taken into account and investigated to empirically estimate Ω_{arag} (prior to and while building the predictive relationship) are described. In section 2.4, each of the steps that led me to the final multiple linear regression (MLR) predictive model form are presented. In section 2.5, a series of statistical robustness tests and model performance assessments (which ultimately justified the use of the final model form) are presented and discussed. In section 2.6, the results of applying the final algorithm to a small portion of the regional historical data archive are presented in order to show the usefulness of the predictive relationship. In section 2.7, the conclusions of my study are presented.

2.2. Data

For this study, I used all available dissolved inorganic carbon (DIC) and alkalinity data from the western Canadian shelf. In addition to alkalinity and DIC, all data sets include T, S, pressure, O_2, phosphate (PO_4^{3-}), nitrate (NO_3^-) and silicic acid (H_4SiO_4) observations. Data from a total of 5 hydrographic surveys of the Canadian West Coast were used: July 1998 [Ianson *et al.*, 2003], August 2004 [Cullen *et al.*, 2009; Nemçek *et al.*, 2008], May 2007 [Feely *et al.*, 2008], July-August 2010 and February of 2010. Most of the samples (about 98%) were collected between May and August, during the summer upwelling season; only data from February 2010 corresponded to the winter downwelling season. Both 2010 cruises were conducted by the Department of Fisheries and Oceans Canada aboard the *CSS J.P. Tully*. The July-August 2010 cruise covered the British Columbian shelf (7 transects) while the February 2010 cruise sampled only two stations (48.6ºN. 126ºW and 48.6ºN. 126.3ºW) off mid-Vancouver Island during the Line P program [Whitney and Freeland, 1999] (Figure 2.1).

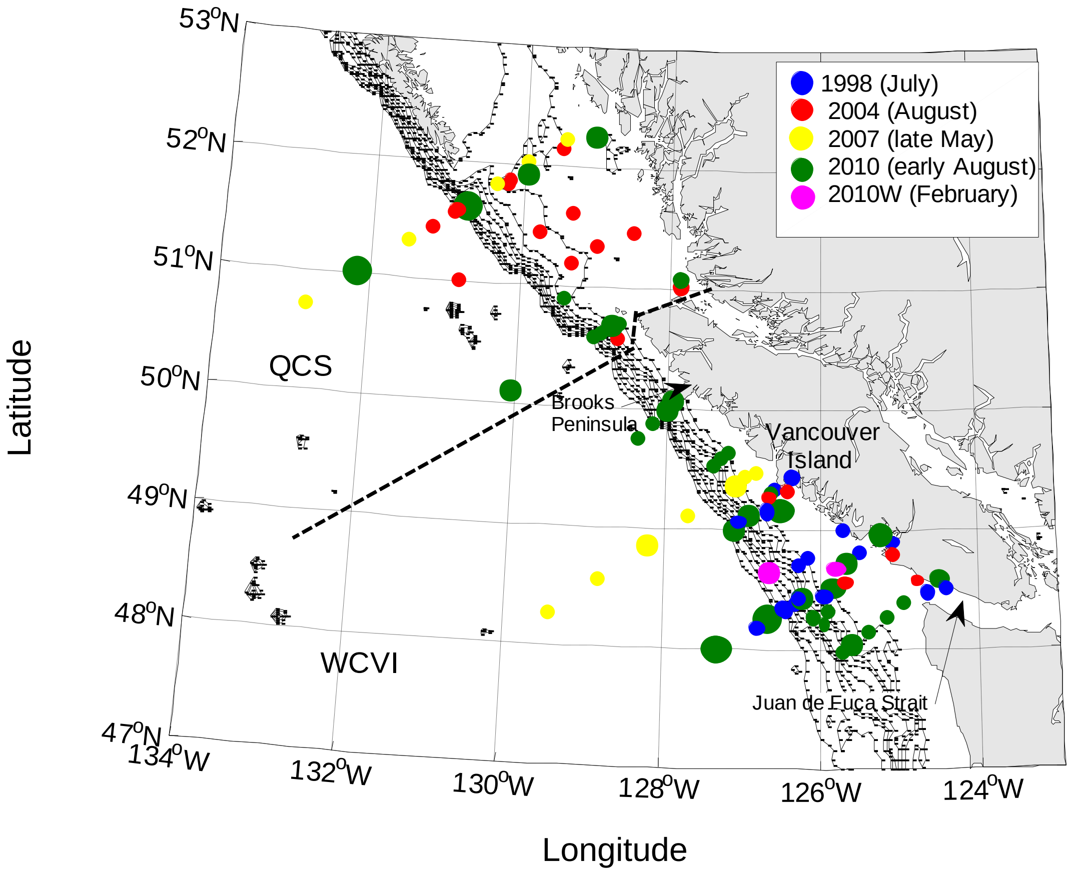

Figure 2.1 Map showing the approximate locations sampled in the five 1998-2010 surveys coloured by individual cruise. The size of the dots indicate sample size: small, 5-10 samples; medium, 11-15 samples and large, 16-20 samples. The bold dashed line shows the separation between QCS and WCVI regions. The 200 m to the 1800 m isobaths are shown at 200 m intervals to mark the slope.

A total of 618 discrete samples were available (Figure 2.1). Some stations were visited on more than one cruise (Figure 2.1). A large fraction of the samples (~40%) were collected in the upper 30 m. On 3 of the surveys, samples were collected from the surface to within 10 m of the bottom; however, 2 surveys sampled only the upper 200 m (1998) and 500 m (2007). Data deeper than 750 m were excluded from the present analysis, as described in detail in the next section.

The majority of the samples were collected on two of the cruises (May 2007 and July-August 2010) and analyzed on board ship. Those from the remaining cruises were refrigerated and stored for later analysis. Collection and analysis of the DIC followed the standard protocol [Dickson and Goyet, 1994]. Collection and analysis of alkalinity followed the closed-cell alkalinity system standard protocol [Dickson and Goyet, 1994] in samples collected in 1998 and the open-cell alkalinity system [Dickson *et al.*, 2007] in samples collected from 2004 to 2010. Salinity values are expressed using the PSS-78 salinity definition [UNESCO, 1981].

2.3. Considerations for algorithm development

The aragonite saturation ratio is a function of the stoichiometric solubility product (K'_{sp}) and the concentrations of Ca^{2+} and CO_3^{2-} in seawater (equation 2.1.1). The solubility product constant is dependent on T, S, pressure and the particular mineral phase (calcite or aragonite). The calcium ion concentration ($[Ca^{2+}]$) varies linearly with salinity [Mucci, 1983] because Ca^{2+} is a conservative element in seawater (i.e. neither S nor Ca^{2+} are greatly affected by processes other than precipitation and evaporation). Because changes in $[Ca^{2+}]$ are proportionally small and variations in K'_{sp} are sufficiently small under the depth range we are considering, variations in Ω_{arag} are largely determined by changes in CO_3^{2-} (discussed further in Section 2.5.5). In this study, I calculated seawater Ω_{arag} values using the Matlab® program CO2SYS [van Heuven *et al.*, 2011] from DIC and alkalinity (along with T, S, pressure, $[SiO_4^{4-}]$ and $[PO_4^{3-}]$). I used dissociation constants proposed by Mehrbach *et al.* [1973] refit by Dickson and Millero [1987]. These constants have been widely preferred because their accuracy has been tested in most ocean basins [Lee *et al.*, 2000].

Dissolved inorganic carbon dynamics and variability are governed by four processes: advection, mixing, biological production/respiration and air-sea CO_2 exchange. Although $CaCO_3$

dissolution/ precipitation and organic matter formation and/or degradation (e.g. oxic degradation of ammonia (NH_3) into NO_3-) play a role in the alkalinity variability [e.g. Sarmiento *et al.*, 2002], the dominant processes controlling its variability in my study region are physical (e.g. advection, mixing, riverine inputs etc.) [Ianson *et al.*, 2003].

T and S are important in the empirical estimation of any CO_2 system parameter since they characterise the water mass properties [e.g. Mackas *et al.*, 1987]. O_2 and nutrient concentrations are strongly influenced by advection, mixing and biological production and therefore are related to variations in carbon parameters. CO_2 production and O_2 consumption are linked through the stoichiometry of the respiration reaction. During respiration, a given amount of O_2 is required to remineralize organic matter (in this case organic carbon) and this ratio has been quantified in several studies (e.g. Redfield *et al.* [1963]; Hedges *et al.* [2002]). Traditionally, the Redfield-Ketchum-Richards [1963] equation has been used to describe the stoichiometry of marine organic matter respiration. However, in more recent studies (e.g. Körtzinger *et al.* [2001]), the Redfield ratios have been recalculated (O_2: Organic carbon=1.34±0.06) based on chemical data analysis on isopycnal surfaces in order to remove the biasing effect of anthropogenic CO_2.

Given these mechanistic relationships and the nature of the available data, I chose to search for an empirical model for Ω_{arag} that involves the inclusion of some or all of T, S, pressure, O_2 and "nutrients". These quantities are expected to covary as the processes that control them are not independent from one another (e.g. T changes affect the degree of O_2 saturation). Furthermore, covariance in these quantities is also expected in cases where two variables coincidently have the same similar trends in their distribution in the water column despite being controlled by dissimilar processes (e.g. low values of salinity and "nutrients" at shallow depths which occur as a result of dilution and nutrient consumption respectively).

Mechanistic relationships between Ω_{arag} and the variables that determine it were explored by conducting sensitivity analyses over the ranges of these variables characteristic of waters off the Canadian west coast. In the first of these, I examined the sensitivity of Ω_{arag} to pressure changes over a depth range of 0 to 1000 m, at three different temperatures (Figure 2.2). As my statistical model will use the natural logarithm of Ω_{arag} (Section 2.4), I also examined the sensitivity of this quantity (Figure 2.3). I repeated the same procedure for each of the variables T (at three alkalinities), alkalinity (at three temperatures) and DIC (at three temperatures). The results showed that the direct dependences of Ω_{arag} on T and pressure are not strong relative to the dependence on DIC or alkalinity for the range of values considered (Figure 2.2a-b, Figure 2.3a-b). In addition, the variations of Ω_{arag} with DIC and alkalinity are non-linear and have a pronounced curvature (Figure 2.2b-c). However, the relationships between $\ln\Omega_{arag}$ and both alkalinity and DIC are much closer to being linear (Figure 2.3c-d). It follows that the relationship between Ω_{arag} and O_2 (through its stoichiometric relationship with DIC) is not linear and therefore it is unreasonable to expect O_2 to be a good linear predictor of Ω_{arag}. In contrast, O_2 should be a comparatively good linear predictor of $\ln\Omega_{arag}$.

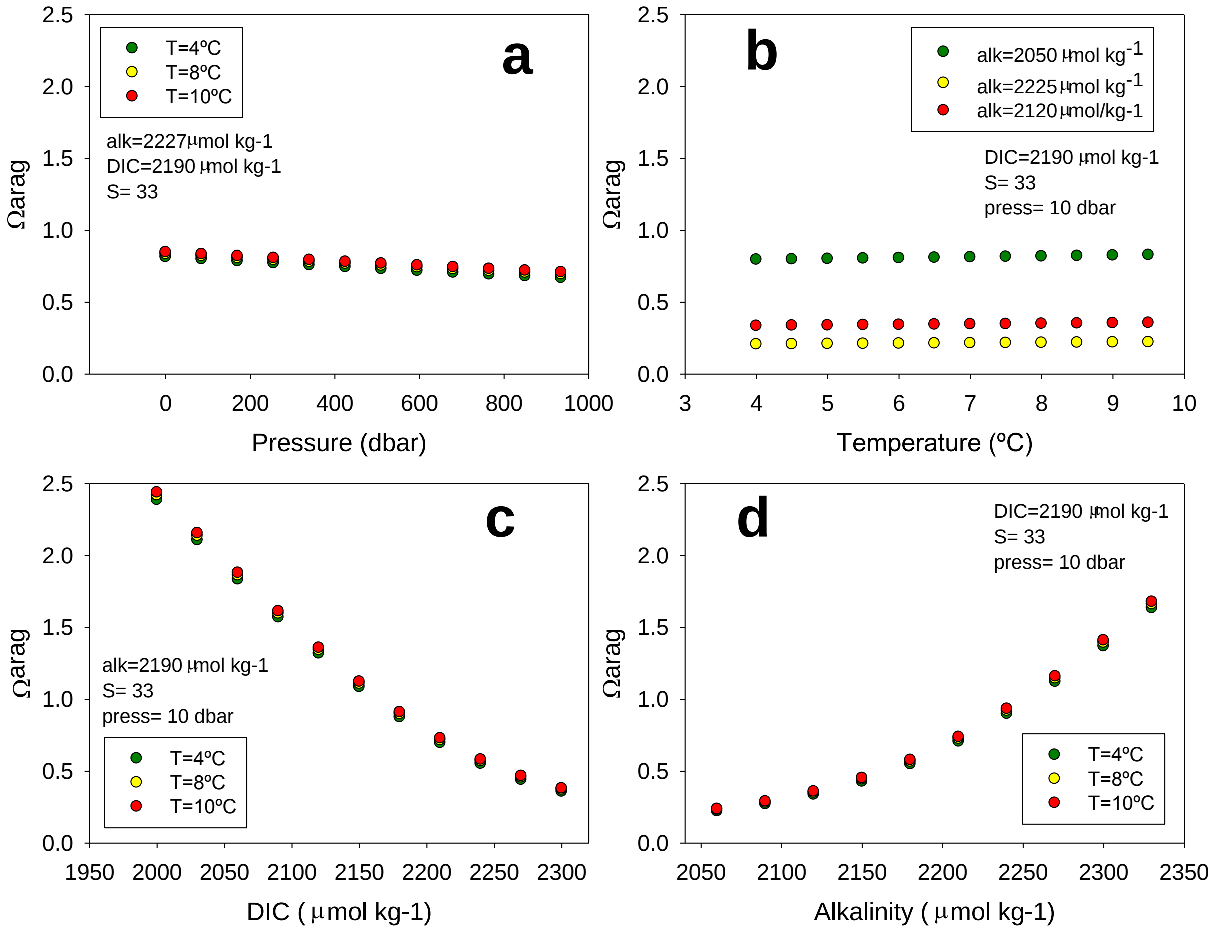

Figure 2.2 Sensitivity of Ω_{arag} to variations in a) pressure, b) temperature, c) DIC and d) alkalinity. These sensitivity tests were derived from the mechanistic model of Ω_{arag}. Each of these variables was independently varied over the approximate range that each experiences from the surface to 1000 m depth in our study region, while the rest were constant to see the effect on Ω_{arag}. The values that were held constant are specified in each panel.

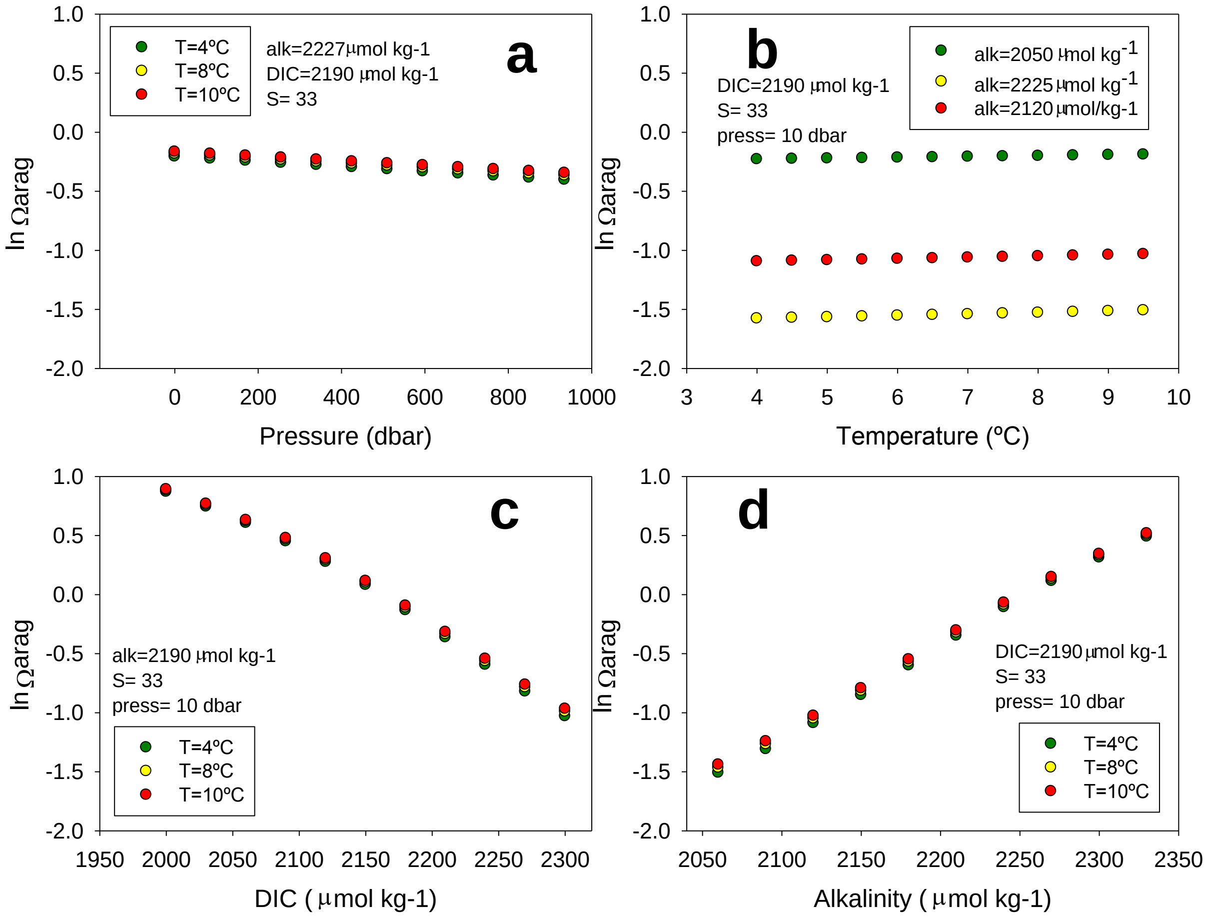

Figure 2.3 Sensitivity of $\ln\Omega_{arag}$ to variations in a) pressure, b) temperature, c) DIC and d) alkalinity. These sensitivity tests were derived from the mechanistic model of Ω_{arag}. Each of these variables was independently varied over the approximate range that each experiences from the surface to 1000 m depth in our study region, while the rest were constant to see the effect on Ω_{arag}. The values that were held constant are specified in each panel.

Two primary factors entered into the selection of the depth range of the data used to construct the predictive model. First, I considered exclusion of data influenced by localized effects that take place in surface waters (e.g. surface warming and gas exchange) and data associated with specific water masses such as the Vancouver Island Coastal Current (VICC) and the California Undercurrent (CUC) (Section 2.5). Second, the effect of the dissolution of $CaCO_3$ was considered. Evidence of dissolution of $CaCO_3$ was observed in data off the continental shelf below ~750 m. A plot of vertical profiles of O_2 of two near-shore stations showed the presence of the O_2 minima (~750 m; Figure 2.4b). Associated with these minima, there was an obvious change in property-property trends below this zone (the hook in the DIC vs O_2 scatter plot of Figure 2.4a). This change is attributed to dissolution of $CaCO_3$ shells at depth, which increases the concentrations of DIC but not O_2. These changes would affect the performance of the linear model; thus, only data from 0-749 m (~96% of the observations) were used to build the regression model.

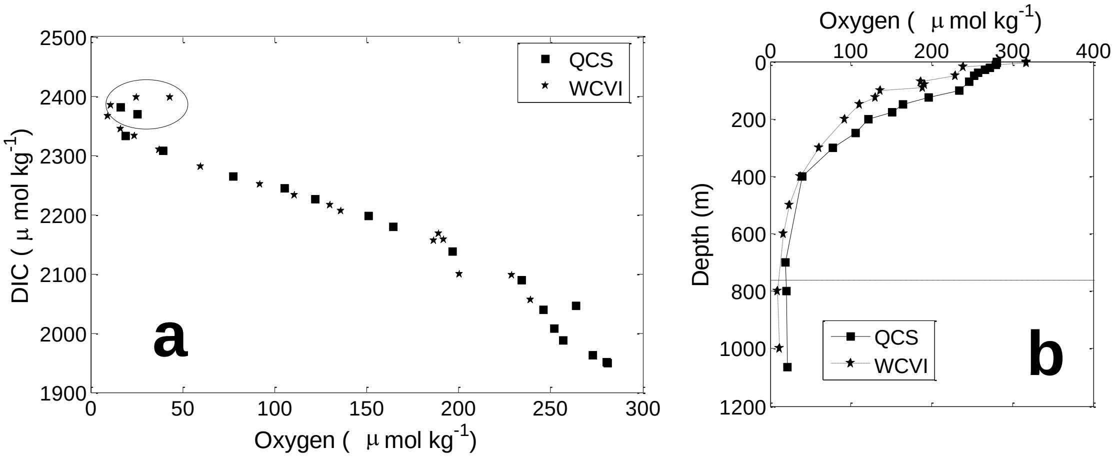

Figure 2.4 a) Typical DIC-O_2 property-property plot and b) oxygen depth profile from the near slope of Queen Charlotte Sound (■) and mid Vancouver Island region (★) (51.62N 130.77W and 48.99N 127.12W respectively, July 2010).The horizontal line shows the 750 m boundary associated with the 'hook' highlighted with an ellipse in panel a.

2. 4. Model development

In this section, the use of Multiple Linear Regression (MLR) (section 2.4.1 below) to empirically estimate Ω_{arag} from measurements of T, S, pressure, O_2, and NO_3^- from the five hydrographic surveys of the Canadian West Coast will be explored (section 2.2). Strength in the correlation of these variables with one another will be tested in order to assess the degree of dependence between potential model predictors (as MLR models are generally most robust when the predictors are independent). It will be demonstrated that the performance of the MLR model is superior when lnΩ_{arag} is the predictand, rather than simply Ω_{arag}. The ability of the various regression models to correctly predict undersaturation (with respect to Ω_{arag}) will be assessed.

2.4.1. Model construction- MLR

Multiple linear regression (MLR) has been previously used to model the regional-scale seasonal evolution of Ω_{arag} [Juranek *et al.*, 2009], six-year [Alin *et al.*, 2012] and decadal variations of Ω_{arag} [Kim *et al.*, 2010]. Additionally, a five-year mooring time series study examined the 2009-2011 La Niña events and the approximated values of pH (a relevant measure of the carbon system) obtained by a local DO-pH relationship [Nam *et al.*, 2011].The MLR expression for Ω_{arag} using the full set of potential linear predictands is represented by:

$$\Omega_{arag}^{pred} = \beta_0 + \beta_1 P + \beta_2 T + \beta_3 S + \beta_4 O_2 + \beta_5 NO_3{}^{-} \qquad (2.4.1)$$

where Ω_{arag}^{pred} is the statistically modeled aragonite saturation state, β_i is the regression coefficient for the i^{th} predictor (≥1) and β_0 is the intercept [Wilks, 1995]. The MLR model parameters are estimated by minimizing the sum of squared differences between the observed and empirically estimated Ω_{arag}. For independent predictors, the regression coefficients determine the independent contributions of each distinct variable to the prediction.

Interpretation of the regression coefficients (β_i) in terms of the respective importance of each predictor is difficult when the predictors have different units and different ranges of variability. To facilitate interpretation of the regression coefficients, the predictors and the predictand (Ω_{arag}) were first standardized by subtracting the respective mean and then dividing the centred residuals by one standard deviation ($x_{standardized} = \frac{x - mean(x)}{std.dev.(x)}$). The means and standard deviations for T and O_2 in each data set are presented in Appendix A. This transformation converted values for the now dimensionless variables to the same scale, allowing direct quantitative comparison of the contributions by each predictor.

2.4.2. Choice of predictors

By definition, the regression coefficients represent the optimal statistical model for the data from which they are estimated. However, spurious correlations between predictors and predictand can inflate estimates of the model predictive skill by producing a statistical model that does not generalize to other independent data. This problem of overfitting generally becomes worse as the number of predictors increases, particularly if the predictors are mutually correlated [Wilks, 2005]. When presented with a suite of correlated potential predictors, more robust models will generally be produced by using only a subset of these.

I computed the correlation coefficients among all the various potential predictors (Table 2.1) in data from above 750 m. In general, these cross-correlations were generally high (all of them above 0.63 and as high as 0.94). Thus, as expected, these variables were not independent from one another, demonstrating the importance of considering a subset of these predictors in the statistical model.

Table 2.1. Table of correlation coefficients of five predictors for the full data used (0-750 m).

	0-750 m depth					
	Pressure	T	S	O_2	NO_3	Ω_{arag}
Pressure	1.00	-0.63	0.71	-0.76	0.73	-0.63
T	-0.63	1.00	-0.83	0.72	-0.84	0.90
S	0.71	-0.83	1.00	-0.88	0.89	-0.85
O_2	-0.76	0.72	-0.88	1.00	-0.94	0.89
NO_3	0.73	-0.84	0.89	-0.94	1.00	-0.93
Ω_{arag}	-0.63	0.90	-0.85	0.89	-0.93	1.00

To assess the relative predictive skill and robustness of each potential predictor individually, I generated one regression for each pair of predictor and survey (25 regression coefficients), and compared the resulting regression coefficients (Figure 2.5). In general, the lowest regression coefficient values were those corresponding to salinity (from 0.015 to 0.35), pressure (from -0.02 to 0.3) and NO_3^- (-0.4 to 0.06). Regression coefficient values for NO_3^- also were characterized by sign inconsistency. In contrast, the fact that T and O_2 regression coefficients were consistently high (0.32 to 0.69 and 0.28 to 1.53, respectively) and consistent in sign identified these as more robust predictors. I therefore decided to keep T and O_2 as the only two predictors and to disregard the rest.

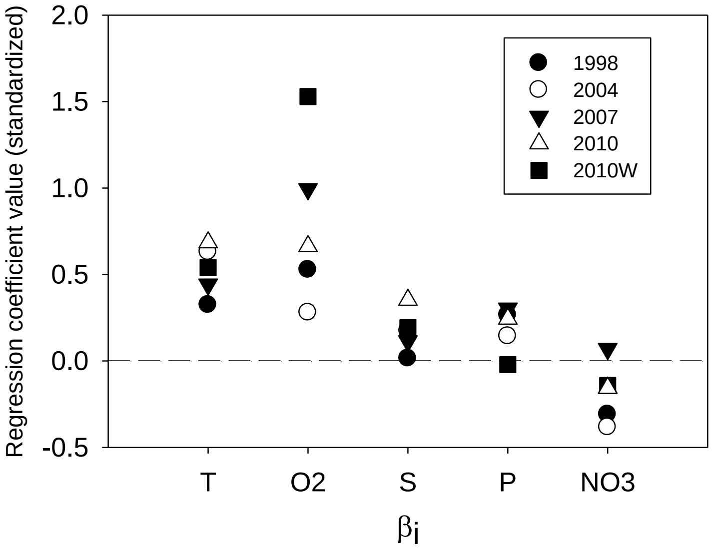

Figure 2.5 Five-parameter MLR (equation 2.4.1) standardized regression coefficients (dimensionless) for each of the five surveys. The dashed line shows the zero line. Coefficients may be greater than one in the case where at least one of the predictors is negative.

2.4.3. Cross-validation

Having selected the predictors to be used, the robustness of the MLR model was assessed. The cross-validation technique [Picard and Cook, 1984] is commonly used to assess the accuracy of statistical models. It minimizes the artificial skill due to overfitting and evaluates how the results of a statistical analysis will generalize to an independent data set. In this study, the complete data set used to build the regression model was divided into five subsets (each corresponding to one of the 5 cruises in which samples were collected (Appendix A)). Each data subset was individually withheld and a corresponding regression model created from the remaining four data sets. This regression model was then used to make a prediction of Ω_{arag} from the withheld data, which was assessed by comparison to the known values. This calculation procedure was repeated for each of the individual subsets, so that at the end I generated predictions of each data subset using statistical models that were not fit to the data being predicted. Each of the 5 out-of-sample predictions of Ω_{arag} was then compared with the corresponding observations of Ω_{arag} (Figure 2.6a).

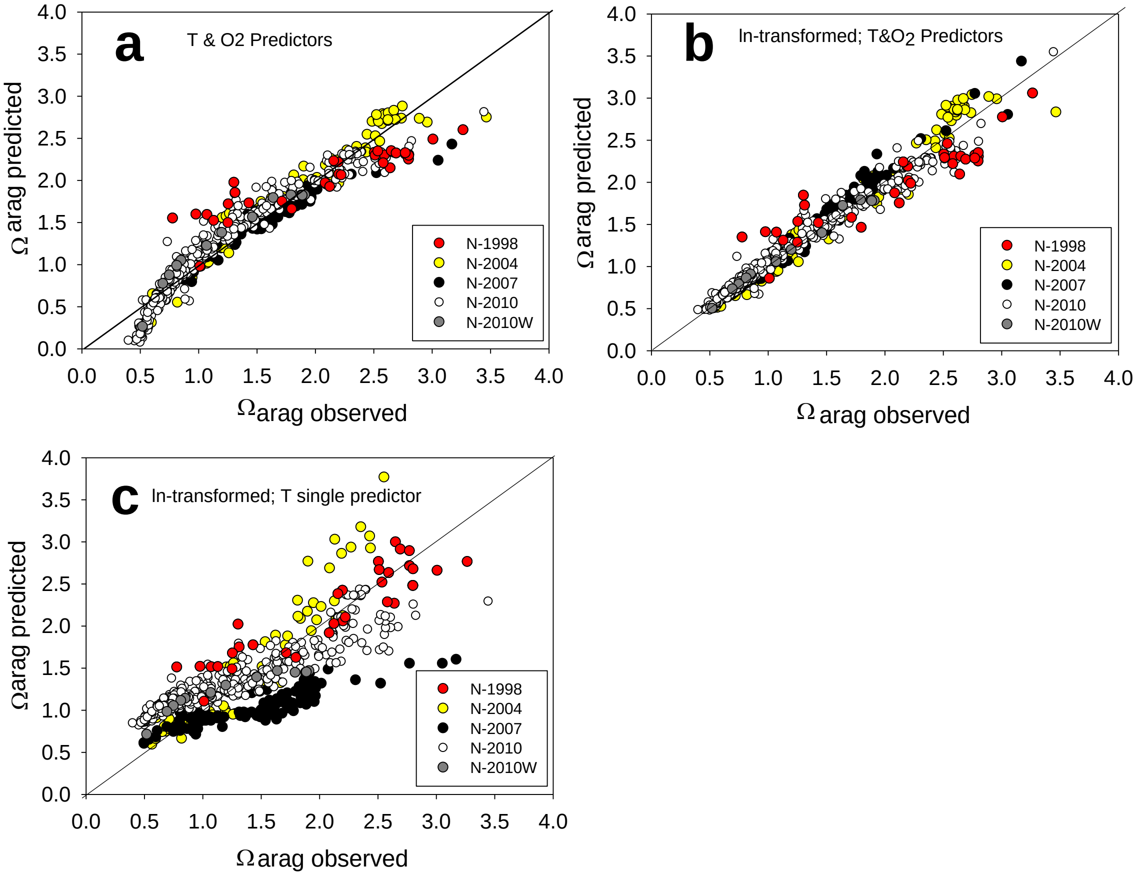

Figure 2.6 Plots of MLR predicted Ω_{arag} vs. observed Ω_{arag} (calculated from DIC and alkalinity observed). Five regression models were built for four of the five surveys minus one withheld (N-'x') survey. The resulting regression model was used to make a prediction of the withheld data set and compared to the true values of each data set (section 4.3). a) Ω_{arag} predicted with two predictors (T and O_2); b) ln-

transformed predictions of Ω_{arag} with two predictors (T and O_2);c) ln-transformed predictions of Ω_{arag} with one predictor (T).

In general, the agreement between empirically estimated ('predicted') and observed Ω_{arag} is reasonably good. However, a striking feature of this initial regression prediction is the bias that exists at low Ω_{arag} values (roughly observed Ω_{arag} values of ~0.75 and less, Figure 2.6a) in each of the 5 predictions. Not only are predictions of Ω_{arag} biased to low values, but further analyses using different partitions of the data (not shown) demonstrated that the linear regression model sometimes predicted negative values of Ω_{arag}, which are meaningless by definition. To remove this bias and avoid predicting negative values of Ω_{arag}, the MLR model was changed to predict $\ln\Omega_{arag}$. The addition of this transformation was also justified by the fact that $\ln\Omega_{arag}$ is more closely a linear function of the variables that determine it than it is Ω_{arag} (Figure 2.3; section 2.3). This model form is expressed as follows:

$$ln\Omega_{arag}^{pred} = \alpha_0 + \alpha_1 T^S + \alpha_2 {O_2}^S \qquad (2.4.2)$$

(or alternately $\Omega_{arag}^{pred} = \exp(\propto_0 + \propto_1 T^S + \propto_2 {O_2}^S)$) where α_i are regression coefficients for the new model and T^S and ${O_2}^S$ are standardized values of T and O_2, respectively (see below; equation 2.4.3). By construction, the predicted values of Ω_{arag} are non-negative. The cross-validated predictions of this model were compared with observations of Ω_{arag} (Figure 2.6b). Clearly, the observed and predicted values of Ω_{arag} fall tightly around the 1:1 line and the bias of Ω_{arag} at low values associated with the model that predicts linear Ω_{arag} (Figure 2.6a) has been eliminated.

In order to quantify the deviations I used the mean-squared error (RMSE) which accounts for the biases that may be present in predictions. RMSE measures root square of the average of the squared "errors" or residuals and because the residuals are squared it is sensitive to large residuals [Wilks, 1995]. As expected, predictions that deviated less from the 1:1 line had smaller

RMSE values than those with larger scatter (Table 2.2). The single predictor relationship (T), for instance, had the highest RMSE values, while the T and O_2 relationship and the ln-transformed T and O_2 relationship had the lowest RMSE values.

The proportion of variability of the true values of Ω_{arag} that is accounted for by the statistical models (the non-ln transformed and the ln-transformed) is measured by the square of the correlation coefficient (R^2) between empirical estimates and observations. The five regressions yielded, in most cases, high R^2 in both models that use T and O_2 as predictors (Table 2.2). These values ranged from the lowest for 'N-1998' (0.84 and 0.85) to the highest for 'N-2010W' (0.93 and 0.99). As shown here, the R^2 of the ln-transformed model is not meaningfully different from that directly predicting Ω_{arag}. However, this transformation was effective in linearizing the relationship between the predicted and observed Ω_{arag} values and eliminated the estimated bias at low values of Ω_{arag} seen in Figure 2.6a. An increase in the scatter (i.e. larger deviations from the 1:1 line) can be observed at the high end of the Ω_{arag} estimate ('prediction') relative to the low Ω_{arag} estimate (Figure 2.6b). However, this scatter in the upper end of the Ω_{arag} range falls out the range of main practical concern, as risk assessment studies' generally focus on regions with Ω_{arag} around and below 1 (low levels of pH) [e.g. in fresh water Magnuson *et al*., 1984; Beamish and Harvey, 2011].

Table 2.2. Five regression models were built with full subsets (one for each survey) minus one withheld subset (column 1). The resulting regression model was used to make a prediction of the withheld data set and compared to the true values of each data set (section 4.3). The values of false negatives, false positives, RMSE and R^2 were computed for model with two predictors (T and O_2) (a), the ln-transformed model with two predictors (T and O_2) (b) and the ln-transformed model with one predictor (T) (c). N is the total number of samples per survey.

	a. T and O_2				b. T and O_2 +ln transformation				c. Only T + ln transformation				N
	False+ (%)	False - (%)	R^2	RMSE	False+ (%)	False - (%)	R^2	RMSE	False+ (%)	False - (%)	R^2	RMSE	
1998	0	6.3	0.84	0.045	0	6.3	0.85	0.34	0	0.61	0.82	0.33	33
2004	0	1.4	0.96	0.16	7.7	0	0.97	0.53	25	0	0.81	1.73	93
2007	4.4	4	0.93	0.15	3.7	0	0.98	0.11	52	0	0.804	0.58	123
2010	0.7	11.4	0.92	0.18	2.4	6.4	0.97	0.11	0	42	0.87	0.32	356
2010-W	0	11.1	0.93	0.14	0	0	0.99	0.069	0	27.3	0.88	0.29	13
TOTAL	1.5	8.2	0.93	0.17	3.9	3.1	0.96	0.13	8	15.4	0.85	0.45	618

I assessed the ability of the ln-transformed model (equation 2.4.2) to accurately empirically estimate both undersaturation and supersaturation using contingency tables that characterise: 1) the percentage of times in which Ω_{arag} was incorrectly estimated to be undersaturated (the false positive rate) and 2) was incorrectly estimated to be supersaturated (the false negative rate). In general, both of the rates were low (between 0% and 6.4% and between 0 and 7.7%, respectively) (Table 2.2).

I chose the ln-transformed model as the final model form (equation 2.4.2) because it provides strong, unbiased estimates. The final model form using all five data sets is:

$$\Omega_{arag}^{pred} = \exp(2.15 \times 10^{-1} + 1.82 \times 10^{-1} * T^S + 3.40 \times 10^{-1} * {O_2}^S) \qquad (2.4.3)$$

where T^S and ${O_2}^S$ (standardized T and O_2) are:

$T^S = \frac{T(°C) - 8.78°C}{2.87°C}$ and ${O_2}^S = \frac{O_2(\mu mol\ kg^{-1}) - 198 \mu mol\ kg^{-1}}{88.6 \mu mol\ kg^{-1}}$ and estimated uncertainty of 0.13 using ±1 standard deviation of the residuals (Figure 2.8b) as an error estimate.

2.5 Sensitivity and stationarity

In this section, I consider the sensitivity of the empirical estimates to the specific data used to build the regression (e.g. inclusion of data from depth ranges associated with distinct water masses). I also explore the potential for use of T as a single predictor of Ω_{arag} and investigate the prediction skill of a single variable that resulted from the implementation of Principal Component Analysis (PCA). I then perform a regional partition of our data and build regressions with these subsets in order to explore the stationarity (i.e. variations in mean and variance) of the statistical model. I next look at whether the prediction could be improved by correcting the biases that result from the fact that the variance in the empirical estimates of Ω_{arag} is lower than that of the observed values of Ω_{arag} by definition. This difference could potentially lead to higher false negative rates. Thus, the variance in the empirical estimates was inflated

(VIF) to match that in the observations. I then assess the use of alternate predictors such as apparent oxygen utilization (AOU). Finally, I compare the results from this study with those obtained by Juranek *et al*. [2009].

2.5.1. Investigation of single-predictor relationships

As T is more easily and commonly measured in seawater than is O_2 and correlations between Ω_{arag} and T were generally strong (Table 2.1), I investigated the predictive skill of Ω_{arag} obtained from a model using this single predictor:

$$ln\Omega_{arag}^{pred} = \gamma_0 + \gamma_1 T^S \qquad (2.5.1)$$

where γ_i are regression coefficients and T^S is standardized T. The parameters of this model were estimated and the prediction skill assessed in exactly the same way as discussed in Section 4.3. The single-predictor (T) relationship (figure 2.6c) did not yield as strong an empirical relationship as that obtained using both T and O_2 (Table 2.2). The curves of empirical estimates of Ω_{arag} (Figure 2.6c) display considerably more scatter around the 1:1 line than those of the two-predictor relationships (Figures 2.6a and 2.6b). Moreover, for empirical estimates of individual cruises, R^2 values were as low as ~0.8 and the false negative rates as high as 42% (i.e. close to 1/2 of the times the relationship failed to accurately estiumate undersaturation of Ω_{arag}). These results clearly indicate the benefit derived from the use of the two predictors (section 2.4.2) relative to the use of the strong single predictor T. The utility of O_2 as a predictor is not surprising as it is directly related to DIC (and ultimately Ω_{arag}) through photosynthesis and respiration and the influence of DIC on Ω_{arag} is strong (Figure 2.3c).

The residuals or errors of a fitted model are the differences between the observed values and the correspondent prediction of the response computed using the regression function ($\Omega_{arag} - \Omega_{arag}^{pred}$). In general terms, if the fit of the data is unbiased, the errors are scattered randomly. Given the reduced quality of the model performance observed in the cross-validation procedure when using only T as predictor (low R^2, high false negative rate, Table 2.2), it was expected that the error associated with this model form would be higher than that of the two-predictor (T and O_2) model form. Large deviations (some residuals are higher than 1.5 for high Ω_{arag} (i.e. $\Omega_{arag} > 2.0$); Figure 2.7b) were observed at high values of Ω_{arag} (Figure 2.7a) while

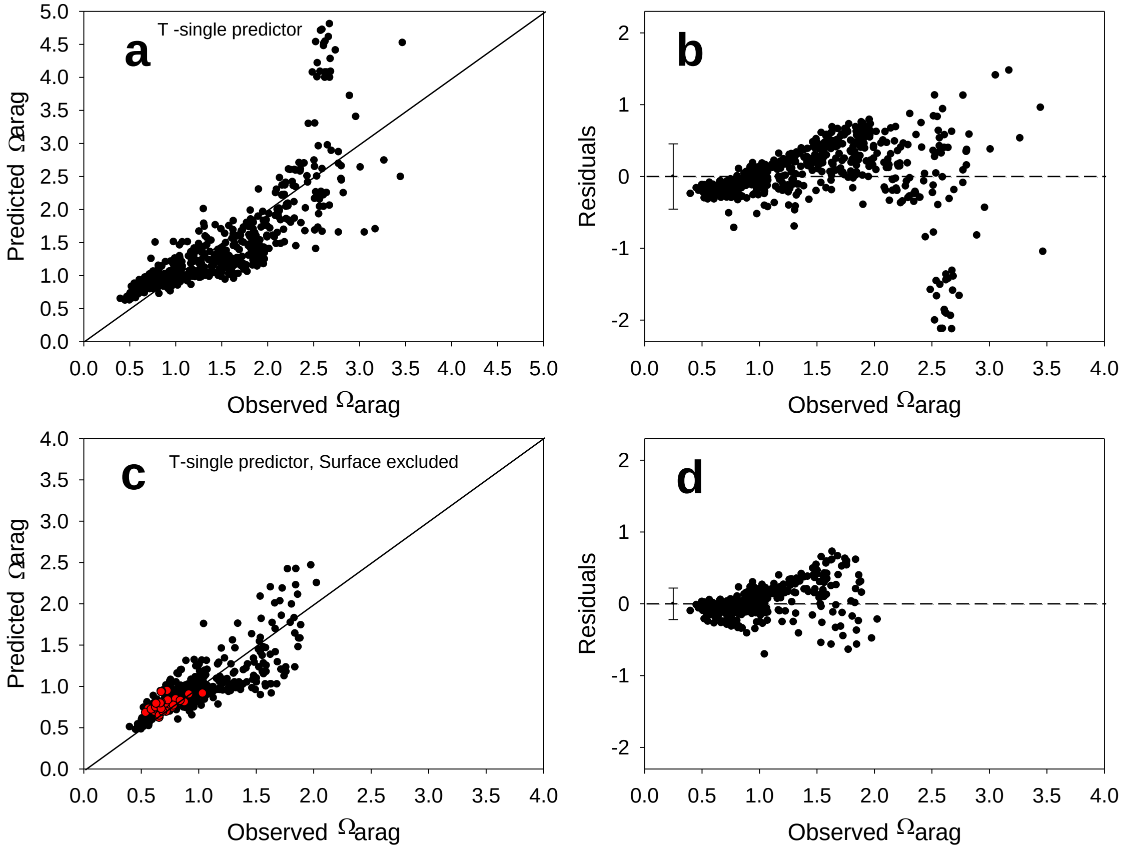

Figure 2.7 a) Measured Ω_{arag} vs. predicted Ω_{arag} with single-predictor (T) relationship (0-750 m) and b) corresponding residuals (Ω_{arag} observed - Ω_{arag} predicted) vs. Ω_{arag} observed and error bars for +/- 1 standard deviation (0.45); c) measured Ω_{arag} vs. predicted Ω_{arag} with single-predictor (T) relationship with surface (<30 m) data excluded (red points represent water from CUC) and d) corresponding residuals (Ω_{arag} observed - Ω_{arag} predicted) vs. Ω_{arag} observed and error bars for +/- 1 standard deviation (0.22).

the largest residual observed in the two-predictor (T and O_2) model form was always <0.7 (Figure 2.8b). The majority of these high residual points correspond to surface data from the

2004 survey. These data had unusually high temperatures (~16.5-17.7ºC) and low salinity (~30.9-31.5). The resulting RMSE value in the single-predictor (T) model was ~4 times larger than that of the T and O_2 relationship (Table 2.2), indicating larger errors in the single-predictor model. To assess the sensitivity of the Ω_{arag} estimates to the elevated temperatures from the 2004 survey, the T-only regression model was recomputed excluding surface data (>30 m). In this model, the largest residuals are avoided, yielding a value for RMSE ~2 times smaller than that of the full subset (Figure 2.7c-d, RMSE=0.22), but the predictive skill remained lower than that of the two-parameter model.

The relationship that used only T as predictor is clearly much weaker than the two-predictor model (T and O_2) even when its regression parameters are recalculated without the surface data. However, it represents an alternative when O_2 is not available:

$$\Omega_{arag}^{pred} = exp(2.14 \times 10^{-1} + 4.28 \times 10^{-1} * T^S) \qquad (2.5.2)$$

where T^S is calculated in the same manner as in equation 2.4.3 and the estimated uncertainty (±1 standard deviation of the residuals) is ± 0.45. This particular equation corresponds to the full data set (surface included).

Although predictive skill is increased through inclusion of both T and O_2 as predictors, it is possible that this skill comes from information held in common between these two correlated quantities. In general, the robustness of predictive models is improved by the exclusion of redundancy among a set of correlated predictors. The chances of building an overfit regression model could potentially be reduced by preprocessing the predictors using Principal Component Analysis (PCA) [Wilks, 1995]. This is a common technique that converts a set of partially correlated variables into a set of values of uncorrelated variables that characterize the dominant features of their variability (i.e. account for most of the variance in the observed values). The leading PCA variable is the linear combination between the predictors' original quantities (in this case, T and O_2) which accounts for the most variance of any such linear combination. This combined variable, which describes the majority of the variance held in common by the two original predictors (this combination is expressed as $pcaX = \frac{T^S - {O_2}^S}{\sqrt{2}}$ when T and O_2 are standardized), can then be used as a single predictor in a regression model.

Predictions computed with the PCA-based predictor carrying common information in the T and O_2 combination were compared to those derived from the final relationship (equation 2.4.3, Appendix B). Results for the PCA-based model (Figure 2.8c, d) showed a reduced quality in the estimates relative to the full model (Figure 2.8a, b).

Residuals of the PCA-based relationship (Figure 2.8d) were slightly larger than those of the final relationship in Ω_{arag} (with an error of 0.18) (Figure 2.8d). In addition, the regression R^2 (Appendix C) from the single predictor, PCA-based model, was essentially the same as that of the final two-predictor model (with a difference of 0.03). False negative and false positive rates of the final relationship and the PCA-based model were essentially the same (Table 2.2; Appendix C). I conclude that, for the prediction of Ω_{arag} along the coast of British Columbia, sufficiently different predictive information is carried by T and O_2 that the presence of both these two predictors is important.

2.5.2. Biases carried by observed-predicted variance differences

By construction, the variance of the MLR-based Ω_{arag} prediction will be smaller than the true value [Wilks, 1995]. This low variance bias could potentially lead to a substantial bias in estimated rates of super- and undersaturation. To offset this bias, the variances of the predicted values of Ω_{arag} were increased to match those of observations by multiplying the deviations from the mean by a constant factor (the Variance Inflation Factor; VIF).

However, the use of this VIF did not significantly affect the quality of the model predictions. In fact, estimates computed with the VIF-model had increasingly higher biases towards the high Ω_{arag} values (Figure 2.8e) than those of the final relationship (equation 2.4.3) and the rates of false negatives did not have a substantial improvement (Appendix C). Moreover, most VIF-model predictions resulted in larger values with respect to the observations of Ω_{arag} (i.e. overestimation of true Ω_{arag}). Associated with this result, there were also large biases (in magnitude) towards negative values in the residuals (Figure 2.8f). As a result, no correction of variance bias was applied to our predictions.

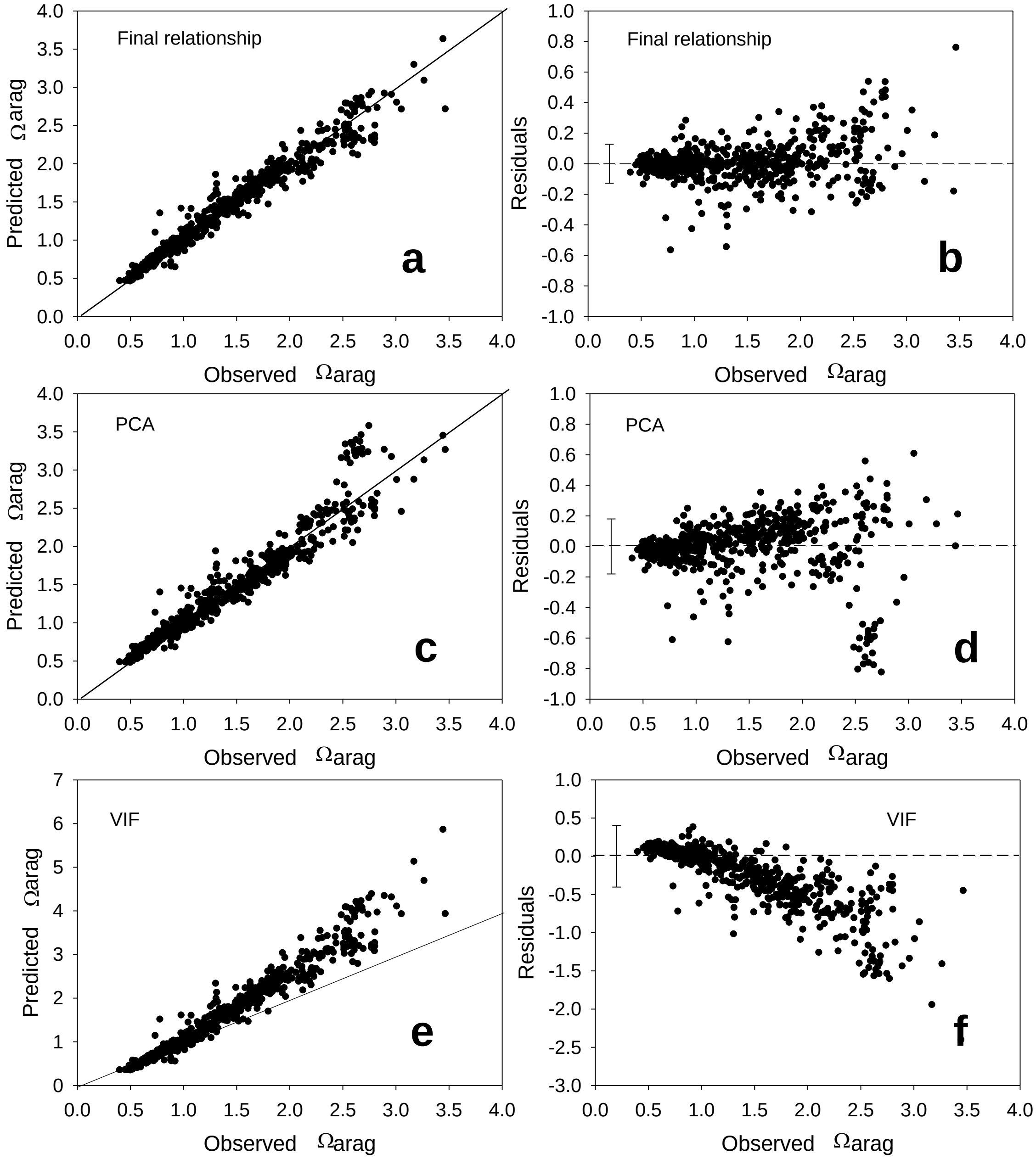

Figure 2.8 Measured vs. predicted Ω_{arag} for the final relationship (a), the PCA-based relationship (c), and the VIF-based relationship (e). The corresponding residual plots are shown in (b), (d), and (f), respectively. Errors (± 1 standard deviation of the residuals) for the final relationship (0.13), the PCA-based relationship (0.18) and the VIF-based relationship (0.40) are also shown.

2.5.3 Sensitivity to region

The cross validated predictions presented in Section 2.4 demonstrate the predictive skill of Ω_{arag} for individual cruise data using models obtained from other cruises, but do not directly address the stationarity (i.e. the possibility of generalizing the model to different regions or different subsets of data) of these statistical relationships. To assess the stationarity of the statistical relationship, we considered one further data partition that was based on the region where the samples were collected.

Prior to partitioning the data, T-S (Figure 2.9a) and T-O_2 (Figure 2.9b) curves were analyzed. Results from these observations indicated that all data sets clustered within similar T-S and T-O_2 space and no inconsistencies were found except for a few warm (or salty) data points (highlighted with an ellipse; Figure 2.9a) that fell outside the range of the rest of the T-S data points. However, these highlighted observations fell within the main cluster of data in T-O_2 space (Figure 2.9b) and they represent less that 1% of the full set and thus, they were not considered for further analysis.

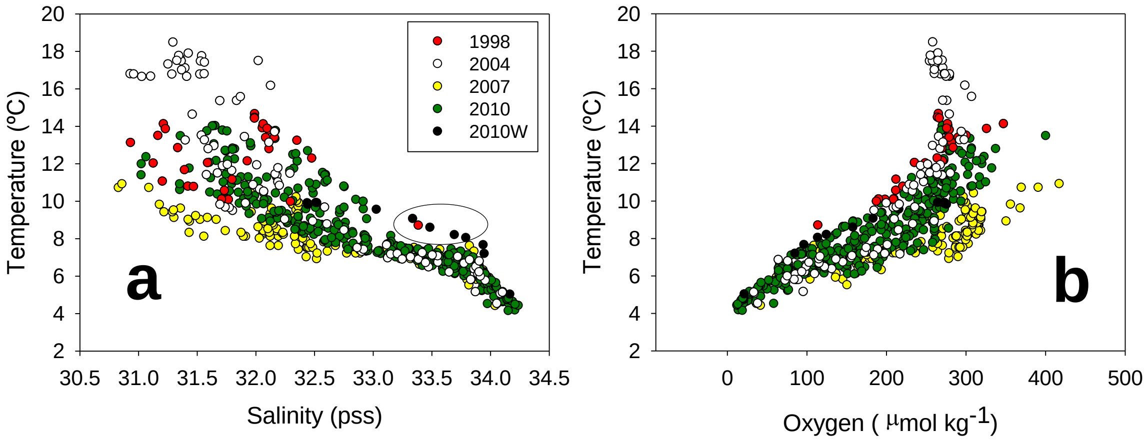

Figure 2.9 Property-property plots for all data from the 5 surveys considered. a) T-S and b) T-O_2. The oval in T-S plot highlights the two populations that fell outside the common range.

I partitioned the data by geographical region. The first subset included only data collected on the west coast of Vancouver Island (WCVI) (~60% of the data) and the second subset included only data from the Queen Charlotte Sound (QCS) region (Figure 2.1). Regression models for $\ln\Omega_{arag}$ were generated from each of the two data sets (WCVI and QCS). For this data partition, a linear regression model (for $\ln\Omega_{arag}$) was generated from each data subset and applied to: the data subset from which the regression was generated ('in sample' estimates; red points, Figure 2.10) and the other data subset ('out of sample' estimates; olive green points, Figure 2.10).

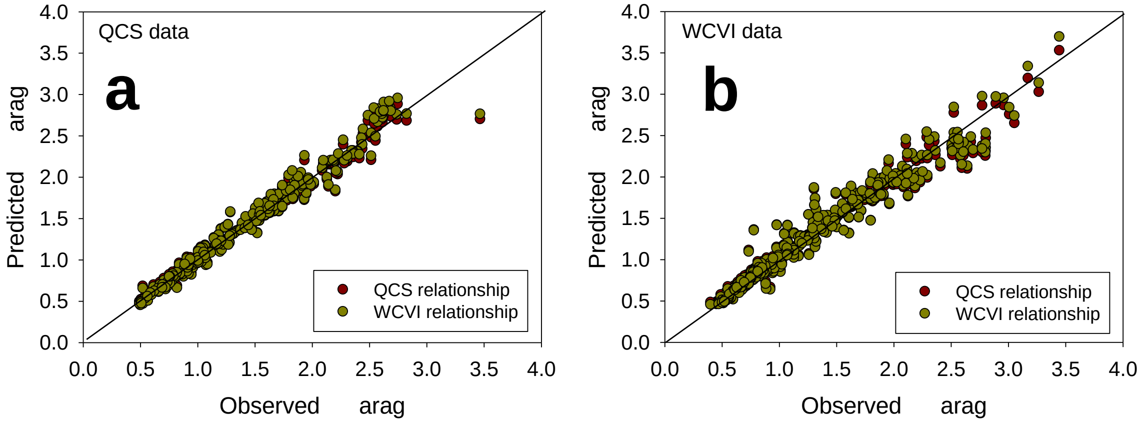

Figure 2.10 Plots of predicted and observed Ω_{arag} (after ln-transformation) for a) the two relationships derived from the QCS region and the WCVI region applied to the QCS data; b) the two relationships derived from the QCS region and the WCVI region applied to the WCVI data.

No substantial differences were observed between the two pairs of signals (Figure 2.10a-b). Thus, the robustness of the statistical model does not appear to be sensitive to changes in spatial location across the coast of British Columbia. This result further justifies the use of all the available data to generate a single model. The 'in sample' relationships, however, did yield slightly smaller RMSE values than the 'out of sample' predictions (Table 2.3). This difference was unsurprising because, by definition, less bias is expected on average when estimated parameters are applied in the same subset from which they were estimated and more bias is expected when estimated parameters are applied in 'out-sample' subsets.

Table 2.3 Calculated RMSE values for the two relationships derived from the QCS region and the WCVI region applied to the QCS data; and the two relationships derived from the QCS region and the WCVI region applied to the WCVI data.

		Relationship applied	
		QCS	WCVI
Subset	QCS	0.10	0.12
used	WCVI	0.15	0.14

2.5.4 Sensitivity to depth ranges

Because the data used in this study were sampled from across a large portion of the water column, the possibility of non-stationarity in the statistical relationships caused by different water masses, which occupy distinct depth ranges, and geographic regions [Mackas *et al*., 1987], was considered. The Vancouver Island coastal current (VICC) is a buoyancy current that flows northward along the inner shelf throughout the year along coast of Vancouver Island [Freeland *et al*., 1984] (northward from the entrance to Juan de Fuca Strait to beyond Brooks Peninsula [Thomson, 1981]). The main source of the VICC in summer is the relatively fresh and nutrient rich outflow from Juan de Fuca Strait, where snowmelt carried by the Fraser River is tidally mixed with deeper ocean waters entering the Strait [Thomson *et al*., 1989; Crawford and Dewey, 1989].This coastal current occupies the whole water column inshore of the ~100 m isobath [Thomson *et al.*, 1989] and is rich in DIC and nutrients [Ianson *et al.*, 2003; Crawford and Dewey, 1989]. Another important water mass is the California undercurrent, a sometimes strong northward subsurface current that flows along the continental shelf-slope [Hill *et al.*, 1998; Thomson, 1981]. The core of this current has been found at depths of 125-325 m off Vancouver Island during summer [Thomson and Krassovski, 2010]. The CUC is characterized by relatively high T and S, low dissolved oxygen concentration and high nutrient (nitrate, phosphate and silicate) content [e.g. Hickey, 1979]. Locations and depth ranges that were considered as VICC and CUC are listed in Appendix D.

Despite its unique T-S, O_2 and nutrient signatures, the CUC water masses did not appear to introduce any bias in our data. Residuals were not higher in the CUC; i.e. at depth ranges of ~125-325 m of the slope and shelf (Figure 2.11) than elsewhere at comparable depths, as it would be expected if the CUC had distinctly different influences on the predictions of Ω_{arag}. As well, the residuals from the T-only regression model for CUC waters are similar to those for the rest of the water mass (Figure 2.7c).

At shallow depths, residuals were substantially larger on average than residuals deeper than ~ 30 m (Figure 2.11). Residuals from the few available data in the surface of the southernmost portion of the VICC, near the mouth of the Juan de Fuca Strait, where the chemical signature is expected to be strongest [Ianson *et al.*, 2003]; Appendix D) showed strong negative bias (grey dots, Figure 2.11) indicating that the relationship overestimated Ω_{arag} by as much as 0.6 in this relatively small region. DIC values in the south VICC are exceptionally high, and do not vary with nutrients, such as nitrate, as expected [e.g. Redfield *et al.*, 1963] and seen elsewhere in the study region [Ianson *et al.*, 2003]. On the other hand, T and O_2 values fall within the range of the rest of the data used to build the model, on the low O_2 side of the scatter.

Beyond 48.8ºN 125.5ºW (Chapter 3; Appendix D) in the central VICC there was no apparent bias and residuals were not as large. The strong chemical signature is expected to have been erased, as the VICC flows northward, by mixing, biological uptake of carbon and gas exchange.

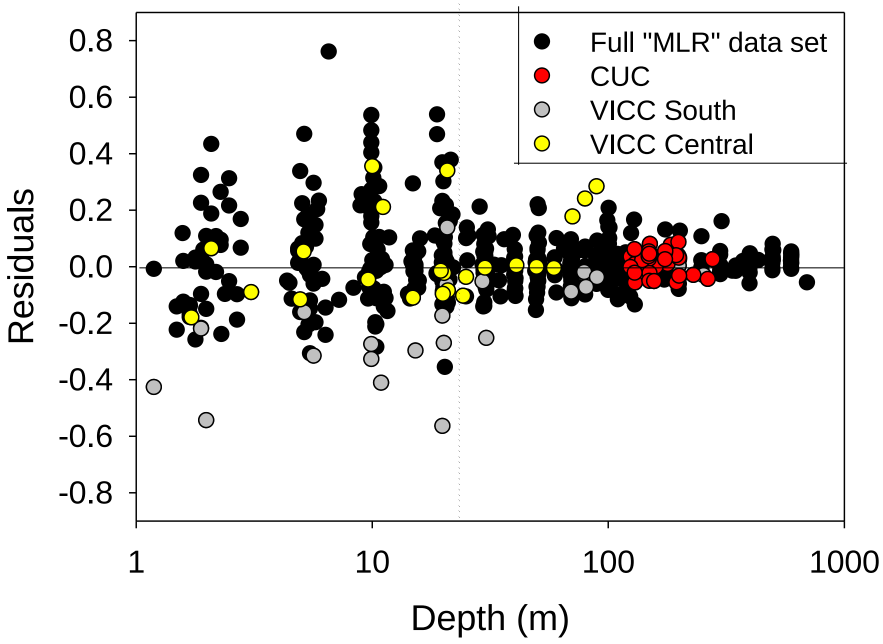

Figure 2.11 Residuals for the final relationship as a function of depth. Relevant water masses are colour coded (CUC, red; south VICC, gray and central VICC, yellow, Appendix D). Depth is expressed on a logarithmic scale. The vertical dotted line indicates 30 m depth.

Surface warming and gas exchange may also be a source of bias in the derived linear model because they can affect the values of T, O_2 and Ω_{arag} and could therefore complicate the relationship between them in surface and subsurface waters. When CO_2 is transferred across the air-sea interface the oceanic DIC (and ultimately Ω_{arag}) changes accordingly. Surface exchange of O_2 also affects the Ω_{arag} estimates. The transfer rate of O_2 is faster than that of CO_2, which could further complicate accurate estimations of Ω_{arag} in surface waters. Furthermore, warming or cooling at the surface could decrease the high correlation between T and Ω_{arag} (largely spatial). Temperature changes would also have a small direct effect on the value of K'_{sp} (i.e. increasing T decreases the solubility of both minerals [Mucci, 1983]) and on the carbonate equilibrium coefficients. In order to identify potential biases caused by both, surface warming and gas exchange, subsets that contained only data from ≥ 30 m depth (the same as the upper cut-off used by Juranek *et al.*, [2009]) were generated.

There were no differences in the model R^2 values (R^2 unchanged ~ 0.97) of the final relationship and that with the surface data excluded (experiments 4 and 7, respectively; Appendix B and Figures 2.8 and 2.11, respectively). As anticipated, the residual error (estimated from the standard deviation of the residual) was significantly lower (0.062, reduced from 0.13, Appendix B) when surface data were excluded. However, there was a shift in the regression coefficients such that the relative importance of T decreased when surface data were excluded (Appendix B). This shift can be a result of the standardization (i.e. when surface data are removed, the standard deviation of T decreases and therefore T is divided by a smaller value that could yield a smaller regression coefficient). In order to address the potential effect of the standardization in both relationships, I compared the values of the regression coefficients of T relative to their corresponding standard deviations. This comparison indicated that, although not as prominent, there was still a shift in the T regression coefficients of the final relationship and that with the surface data excluded which yielded values of 6.34×10^{-2} and 5.5×10^{-2}, respectively.

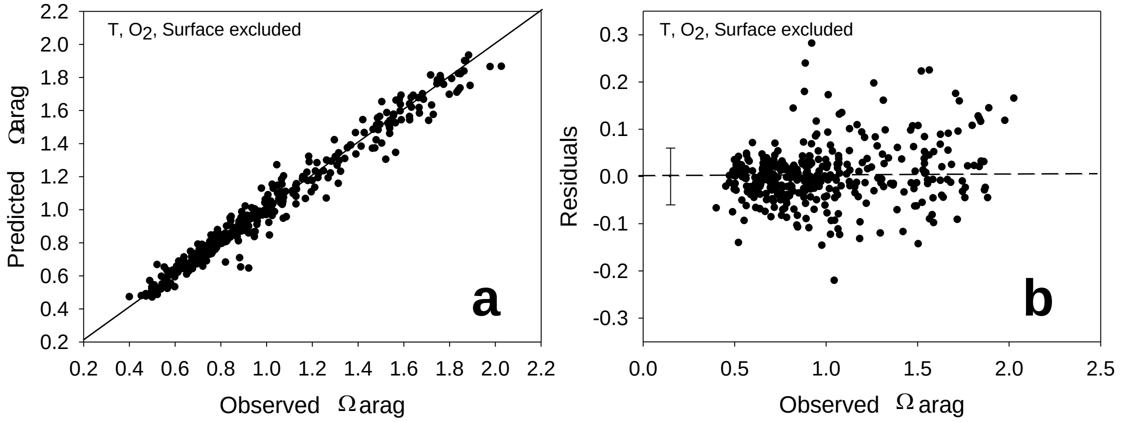

Figure 2.12 a) Measured Ω_{arag} and predicted Ω_{arag} (ln-transformed) with two-predictor (T and O_2) with surface data (<30 m) excluded and b) the corresponding residuals and error bars for ± 1 standard deviation of the residuals (0.06).

2.5.5 Other predictors

In addition to the two-predictor (T and O_2) relationship, I also tried the predictor pair T and AOU (Apparent Oxygen Utilization) (experiment 5, Appendix B). Because AOU is an estimate of the remineralized component of O_2 (i.e. the change in O_2 that has occurred since the water parcel was last in contact with the atmosphere), it was hypothesized that predictions of Ω_{arag} could be improved as the influence of T (and hence O_2 saturation) is removed from the O_2 observed value. However, no substantial improvement resulted from this analysis (Figure 2.13). Predictions with T-AOU yielded essentially the same error as that from T-O_2 (0.13 in both cases, Appendix B).

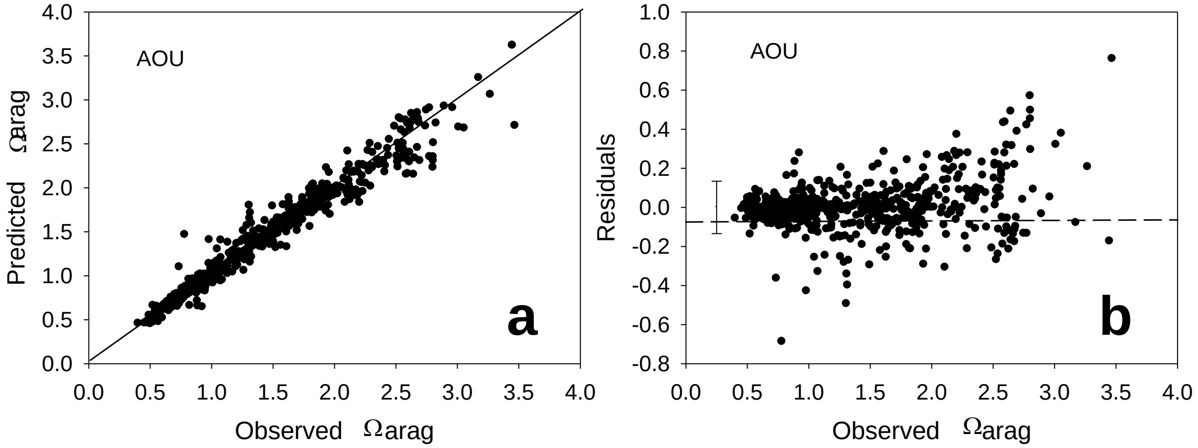

Figure 2.13 a) Measured Ω_{arag} and predicted Ω_{arag} (ln-transformed) with the two-predictor (T and AOU) relationship (0-750 m) and b) corresponding residuals and error bars for ± 1 standard deviation of the residuals (0.13).

One more approach for predicting Ω_{arag} was tested. I considered building a relationship that constructs predictions of Ω_{arag} from predictions of CO_3^{2-} (Figure 2.14a-b):

$$ln\left[{CO_3}^{2-}\right]^{pred} = \delta_0 + \delta_1 T + \delta_2 O_2 \qquad (2.5.3)$$

where δ_i are the regression coefficients. This relationship was built using the same predictors, T and O_2 as the model directly predicting Ω_{arag}. The carbonate ion concentration ($[CO_3^{2-}]$) is a function of the same variables as Ω_{arag}, and these two quantities are related through Eqn. (2.1.1). Because $[Ca^{2+}]$ and K'_{sp} are well known and are unrelated to $[CO_3^{2-}]$, I hypothesized that predicting $[CO_3^{2-}]$ could improve the results.

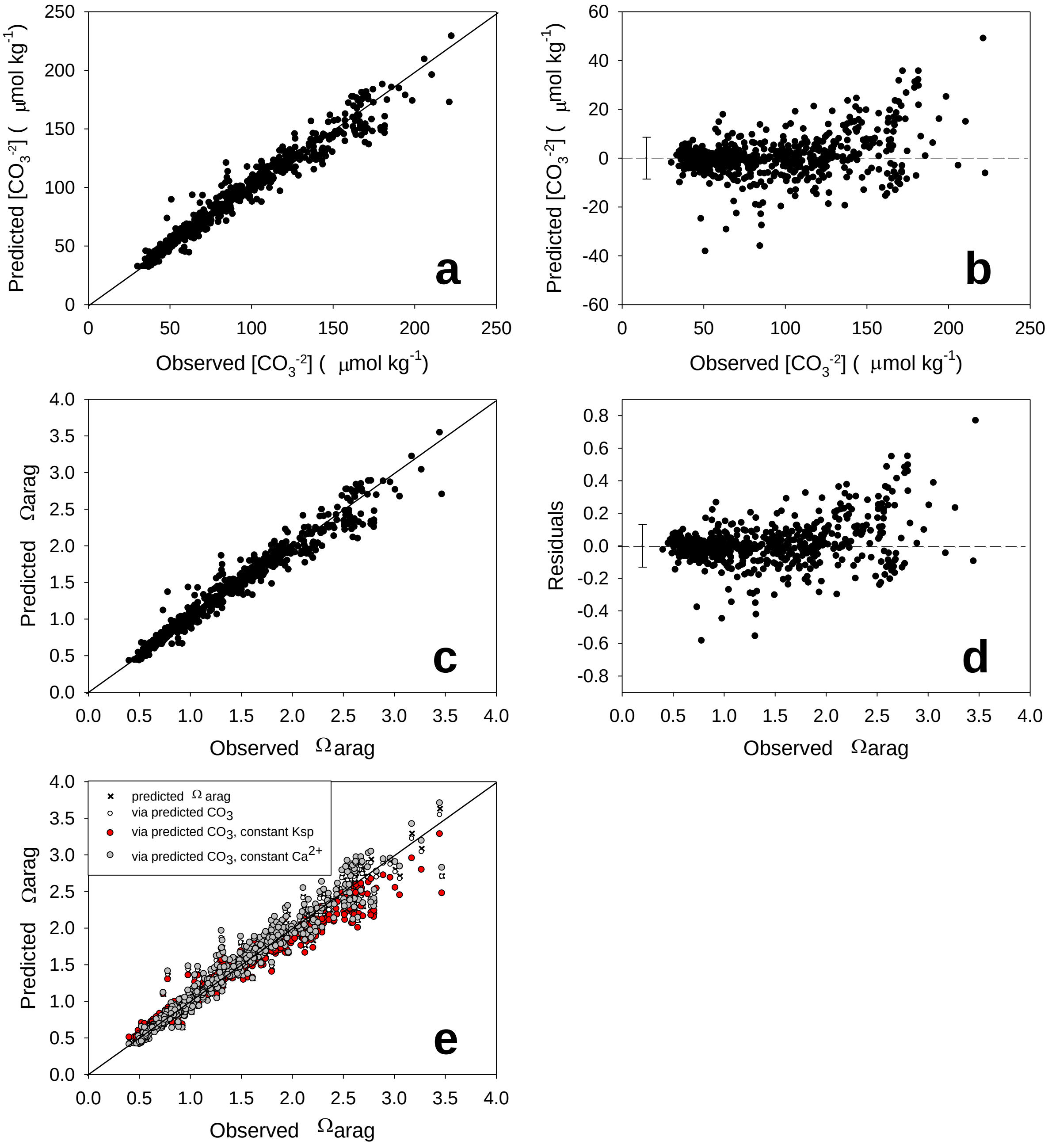

Figure 2.14 a) Measured vs. predicted [CO$_3^{2-}$] from two-predictor (T and O$_2$); c) Measured Ω_{arag} vs. calculated Ω_{arag} (using predicted [CO$_3^{2-}$] above) from two-predictor (T and O$_2$) relationship. The corresponding residual plots are shown in (c) and (d); e) Measured Ω_{arag} vs. Ω_{arag} of final relationship: T and O$_2$ predictors (black), calculated via predicted [CO$_3^{2-}$] (white) (ln transformed), at K'$_{sp}$ (red) and at [Ca^{2+}](gray) constant values of 6.43X10^{-7} mol^2 kg^{-2} and 0.000963 mol kg^{-1} respectively. All predictions a-e were obtained from full data set (0-750 m). Errors bars (± 1 standard deviations of the residuals) for predicted [CO$_3^{2-}$] (8.6) and predicted Ω_{arag} (0.13) are also shown in the corresponding residual plots.

However, empirical values of Ω_{arag} calculated via predicted $[CO_3^{2-}]$ (Figure 2.14c) and direct estimation of Ω_{arag} were almost the same (Figures 2.8a, 2.14a) as were their corresponding residuals (Figures 2.8b, 2.14d). These strong similarities are attributed to a dominant influence of $[CO_3^{2-}]$ in variations of Ω_{arag} relative to those of $[Ca^{2+}]$ and K'_{sp}.

To demonstrate the secondary role of variations in Ca^{2+} in controlling variations in Ω_{arag}, I fixed the variables K'_{sp} and $[Ca^{2+}]$ to constant average values ($6.43X10^{-7}$ mol^2 kg^{-2} and 0.00963 mol kg^{-1}, respectively) and using these (along with predicted $[CO_3^{2-}]$) to calculate Ω_{arag}. The results from these computations (Figure 2.14e) showed that the variations in K'_{sp} and $[Ca^{2+}]$ are small enough to not cause large changes when they are held constant (particularly $[Ca^{2+}]$), consistent with the fact that predictions of Ω_{arag} and calculations of Ω_{arag} via predicted $[CO_3^{2-}]$ were nearly identical. I concluded that there is no advantage given our data to predicting Ω_{arag} indirectly from predictions of CO_3^{2-}, relative to direct predictions of Ω_{arag}.

2.5.6. Comparison with previous study by Juranek *et al*. [2009]

Juranek *el al.* [2009] also used MLR to empirically estimate Ω_{arag} from hydrographic data. Their study was conducted ~250 km south of the present study area, on the shelf of central Oregon.

For comparison, I generated contour lines of the model predictions of Ω_{arag} from Juranek *et al*.'s relationship and my relationship in T-O_2 space along with the observed Ω_{arag} scatter of my study (Figure 2.15). In general, contour lines for estimates computed with both

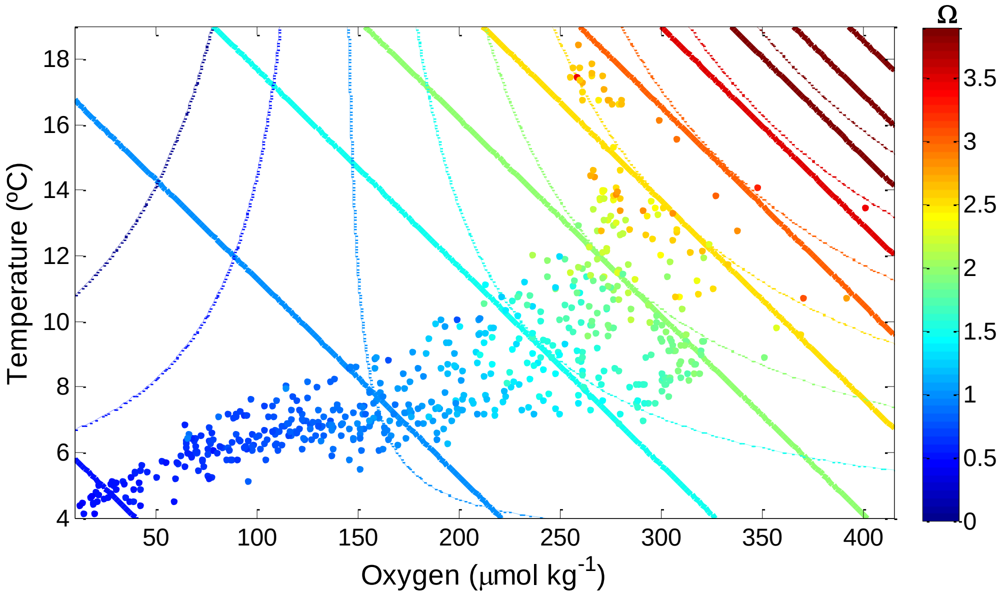

Figure 2.15 The distribution of observed Ω_{arag} through T-O_2 space (coloured dots). The thick contour lines correspond to final model predictions (equation 4, Figure 8a-b) from present study while the thin lines correspond to predictions of the model of Juranek et.al., [2009].

relationships (Figure 2.15) were similar over the range of T and O_2 of my data despite the differences in the geographical regions of the two studies and the fact that Juranek *et al.* [2009] limited their study to a narrower portion of the water column. However, outside of the T and O_2 space occupied by my data, the relationships were not similar.

Despite Juranek *et al.*'s exclusion of surface data (<30 m), the contour lines (Figure 2.15) of the two estimates are close together at high values of Ω_{arag}. However, at lower values of Ω_{arag}, which are generally found deeper in the water column, both predictions increasingly diverge. This characteristic can also be observed in predictions of my data, computed using Juranek *et al.* [2009]'s relationship (Figure 2.16). Evident biases are found at low values of Ω_{arag} and their corresponding residuals (which tend to be negative).

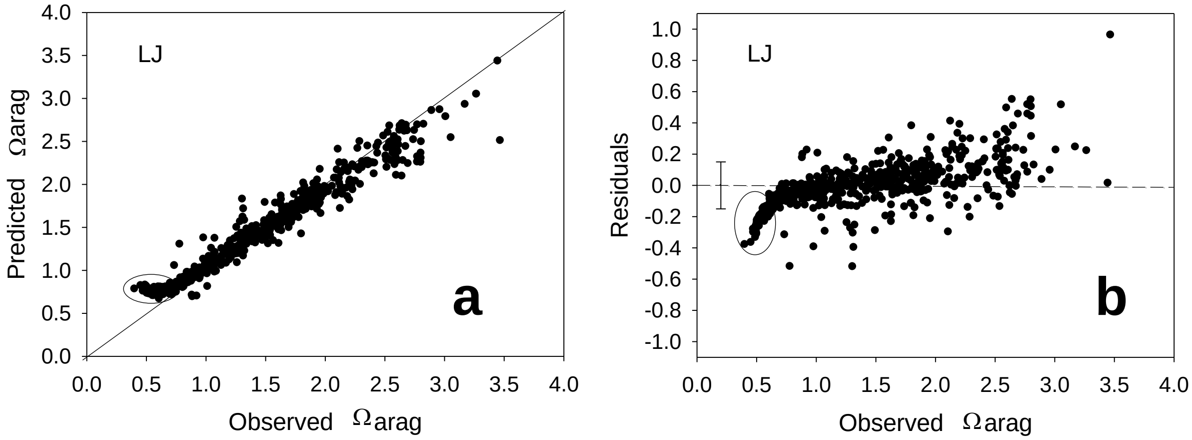

Figure 2.16 a) Measured Ω_{arag} and predicted Ω_{arag} computed with Juranek *et al.* [2009]'s relationship (0-750 m of my data) and b) corresponding residuals. The ovals in both plots highlight the tendency of larger errors at low values of Ω_{arag}. A calculation of the error (± 1 standard deviation of the residuals; 0.15) for the final relationship is shown (b).

Juranek *et al.* [2009] limited their study to the 30-300 m depth range whereas in my study, the depth range was 0-750 m. My relationship minimizes biases at low Ω_{arag} by applying the ln-transformation (section 2.4). An estimate bias at low Ω_{arag} values similar to that which was observed for the non-log transformed model in this study was also observed in Juranek *et al.* [2009]'s study. They addressed this problem by choosing a regression model form that involved the inclusion of interaction terms (while I chose the ln-transformation to correct the bias at low Ω_{arag}). Their model still produces biased predictions at low Ω_{arag} for the present data. For my analysis, I chose to use the relatively simple model form without interaction terms in order to limit the number of statistical parameters to be estimated, in an effort to avoid model overfitting.

2.6 Model utility

In this section, I will use the final predictive model to predict Ω_{arag} in one transect from the historical data sets (Institute of Ocean Sciences (IOS) archive) to illustrate its utility. I will also present the advantages of using this model.

Historically, the sampling of carbonate system parameters in the WCVI has been temporally and spatially sparse owing to difficulty and cost of making such measurements and a previous lack of interest. In this study I have derived an empirical relationship that can reconstruct Ω_{arag} from available measurements of hydrographic parameters (T, O_2). These two parameters have been collected semi-regularly in my study region over the past few decades (starting regularly in 1979).

I applied the final algorithm to data sampled in May of 1998 along transect LC (Figure 2.1) of the regional historical archive (Figure 2.17), which is a line that has been occupied frequently by the La Perouse program [Ware and Thomson, 1986; Mackas, 1992; Ware and Thomson, 1993] and it is also located near a site where continuous sampling has recently begun [NEPTUNE, Barnes *et al.*, 2008]. These data were collected at the end of the spring. With simple linear interpolation from our computations, I also estimated the saturation horizon (the depth at which $\Omega_{arag} = 1$) (Figure 2.17) below which the first indications of $CaCO_3$ dissolution are found [Mucci, 1983; Orr *et al.*, 2005].

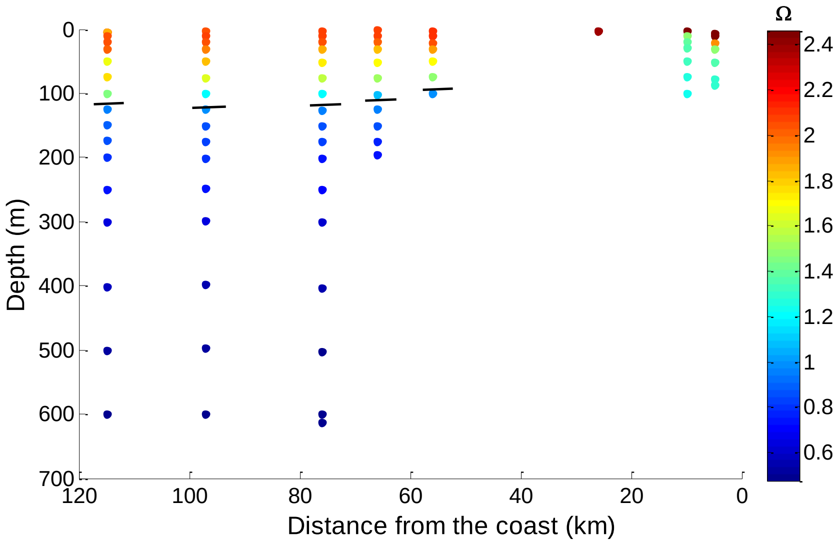

Figure 2.17 Predicted Ω_{arag} from T and O_2 data collected in May of 1998 along LaPerouse transect LC of the regional (west coast of Vancouver Island, Figure 2.1) archive using the final predictive algorithm. The black horizontal lines represent the Ω_{arag}=1 saturation horizon.

Undersaturation with respect to Ω_{arag} does not occur close to the shore in this example (Figure 2.17; 0-40 km from the coast). In the offshore portion of this transect, undersaturation with respect to Ω_{arag} occurs between ~100 – 130 m. Understanding the variability in the distribution of the saturation horizon throughout the water column can potentially be used to help understand how organisms may respond to sporadic presence of waters low in Ω_{arag} along the WCVI region. The use of this approach may provide insight into natural variability and the key controls of Ω_{arag} in future studies at a low cost. However, it is important to keep in mind that this empirical relationship cannot hind-cast data to evaluate the presence of the anthropogenic signal.

Uptake of anthropogenic CO_2 by the ocean results in a decrease of seawater Ω_{arag} values. Previous studies [Feely *et al.*, 2008] have suggested that the upwelled water along the northern California coast (most northern stations of the Sabine *et al.* [2002] data sets) currently contain

31±4 µmol kg^{-1} anthropogenic CO_2. Removing this signal from the DIC concentration increases the aragonite saturation state by ~0.2 units. Since the addition of anthropogenic CO_2 can cause the slope of the O_2:CO_2 ratio to change my relationship may become biased if applied to data from time periods much earlier than our calibration data was collected (1998-2010) when the anthropogenic signal in my study region was presumably less. For example Juranek *et al.* [2011] limit the application of their open ocean relationship to ±10 years. However, evaluation of the anthropogenic signal in the California current system is difficult [Min *et al.*, 2002] and it remains unquantified. The natural variability in our region is high and so we anticipate that it will be may decades before it is possible to identify an anthropogenic increase in CO_2 [*sensu* Easterling and Wehrner, 2009; J. Christian, pers. comm.]. Thus, I assume that it is appropriate to apply the relationship over a time frame of 30 years outside of my calibration set.

In a recent study [Borges and Gypens, 2010] it was shown (through the use of a numerical model) that in highly productive nearshore environments, the effect of eutrophication on carbon cycling can counter the effect of ocean acidification on the carbon chemistry of surface waters. With the use of a coupled numerical model, the authors show the importance of a variety of anthropogenic influences on carbon cycling in coastal zones and highlight the potential difficulty in isolating the atmospheric anthropogenic carbon signal in these regions. In other studies [e.g. Cai *et al*., 2011] it has been suggested that eutrophication could increase the susceptibility of coastal waters to ocean acidification as a result anthropogenic inputs of nutrients that have caused massive algal blooms which deplete O_2 in bottom waters and release CO_2 when the organic matter from these blooms is respired by bacteria.

2.7 Conclusion

In this study, I explored a sequence of statistical procedures that led to the construction of a multiparameter linear regression model relating seawater Ω_{arag} to hydrographic measurements. Two key variables, T and O_2, were identified as strong predictors of the variability of seawater Ω_{arag}. I found that a logarithmic transformation of the relationship derived with the predictors T and O_2 improved the predictions. While this transformation did not substantially improve the model performance globally (R^2 from 0.93 to 0.96 and false negative from 8.2 to 3.1%; Table 2.1), it was effective in linearizing the curvature of the two-predictor relationship and eliminating a strong bias in predictions at low Ω_{arag} (Figure 2.6). Observations of Ω_{arag} shown in the T-O_2 space (Figure 2.15) agreed well with the contour lines of our predictions. In the deep ocean (low T, low O_2 region) Ω_{arag} is less variable relative to surface waters (Figure 2.15) and its variability is not affected by dynamic processes such as upwelling and biological uptake of carbon in response to this upwelling, as it is in the surface. My statistical model was able to simulate variations from high to low Ω_{arag} in the surface. This good agreement between the observations and the empirical estimates occurred despite the inclusion of data from the surface (a highly variable region of the water column) and distinct water masses (such as the CUC and VICC). The relationship appears to overestimate omega in the surface in the limited region where the VICC begins, at the southern tip of the island; however, residual error in this region is of the same order as the largest surface residuals in the rest of the study region. Other model forms, which included the use of both a single PCA-based predictor and a VIF factor, were built and compared with the ln-transformed model form, but the quality of the empirical estimates using these two model forms did not improve with respect to the ln-transformed model form. Perturbations caused by anthropogenic carbon cannot be accounted for when only knowing T and O_2, therefore our relationship may become biased as the ocean absorbs more carbon, regardless on how T and O_2 also change.

Chapter 3
Determination of the acidification state of Canadian Pacific coastal waters using empirical relationships with hydrographic data

3.1 Introduction

On the northeastern Pacific continental shelf, a region that boasts high primary production and as a result hosts a lucrative fishery [Ware and Thomson, 2005], corrosive conditions with respect to aragonite, a direct result of absorption of large emissions of anthropogenic CO_2 by the oceans, have recently been reported [e.g. Feely *et al.*, 2008]. Continental margin carbon cycling is complex and highly variable over a range of space and time scales, and influenced by multiple physical and biogeochemical processes [e.g. Borges *et al.*, 2005]. It is uncertain how increases in ocean carbon concentration will affect fisheries, but a negative economic impact is anticipated [e.g. Cooley and Doney *et al.*, 2009]. Until relatively recently, direct [e.g. Ianson *et al.*, 2003; Hales *et al.*, 2005; Feely *et al.*, 2008] and indirect [e.g. Juranek *et al.*, 2009; Kim *et al.*, 2010, Alin *et al.*, 2012] measurements of carbonate system parameters in these regions have been scarce, and consequently natural variations in ocean carbon chemistry on short time scales (i.e. seasonal, interannual) are still poorly understood.

A relevant parameter for marine shell-forming organisms is the $CaCO_3$ saturation state (Ω) defined as: $\Omega = \frac{[{CO_3}^{2-}][Ca^{2+}]}{K'_{sp}}$, where $[Ca^{2+}]$ is the concentration of calcium, $[{CO_3}^{2-}]$ is the carbonate ion concentration and K'_{sp} is the apparent solubility product dependent on temperature (T), salinity (S) and pressure. Calcium carbonate occurs in two common polymorphs: aragonite and calcite. Aragonite has a 50% higher solubility than calcite [Mucci, 1983]. Thus, organisms that produce aragonite are likely more vulnerable to changes in ocean acidity than those that produce calcite [Ohde and Mozaffar-Hossain, 2004]. Because Ω_{arag} depicts the most acute extent to which organisms are affected by the changes in $[CO_3^{2-}]$, it is commonly used as one of the variables that describe corrosive water [e.g. Feely *et al.*, 2008; Doney *et al.*, 2009]. In thermodynamic equilibrium, $\Omega_{arag}>1$ indicates that mineral precipitation (as aragonite) is favoured. For values $\Omega_{arag}<1$, seawater is corrosive to $CaCO_3$ and, in the absence of biologically-mediated protective mechanisms such as organic coatings, dissolution will begin [Fabry *et al.*, 2008]. It has been reported [e.g. Feely *et al.*, 2004; Hoegh-Guldberg *et al.*, 2007; Doney *et al.*, 2009] that such undersaturated conditions can cause marine organisms to reduce calcification rates.

Along the Canadian portion of the northeastern Pacific coast, several process studies and semi-regular hydrographic surveys have been carried out over the past decades resulting in nearly twelve thousand observations of T, S and oxygen (O_2). However, relatively few of these coastal data sets contain carbon system parameters to assess natural variations in Ω_{arag}.

In this study, I use a previously-constructed statistical model to predict Ω_{arag} from historical observations of T and O_2 (1979-2009) along the Canadian Pacific coast. Seasonal distributions of Ω_{arag} are contrasted and regional patterns are described. The effect of seasonal upwelling on these distributions is discussed.

3.2 Data and Methods

3.2.1 Data

The majority of the Canadian Pacific hydrographic (T, S, O_2) data, hereafter called 'historical' data, were collected along the coast adjacent to Vancouver Island (primarily as part of the La Perouse program [Ware and Thomson 1986; Mackas, 1992]) but a small percentage (~10%) were collected north of ~50º (Figure 3.1) in the Queen Charlotte Sound region. The LaPerouse program began in 1985 and is on-going with two surveys per year in spring and late summer. The Line P program [Whitney and Freeland, 1999] also contributes to this record, although the majority (more than ~ 75%), of this record is not coastal (i.e. over ~90 km offshore, water depth > 1300 m) and is not relevant in this study.

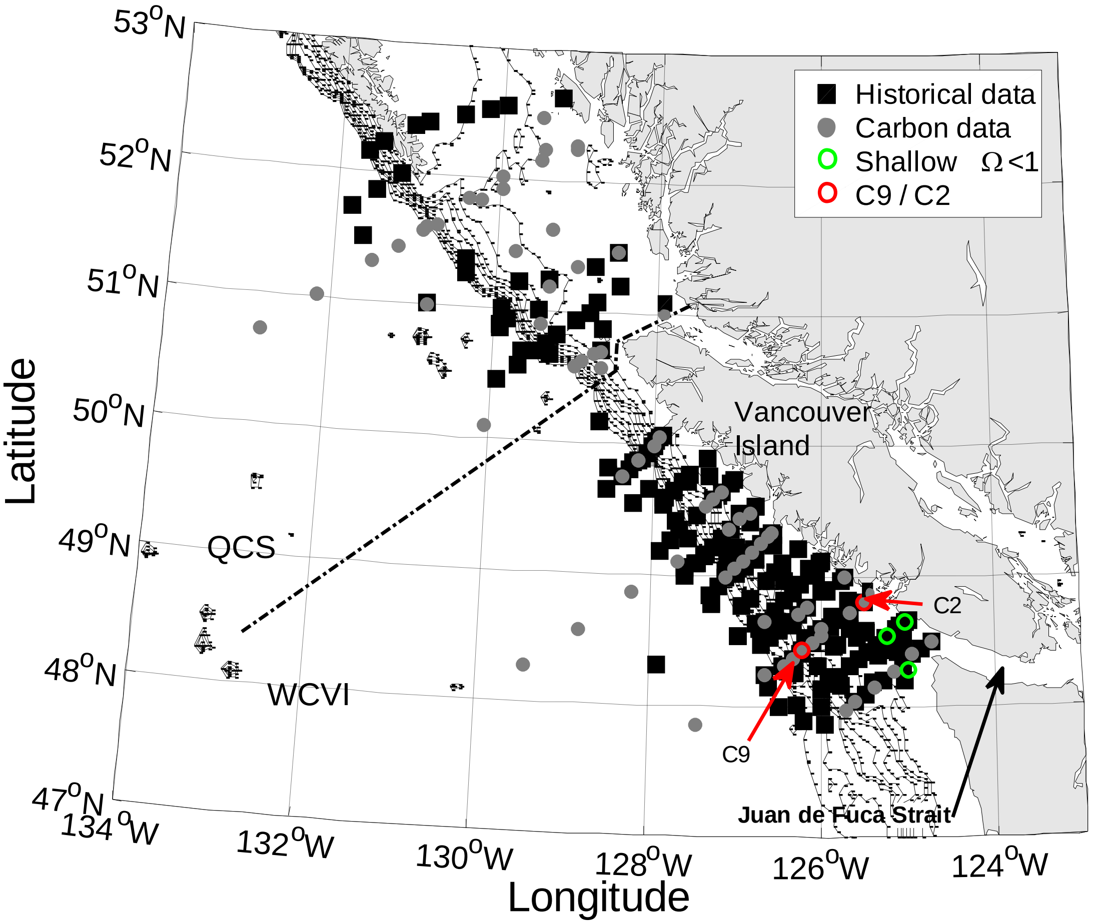

Figure 3.1 Map of stations with carbon measurements (1998-2010, gray circles), the historical stations (1979-2009, black squares) and those stations that had exceptionally shallow undersaturation of Ω_{arag} (open green circles). Bold dashed line shows the separation between QCS and WCVI regions. The 200 m to the 1800 m isobaths are shown at 200 m intervals to mark the slope. Red arrows and red open circles highlight stations C2 and C9.

Recent O_2 samples (1994 to date) were analyzed on an automated Winkler titration system with colorimetric detection of endpoint to titrate the oxygen samples following Carpenter [1965]. O_2 samples taken before these dates were collected into oxygen flasks and titrated on board with a visual endpoint by using a modified Winkler method [Carpenter, 1965]. Conductivity, temperature, depth (CTD) casts were made concurrently using the following instruments: Guildline CTD (1979 to late 1987), Guildline Salinometer

45055/41226/53501/8705 (1988-1990), Guildline Salinometer 8715 (1991-1996), Seabird model SBE 911+ (1997 to present).

A total of 618 records that include dissolved inorganic carbon and total alkalinity from 5 hydrographic surveys (1998—2010) off the Canadian West Coast, hereafter referred to as 'carbon data', were used to construct the statistical model (Chapter 2) relating hydrographic quantities and the aragonite saturation state. Data coverage was denser in the surface with about 40% of the samples collected at or above 30 m. Most data (~ 98%) were collected between May and August, during the summer upwelling season. Only data from February 2010 corresponded to the winter downwelling season. The historical data include data collected in all seasons, although there are fewer winter data (only 10% of the total). Spring transition, summer upwelling and fall transition (as defined by oceanographic season; Bylhouwer *et al.*, submitted; Appendix E) make up 36, 24 and 30% of the record, respectively.

Generally, the carbon data (N=618) occupied similar regions in T-S and T-O_2 space as the majority of the historical data (N=11392) (Figure 3.2). The 2004 carbon data contain a cluster of data that falls at the warmest edge of the historical scatter in both T-S and T-O_2 space. In addition, the freshest waters in the historical data set (S ~28-30) were not represented by the carbon data; however, they correspond to less than 1% of the historical record. These results suggest that the most relevant water masses in the historical data and their variability across space and season were captured in the data used to construct the model.

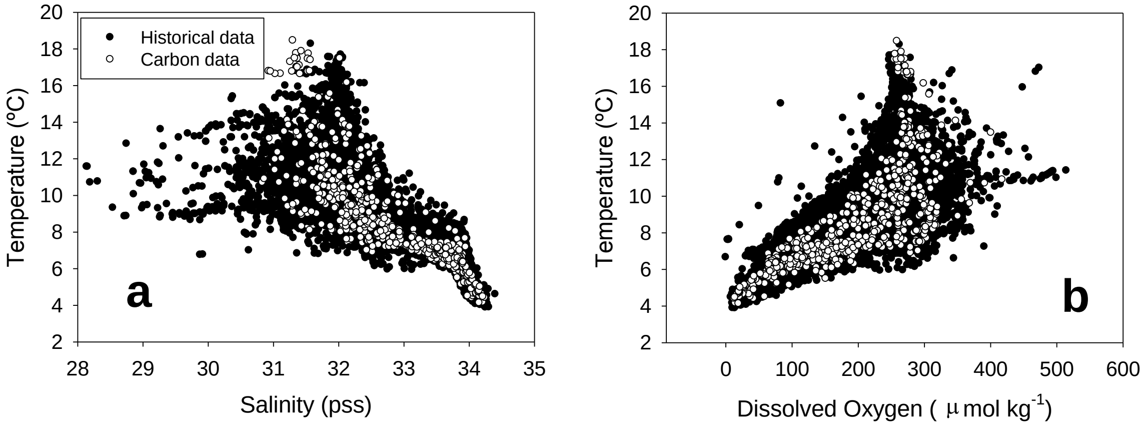

Figure 3.2 a) T-S and b) T-O_2 diagrams of model data (white circles) overlaid on historical data (black circles).

The shelf was distinguished from slope or 'offshore' waters by nominally using the 200 m isobath as the shelf-break. Thus, all sampling locations with bottom depths of 200 m (or shallower) were grouped as 'shelf', while the remainder were considered 'offshore'. Two exceptions (52°N 131.23°W and 51.24°N 129.35°W) occurred in the QCS where troughs that are deeper than 200 m cut the shelf. These exceptions were grouped with shelf stations.

3.2.2 Predictions of Ω_{arag} from Historical data

The statistical model was derived using a multiple linear regression approach to determine the (Ω_{arag}) from observed T and O_2 standardized anomalies:

$$\Omega_{arag}^{pred} = \exp(2.14 \times 10^{-1} + 1.82 \times 10^{-1} * T^S + 3.40 \times 10^{-1} * {O_2}^S) \qquad (3.2.1)$$

where the standardized anomalies of T and O_2 (T^S and ${O_2}^S$) were

$T^S = \frac{T(°C) - 8.78°C}{2.87°C}$ and ${O_2}^S = \frac{O_2(\mu mol\ kg^{-1}) - 198\ \mu mol\ kg^{-1}}{88.6\ \mu mol\ kg^{-1}}$, respectively.

The standardized anomalies were produced by first subtracting the mean from the data, and then dividing by the standard deviation. The estimated error for this relationship was 0.13. Model development is fully described in Chapter 2. Briefly, predicting the natural logarithm of Ω_{arag} avoided prediction biases that occurred at low values of Ω_{arag} when the model was constructed without this transformation. The model was applied to data from the surface to a depth of 750 m. Residuals (differences between the observed and predicted) in the upper 30 m were larger than below 30 m (which is expected due to surface warming and gas exchange in the upper ocean). However, the residuals in general were symmetrically scattered about zero and displayed no bias (discussed in detail in Chapter 2; section 2.5.4). Since interest in surface layer carbon prediction is high and a substantial portion (~40%) of the carbon data fall in this region, I included them. Data below 750 m are subject to inputs of alkalinity and DIC as a result of $CaCO_3$ dissolution and so the carbon system is not predicted well using changes in T and O_2 alone (Chapter 2, sect. 3).

Comparison of the relative frequencies showed that the overall range in Ω_{arag} predicted from historical data (~0.2-4.0) was the same as in the observed Ω_{arag} from the carbon data and that the distributions were almost identical (Figure 3.3a). These distributions are both bimodal. Relative frequencies of the sampling depths in the historical and carbon data are similar (Figures 3.3b and c) and show that there is no bias in the comparison of historical and model Ω_{arag} distribution due to choice of sampling depth.

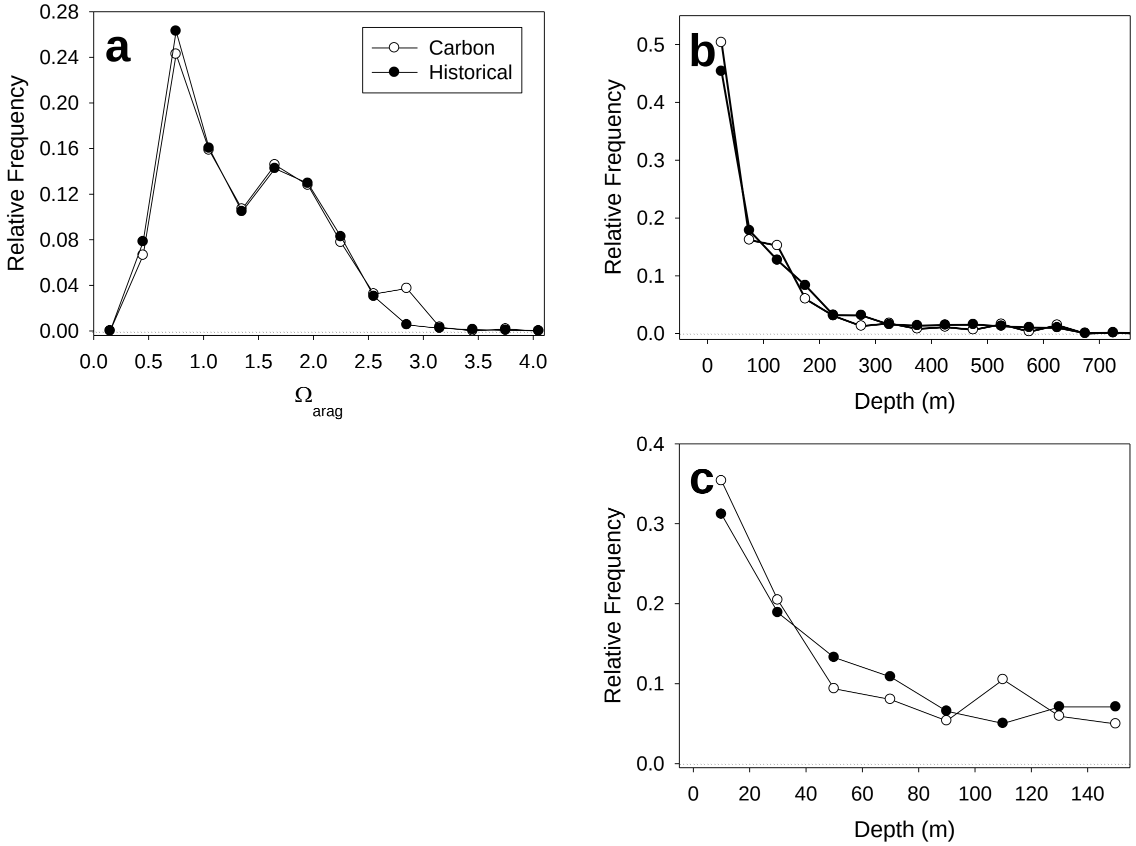

Figure 3.3 a) Relative frequency (number of observations/ total observations,) for data used to build the model (total number of observations N=618) and historical data (N=11392) (a) Ω_{arag} (0.3 bins) (b) sampling depth from 0 to 750 m (50 m bins) (c) sampling depth from 0 to 150 m (20 m bins).

3.3 Variability in the Ω_{arag}=1 saturation horizon

From a thermodynamic perspective, Ω_{arag}=1 represents a transition between preservation and dissolution of the aragonitic form of $CaCO_3$ [Orr *et al.*, 2005; Fabry *et al.*, 2008]. In seawater, a natural horizontal boundary between corrosive and non-corrosive waters, known as the saturation horizon, is formed primarily as a result of increasing pressure and dissolved inorganic carbon (DIC) concentrations with depth [Mucci, 1983]. Above this saturation horizon, Ω_{arag} has a value greater than 1, and mineral $CaCO_3$ does not readily dissolve. Most calcifying organisms live in such waters [Raven *et al.*, 2005]. Below this depth, Ω_{arag} has a value less than 1, and $CaCO_3$ will dissolve (although biological processes can inhibit dissolution). Assuming that Ω_{arag} varies linearly between sampling depths at a given sampling station, I calculated the horizon depth (Ω_{arag}=1) by linear interpolation.

3.3.1 Vertical profiles of estimated Ω_{arag} on a typical transect

Two locations from a well-visited transect (C2 and C9, Figure 3.1) off SW Vancouver Island were used to illustrate typical profiles and vertical sampling density, as well as, seasonal variations in Ω_{arag} on-shelf and off-shelf (Figure 3.4a and b, respectively). The reconstructed Ω_{arag} at these two stations decreased with depth as expected and was rarely undersaturated at depths of less than 40 m. In shelf profiles, undersaturation occurred most often during summer upwelling and fall transition, although surface Ω_{arag} was considerably lower during the winter season. During winter, the full data showed that supersaturation with respect to aragonite throughout the entire water column over the shelf was more likely (73%, Table3.1) than during any other season. In contrast, during summer and fall undersaturation in the lower portion of the water column (as at C2, Figure 3.4a) was common over the shelf (83% and 79%, respectively; Table 3.1).

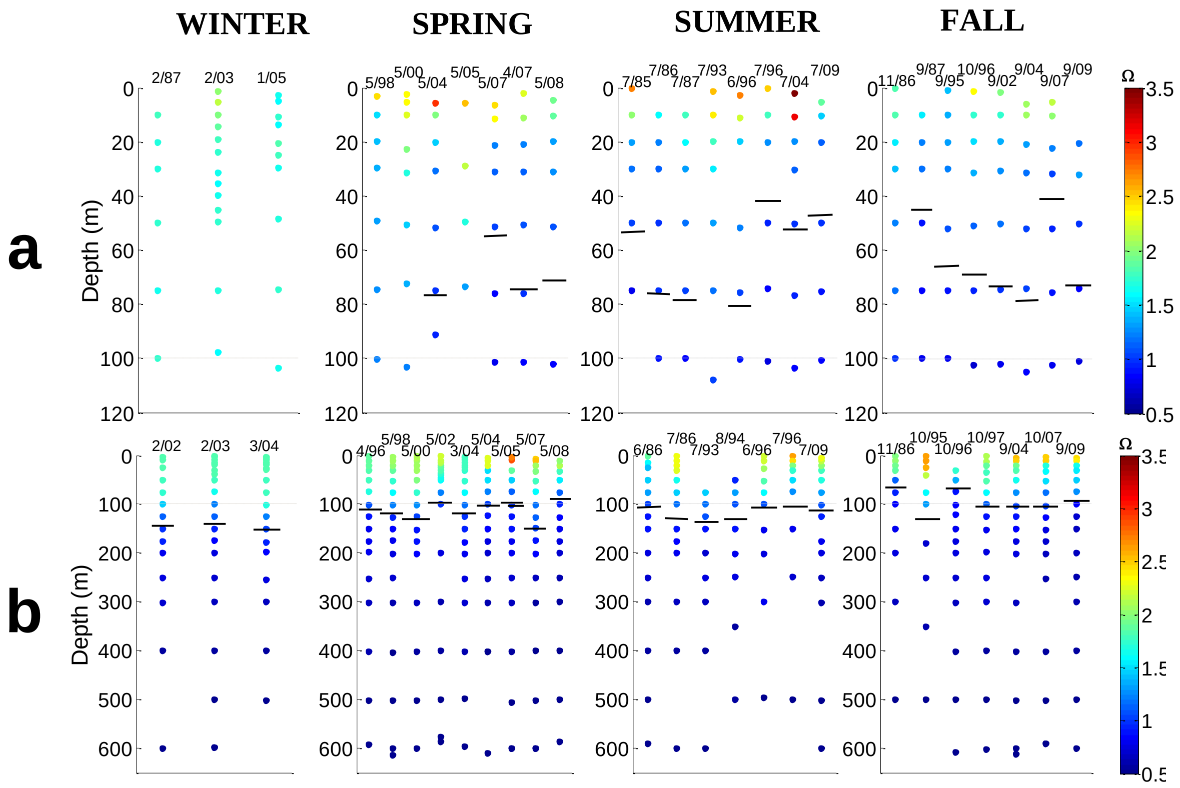

Figure 3.4 Vertical profiles of predicted Ω_{arag} off Vancouver Island (a) on the shelf C2 (48.8N. 125.5ºW; 10 km from the coast; depth of ~120 m) and b) off the shelf C9 (48.4N. 126.2W; 76 km from the coast; ~600 m depth). Horizontal black lines show the Ω_{arag}=1 saturation horizons. Numbers on top indicate month and year, respectively when data were collected. Gray dashed line indicates the 100 m isobath.

Table 3.1 Total number of vertical profiles of historical data for the shelf and slope regions divided by season; winter downwelling (WD), spring transition (ST), summer upwelling (SU) and fall transition (FT) [Thomson and Ware, 1996; Bylhouwer *et al*. submitted]. The total number is given followed in parentheses by the number of profiles in which the entire water column was saturated with respect to Ω_{arag}. The number of profiles in QCS is also listed for each depth region. There are three shallow profiles sampled only from the surface to < 100 m, from the offshore region that were not included.

	Shelf		Offshore	
	Total	QCS	Total	QCS
WD	49 [36]	2[1]	39[0]	4[0]
ST	212 [86]	22[12]	160[0]	18[0]
SU	229 [38]	11[2]	134[0]	20[0]
FT	207 [43]	15[7]	136[0]	11[0]

Off the shelf, results were similar but less variable (Figure 3.5) with the shallowest saturation horizons at C9 occurring in fall (Figure 3.4b). The saturation horizon was never as shallow as on the shelf on this transect (Figure 3.4). Similar to the shelf station (C2), surface Ω_{arag} was lowest at this offshore station (C9) during the winter (Figure 3.4b). The saturation horizon occurred above 200 m in all offshore profiles (Table 3.1). These results agree with recent observations by Feely *et al.* [2012]. They show that Ω_{arag}=1 horizon was located relatively deep (~160 m) in the water column during winter offshore of my study region.

Vertical sampling density in the historical data is particularly high above 50 m (e.g. Figure 3.4) where vertical gradients in water properties can be significant (e.g. Appendix F). Thus, errors associated with estimating the depth of the saturation horizon by linear interpolation are minimal. Vertical gradients in Ω_{arag} were generally highest at the base of the mixed layer over the shelf during all seasons (e.g. Appendix F), where samples were collected at least every 20 m and often at intervals of 10 m or less (e.g. Figure 3.4a).

3.3.2 Changes in the distribution of Ω_{arag} due to seasonal cycles

While four distinct oceanographic seasons can be defined in terms of local wind-forcing; i.e., summer upwelling , fall transition, winter downwelling, and spring transition (Bylhouwer *et al*. submitted; Appendix E), the observed Ω_{arag} saturation horizon distributions from the complete historical data record can be characterised by two dominant seasons; winter-spring and summer-fall (Figure 3.5a). Relative frequencies of the Ω_{arag}=1 horizon depth in profiles experiencing undersaturation (83% of the 1166 total profiles, Table 3.1) (Figure 3.5a) occupied similar ranges in winter and spring with a relatively narrow distribution around the highest relative frequencies between 80 and 140 m, showing that most often waters became undersaturated with respect to Ω_{arag} within this depth range (Figure 3.5a) where undersaturated conditions were present (Table 3.1). In contrast, in summer-fall the distribution (based on profiles experiencing some undersaturation, 89% of the total, Table 3.1) was flatter showing that the saturation horizon was found over a wider range of depths (~20-120 m) and was more likely to occur at shallower depths compared to winter/spring (Figure 3.5a). The Ω_{arag} profiles from station C2 (Figure 3.4a) illustrate these trends; undersaturation of Ω_{arag} in summer and fall was present in most cases yet variable in depth (with the saturation horizon ranging from ~40 to ~80 m).

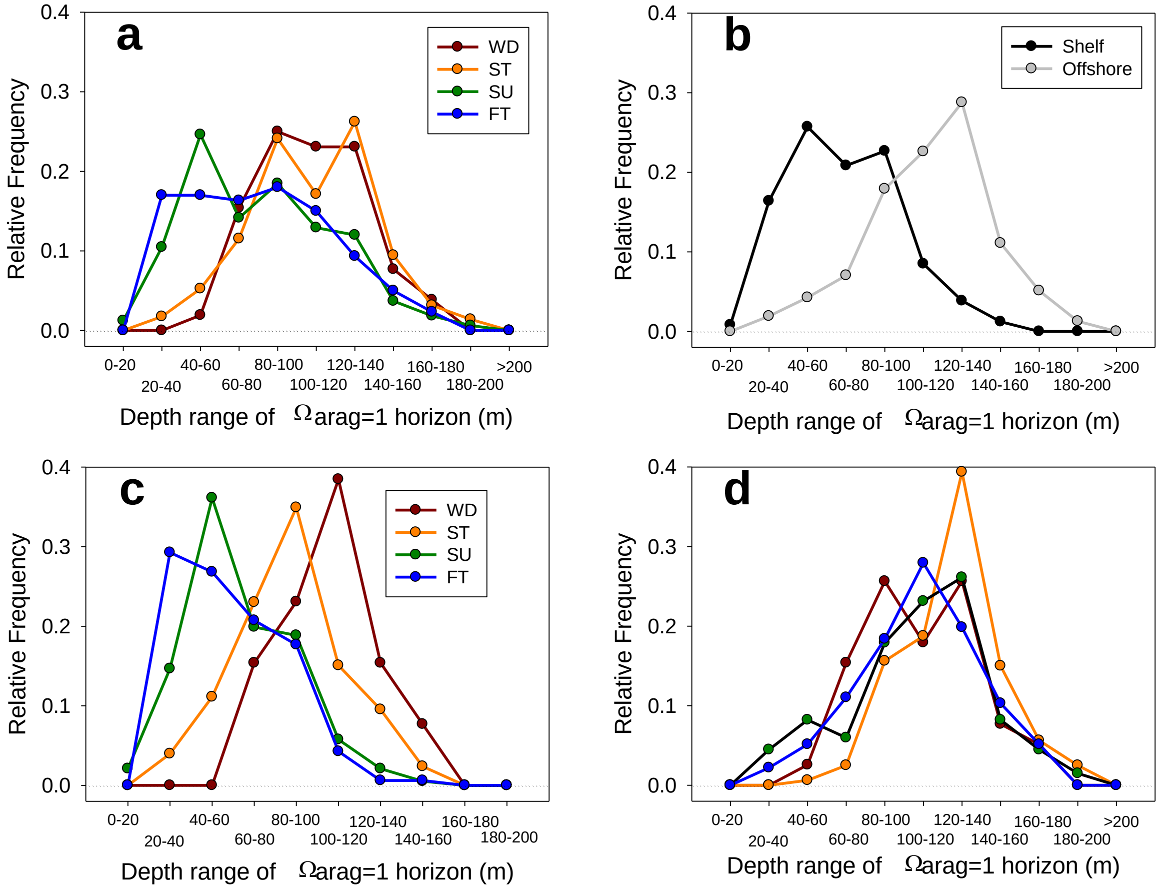

Figure 3.5 Relative frequency of depth of Ω_{arag}=1 horizon (only profiles in which undersaturation occurs; Table 3.1) from a) the historical record (20 m bins) for winter (N=52; red), spring (N=286; orange), summer (N=325; green) and fall (N=300; blue); b) the historical record (20 m bins) for shelf (shallower than 200m; N=494; black) and offshore (deeper than 200m; N=469; gray) waters; c) the historical record (20 m bins) for winter, spring, summer and fall for shelf waters and d) from the historical record (20 m bins) for winter, spring, summer and fall for offshore waters. The horizontal dotted lines show the zero lines.

The chances of a shallow transition to undersaturation of Ω_{arag} are high in summer, likely as a result of shoaling of deep carbon-rich isopycnals due to summer upwelling [Smith, 1994; Ianson *et al*. 2003]. During the fall transition season, relatively shallow saturation horizons are also observed (20 -- 60 m) (Figure 3.5) likely due to late season upwelling events which may occur later than local winds suggest [Thomson and Ware, 1996; Bylhouwer *et al*., submitted] augmented by *in situ* remineralization of organic matter produced during spring and summer [Ianson *et al.*2003*;* Ianson *et al*. 2009]. Substantial oscillations of pH and Ω_{arag} surfaces through an annual cycle and particularly low Ω_{arag} water masses near the surface in the summer were also observed in southern California (between Punta Eugenia and San Diego) in 2005 [Alin *et al.*, 2009] and near central Oregon in 2007 [Juranek *et al.*, 2009], where outcropping of the 0.8 Ω_{arag} horizon to the upper 30 m in the mid-shelf was associated with a strong upwelling event in mid-May.

Unlike the summer -fall, undersaturation with respect to aragonite was not common at shallow depths (<80 m) during the spring and winter (Figure 3.5a). This absence of undersaturated conditions is likely a result of winter downwelling, during which offshore surface waters flow onshore and isopycnals become depressed at the coast, forcing lower layer shelf water offshore [Smith, 1994]. During winter, the deeper DIC concentration decreases (increasing Ω_{arag}) (i.e. deeper Ω_{arag}=1 horizons, Figure 3.5a) due to dilution with surface water caused by this off-shelf transport and by mixing during winter storms [Ianson *et al.*, 2009]. In the spring, surface $p\mathrm{CO}_2$ is drawn down by phytoplankton [e.g. Evans *et al*. 2012] and surface Ω_{arag} increases (e.g Figure 3.4), but the more carbon-rich isopycnals remain relatively deep until the onset of upwelling in summer (e.g. Figure 3.5).

3.3.3 The role of the continental shelf

Distributions of Ω_{arag} saturation horizon over the shelf (i.e. waters with a bottom depth of 200 m or shallower; section 3.2.1) were substantially different than off the shelf (Figure 3.5b). Undersaturated waters were always present in offshore waters (Table 3.1), but the transition to undersaturation most often occurred between 80 and 160 m, which is the depth of the majority of the shelf (Figure 3.5b). More importantly, when present, undersaturation with respect to aragonite was more likely to occur at shallower depths (between 20 and 100 m) over the shelf relative to adjacent offshore waters (Figure 3.5b) likely as a result of upwelling (discussed above). Also, the distribution in Ω_{arag} was broader over the shelf than offshore, indicating greater variability in the depth of Ω_{arag} undersaturation over the shelf (Figure 3.5b). In the offshore region, undersaturation rarely occurs shallower than 80 m.

Undersaturation rarely occurred in the upper 20 m over the shelf (Figure 3.5b, Figure 3.6a). These rare instances of near-surface undersaturation all occurred off the mouth of Juan de Fuca Strait (48.29ºN 125.29ºW, 48.57ºN 125.26ºW and 48.64ºN 125.37ºW; Figure 3.1).The first location in 1985 and 1994 (summer), the second location in 1994 (summer) and the third location in 1979 (summer). In all cases a depth of 15-19 m was reached, not the surface. These locations are all along the southernmost west coast of Vancouver Island, where the influence of the inner shelf buoyancy current (Vancouver Island Coastal Current, VICC) and its source, the Juan de Fuca Strait, are strong [Thomson *et al.*, 1989]. Errors in my method are expected to be the greatest in the surface of this region (Chapter 2, section 2.5.4, Figure 2.11) where DIC tends to be unusually high [Ianson *et al*. 2003]. However, the few available data suggest that our relationship tends to over-predict omega above 30 m in this region; thus, it is possible that the true omega was even lower than we have predicted.

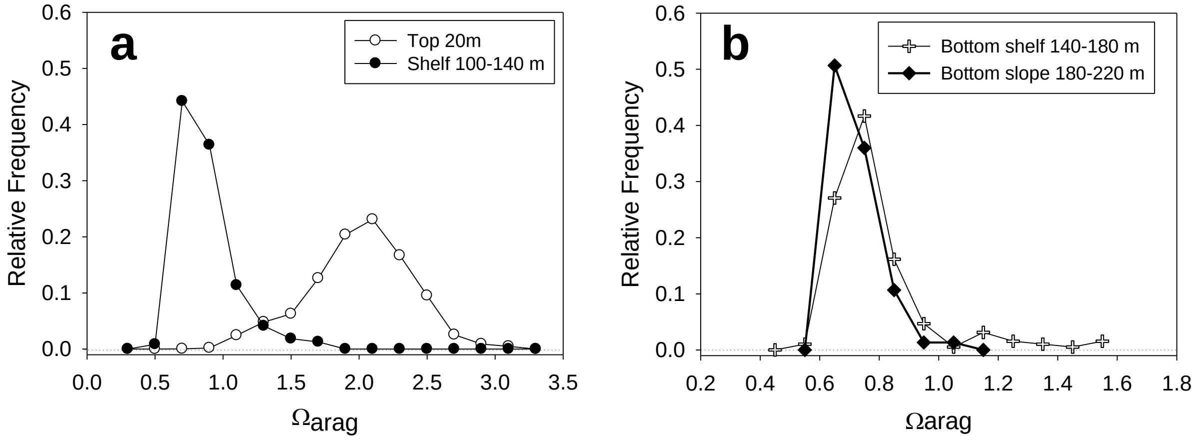

Figure 3.6 Relative frequency of historical record (number of observations/ total observations) for a) Ω_{arag} (0.2-bin) within the upper 20m (N=3044; white circle) and typical shelf depth 100-140 m (N=704; black circles); b) Bottom shelf Ω_{arag} (0.1 bin) depth 140 to 180 m (N=192; black diamonds) and c) Bottom slope Ω_{arag} (0.1 bin), 180-220 depth (N=75; crosses). The horizontal dotted lines show the zero lines.

The surface (0 – 20 m) Ω_{arag} distribution covered a broad range of values, as expected given high productivity and biological uptake at times. Thus, it was most frequently supersaturated and often well above Ω_{arag} =1 (Figure 3.6a). Deeper in the water column, at the depth of much of the shelf and also within the expected depths of upwelling (100 – 140 m) the Ω_{arag} distribution was substantially narrower and most often near, but just below saturation (Figure 3.6a). Similarly, in the shelf-break region (180 – 220 m depth), where most commercially fished groundfish are found [Sinclair *et al.*, 2005; Juan-Jorda *et al*, 2009], Ω_{arag} distributions were relatively narrow and showed that these waters were almost always undersaturated (Figure 6b). Just above the shelf-break (140-180 m) waters were only occasionally supersaturated with respect to Ω_{arag} (Figure 3.6b). A recent simulation study in central California showed similar results [Gruber *et al.*, 2012]. They estimated that, in 2005, ~70% of near bottom region was occupied by undersaturated conditions in in their study area (with water depths ranging between 50 and 120 m).

On the shelf, seasons played an important role in the location of Ω_{arag}=1 saturation horizons. As previously discussed (section 3.3.2), results from this additional separation of data (Figure 3.5c) also showed a strong tendency of shallow undersaturation (most often between 20-80 m) during summer-fall. Winter distributions were similar to spring on the shelf; however, undersaturated waters during winter tended to occur at deeper (~ 20 m) depths than in spring. Because surface $p\mathrm{CO}_2$ is drawn down by phytoplankton during spring [e.g. Evans *et al*. 2012] (and thus, surface Ω_{arag} increases) (see section 3.3.2), no shallow undersaturation was expected during this season. Although not common, undersaturation did occur at shallow depths such as 20-60 m during spring which is likely the result of early transitions from spring to summer upwelling. I used the average transition time, but variability in timing is large (~28 d^{-1}; Appendix E from Bylhouwer *et al.*, submitted) in the timing of this transition.

There were substantial differences between winter-spring and summer-fall in shelf waters (Figure 3.5c). However, differences between the seasons were not as obvious in offshore waters (Figure 3.5d). Off the shelf, distributions of the Ω_{arag}=1 horizon were similar regardless of time of year. The saturation horizon usually occurred between 80-140 m depth.

Vertical gradients in Ω_{arag} were often substantially larger over the shelf compared to offshore, particularly at the base of the upper mixed layer (by ~ 1 order of magnitude, Appendix F), during the summer-fall season. These gradients emphasize the variability in conditions over the shelf. Where vertical gradients are strong in the vicinity of the Ω_{arag} saturation horizon, even relatively small vertical fluctuations in isopycnals may cause an appreciable change in the depth of the horizon.

At deeper depths (100-250 m), a region of high interest in fisheries due to the considerable biomass of groundfish [Sinclair *et al.*, 2005; Juan-Jorda *et al.*, 2009] and the region in which the saturation horizon is most often found (Figures 3.5), gradients were weaker in most cases (Appendix F) and conditions were generally more stable (e.g. Figure 3.6b). Similar results were observed in southern California [Alin *et al.*, 2012], where seasonal cycle and interannual variability were pronounced in the upper water column and more stable below.

3.3.4 Regional differences

In the northernmost portion of the study region, Queen Charlotte Sound, the continental shelf is wide (roughly 140 km) and partially enclosed by Vancouver Island to the south and Haida Gwaii to the north (Figure 3.1), resulting in significantly different circulation patterns than along the west coast of Vancouver Island (WCVI). In addition, upwelling favourable winds are less prevalent than along the WCVI [Foreman *et al.*, 2011]. Surface currents in the QCS during summer are weak and may be influenced by eddies [Crawford *et al.*, 1985], while the WCVI is part of the California Current System, with a strong equatorward current flowing along the shelf-break during summer [Freeland *et al.*, 1984].

In the WCVI region, summer winds undergo episodic direction reversals, so upwelling favourable winds are not steady [Ianson *et al.*, 2003]. On the other hand, upwelling in QCS is weak or absent; instead the summer is characterized by a relaxation from winter downwelling whereas upwelling along WCVI is moderate [Ware and Thomson, 2005]. Thus, regional on-shelf variations in Ω_{arag} are expected.

Distributions of Ω_{arag} from the two regions were similar, particularly in offshore waters (Figure 3.7b). Shelf distributions in Ω_{arag} were less similar, i.e. in the more southern WCVI offshore waters lower Ω_{arag} values (near saturation, 0.8 – 1.2, Figure 3.7a) were more common than they are in QCS. Sampling depth did not appear to introduce any bias in the regional comparison (Figures 3.7c-e) however, there are fewer samples in QCS (~10% that off WCVI). Similarly, the proportion of samples from each season was similar in the two regions (e.g. 6.3% and 9.7% of the samples correspond to winter QCS and WCVI, respectively; 39% and 35% to summer; 31% and 27% to summer and 23% and 29% to fall) so bias due to season is not expected.

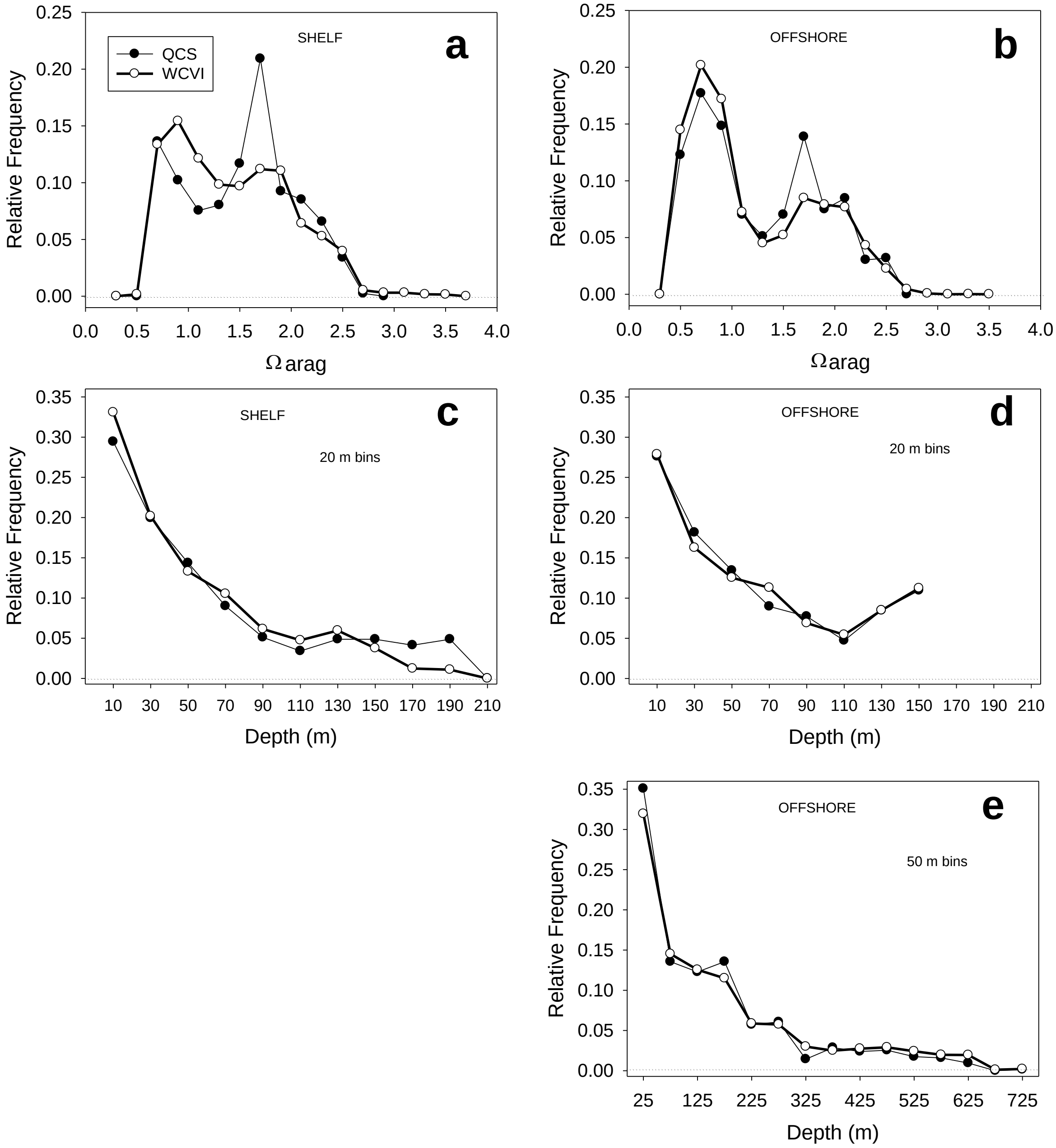

Figure 3.7 Relative frequency (number observations/ total observations) of Ω_{arag} for a) shelf WCVI (N=4788) and QCS (N=411) (bin 0.2); b) offshore WCVI (N=5566) and QCS (N=627) (bin 0.2), c) shelf (20 m bins). d) Relative frequency of sampling depth in offshore waters (20m bins) from 0 to 150 m and e) in offshore waters (50 m bins) from 0 to 750 m. The horizontal dotted lines show the zero line.

Over the shelf, the mean saturation depth is shallower on the WCVI (~70 m) than in QCS (~90 m) (Figure 8b) where undersaturation was present (70% of the WCVI shelf profiles and in only in 60% of the QCS shelf profiles, Table 3.1). In offshore waters, the mean saturation depth was also ~20 m shallower in the WCVI with respect to QCS (~110 m in WCVI and ~130 m in QCS, Figure 8c).

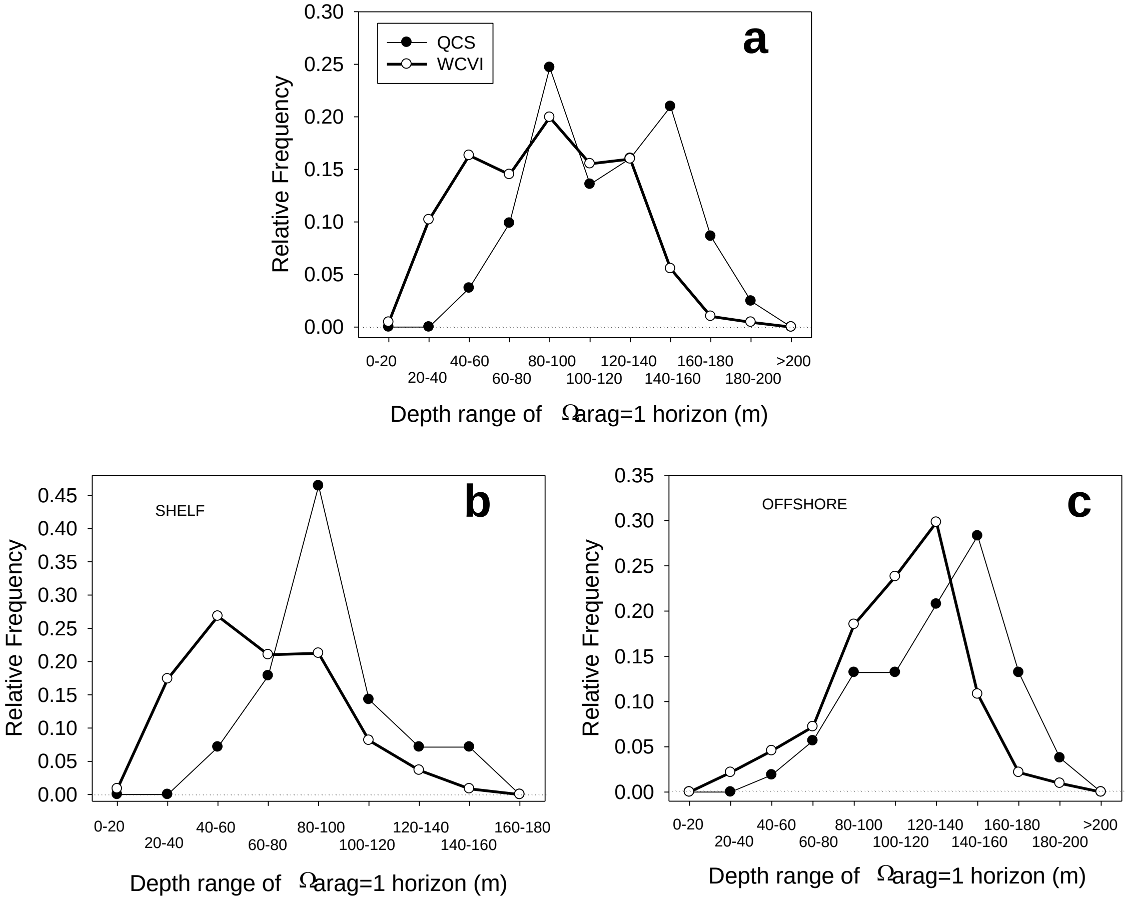

Figure 3.8 a) Relative frequency of a) depth of the Ω_{arag}=1 horizon (20 m bins) in WCVI (N=882, white circle) and QCS (N=81, black circle); b) over the shelf (shallower than 200 m; 20 m bins) for QCS (black circle) and WCVI (white circle) (shallower than 200m; N=28 and N=466, respectively) and c) offshore (deeper than 200 m; 20 m bins) QCS (black circle) and WCVI (white circle) (N= 53 and N=416 respectively) waters. The horizontal dotted lines show the zero lines.

Undersaturation with respect to Ω_{arag} is rare at shallow depths (shallower than 40-60 m) in QCS. The shallowest horizons occurred in the inside passage near the northern tip of the island (51ºN 127.8ºW) at the entrance to QCS, where tidal mixing is strong, causing nutrient and carbon supply to the surface [e.g. Nemçek *et al.*, 2008]. Because upwelling is stronger in WCVI than it is in QCS, it is likely that large concentrations of DIC are brought into the WCVI shelf surface during the upwelling season. Consequently, undersaturated waters with respect to Ω_{arag} were more common at shallower depth ranges in WCVI near the shelf-break (as shallow as 20-60 m) with respect to those in QCS (Figure 3.8) where conditions were in general more variable. These results suggested that WCVI may experience adverse effects of ocean acidification more acutely than QCS.

3.4 Conclusion

I have used a regression model (Chapter 2) in combination with historical T and O_2 observations to reconstruct seasonal cycles of Ω_{arag} and to investigate the local and regional patterns in Ω_{arag} distribution. This work has allowed me to generate a large number of empirical predictions, close to 20 times more data than the number of real measurements, which are spatially and temporally sparse owing to the difficulty and high cost involving in making them and a previous lack of interest.

Relative distributions of the Ω_{arag}=1 horizon showed that winter and spring appear to occupy essentially the same ranges and had similar probability distributions while summer and fall occupied similar ranges and have similar probability distributions (Figure 3.5). The Ω_{arag} saturation horizon over the continental shelf tended to be more stable in winter and spring and more variable and shallow (as shallow as 10-20 m in a few instances) in summer and fall during and following the upwelling season. Although it is unclear how organisms in our study region are affected by these conditions, it is known that the overall calcification rates for many species of marine calcifiers including corals, coccolithophores, foraminifera and pteropods are impacted by undersaturation of aragonite [Fabry *et al.*, 2008; Doney *et al.*, 2009]. Equivalent levels of dissolution of intact samples of pteropods from the Southern Ocean under Ω_{arag} saturation levels of 0.94-1.12 have also been observed in laboratory incubations [Bednaršek *et al.*, 2012] however cold water corals appear to be more susceptible and experience decreased calcification rates at higher values of Ω_{arag} [Bednaršek *et al.*, 2012]. It is also likely that increased carbon dioxide concentrations cause stress for higher trophic organisms such as finfish [e.g. Brewer and Peltzer, 2009].

The saturation horizon (Ω_{arag}=1) generally tended to occur at deeper depths over the QCS shelf and be more stable with respect to WCVI. These results suggest that shelf along the WCVI, a highly productive region with a major fishery, may experience adverse effects of ocean acidification more acutely than QCS.

Bibliography

Alin S. R., R. A. Feely, A. G. Dickson, J. M. Hernández-Ayón, L. W. Juranek, M. D. Ohman, and Ralf Goericke (2012). Robust empirical relationships for estimating the carbonate system in the southern California Current System and application to CalCOFI hydrographic cruise data (2005–2011). Journal of Geophysical Research, 117, C05033, doi:10.1029/2011JC007511

Barnes, C., Best M. and Zielinski A. (2008). The NEPTUNE Canada Regional Cabled Ocean Observatory. Sea Technology.

Beamish R. J., H. H. Harvey (1972). Acidification of the La Cloche Mountain Lakes, Ontario, and Resulting Fish Mortalities, Journal of the Fisheries Research Board of Canada, 29:(8) 1131-1143, doi: 10.1139/f72-169.

Bednaršek, N., G. A. Tarling, D. C. E. Bakker, S. Fielding, E. M. Jones, H. J. Venables, P. Ward, A. Kuzirian, B. Lézé, R. A. Feely & E. J. Murphy (2012). Extensive dissolution of live pteropods in the Southern Ocean. Nature Geoscience 5, 881-885 pp. doi:10.1038/ngeo1635

Brewer P.G and E. T. Peltzer (2009). Limits to Marine Life, *Science*, 324:347–48

Borges, A. V., Delille, B., and Frankignoulle, M. (2005). Budgeting sinks and sources of CO_2 in the coastal ocean: diversity of ecosystems counts, Geophysical Research. Letters, 32, L14601, doi:10.1029/2005GL023053.

Borges, A.V., N. Gypens. (2010). Carbonate chemistry in the coastal zone responds more strongly to eutrophication than ocean acidification. Limnology and Oceanography, 346-353, doi: 10.4319/lo.2010.55.1.0346.

Bylhouwer, B., Ianson, D. and Kohfeld K., Submitted for publication Investigating alongshore wind trends in the Pacific Northwest, Journal of Geophysical Research-Oceans.

Cai, W.-J., Hu, X, Huang, W.-J, Murrell M.C., Lehrter J.C., Lohrenz S.E., Chou W.-C *et al.* (2011), Acidification of subsurface coastal waters enhanced by eutrophication, Nat. Geosci., 4, 766‑770, doi:10.1038/ngeo1297.

Caldeira, K., and Wickett, M.E (2005). Ocean model predictions of chemistry changes from carbon dioxide emissions to the atmosphere and ocean. Journal of Geophysical Research, 110: C09S04. doi: 10.29JC002671.

Carpenter, J.H. (1965). The Chesapeake Bay Institute Technique for the Winkler Dissolved Oxygen Method, Limnology & Oceanography, 10: 141-143.

Crawford W. R., W.S. Huggett, M.J. Woodward. (1985). Summer circulation of the waters in Queen Charlotte Sound, Atmosphere-Ocean, 23: 393-413.

Crawford, W. and Dewey, R. (1989). 'Turbulence and mixing: Sources of nutrients on the Vancouver Island continental shelf', Atmosphere-Oceans, 27(2), 428–442.

Cullen J. T., M. Chong, D. Ianson (2009). British Columbian continental shelf as a source of dissolved iron to the subarctic northeast Pacific Ocean, Global Biogeochemical Cycles 23 (4), 1-12, doi: 10.1029/2008GB003326.

Cooley S.R. and S. C. Doney (2009). Anticipating ocean acidification's economic consequences for commercial fisheries. Environmental Research Letters, 4(2), doi:10.1088/1748-9326/4/2/024007

Dickson, A. G., and F. J. Millero (1987). A comparison of the equilibrium constants for the dissociation of carbonic acid in seawater media, Deep Sea Research, Part I, 34, 1733–1743, doi:10.1016/0198-0149(87)90021-5.

Dickson, A. G., and Goyet, C. (Eds.), DOE, U.S. Depertment of Energy (1994). Handbook of methods for the analysis of the various parameters of the carbon dioxide system in seawater: Version 2.0, ORNL CDIAC-74.

Dickson, A. G., C. L. Sabine, and J. R. Christian (2007). Guide to best practices for ocean CO_2 measurements, North Pac. Mar. Sci. Organ., Sidney, B. C., Canada, PICES Spec. Publ. 3, 191 pp.

Doney, S.C., V.J. Fabry, R.A. Feely and J.A. Kleypas (2009). Ocean acidification: the other CO_2 problem. Annu. Rev. Mar. Sci, 1, 169-192, doi: 10.1146/annurev.marine.010908.163834.

Easterling D.R. and Wehrner, M.F. (2009). Is the climate warming or cooling?. Geophysical Research Letters, 36. L08706, doi:10.1029/2009GL037810

Evans W., B. Hales, P. G. Strutton and D. Ianson (2012). Sea-air CO_2 fluxes in the western Canadian coastal ocean, Progress in Oceanography,101, 78-91pp. doi: 10.1016/j.pocean.2012.01.003.

Fabry V.J., B.A. Seibel, R.A. Feely, and J.C. Orr (2008). Impacts of ocean acidification on marine fauna and ecosystem processes. ICES Journal of Marine Science, 65, 414-432, doi: 10.1093/icesjms/fsn048.

Feely R.A., Boutinb J., Coscaa C. E., Dandonneaub Y., Etchetob J., Inouec H.Y, Ishiic M., Le Quéréd C., Mackeye D. J., McPhadena M., Metzlf N., Poissonf A., Wanninkhof R. (2002). Seasonal and interannual variability of CO_2 in the equatorial Pacific, Deep Sea Research, part II, 49, 2443-2469.

Feely R.A., C.L. Sabine, K. Lee, W. Berelson, J. Kleypas, V.J. Fabry and F.J. Millero (2004). Impact of anthropogenic CO_2 on the $CaCO_3$ system in the oceans. Science, 305, 362-366, doi: 10.1126/science.1097329.

Feely R.A., Sabine C.L., Hernández-Ayón J.M., Ianson D. and Hales B. (2008). Evidence for upwelling of corrosive “acidified” water onto the continental shelf, Science, 320, 1490-1492, doi: 10.1126.science.1155676.

Feely R. A., C.L. Sabine, R. H. Byrne, F. J. Millero, A. G. Dickson, R. Wanninkhof, A. Murata, L.A. Miller and D. Greeley (2012). Decadal changes in the aragonite and calcite saturation state of the Pacific Ocean, Global Biogeochemical Cycles, 26, 15pp, doi:10.1029/2011GB004157

Foreman M. G. G., B. Pal and W. J. Merryfield (2011). Trends in upwelling and downwelling winds along the British Columbia shelf, Journal of Geophysical Research, 116, C10023, 11 PP., 2011, doi:10.1029/2011JC006995

Freeland, H., W. Crawford, and R. Thomson (1984). Currents along the Pacific coast of Canada, Atmosphere- Oceans, 22, 151–172.

Gruber, N., C. Hauri, Z. Lachkar, D. Loher, T.L. Frölicher, Plattner G.-K. (2012). Rapid Progression of Ocean Acidification in the California Current System, Science, 337, 220-223, doi: 10.1126/science.1216773

Hales, B., Takahashi T. and Bandstra L.(2005). Atmospheric CO_2 uptake by a coastal upwelling system, Global Biogeochemical Cycles 19, 11pp, doi:10.1029/2004GB002295.

Hedges, J.I., Baldock, J.A., Gélinas, Y., Lee, C., Peterson, M.L., Wakeham, S.G., (2002). The biochemical and elemental compositions of marine plankton: a NMR perspective. Marine Chemistry 78, 47 – 63.

Hickey, B. M. (1979). The California Current system-Hypotheses and facts, Prog. Oceanogr., 8, 191–279, doi:10.1016/0079-6611(79)90002-8.

Hill, A. E., B. M. Hickey, F. A. Shillington, P. T. Strub, K. H. Brink, E. D. Barton, and A. C. Thomas (1998). Eastern ocean boundaries coastal segment (E), in The Sea, vol. 11, The Global Coastal Ocean: Regional Studies and Syntheses, edited by A. R. Robinson and K. H. Brink, 416 pp. 29–67, John Wiley, New York.

Hoegh-Guldberg, O., P. J. Mumby, A. J. Hooten, R. S. Steneck, P. Greenfield, E. Gomez, C. D. Harvell, P. F. Sale, A. J. Edwards, K. Caldeira, N. Knowlton, C. M. Eakin, R. Iglesias-Prieto, N. Muthiga, R. H. Bradbury, A. Dubi, M. E. Hatziolos (2007). Coral reefs under rapid climate change and ocean acidification. Science, 318, 1737-1742, doi: 10.1126/science.1152509.

Ianson, D., S. E. Allen, S. L. Harris, K. J. Orians, D. E. Varela, and C. S. Wong (2003). The inorganic carbon system in the coastal upwelling region west of Vancouver Island, Canada, Deep Sea Research. Part I, 50, 1023–1042, doi:10.1016/S0967-0637(03)00114-6.

Ianson, D., R.A. Feely, C.L. Sabine, and L.W. Juranek (2009). Features of coastal upwelling regions that determine net air-sea CO_2 flux, Journal of Oceanography., 65(5), 677–687, doi: 10.1007/s10872-009-0059-z.

IPCC, Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, S. Solomon *et al*., Eds. (Cambridge University Press, Cambridge, UK, and New York, 2007).

Juan-Jordá, M. J., J. A. Barth, M. E. Clarke & W. W. Wakefield, (2009). Groundfish species associations with distinct oceanographic habitats in the northern California Current. Fisheries Oceanography 18(1): 1–19.

Juranek L.W., R.A. Feely, W.T. Peterson, S.R. Aline, B. Hales, K. Lee, C.L. Sabine and J. Peterson (2009). A novel method for determination of aragonite saturation state on the continental shelf of central Oregon using multi-parameter relationships with hydrographic data, Geophysical Research Letters, 36, L24601, doi: 10.1029/2009GL040778.

Juranek L.W., R.A. Feely, D. Gilbert, H. Freeland, and L. Miller (2011). Real-time estimation of pH and aragonite saturation state from Argo profiling floats: Prospects for an autonomous carbon observing strategy, Geophysical Research Letters, 38, L17603, doi:10.1029/2011GL048580.

Kim, T.-W., K. Lee, R. A. Feely, C. L. Sabine, C.-T. A. Chen, H. J. Jeong, and K. Y. Kim (2010). Prediction of Sea of Japan (East Sea) acidification over the past 40 years using a multiparamter regression model, Global Biogeochemical Cycles, 24, GB3005, doi:10.1029/2009GB003637.

Kleypas, J. A., R.A. Feely, V.J. Fabry, C. Langdon, C.L. Sabine and L.L. Robbins (2006). Impacts of ocean acidification on coral reefsand other marine calcifiers: A guide to future research, 88 pp., Univ. Corp. of Atmos. Res., Boulder, Colo., (Available at http://www.ucar.edu/communications/Final_acidification.pdf)

Körtzinger, A., J.I. Hedges, and P.D. Quay (2001). Redfield ratios revisited: Removing the biasing effect of anthropogenic CO_2. Limnology and Oceanography. 46, 964–970.

Lee, K., F. J. Millero, R. H. Byrne, R. A. Feely, and R. Wanninkhof (2000). The recommended dissociation constants of carbonic acid in seawater, Geophysical Research Letters, 27, 229–232, doi:10.1029/1999GL002345.

Lee K., L. T. Tong, F. J. Millero, C. L. Sabine, A. G. Dickson, C. Goyet, G.-H Park, R. Wanninkhof, R. A. Feely, and R. M. Key (2006). Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans, Geophysical Research Letters, 33, L19605, doi: 10.1029/2006GL027207

Levine, N. M., Doney, S. C., Lima, I., Wanninkhof, R., Bates, N. R. and Feely R.A. (2011). The impact of the North Atlantic Oscillation on the uptake and accumulation of anthropogenic CO_2 by North Atlantic Ocean mode waters. Global Biogeochemical Cycles. 25
Li, Y. H., T. Takahashi, and W. S. Broecker (1969). Degree of saturation of $CaCO_3$ in the oceans, Journal of Geophysical Research., 74, 5507– 5525, 1969.

Mackas, D. L., K. L. Denman, and A. F. Bennett (1987). Least squares multiple tracer analysis of water mass composition, Journal of Geophysical Research, 92(C3), 2907-2918, doi: 10.1029/JC092iC03p02907

Mackas, D. (1992). Seasonal cycle of zooplankton off southwestern British Columbia: 1979–89, Canadian Journal of Fisheries and Aquatic Sciences, 49, 903–921, doi: 10.1139/f92-101.

Magnuson J. J., J. P. Baker, E. J. Rahel and J. R. Kramer (1984). A Critical Assessment of Effects of Acidification on Fisheries in North America [and Discussion], Phil. Trans. R. Soc. Lond. B, vol. 305 no. 1124 501-516.

Mehrbach, C., Culberson, C.H., Hawley, J.E., Pytkowicz, R.M. (1973). Measurement of the apparent dissociation constants of carbonic acid in seawater at atmospheric pressure, Limnology and Oceanography, 18, 897–907, doi:10.4319/lo.1973.18.6.0897.

Min D.-H., J. L. Bullister, and R. F. Weiss (2002), Anomalous chlorofluorocarbons in the Southern California Borderland Basins, Geophysical Research Letters, 29(20), 1955, doi:10.1029/2002GL015408.

Mucci, A. (1983). The solubility of calcite and aragonite in seawater at various salinities, temperatures and 1 atmospheric total pressure, Am. J. Sci., 238, 780–799

Nam, S., H.-J. Kim, and U. Send (2011). Amplification of hypoxic and acidic events by La Niña conditions on the continental shelf off California, Geophysical Research Letters, 38, L22602, doi:10.1029/2011GL049549.

Nemçek, N., D. Ianson and P. D. Tortell (2008). A high-resolution survey of DMS, CO_2, and O_2/Ar distributions in productive coastal waters, Global Biogeochem. Cycles, 22, GB2009, doi:10.1029/2006GB002879

Ohde S. and Mozaffar-Hossain M. (2004). Effect of $CaCO_3$ (aragonite) saturation state of seawater on calcification of Porites coral, Geochemical Journal, 38, pp. 613-621

Orr, J. C., Fabry, V. J., Aumont, O., Bopp, L., Doney,S.C., Feely,R.A., Gnanadesikan, A., Gruber, N., *et al*. (2005). Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. Nature 437.7059 (2005): 681-86, doi: 10.1038/nature04095

Picard, R. R., and R. D. Cook (1984). Cross-validation of regression models, J. Am. Stat. Assoc., 79, 575–583, doi:10.2307/2288403

Raven, J., K. Caldeira, H. Elderfield, O. Hoegh-Guldberg; P. Liss, U. Riebesell, J. Shepherd, C. Turley, and A. Watson. (2005). Ocean acidification due to increasing atmospheric carbon dioxide. The Royal Society, London, UK.

Redfield, A.C.; B.H. Ketchum and F.A. Richards. (1963). The influence of organisms on the composition of seawater. In: The Sea, Vol. 2. N.M. Hill (Ed.), WileyInterscience, New York, pp. 26-77.

Sabine, C.L., R. A. Feely, R. M. Key, J. L. Bullister, F. J. Millero, K. Lee, T.-H. Peng, B. Tilbrook, T. Ono, and C. S. Wong (2002). Distribution of anthropogenic CO_2 in the Pacific Ocean, Global Biogeochemical Cycles, 16(4), 1083, doi:10.1029/2001GB001639

Sarmiento J. L., J. Dunne, A. Gnanadesikan, R. M. Key, K. Matsumoto, and R. Slater (2002). A new estimate of the $CaCO_3$ to organic carbon export ratio, Global Biogeochemical Cycles, 16(4), 1107, doi:10.1029/2002GB001919

Sinclair, A. F., K.W. Conway, W.R. Crawford (2005), Associations between bathymetric, geologic and oceanographic features and the distribution of the British Columbia bottom trawl fishery, unpublished

Smith, R. L. (1994). The physical processes of coastal ocean upwelling systems. p. 39–64. In Upwelling in the Ocean: Modern Processes and Ancient Records, ed. by C. P. Summerhayes, K.-C. Emeis, M. V. Angel, R. L. Smith andB. Zeitzschel, Wiley and Sons Ltd.

Thomson, R. E. (1981), Oceanography of the British Columbia coast. Canadian Special Publication Fisheries and Aquatic Sciences 56, 291 pp. [Available from Canadian Government Publishing Centre, Supply and Service Canada, Ottawa K1A OS9, Canada.]

Thomson, R. E., B. M. Hickey, and P. H. LeBlond (1989). The Vancouver Island Coastal Current: fisheries barrier and conduit, in effects of ocean variability on recruitment and an evaluation of parameters used in stock assessment models, vol. 108, edited by R. J. Beamish and G. A. McFarlane, pp. 265–296, Dept. of Fish. and Oceans, Ottawa, Ont., Canada.

Thomson, R.E., and M. Krassovski. (2010). The poleward reach of the California Undercurrent extension. Journal Geophysical Research, 115, C09027, 9PP, doi:10.1029/2010JC006280.

Thomson, R.E. and D.M. Ware (1996). A current velocity index of ocean variability, Journal Geophysical Research, 101, 14,310.

Unesco/ICES/SCOR/IAPSO joint panel on oceanographic tables and standards: Background papers and supporting data on the Practical Salinity Scale 1978, Unesco technical papers in marine science 37, 1981.

van Heuven, S., D. Pierrot, J.W.B. Rae, E. Lewis and D.W.R. Wallace. (2011). MATLAB program developed for CO_2 system calculations, Rep. ORNL/CDIAC-105b. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee. doi: 10.3334/CDIAC/otg.CO2SYS_MATLAB_v1.1

Ware, D. M and Thomson, R.E. (1986). La Perouse project first progress report 1985. DFO report. 25p.

Ware, D.M. and Thomson R. E. (1993). La Perouse Project Seventh Annual Report: 1992. Fisheries and Oceans Canada, 74 p.

Ware, D.M. & Thomson, R.E. (2005). Bottom-up ecosystem trophic dynamics determine fish production in the Northeast Pacific. Science, 308, 1280–1284, doi: 10.1126/science.1109049

Wei G., M.T. McCulloch, G. Montimer, W. Deng and L. Xie (2009). Evidence for ocean acidification in the Great Barrier Reef of Australia. Geochemica et Cosmochemica Acta, 73, 2332-2346, doi: 10.1016/j.gca.2009.02.009

Whitney F.A. and H.J. Freeland (1999).Variability in upper-ocean water properties in the NE Pacific Ocean. Deep-Sea Research Part II: Topical Studies in Oceanography, 46 (11-12), pp. 2351-2370.

Wilks, D. S. (1995) Statistical methods in the atmospheric sciences, second edition, International Geophysics Series, Vol 59, Academic Press, 464pp. ISBN-10: 0127519653.

Appendix

Appendix A.

Mean (M) and standard deviation (Sd) of T, S, pressure, O_2 and NO_3^- for each of the 5 individual data sets.

	T (ºC)		O_2 (µmol kg^{-1})		S		Pressure		NO_3^- (µM kg^{-1})	
	M	Sd	M	Sd	M	Sd	M	Sd	M	Sd
1998	12.3	1.72	253.22	44.96	31.81	0.49	14.75	16.8	7.61	7.88
2004	11.06	4.3	208.19	75.41	32.42	0.98	72.81	105.22	14.4	13.59
2007	8.01	1.32	252.17	81.24	32.54	0.81	71.28	96.43	14.8	12.59
2010	8.12	2.32	171.8	86.2	33.04	0.82	71.28	139.38	22.68	13.01
2010W	8.65	1.46	182.03	88.18	33.21	0.68	159.33	161.34	19.73	12.11

Appendix B Summary of coefficients for the 6 regression models built[a]. In all of these the predictors (except the $\ln[CO_3^{2-}]$) are the dimensionless standardized anomalies of the original dimensional variables (with means subtracted and divided by the standard deviation). Error (1 standard deviation of the residuals) and R^2 for each model is also shown

Experim. #	Predicted variable	Depth range (m)	Predictors	Regression coefficients (dimensionless)					Error[b]	R^2
				Intercept	T	O_2	AOU	x_{PCA}		
1	$\ln\underline{\Omega}_{arag}$	30-750	$T^{S\,\S}$	-7.91×10^{-2}	3.15×10^{-1}	--	--	--	0.22	0.67
2	$\ln\underline{\Omega}_{arag}$	0-750	$T^{S\,\dagger}$	2.14×10^{-1}	4.28×10^{-1}	--	--	--	0.45	0.85
3	$\underline{\Omega}_{arag}$	0-750	$T^{S\,\dagger}$, $O_2^{S\,\dagger}$	1.39×10^{-15}	5.3×10^{-1}	5.07×10^{-1}	--	--	0.18	0.93
4	$\ln\underline{\Omega}_{arag}$	0-750	$T^{S\,\dagger}$, $O_2^{S\,\dagger}$	2.14×10^{-1}	1.82×10^{-1}	3.40×10^{-1}	--	--	0.13	0.96
5	$\ln\underline{\Omega}_{arag}$	0-750	$T^{S\,\dagger}$,AOU^{S}	2.14×10^{-1}	1.26×10^{-1}	--	3.81×10^{-1}	--	0.13	0.96
6	$\ln\underline{[CO_3^{2-}]}$	0-750	T, O_2	3.16	5.94×10^{-1}	3.66×10^{-3}	--	--	0.13	0.96
7	$\ln\underline{\Omega}_{arag}$	30-750	$T^{S\,\S}$, $O_2^{S\,\S}$	-7.91×10^{-2}	6.97×10^{-2}	3.04×10^{-1}	--	--	0.062	0.97
8	$\ln\underline{\Omega}_{arag}$	0-750	x_{PCA}	2.14×10^{-1}	--	--	--	3.69×10^{-1}	0.18	0.93

[a] Model 4 was chosen as the optimal model in this study

$T^{S\,\S}$ = (T -7.05ºC)/1.38ºC

$T^{S\,\dagger}$ = (T -8.78ºC)/2.87ºC

$O_2^{S\,\dagger} = (O_2\ \mu mol\ kg^{-1} - 197.84\ \mu mol\ kg^{-1}) / 88.59\ \mu mol\ kg^{-1}$

$AOU^{S} = (AOU\ \mu mol\ kg^{-1} - 89.5\ \mu mol\ kg^{-1}) / 100.53\ \mu mol\ kg^{-1}$

$O_2^{S\,\S} = (O_2\ \mu mol\ kg^{-1} - 148.95\ \mu mol\ kg\text{-}1) / 76.67\ \mu mol\ kg^{-1}$

[b] +/-1 standard deviation of the residuals

Appendix C.

Five regression models were built with full subsets (one for each survey) minus one withheld subset (column 1) and then each regression model was used to make a prediction of the withheld data set and compared to the true values of each data set (section 2.5.1). Then, values of false negatives, false positives and R^2 were computed for ln-transformed model with two predictors (T and O_2) (Appendix C, experiment 4) with VIF applied and ln-transformed model with one PCA-derived predictor (Appendix C, experiment 8).

	T and O_2+ VIF			PCA
	False + (%)	False – (%)	R^2	R^2
1998	1	6.3	0.79	0.33
2004	7.7	0	0.84	0.57
2007	4	2	0.16	0.69
2010	0	1.4	0.55	0.93
2010-W	4.4	0	0.98	0.98
TOTAL	5.8	1.5	0.91	0.75

Appendix D

Station names, latitudes and longitudes of the sites included in the water masses CUC (which is located at depth ranges of ~125-325 m of the slope and shelf [Thomson and Krassovski, 2010]) and VICC (which occupies the whole water column inshore of the ~100m isobath [Thomson *et al.*, 1989]; except for C2 that is deeper than 100 m). VICC was divided in two parts: south[a] (i.e. the mouth of the Juan de Fuca, where the chemical signature was expected to be strong [Ianson *et al*. 2003]) and central (i.e. the central portion of the WCVI where the chemical signature was expected to have been erased by gas exchange, primary production and mixing) (section 2.5.4).The station names are those used during the surveys in which the samples were collected for reference. Some of the stations were visited on more than one survey, but only the first survey is mentioned.

CUC			VICC		
Stn	Lat N	Long E	Stn	Lat N	Long E
QCS-2 (2010)	52.48	-129.5	L2_16 (2007)	49.53	-126.7
MB6 (2004)	52.27	-129.05	L2_15 (2007)	49.48	-126.84
QCS-1 (2010)	52.25	-129.05	G1 (1998)	49.34	-126.59
SP4 (2004)	52.15	-129.5	G2 (2004)	49.31	-126.64
MB4 (2004)	51.92	-129.97	D1 (2010)	49	-125.73
MT2 (2004)	51.6	-130.93	C1 (1998/2010)[a]	48.84	-125.47
SS7 (2004)	51.47	-128.5	C2 (1998/2010)[a]	48.81	-125.51
GB4 (2004)	51.45	-129.78	A1 (1998/2010)[a]	48.5	-124.73
SS6 (2004)	51.35	-129			
L1_3 (2007)	51.31	-131.53			
SS5 (2004)	51.19	-129.34			
CS3 (2010)	50.9	-129.42			
CS2 (2010)	50.7	-128.76			
LJ4 (2010)	49.59	-127.28			
L2_13 (2007)	49.36	-127.1			
G4 (1998/2004)	49.2	-126.83			
G5 (2004)	49.12	-126.92			
C4 (2010)	48.7	-125.68			

[a] SouthernVICC stations (the remainder are central stations)

Appendix E

Average transition time for the onset of each season (Julian Day) and associated standard deviation (Sd) estimated from wind data reanalysed (1950-2010) [Bylhouwer *et al*. submitted]. The summer represents the upwelling season (SU) which begins earlier than the summer equinox and winter represents downwelling (WD). Both spring (ST) and fall (FT), on the other hand, represent the periods of transition between winter downwelling and summer upwelling.

Season	Onset (Julian day)	STD (d)	Calendar day of onset
ST	65	26	6 March
SU	148	28	29 May
FT	239	16	26 August
WD	311	19	7 November

Appendix F

Vertical gradients in Ω_{arag} calculated at stations C2 and C9 for each of the times they were visited (Figure 4). The maximum gradient within the upper layer (0 - 30 m) is shown for each profile. This maximum occurred most often between 20 and 30 m with several exceptions (footnote a, b and c) primarily as a result of sampling at different depths. Similarly, maximum gradients deeper in the water column occurred between 75 and 100 m over the shelf (no exceptions), and most often between 174 and 200 m offshore with several exceptions (footnote d) where sampling resolution was more coarse. Means are also shown. Where no data were available a "---" was entered.

		C2			C9	
		20-30 m	75-100 m		20-30 m	174-200 m
Season	Month/ Year	Gradient (m^{-1})	Gradient (m^{-1})	Month/ Year	Gradient (m^{-1})	Gradient (m^{-1})
WD	02/1987	$-1.1\text{X}10^{-2}$[a]	$-3.9\text{X}10^{-3}$	02/2002	----	$-4.6\text{X}10^{-3}$
	02/2003	$-1.4\text{X}10^{-2}$	$-2.6\text{X}10^{-3}$	02/2003	$-1.6\text{X}10^{-3}$	$-2.3\text{X}10^{-3}$
	01/2005	$-2.8\text{X}10^{-2}$	$-2.0\text{X}10^{-3}$	03/2004	$-2.8\text{X}10^{-3}$	$-3.9\text{X}10^{-3}$
	Mean	**$-1.8\text{X}10^{-2}$**	**$-2.8\text{X}10^{-3}$**	**mean**	**$-2.2\text{X}10^{-3}$**	**$-3.6\text{X}10^{-3}$**
ST	05/1998	$-1.1\text{X}10^{-2}$ [a]	$-1.8\text{X}10^{-3}$	04/1996	$-2.8\ \text{X}10^{-3}$ [c]	$-7.2\text{X}10^{-4}$
	05/2000	$-2.9\text{X}10^{-2}$	$-4.9\text{X}10^{-3}$	05/1998	$-1.5\ \text{X}10^{-2}$	$-2.5\text{X}10^{-3}$
	05/2004	$-5.1\text{X}10^{-2}$[a]	$-3.8\text{X}10^{-3}$	05/2000	$-3.4\ \text{X}10^{-3}$ [c]	$-1.6\text{X}10^{-3}$
	05/2005	$-1.6\text{X}10^{-2}$ [b]	----	05/2002	$-1.2\ \text{X}10^{-2}$	---
	05/2007	$-1.2\text{X}10^{-1}$[a]	$-2.0\text{X}10^{-3}$	03/2004	$-1.5\ \text{X}10^{-3}$	$-2.0\text{X}10^{-3}$
	04/2007	$-8.5\text{X}10^{-2}$[a]	$-3.1\text{X}10^{-3}$	05/2004	$-7.7\text{X}10^{-2}$	$-1.5\text{X}10^{-3}$
	05/2008	$-6.2\text{X}10^{-2}$[a]	----	05/2005	---	$-2.7\text{X}10^{-3}$ [d]
	Mean	**$-5.3\text{X}10^{-2}$**	**$-3.1\text{X}10^{-3}$**	05/2007	$-4.5\text{X}10^{-2}$	$-2.1\text{X}10^{-3}$
				05/2008	$-2.5\text{X}10^{-2}$	$-1.6\text{X}10^{-3}$
				mean	**$-2.3\text{X}10^{-2}$**	**$-1.8\text{X}10^{-3}$**
SU	07/1985	$-6.8\text{X}10^{-2}$ [a]	----	06/1986	$-2.2\text{X}10^{-2}$	$-2.6\text{X}10^{-3}$[d]
	07/1986	$-3.8\text{X}10^{-2}$[a]	$-4.7\text{X}10^{-3}$	07/1986	$-2.9\text{X}10^{-2}$[c]	$-3.3\text{X}10^{-3}$
	07/1987	$-2.6\text{X}10^{-2}$	$-6.3\text{X}10^{-3}$	07/1993	---	$-3.6\text{X}10^{-3}$ [d]
	07/1993	$-1.9\text{X}10^{-2}$	$-4.9\text{X}10^{-3}$	08/1994	---	$-2.6\text{X}10^{-3}$ [d]
	06/1996	$-8.6\text{X}10^{-2}$ [a]	$-5.5\text{X}10^{-3}$	06/1996	---	$-1.2\text{X}10^{-3}$ [d]
	07/1996	$-5.2\text{X}10^{-2}$ [a]	$-3.3\text{X}10^{-3}$	07/1996	$-4.9\text{X}10^{-2}$	----
	07/2004	$-2.0\text{X}10^{-1}$[a]	$-2.7\text{X}10^{-3}$	07/2009	$-2.2\ \text{X}10^{-2}$	$-2.3\text{X}10^{-3}$
	07/2009	$-2.9\text{X}10^{-2}$ [a]	$-1.9\text{X}10^{-3}$	**mean**	**$3.1\text{X}10^{-2}$**	**$-2.6\text{X}10^{-3}$**
	mean	**$-7.3\text{X}10^{-2}$**	**$-4.2\text{X}10^{-3}$**			
FT	11/1986	$-2.7\text{X}10^{-2}$[a]	$-6.2\text{X}10^{-3}$	11/1986	$-3.7\text{X}10^{-2}$	$-3.9\text{X}10^{-4}$[d]
	09/1987	$-2.9\text{X}10^{-2}$	$-1.6\text{X}10^{-3}$	10/1995	$-2.8\text{X}10^{-2}$	$-7.3\text{X}10^{-3}$
	09/1995	$-1.1\text{X}10^{-2}$	$-3.9\text{X}10^{-3}$	10/1996	---	$-7.5\text{X}10^{-4}$
	10/1996	$-1.2\text{X}10^{-2}$	$-7.7\text{X}10^{-3}$	10/1997	$2.6\text{X}10^{-2}$[a]	$-8.4\text{X}10^{-4}$[d]
	09/2002	$-6.1\text{X}10^{-3}$	$-6.3\text{X}10^{-3}$	09/2004	$-1.0\ \text{X}10^{-2}$ [c]	$-1.3\text{X}10^{-3}$
	09/2004	$-1.3\text{X}10^{-2}$	$-5.7\text{X}10^{-3}$	10/2007	$-8.8\text{X}10^{-2}$[a]	$-1.9\text{X}10^{-3}$
	09/2007	$-6.9\text{X}10^{-2}$[a]	$-3.8\text{X}10^{-3}$	09/2009	$-7.8\ \text{X}10^{-2}$[a]	$-3.4\text{X}10^{-4}$ [d]
	09/2009	$-1.5\text{X}10^{-2}$	$-3.2\text{X}10^{-3}$	**mean**	**$-4.5\text{X}10^{-2}$**	**$1.83\text{X}10^{-3}$**
	mean	**$-2.3\text{X}10^{-2}$**	**$-4.8\text{X}10^{-3}$**			

[a] 10-20 m [b] 5-30 m [c] 30-50 m [d] 150- 200 m