**TCO2 and pCO2 measurement quality control and uncertainty assessment**

The following describes the quality control (QC) procedure for measurements of total dissolved inorganic carbon (TCO2) and carbon dioxide partial pressure (pCO2) made on discrete seawater samples collected using Niskin bottles and analyzed using a Burke-o-Lator TCO2/pCO2 Analyzer1. The approach to evaluate measured and derived marine CO2 system parameter uncertainties is also described. The below steps extend from the evaluation of Excel worksheet calculations completed during sample analysis1 and routines described in provided code on GitHub (https://github.com/HakaiInstitute/Discrete-CO2-Processing-Code?tab=readme-ov-file). Note that TCO2 measurements would have been evaluated and deemed “good” or “questionable” during the evaluation of Excel worksheets by examining correction factors (acceptable values within 1±0.01) and the standard deviation of triplicate reference material analyses.

*Concatenation and initial quality checks using CTD data*

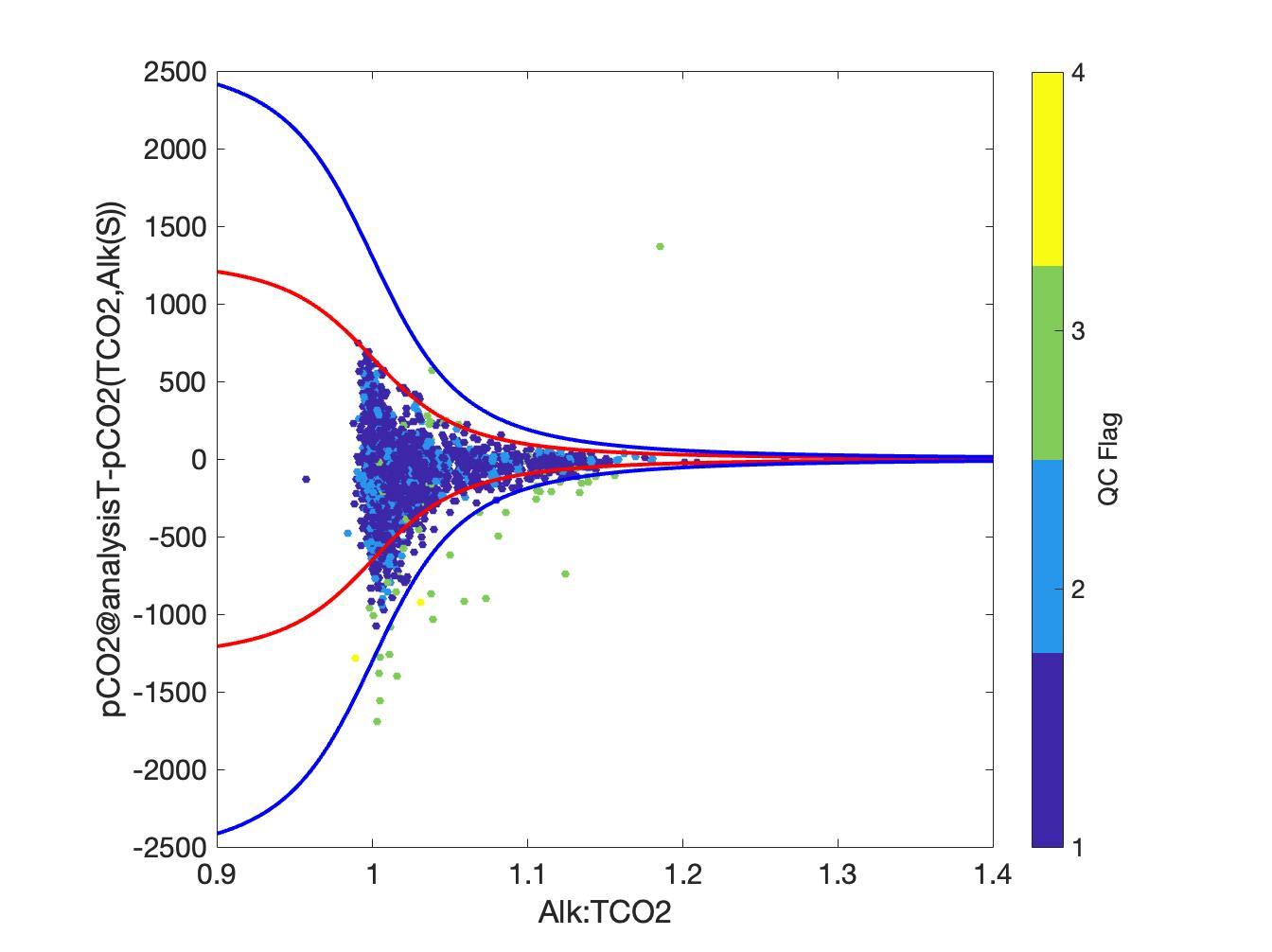
The first step in examining the quality of TCO2 and pCO2 measurements from an oceanographic field campaign was to produce a “master file” of concatenated data from the Excel worksheets generated during sample analysis and the corresponding sample logs completed during sample collection. Information generated during sample collection included collection date and time, latitude, longitude, depth, sample temperature at the time of collection, and any notes related to Niskin bottle integrity checks and sample collection. Each sample in the master file was initially given a data quality flag of “1” if considered a good measurement, or “2” if part of a replicate series (usually a triplicate set), or “4” if not analyzed. Further evaluation of data quality, and the adjustment of flags “1” or “2” to flag “3”, representing a questionable measurement, occurred in the following steps.

The next data-handling step involved merging conductivity-temperature-depth (CTD) profile data with the master file and then doing comparisons between select measurements. Note that available nutrient data were also merged with the master file, although these data were not used in CO2 system calculations. Data merging consisted of matching date, latitude, longitude, and depth between the discrete TCO2/pCO2 samples collected with Niskin bottles and 1-m bin-averaged CTD data. Note that CTD profiling and Niskin bottle sampling occurred sequentially during our small boat surveys. After CTD data were merged with the master file, the sample temperature at the time of collection (recorded using a NIST-traceable probe) and the sample salinity (measured in the laboratory with a YSI conductivity cell that was calibrated using certified reference materials from Andrew Dickson at UCSD) were compared with the CTD temperature and salinity recorded at the depth the Niskin bottle was tripped (determined using an RBR Solo pressure sensor attached to each Niskin bottle, or line out markers if no pressure measurement was available). All salinity units were reported on the Practical Salinity Scale - 1978. The comparison between NIST temperature and CTD temperature was useful for evaluating the degree of warming that may have occurred between recovering the Niskin bottle and drawing the sample; however, this warming was typically minimal. The comparison between sample and CTD salinity was examined for large differences that would indicate a sample collection issue, such as a sample drawn from the wrong Niskin bottle or a Niskin bottle tripped at the wrong depth. Samples that exceeded 8 times the root mean square error (RMSE) of the linear relationship between sample and CTD salinity (a typical RMSE value is 0.12) received a flag of “3”. This large exceedance threshold was used to account for the expected high variance between CTD and discrete sample salinity within the surface layer.

*Quality control of pCO2 measurements*

Following the comparison to CTD data, we utilized a regional alkalinity-salinity relationship2 to help identify questionable measurements of pCO2 recorded at the analysis temperature (pCO2@analysisT). Implicit here was the assumption that analytical errors in the TCO2 measurement will have been identified during the evaluation of Excel worksheet calculations1. Therefore, flag “1” TCO2 measurements could be combined with salinity-determined alkalinity to derive pCO2, and this derived pCO2 could then be compared with pCO2@analysisT to isolate questionable measurements. A second assumption here is that the regional alkalinity-salinity relationship was robustly constrained with minimal non-conservative character in alkalinity2. If a pCO2 measurement appeared questionable, a large difference manifested between pCO2@analysisT and the derived value; however, the extent of this difference was dependent on the buffering state of the sample (*i.e.*, the ratio between alkalinity and TCO2).

To aid in evaluating the difference between measured and derived pCO2 across a wide range of buffering states, we utilized the uncertainty of the regional-alkalinity salinity relationship (characterized by the root mean square error; RMSE) to create a hypothetical pCO2 threshold. We considered two cases: (1) seawater containing high TCO2 (nominally 2100 µmol kg-1 for oceanographic station QU39), and (2) seawater containing low TCO2 (nominally 1600 µmol kg-1). Constant temperature and salinity values of 18.7°C and 29.6 (median analysis temperature and sample salinity in the QU39 dataset), respectively, were used to create the hypothetical pCO2 thresholds for each case. Alkalinity was calculated for a range of alkalinity and TCO2 ratios (Alk:TCO2 from 0.9 to 1.4) for both cases. Then the RMSE of the regional alkalinity-salinity relationship (22.47 µmol kg-1) was either added to or subtracted from alkalinity to derive high or low pCO2 values across the range of Alk:TCO2 ratios for the high and low TCO2 cases. The difference between high and low pCO2 values reflects uncertainty in derived pCO2 across the range of Alk:TCO2 ratios, and was similar between the high and low TCO2 cases. The difference between high and low pCO2 values across the range of Alk:TCO2 ratios was then scaled up by a factor of 2 in order to account for divergences from the regional alkalinity-salinity relationship and the constant temperature and salinity values, and then these scaled values were used as the hypothetical pCO2 threshold to isolate questionable measurements (Figure 1).



**Figure 1:** The flagging of pCO2@analysisT that exceeded a hypothetical pCO2 threshold across a range of Alk:TCO2 ratios (blue lines). The red lines are the theoretical differences between pCO2(TCO2, Alk(S)+RMSE) and pCO2(TCO2, Alk(S)-RMSE) and the blue lines are those differences scaled up by 2. Data points are pCO2@analysisT minus pCO2(TCO2,Alk(S)), and values outside of the range of the blue lines were identified as questionable and given a flag of “3”.

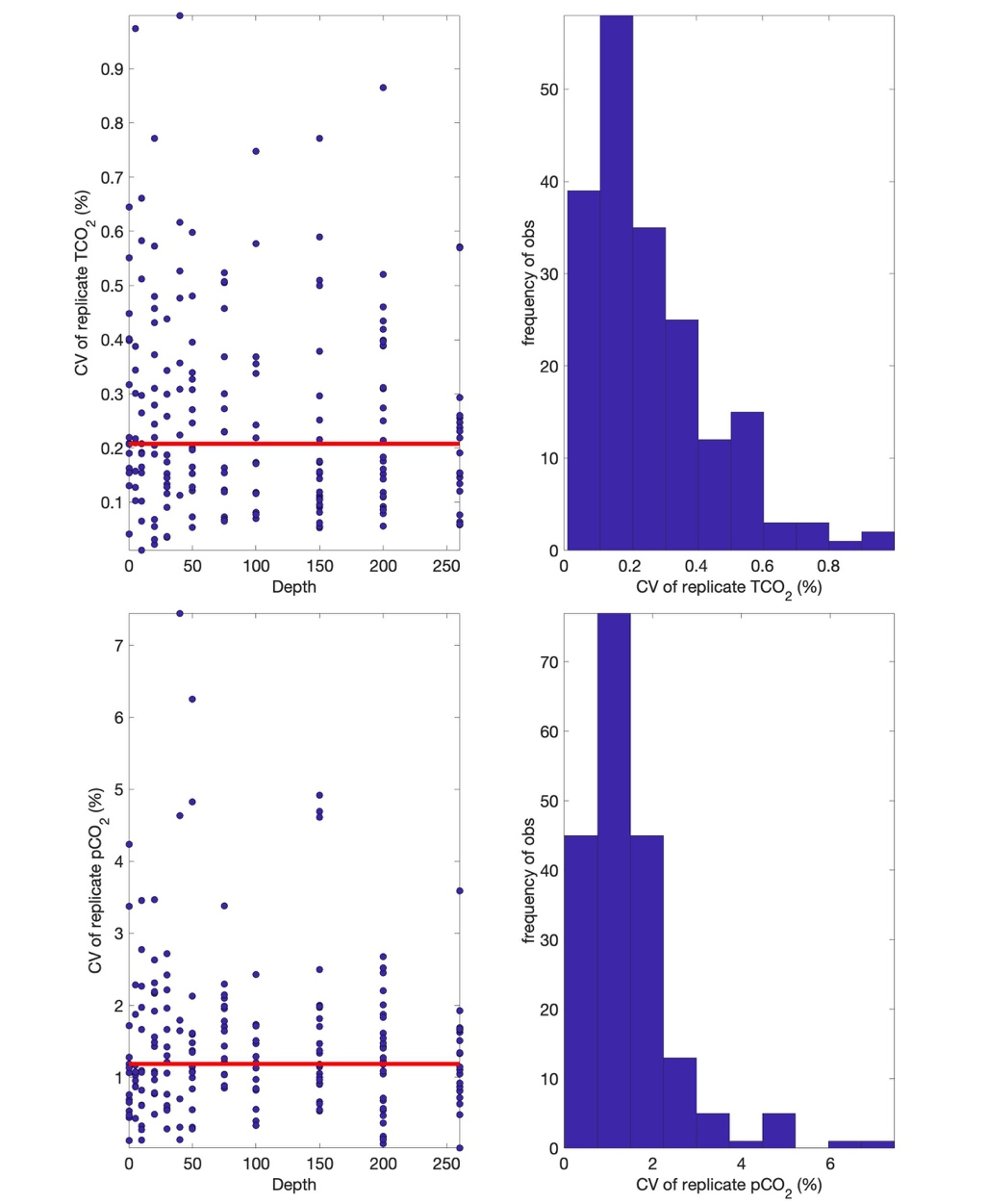
*Final quality control evaluation*

Final checks of the data quality involved plotting the derived alkalinity versus salinity and examining assorted property plots. In a region where the relationship between alkalinity and salinity is quasi-linear (sites or surveys that encompass deep samples from off the continental shelf may have two linear regions in an alkalinity-salinity plot; *i.e.*, a “hockey stick” curve) and robust, this check can be used to identify outliers that may represent questionable data not captured by the above steps. Property versus time plots (for pCO2 and alkalinity) for each sampling depth (*i.e.*, treating each sampling depth as a distinct time series), property versus depth plots (for pCO2, TCO2, and alkalinity), and property versus property plots (TCO2 versus oxygen, pCO2 versus oxygen, and TCO2 versus nitrite plus nitrate) were used to evaluate the dataset for missed outliers. Finally, triplicate measurements were statistically evaluated (below) with medians retained for the final dataset, and quality control flag values were adjusted to conform to conform to a WOCE-style flagging scheme3 (2 = acceptable, 3 = questionable, 6 = median of replicates, 9 = missing value).

*Evaluating TCO2 and pCO2 measurement uncertainty*

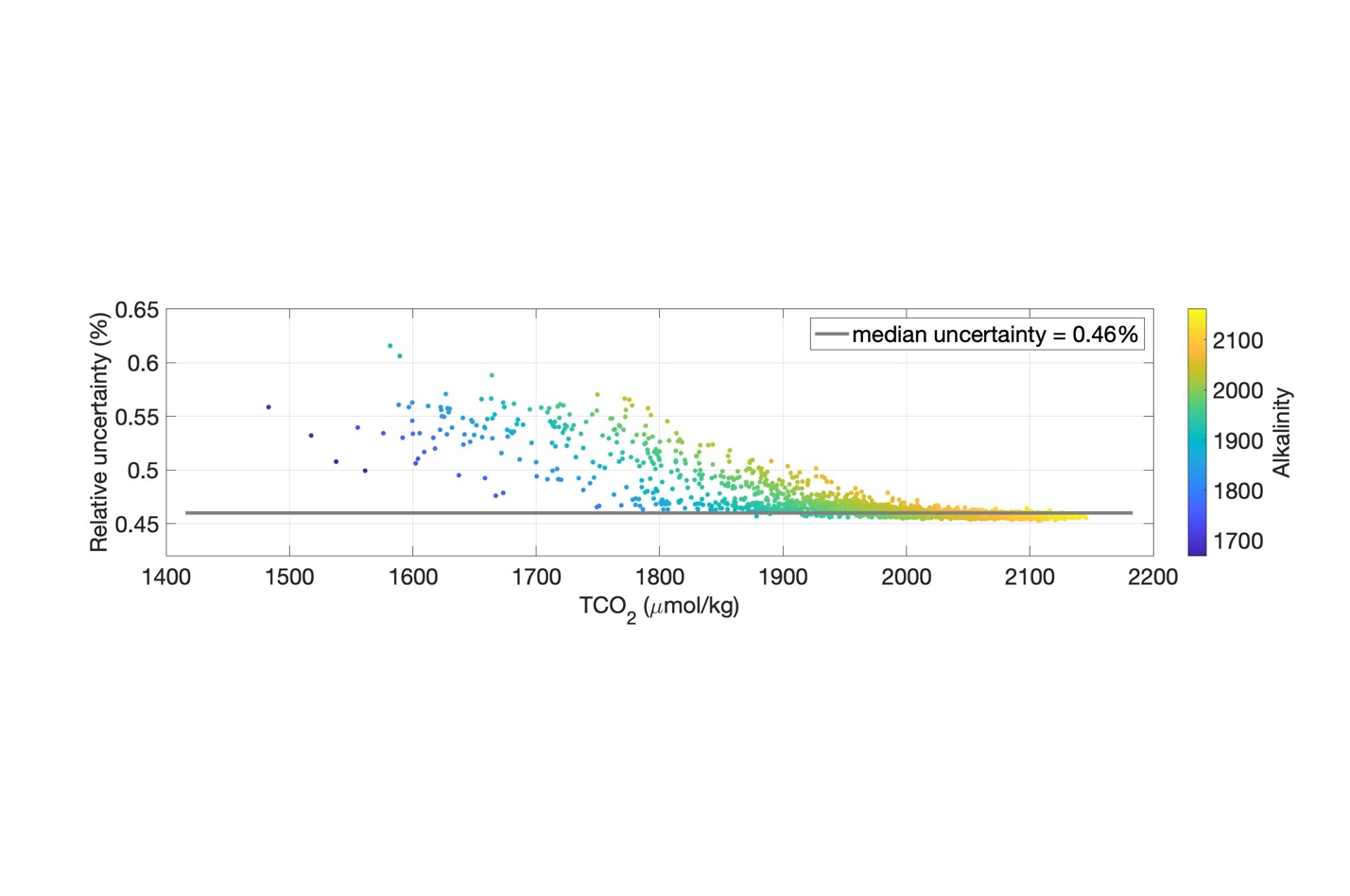
As mentioned above, TCO2 measurements were corrected using results from reference material analyses done in triplicate at the start and end of a series of seawater sample analyses consisting of approximately 20 samples1. Correction factors were applied to each seawater measurement by first dividing the reported TCO2 content of the reference material by the average from each triplicate analysis of certified reference material that bracketed the seawater sample analyses and then interpolating those values to the seawater measurement time points. Reference materials used were either certified reference material (CRM) from Andrew Dickson (UCSD) or internal reference material (IRM) consisting of mixtures of sodium carbonate and sodium bicarbonate in deionized water that had been analyzed alongside CRMs to determine their TCO2 content. The standard deviation of each triplicate analysis series of reference materials provided a metric for the analytical uncertainty of the TCO2 measurement. The median standard deviation of 582 triplicate analyses of CRMs was 6.2 µmol/kg with a median-centered coefficient of variation (CV) of 0.3%. This is identical to the relative uncertainty for TCO2 measurements reported by Evans et al2. The median standard deviation of 294 triplicate analyses of IRMs was 3.3 µmol/kg with a median-centered CV of 0.2%. Combining the median-centered CVs from CRM and IRM analyses in quadrature results in a value of 0.36%, which we take to be the relative uncertainty for our TCO2 analyses (*i.e.*, the analytical uncertainty).

We currently lack reference materials for seawater pCO2 analysis. The uncertainty of pCO2 analyses was previously evaluated by analyzing CRMs for both TCO2 and pCO2, then deriving alkalinity, and finally comparing those derived values to the certified CRM values2. The difference between derived and certified alkalinity was typically within 1.5% and taken to be the relative uncertainty in the measurement. However, this value includes uncertainties in both of the TCO2 and pCO2 measurements and the CO2 equilibrium constants. Additionally, the TCO2 analytical uncertainty reported above does not include uncertainty that can occur during sample collection. We therefor incorporate results from 198 triplicate sets of seawater samples collected randomly across sampling depths at oceanographic station QU39 to provide an additional assessment of TCO2 measurement uncertainty and a more robust estimate of pCO2 measurement uncertainty. Figure 2 shows CV values from the 198 triplicate sets of seawater samples with the median value for each record taken as a measure of the combined uncertainty from sampling and analysis. This value for pCO2 was 1.2%, slightly lower than from the value reported previously2. The value for TCO2 was 0.21% and was slightly lower than the assessment from reference material triplicates above. In order to produce a more conservative estimate of TCO2 measurement uncertainty, we combined this value with the median CV from the triplicate reference material assessment in quadrature to produce a measure of total (*i.e.*, analytical plus sampling) relative uncertainty; this value was estimated to be 0.42% for our discrete seawater TCO2 measurements and deemed to be “weather quality”4.



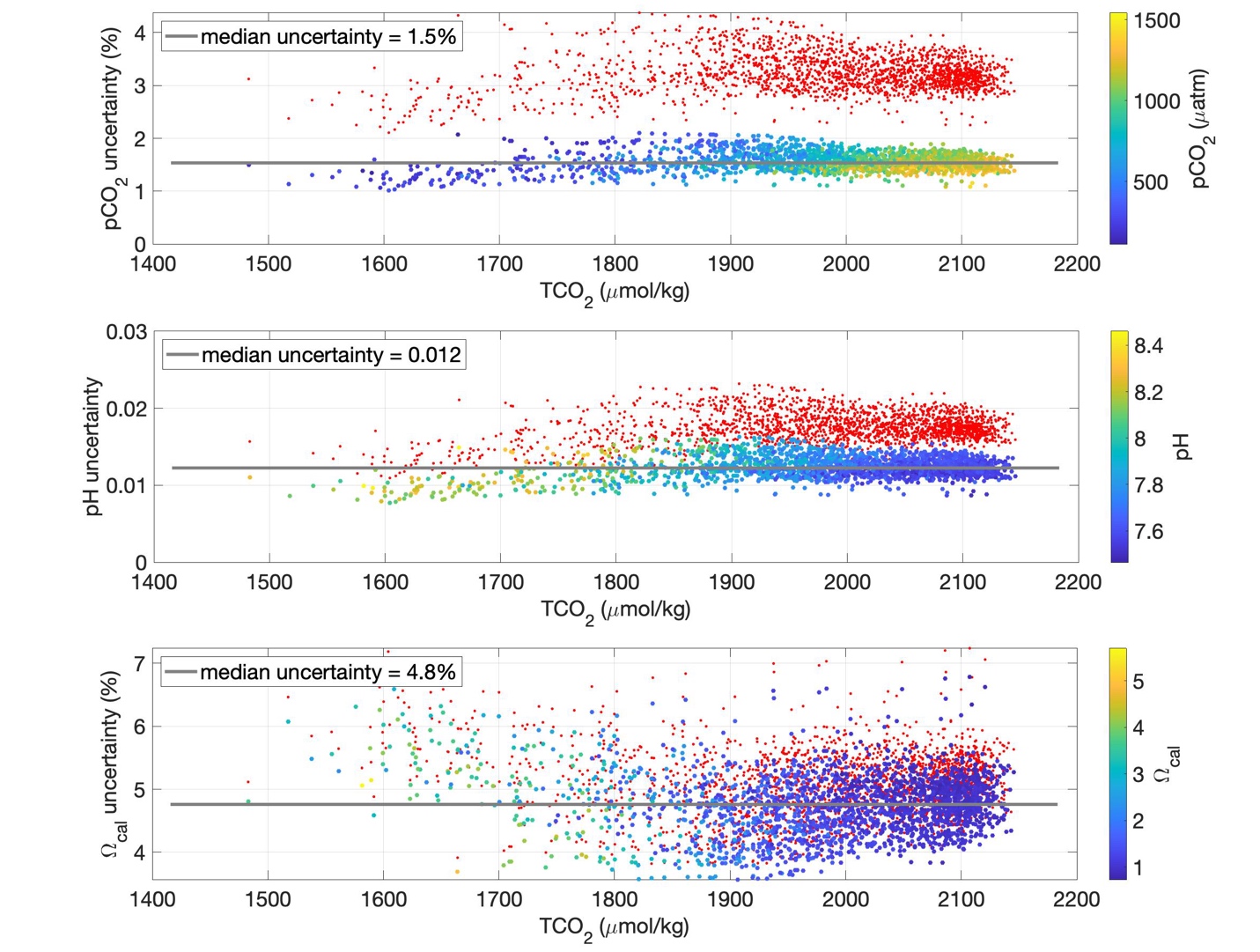
**Figure 2:** Median-centered coefficient of variation (CV) values from 198 triplicate sets of seawater TCO2 (top) and pCO2 (bottom) measurements. Left panel shows CV values versus sampling depth with the red horizontal line representing the median CV value in the record. The right panel shows the frequency distribution of CV values. The median CV value for each record is taken as a measure of uncertainty (0.21% for TCO2; 1.2% for pCO2).

*Derived marine CO2 parameter uncertainty*

The conversion to *in situ* pCO2 conditions (*i.e.*, at the temperature and pressure the sample was collected) and the calculation of the remaining marine CO2 system parameters, *e.g.*, alkalinity, pH reported here on the total scale (pHT), and aragonite and calcite saturation states, required a two-step process owing to the fact that the pCO2 measurement utilized equilibration with a headspace gas and this process slightly alters the TCO2 content of the sample. A headspace gas correction was applied to the TCO2 measurements5, which incorporated measurements of the headspace gas prior to equilibration (*i.e.*, an atmospheric pCO2 measurement, nominally near 400 µatm), the seawater volume of the sample (325 ml), and the headspace gas volume (90 ml). Headspace gas-adjusted TCO2 data were then used with the pCO2 measurements to derive the seawater alkalinity using the carbonic acid dissociation constants from Waters et al6, the bisulfate dissociation constant from Dickson et al7, the fluoride and hydrogen association constants from Perez and Fraga8, and the boron / chlorinity ratio from Uppström9. Alkalinity was considered here to consist of carbonate, bicarbonate, borate, hydroxide and hydrogen ions as determined using CO2SYS software (Matlab v3.0)10. Uncertainty in all derived parameters was evaluated using the error.m routine from Orr et al11. To estimate the uncertainty of derived alkalinity, uncertainties were propagated from the pCO2 measurement and the headspace gas-adjusted TCO2. Uncertainty in headspace gas-adjusted TCO2 was taken to be equivalent to the TCO2 measurement uncertainty despite the headspace gas adjustment having uncertainty associated with primarily the atmospheric measurement (~2 µatm) and the headspace gas volume (~2%). Typical adjustments are on the order of 1 µmol/kg and the uncertainty amounted to roughly 2% of the adjustment, and therefor was considered negligible. Figure 3 shows the resulting uncertainty in derived alkalinity as a function of TCO2. We take the median value to represent our typical uncertainty for derived alkalinity, albeit note larger spread in the uncertainty in derived alkalinity at lower TCO2 content.

**Figure 3:** Relative uncertainty in seawater alkalinity (%) as a function of TCO2 content (µmol/kg) with color as the derived alkalinity content. The typical uncertainty of derived alkalinity was 0.46% and was represented by the median value.

After deriving alkalinity, the pCO2, pHT, and carbonate mineral saturation states were calculated at *in situ* temperature, salinity, and pressure using the same suite of marine CO2 system constants. This calculation was done using CO2SYS with derived alkalinity and measured TCO2. However, it is important to note that evaluating the uncertainty of these terms using the derived alkalinity and measured TCO2 uncertainties would result in double-counting the measured TCO2 uncertainty. Therefore, we evaluate the uncertainty in pCO2, pHT, and carbonate mineral saturation states at in situ conditions using the uncertainties in our direct measurements, pCO2 and TCO2, in addition to uncertainty in the dissociation constants. Uncertainty in these derived parameters was evaluated as before by propagating the uncertainties of the input parameters. The median uncertainty values for pCO2, pH, and calcite saturation state were 1.5%, 0.012, and 4.8%, respectively. Note that we report the absolute uncertainty for pHT because it is a logarithmic variable; an absolute uncertainty in pH represents a relative uncertainty in hydrogen ion content11. We take the median values to reflect our typical uncertainties for these variables. To provide additional context for these estimates of uncertainty, we prescribed the “weather quality” threshold of 10 µmol/kg for TCO2 and alkalinity4 to the data used to create Figure 4. This assessment highlights that our data quality is within the “weather quality” threshold, as well as our assessment of calcite saturation state uncertainty is consistent with the 4% uncertainty in aragonite saturation states reported in another study using the same instrumentation12.



**Figure 4:** Uncertainty in pCO2 (top; relative %), calcite saturation states (Ωcal; middle; relative %), and pHT (bottom) as a function of TCO2 content (µmol/kg) with color indicating the corresponding derived parameter values. Also shown as red dots are parameter uncertainties calculated with prescribed GOA-ON “weather” thresholds for pCO2 and TCO2, respectively. As indicated using median values, typical pCO2, Ωcal, and pHT uncertainty was 1.5%, 4.8% and 0.012, respectively.

*Examining deviations in the global calcium-salinity relationship*

We wanted to consider the potential for local deviations in the global calcium-salinity relationship used within the CO2SYS software as an additional source of uncertainty for saturation state determinations13. Calcium samples were collected during two surveys at oceanographic station QU39 and used to examine this potential. We lack local freshwater data to build a robust regional relationship, but we do note from this comparison that there is a slight tendency to over-estimate saturation state without better regional constraint of the calcium content. The degree of this over-estimate is approximately 0.04 based on these limited data.

**Table 1:** The difference between measured and calculated calcium content using CO2SYS and the impact on derived aragonite saturation states. The measurements were made on December 1, 2020 (Date 1) and December 17, 2020 (Date 2). The average difference between derived aragonite saturation states using measured calcium or calcium calculated in CO2SYS was -0.04, suggesting a slight offset in saturation states derived using the global calcium-salinity relationship within CO2SYS.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **[Ca2+], mol kg-1** | | | **Calculated ΩAr** | | |
| Date | Depth | YSI\_S | Measured | CO2SYS | Meas.[Ca2+] | CO2SYS | Diff |
| 1 | 0 | 28.1 | 0.00830 | 0.00827 | 0.919 | 0.914 | 0.005 |
| 1 | 0 | 28.1 | 0.00806 | 0.00827 | 0.892 | 0.914 | -0.022 |
| 1 | 5 | 28.2 | 0.00806 | 0.00830 | 0.854 | 0.879 | -0.024 |
| 1 | 10 | 28.5 | 0.00830 | 0.00839 | 0.889 | 0.897 | -0.008 |
| 1 | 20 | 28.6 | 0.00805 | 0.00842 | 0.842 | 0.878 | -0.037 |
| 1 | 30 | 28.7 | 0.00805 | 0.00845 | 0.833 | 0.872 | -0.039 |
| 1 | 40 | 28.8 | 0.00805 | 0.00848 | 0.808 | 0.849 | -0.041 |
| 1 | 50 | 29.1 | 0.00805 | 0.00856 | 0.671 | 0.713 | -0.042 |
| 1 | 75 | 30.2 | 0.00853 | 0.00889 | 0.579 | 0.602 | -0.023 |
| 1 | 100 | 30.7 | 0.00877 | 0.00903 | 0.598 | 0.615 | -0.017 |
| 1 | 150 | 30.9 | 0.00877 | 0.00909 | 0.607 | 0.629 | -0.022 |
| 1 | 200 | 31 | 0.00876 | 0.00912 | 0.635 | 0.660 | -0.025 |
| 1 | 260 | 31.1 | 0.00900 | 0.00915 | 0.645 | 0.655 | -0.010 |
| 2 | 0 | 27.5 | 0.00782 | 0.00809 | 0.821 | 0.849 | -0.028 |
| 2 | 0 | 27.5 | 0.00757 | 0.00809 | 0.795 | 0.849 | -0.053 |
| 2 | 5 | 27.7 | 0.00782 | 0.00815 | 0.803 | 0.836 | -0.033 |
| 2 | 10 | 28.1 | 0.00781 | 0.00827 | 0.776 | 0.821 | -0.044 |
| 2 | 20 | 28.3 | 0.00757 | 0.00833 | 0.756 | 0.831 | -0.075 |
| 2 | 30 | 28.4 | 0.00757 | 0.00836 | 0.747 | 0.824 | -0.077 |
| 2 | 40 | 28.5 | 0.00757 | 0.00839 | 0.751 | 0.831 | -0.080 |
| 2 | 50 | 28.8 | 0.00756 | 0.00848 | 0.645 | 0.722 | -0.077 |
| 2 | 75 | 30.3 | 0.00804 | 0.00892 | 0.566 | 0.627 | -0.061 |
| 2 | 100 | 30.6 | 0.00828 | 0.00901 | 0.567 | 0.615 | -0.048 |
| 2 | 150 | 30.9 | 0.00828 | 0.00909 | 0.582 | 0.638 | -0.056 |
| 2 | 200 | 31 | 0.008 | 0.578 | 0.009 | 0.636 | -0.058 |
| 2 | 260 | 31.1 | 0.008 | 0.582 | 0.009 | 0.643 | -0.061 |

1 Campbell, K., Weekes, C., Evans, W., Gimenez, I. & Hales, B. Hakai Institute’s Burke-o-Lator TCO2/pCO2 Analyzer Discrete Sample Analysis Protocols (v2.0). doi:10.21966/1.521066 (2023).

2 Evans, W. *et al.* Marine CO2 Patterns in the Northern Salish Sea. *Frontiers in Marine Science* **5**, doi:10.3389/fmars.2018.00536 (2019).

3 Jiang, L.-Q. *et al.* Best Practices Data Standards for Discrete Chemical Oceanographic Observations. *Frontiers in Marine Science* **8**, 10.3389/fmars.2021.705638 (2022).

4 Newton, J. A., Feely, R. A., Jewett, E. B., Williamson, P. & Mathis, J. Global Ocean Acidification Observing Network: Requirements and Governance Plan [*http://goa-on.org/docs/GOA-ON\_2nd\_edition\_final.pdf*](http://goa-on.org/docs/GOA-ON_2nd_edition_final.pdf), 60 (2015).

5 Wanninkhof, R. & Thoning, K. Measurement of fugacity of CO2 in surface water using continuous and discrete sampling methods. *Marine Chemistry* **44**, 189-204 (1993).

6 Waters, J., Millero, F. J. & Woosley, R. J. Corrigendum to “The free proton concentration scale for seawater pH”, [MARCHE: 149(2013) 8-22]. *Marine Chemistry* **165**, 66-67 (2014).

7 Dickson, A., Wesolowski, D. J., Palmer, D. A. & E., M. R. Dissociation constant of bisulfate ion in aqueous sodium chloride solutions at 250 ℃. *Journal of Physical Chemistry* **94**, 7978-7985, doi:doi: 10.1021/j100383a042 (1990).

8 Perez, F. F. & Fraga, F. Association constant of fluoride and hydrogen ions in seawater *Marine Chemistry* **21**, 161-168 (1987).

9 Uppström, L. R. The boron/chlorinity ratio of deep-sea water from the Pacific Ocean. *Deep-Sea Research and Oceanographic Abstracts* **21**, 161-162, doi:doi: 10.1016/0011-7471(74)90074-6 (1974).

10 Sharp, J. D. *et al.* CO2SYSv3 for MATLAB (Version v3.2.0). *Zenodo*, <http://doi.org/10.5281/zenodo.3950562> (2021).

11 Orr, J. C., Epitalon, J.-M., Dickson, A. G. & Gattuso, J.-P. Routine uncertainty propagation for the marine carbon dioxide system. *Marine Chemistry* **207**, 84-107, doi:10.1016/j.marchem.2018.10.006 (2018).

12 Hales, B., Suhrbier, A., Waldbusser, G. G., Feely, R. A. & Newton, J. A. The Carbonate Chemistry of the "Fattening Line," Willapa Bay, 2011-2014. *Estuaries and Coasts*, DOI 10.1007/s12237-12016-10136-12237 (2016).

13 Dillon, W. D. N., Dillingham, P. W., Currie, K. I. & McGraw, C. M. Inclusion of uncertainty in the calcium-salinity relationships improves estimates of ocean acidification monitoring data quality. *Marine Chemistry* **226**, 103872 (2020).