LIMNOLOGY
and
OCEANOGRAPHY

Limnol. Oceanogr. 9999, 2024, 1–16

doi: 10.1002/lno.12674

Review

Random and systematic uncertainty in ship-based seawater carbonate chemistry observations

Brendan R. Carter,[1,2*] **Jonathan D. Sharp**,[1,2] **Maribel I. García-Ibáñez**,[3] **Ryan J. Woosley**,[4] **Michael B. Fong**,[5] **Marta Álvarez**,[6] **Leticia Barbero**,[7,8] **Simon L. Clegg**,[9] **Regina Easley**,[5] **Andrea J. Fassbender**,[2] **Xinyu Li**,[10] **Katelyn M. Schockman**,[7,8] **Zhaohui Aleck Wang**[11]

[1]Cooperative Institute for Climate Ocean and Ecosystem Studies, University of Washington, Seattle, Washington, USA
[2]Pacific Marine Environmental Laboratory, National Oceanic and Atmospheric Administration, Seattle, Washington, USA
[3]Instituto Español de Oceanografía, Palma, Spain
[4]Center for Global Change Science, Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
[5]National Institute of Standards and Technology, Chemical Sciences Division, Gaithersburg, Maryland, USA
[6]Instituto Español de Oceanografía, A Coruña, Spain
[7]Cooperative Institute for Marine and Atmospheric Sciences, University of Miami, Miami, Florida, USA
[8]Atlantic Oceanographic and Meteorological Laboratory, National Oceanic and Atmospheric Administration, Miami, Florida, USA
[9]School of Environmental Sciences, University of East Anglia, Norwich, UK
[10]School of Marine Science and Policy, University of Delaware, Newark, Delaware, USA
[11]Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA

Abstract

Seawater carbonate chemistry observations are increasingly necessary to study a broad array of oceanographic challenges such as ocean acidification, carbon inventory tracking, and assessment of marine carbon dioxide removal strategies. The uncertainty in a seawater carbonate chemistry observation comes from unknown random variations and systematic offsets. Here, we estimate the magnitudes of these random and systematic components of uncertainty for the discrete open-ocean carbonate chemistry measurements in the Global Ocean Data Analysis Project 2022 update (GLODAPv2.2022). We use both an uncertainty propagation approach and a carbonate chemistry measurement "inter-consistency" approach that quantifies the disagreement between measured carbonate chemistry variables and calculations of the same variables from other carbonate chemistry measurements. Our inter-consistency analysis reveals that the seawater carbonate chemistry measurement community has collected and released data with a random uncertainty that averages about 1.7 times the uncertainty estimated by propagating the desired "climate-quality" random uncertainties. However, we obtain differing random uncertainty estimates for subsets of the available data, with some subsets seemingly meeting the climate-quality criteria. We find that seawater pH measurements on the total scale do not meet the climate-quality criteria, though the inter-consistency of these measurements improves (by 38%) when limited to the subset of measurements made using purified indicator dyes. We show that GLODAPv2 adjustments improve inter-consistency for some subsets of the measurements while worsening it for others. Finally, we provide general guidance for quantifying the random uncertainty that applies for common combinations of measured and calculated values.

---

*Correspondence: brendan.carter@noaa.gov

Additional Supporting Information may be found in the online version of this article.

Author Contribution Statement: BRC led project administration, formal analysis, and investigation. BRC led visualization with contributions from JDS. BRC led conceptualization, writing, and editing with contributions from all authors. JDS, MIGI, RJW, and MBF provided extensive contributions to investigation, editing, and review. Each author supplied their own funding and resources.

Motivation and background

It is more important than ever to quantify the carbonate chemistry in seawater with well-characterized uncertainty given ongoing and intensifying global ocean carbon accumulation (Gruber et al. 2019; Friedlingstein et al. 2023; Müller et al. 2023), ocean acidification (Doney et al. 2009, 2020), and interest in using and assessing marine carbon dioxide removal strategies (Doney et al. 2022). Well-constrained uncertainties are fundamental for all scientific analyses, and this is especially the case for seawater carbonate chemistry, for which decadal or shorter timescale climate signals are small—often close to analytical detection limits—yet nevertheless quantitatively meaningful for the planetary carbon cycle because they span immense volumes of the ocean.

Uncertainties inherent in seawater carbonate chemistry measurements and calculations have become more quantifiable over time largely due to two developments. First is the increasingly common practice of overdetermining seawater carbonate chemistry with at least three carbonate chemistry variables measured on a single seawater sample, allowing scientists to scrutinize disagreements between measurements and calculations of the variables (McElligott et al. 1998; Kuliński et al. 2014; Chen et al. 2015; Patsavas et al. 2015; Williams et al. 2017; Woosley et al. 2017; Carter et al. 2018; Fong and Dickson 2019; Álvarez et al. 2020; Sulpis et al. 2020; Takeshita et al. 2021; García-Ibáñez et al. 2022; Guallart et al. 2022). Second are the continued improvements in measurement quality; this includes developing and adopting best practices for carbonate chemistry analysis and reporting (Dickson and Goyet 1994; Dickson et al. 2007; Riebesell et al. 2011), the general use of reference materials (RMs; Dickson 2010; Dickson et al. 2003) for total titration seawater alkalinity (A_T) and total dissolved inorganic carbon (C_T) measurements, and improvements in community measurement practices and equipment (e.g., the production of purified spectrophotometric indicator dyes for seawater pH measurements on the total hydrogen ion scale, or pH_T). Because of these developments, consistent patterns in disagreements between measured and calculated values have emerged from the noise of random methodological uncertainties. Importantly, the magnitudes of the disagreements between measured and calculated values rival the magnitudes of contemporary decadal changes in carbonate chemistry from ongoing ocean acidification and other climate change impacts in the open ocean.

Monitoring global change often requires making inferences based on carbonate chemistry calculations, mixing carbonate chemistry measurements with calculations, linking modern measurements to historical measurements (e.g., Carter et al. 2018), and evaluating confidence in measurements and calculations from novel measurement technologies (e.g., Bushinsky et al. 2019). Contending with the uncertainties inherent to carbonate chemistry measurements and calculations is therefore an urgent and perhaps underappreciated challenge for chemical oceanography and metrology. In Supporting Information Text S1, we provide an example of a monitoring challenge where carbonate chemistry uncertainties are a central and urgent challenge.

Definitions for random, systematic, and combined standard uncertainty

As articulated by Newton et al. (2015) (or earlier: Ellison and Williams 2012), uncertainty can be thought of as "a parameter associated with the result of a measurement that permits a statement of the dispersion (interval) of reasonable values of the quantity measured, together with a statement of the confidence that the (true) value lies within the stated interval." Uncertainty is not the same as measurement error, which is the difference between the true and measured values. For this effort, we use the definition of Newton et al. (2015) and further follow precedent and community recommendations to express our uncertainties as standard deviations (BIPM et al. 2008; Ellison and Williams 2012; Newton et al. 2015; Orr et al. 2018). Thus, if we simplistically assume that the sources of uncertainty result in normally distributed errors, then the 95% confidence interval for the true value would be approximately two standard uncertainties around the reported value (i.e., within $\pm 2u$). We define the terms used to describe aspects of uncertainty in Supporting Information Text S2, where possible using the definitions provided by BIPM et al. (2008) (*see* also Taylor and Kuyatt 1994). We use u to refer to the combined standard uncertainty or total uncertainty and append subscripts when referring to individual components of the combined standard uncertainty. All terms are represented as standard uncertainties (i.e., the probability distribution of the true error resulting from this contribution to uncertainty can be approximated as a normal distribution with a standard deviation equal to u).

The chemical complexity and variability of seawater solutions makes overall uncertainty estimation challenging, and analysts and data curators often instead report mean offsets between their measurements of a seawater RM and the reported value (i.e., bias) or statistics from replicate measurements made on a single cruise (i.e., precision). These statistics unfortunately do not capture all sources of variability that contribute to uncertainty (e.g., cruises with different instrumentation, procedures, locations, analysts, etc.). Fortunately, the true value of a variable is not the most critical piece of information for many oceanographic analyses. For example, in a time series of $f\text{CO}_2$ calculated from C_T and A_T, errors in the carbonate chemistry constants may have a meaningful impact on the uncertainty of calculated $f\text{CO}_2$ and therefore the air–sea $f\text{CO}_2$ difference, but a small impact on the uncertainty of the computed $f\text{CO}_2$ trend (this claim is demonstrated with an example calculation in Supporting Information Text S3).

Given this complexity, it is useful to separately quantify "random" components of uncertainty, which result from effects that vary unpredictably within a set of measurements, from the "systematic" components of uncertainty, which are consistent over a set of measurements and lead to offsets relative to the true value or a reference value. The combined standard uncertainty (u) then equals the random uncertainty (u_R) and systematic uncertainty (u_S) components, expressed as standard deviations, added in quadrature (i.e., the square root of the sum of squares). While the combined standard uncertainty of a parameter ideally accounts for all uncertainty contributions (u_R and u_S), some contributions may cancel for an analysis (e.g., Supporting Information Text S3).

Systematic uncertainty is not the same as bias. A bias is a known systematic offset relative to a reference value that should be corrected; however, the uncertainty in the bias correction (e.g., the uncertainty in the mean value of several measurements of a RM used to derive a bias correction) represents a form of systematic uncertainty that should be included in the calculation of u. It can be challenging to distinguish which offsets between measured and calculated or reference values are biases that should be corrected and which should be treated as the symptoms of quantifiable systematic uncertainties.

The dividing line between random and systematic uncertainties can be unclear. As an example, the sources of uncertainty that we consider to be random between cruises (e.g., day-to-day instrument calibration variations, temperature control errors, sample handling errors) can persist over a set of measurements and contribute to the apparent offset and, thus, to u_S. In our analysis, we rely on many different cruises considered collectively, and, with enough randomly distributed cruise offsets, the impact of these offsets should mostly cancel when computing an average offset. We therefore treat these cruise-specific offsets herein as contributions to random uncertainty, but we distinguish between random uncertainty contributions related to the short-term repeatability of measurements ($u_{\text{Repeatability}}$) and contributions that vary on longer timescales between cruises (u_{Cruise}), thus allowing the cruise-specific variations to be considered as systematic uncertainties when data from a smaller number of cruises are considered. Similarly, the sources of uncertainty that we consider to be systematic (including uncertainties from calibrating a method to an uncertain traceable standard measurement, $u_{\text{Calibration}}$, or incomplete knowledge of seawater acid–base species substance content and thermodynamic constants, $u_{\text{Constants}}$) can vary with seawater composition and properties and can therefore contribute to the estimated u_R in a set of measurements. However, we later show also that these contributions to the random variations are secondary to the measurement and cruise-specific variations for our analysis. The simplifications and terms we use to separate interwoven random and systematic uncertainty components are therefore imperfect, but nevertheless give us a foothold to begin considering how components of uncertainty might affect the full range of possible analyses.

In summary, the combined standard uncertainties u can be computed in two ways:

$$u^2 = u_R^2 + u_S^2 \tag{1}$$

$$u^2 = u_{\text{Repeatability}}^2 + u_{\text{Cruise}}^2 + u_{\text{Constants}}^2 + u_{\text{Calibration}}^2 \tag{2}$$

We consider $u_{\text{Repeatability}}$ to be a component of u_R; u_{Cruise} to be a component of u_R when considering many cruises collectively, as we do here, and a component of u_S when considering measurements from a single cruise; and $u_{\text{Constants}}$ and $u_{\text{Calibration}}$ to be primarily contributors to u_S. In Supporting Information Text S4, we generate a synthetic dataset visually illustrating each of these uncertainties and showing how u_{Cruise} can act variably as a random or systematic component depending on the data collection considered.

In this paper, we aim to quantify random and systematic uncertainties in the collection of readily available (primarily) open-ocean carbonate chemistry measurements and in calculations made using these measurements. We further explore how the uncertainties we estimate vary between measurements and measurement combinations and how they change when different sets of constants are used for seawater carbonate chemistry calculations. Finally, we propose some guidelines by which one might estimate the combination of random and systematic uncertainties that is most applicable to their analysis.

Methods

Data

Our analysis uses discrete ship-based A_T, C_T, $\mathrm{pH_T}$, and $f\mathrm{CO_2}$ measurements and calculations from measurements in GLODAPv2.2022 (Lauvset et al. 2022). This data product is a compilation of datasets from 1085 mostly open-ocean cruises adjusted to improve consistency between cruises. However, we are interested in the data being produced by the community before efforts are taken to improve the coherence of measurements between various cruises, so the first step in our analysis is to reverse the adjustments given in the GLODAPv2.2022 adjustment table (both the data product and the adjustment table are available from www.glodap.info). This step produces the "GLODAPv2.2022_unadjusted" data product.

Seawater carbonate chemistry variables are calculated for all combinations of input pairs for the overdetermined carbonate chemistry measurements in the GLODAPv2.2022_unadjusted data product. Except where noted, these calculations are made using the CO2SYSv3 code for MATLAB® (van Heuven et al. 2011; Orr et al. 2018; Sharp et al. 2021) with carbonic acid dissociation constant parameterizations from Lueker et al. (2000), hydrogen fluoride (HF) thermodynamic constant parameterization from Perez and Fraga (1987), the total boron

(B_T) to practical salinity (S_P) ratio (B_T/S_P) from Lee et al. (2010), and the $K_{HSO_4^-}$ from Dickson et al. (1990). As pH_T and fCO_2 are often reported at different temperatures, we follow GLODAPv2 practices and re-compute all measurements and calculations at a standard temperature and pressure of 25°C and 101,325 Pa (i.e., 1 atm, or 0 applied pressure in CO2SYSv3), respectively. The choice of the assumed temperature and pressure has minor impact on the disagreements between measured and calculated values because the conversion affects the measured and calculated values nearly identically. However, the choice does meaningfully impact the propagated uncertainty values because an assumed uncertainty, for example, 0.003 in pH_T, can correspond to a different impact on calculated variables at low temperature and high pressure than it does at standard conditions. We rely on standard conditions for our calculations because most prior measurement uncertainty estimates and benchmarks have been obtained from measurements at standard or near-standard conditions. However, we note that uncertainties at in situ conditions are often of greater interest than uncertainties at standard laboratory conditions and urge consideration of how uncertainties might vary at the conditions of interest for an analysis. Despite this relationship between seawater conditions and carbonate chemistry uncertainties, the findings from this study are qualitatively unchanged for calculations at in situ conditions (i.e., the numbers vary, but the claims we make regarding relative uncertainties remain true).

Uncertainty estimates

We use two approaches to quantify both systematic and random uncertainty: uncertainty propagation and inter-consistency comparisons (Fig. 1). The uncertainty propagation estimates are derived by propagating estimates of individual uncertainties in the many measurements and constants used in carbonate chemistry calculations (Orr et al. 2018), whereas inter-consistency estimates are made by comparing measurements of a variable to calculations of the same variable from other measurements. A third approach that we do not consider uses "inter-laboratory comparisons" of measurements from an identical batch of seawater that are made across multiple laboratories (Bockmon and Dickson 2015; Olsen et al. 2019). This approach quantifies uncertainties from methodological variations that exist across the research community, providing bounds on inter-laboratory measurement consistency. However, it only reflects community precision in laboratory measurements and only for a limited number of seawater compositions and laboratories and does not distinguish between labs that do and do not contribute to open-ocean seawater carbonate chemistry data products, and thus does not directly reveal the combined standard uncertainty for open-ocean ship-based discrete carbonate chemistry measurement data products.

Due to the variety of uncertainty terms discussed and the need for specificity, we refer to uncertainties using several subscripts given in Table 1, for example, $u_{S,C,UP,AT}$, where u with a subscript indicates an uncertainty contribution, "S" indicates the u is systematic, "C" indicates a calculated variable, "UP" indicates estimation through uncertainty propagation, and "AT" means the uncertainty is specific to A_T.

Uncertainty propagation

Significant progress has been made in recent years toward estimating reasonable uncertainties for the measurements and constants used in seawater carbonate chemistry calculations (Table 2, with references therein) and propagating these uncertainties to compute uncertainty in calculated quantities. Notably, Orr et al. (2018) released code that propagates user-provided uncertainties, which has been incorporated into recent releases of carbonate chemistry software (Orr et al. 2018; Sharp et al. 2021; Humphreys et al. 2022). Likewise, inter-consistency has long been used for estimating carbonate chemistry uncertainty, and the GLODAPv2.2022 data product update includes discrete fCO_2 measurements (Lauvset et al. 2022) providing new opportunities for these comparisons (e.g., García-Ibáñez et al. 2022).

We use the CO2SYSv3 uncertainty propagation code (Sharp et al. 2021; adapted from Orr et al. 2018) to quantify how input uncertainties in the carbonate chemistry calculations impact the calculated outputs (Fig. 1). For the systematic uncertainty propagation calculations ($u_{S,C,UP}$), we use the carbonate chemistry and other constant uncertainties suggested by Orr et al. (2018) (i.e., the bolded terms in Table 2). For the random uncertainty propagation calculations ($u_{R,C,UP}$), we use the random uncertainty estimates (i.e., random measurement terms in Table 2 associated with "climate-quality" measurements from Newton et al. (2015)). Newton et al. (2015) defined climate-quality measurements as those capable of resolving changes in carbonate ion contents with a relative standard uncertainty of less than 1% and suggest this translates to standard uncertainties of 2 μmol kg^{-1} in C_T and A_T, 0.003 in pH_T, and a relative uncertainty of 0.5% in fCO_2. Critically, this "climate-quality" definition implicitly omits the contribution of systematic uncertainty in measurements because it is based on resolving changes rather than quantifying true values. It is therefore a metric for random uncertainty (u_R).

We calculate random uncertainties for the differences ($u_{R,\Delta,UP}$) between measurements ($_M$) and calculations ($_C$), assuming the measurement also meets climate-quality ($_{CQ}$) criteria, by taking the positive square root of the sum of the squared uncertainties:

$$u_{R,\Delta,UP} = \sqrt{(u_{R,C,UP})^2 + (u_{R,M,CQ})^2} \quad (3)$$

This equation applies equally to terms later calculated for inter-consistency calculations ($_{IC}$), and allows $u_{R,\Delta,UP}$ estimates to be compared to random uncertainties from the inter-

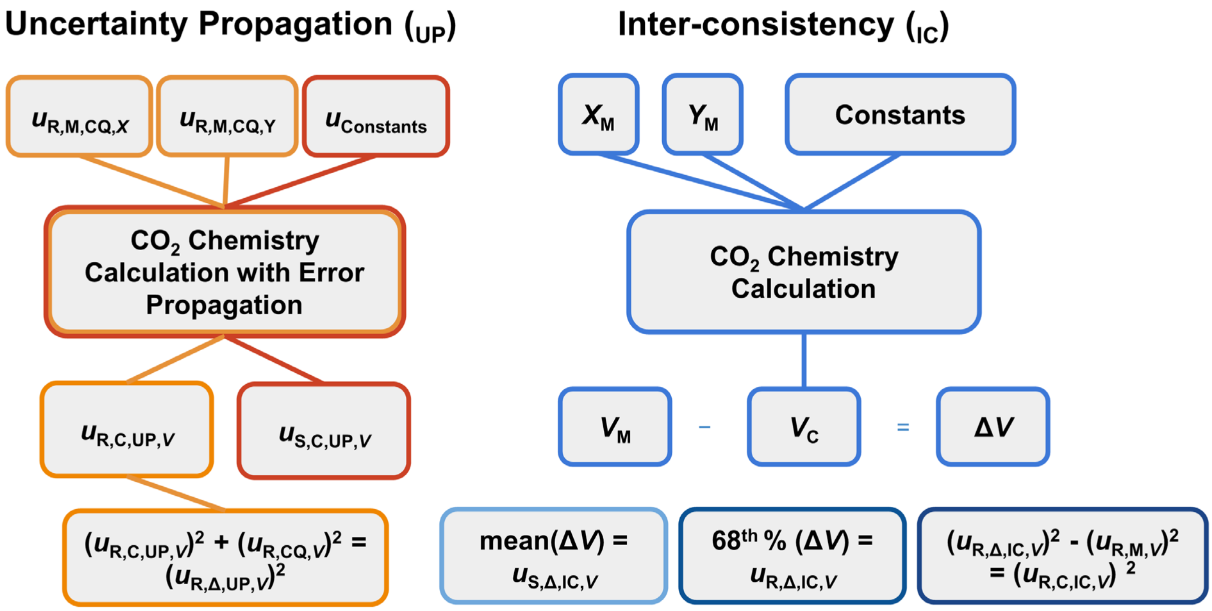

Fig. 1. Conceptual schematic illustrating the process by which random (subscripted as $_R$) and systematic ($_S$) uncertainty (u) estimates are obtained for variable V (V) when calculated (C) from variables X (X) and Y (Y). The process is different when the uncertainty is estimated using uncertainty propagation (warm colors, UP) and inter-consistency (cool colors, IC). In the uncertainty propagation approach, a priori estimates of "climate-quality" (Newton et al. 2015) uncertainties in measured (M) values (i.e., $u_{R,M,CQ}$) and uncertainties in carbonate chemistry constants ($u_{Constants}$) are propagated to generate random (light orange outline) and systematic (dark orange outline) u estimates in the calculations, respectively. The random uncertainties are then added in quadrature to the climate-quality uncertainty of the measured value for V to estimate the expected random u for the difference between the measured and calculated values of V (i.e., ΔV or Δ). In the inter-consistency approach, measurements of X and Y (i.e., X_M and Y_M) are used along with carbonate chemistry constants (constants) to compute a calculated value of V (V_C), which is compared to the associated measured value of V (V_M). The average offset between the two is the inter-consistency (IC) estimate of the systematic uncertainty ($u_{S,\Delta,IC,V}$) in this difference. The 68th percentile of the spread in the residuals ($u_{R,\Delta,IC,V}$) around the $u_{S,\Delta,IC,V}$ is the result of the random measurement uncertainty in V (i.e., $u_{R,M,V}$) and the random uncertainty in the calculated value of V ($u_{R,C,IC,V}$) added in quadrature.

consistency analysis that—by necessity—include uncertainties from both measurements and calculations. As we do not separately account for systematic uncertainty in the measurements, $u_{S,\Delta,UP}$ equals $u_{S,C,UP}$.

Uncertainty propagation estimates themselves vary with sample composition (*see* Woosley and Moon 2023). We propagate uncertainties for all measurements individually, but, for simplicity, present the means of these propagated uncertainties (and, in one instance, their statistics) for collections of measurements in our figures and tables. We justify the omission of $u_{Constants}$ from $u_{R,C,UP}$ and note three nuances that are important for interpreting our uncertainty propagation estimates in Supporting Information Text S5.

Inter-consistency comparisons

We estimate random and systematic contributions to combined uncertainty with an "inter-consistency" analysis using collections of overdetermined carbonate chemistry measurements (Fig. 1). There are several steps for these calculations, and all of them are repeated for all combinations of four variables of interest "V" (A_T, C_T, pH_T, and fCO_2) and for each of the three combinations of choices for variables "X" and "Y" that can be used to calculate each V (for 12 total combinations of V, X, and Y).

1. For each measurement of V (V_M), we calculate values of V (V_C) from X and Y.
2. Residuals (ΔV) are computed between the measured and calculated values:

$$\Delta V = V_M - V_C \tag{4}$$

3. The inter-consistency-based estimate of the systematic component of uncertainty in ΔV (i.e., $u_{S,\Delta,IC,V}$) is computed as the mean ΔV across all samples in the data product. These inter-consistency-based estimates of systematic uncertainty do not distinguish between systematic uncertainty in measurements ($u_{S,M,IC,V}$) and calculations ($u_{S,C,IC,V}$) of V and only reflect their apparent offset. Supporting Information Text S6 offers discussion of the implications of calculating $u_{S,IC}$ in this fashion.
4. The inter-consistency-based estimate of the random component of uncertainty in ΔV (i.e., $u_{R,\Delta,IC,V}$) is a measure of how much spread in ΔV values exists around the mean ΔV (i.e., around $u_{S,\Delta,IC,V}$). It is estimated as the 68th percentile of absolute deviations between ΔV values and $u_{S,\Delta,IC,V}$. We

Table 1. Subscripts used herein for u and their explanations.

Position ($u_{1,2,3,4}$)	Subscript	Meaning
1	R	Random
1	S	Systematic
2	M	Specific to a measurement
2	C	Specific to a calculation
2	Δ	Specific to a difference between a measurement and a calculation
3	UP	Estimated using uncertainty propagation
3	IC	Estimated using inter-consistency calculations
3	CQ	Assumed to be "climate-quality"
4	V	Applies to variable of interest "V"
4	X	Applies to variable "X" used with variable "Y" to calculate variable of interest V
4	Y	Applies to variable "Y" used with variable "X" to calculate variable of interest V
4	AT	Applies to total titration seawater alkalinity (A_T)
4	CT	Applies to total dissolved inorganic carbon (C_T)
4	pHT	Applies to seawater pH on the total hydrogen ion scale (pH_T)
4	fCO_2	Applies to the fugacity of CO_2 in a headspace in gas-exchange equilibrium with the seawater (fCO_2)

use the 68th percentile in place of the standard deviation here because the ΔV differences are not normally distributed. We justify this choice and briefly discuss its implications in Supporting Information Text S7.

5. For discussion purposes, we repeat these calculations using variations and subsets of the full collection of overdetermined measurements (mapped in Supporting Information Fig. S1) including:
    a. the GLODAPv2.2022 data product with the secondary quality control (QC) adjustments included,
    b. three subsets where spectrophotometric pH_T measurements were made using purified indicator dyes, using indicator dyes that have not been purified, and using potentiometric or unknown pH_T measurement methods,
    c. and the subset of GLODAPv2.2022 with a QC flag of 0 indicating secondary QC was not attempted.

Quantifying measurement uncertainties

So far, we have used inter-consistency comparisons only to quantify uncertainties in the differences between measured and calculated seawater carbonate chemistry variables, but we are also interested in using inter-consistency to estimate uncertainties in the measured and calculated values individually. In principle, we can leverage our 12 sets of inter-consistency calculations to go beyond our simplistic assumption of climate-quality measurements: we can rewrite Eq. 3 for the $_{\mathrm{IC}}$ terms and expand the $u_{\mathrm{R,C,IC},V}$ term to reflect contributions from random uncertainties in measurements of the variables "X" and "Y" used to calculate "V" multiplied by sensitivity terms for the calculated value (i.e., $\frac{\partial V_C}{\partial X}$ and $\frac{\partial V_C}{\partial Y}$), and a term u_{Other} that reflects the contributions of random uncertainties in all other inputs to carbonate chemistry calculations (e.g., temperature, salinity, and dissolved inorganic nutrient contents. Note: this u_{Other} term is also specific to the combination of V, X, and Y and the sample composition despite the absence of detailed subscripts in our notation). The sensitivity terms are calculated for each overdetermined sample individually by perturbing the input value for the calculation by (adding) the random uncertainty values in Table 2 and recalculating V. These sensitivities are then averaged across all measurements to obtain mean values used to calculate $u_{\mathrm{R},\Delta,\mathrm{IC},V}$:

$$u^2_{\mathrm{R},\Delta,\mathrm{IC},V} = u^2_{\mathrm{R,M},V} + \left(\frac{\partial V_\mathrm{C}}{\partial X} u_{\mathrm{R,M,X}}\right)^2 + \left(\frac{\partial V_\mathrm{C}}{\partial X} u_{\mathrm{R,M,Y}}\right)^2 + u^2_{\mathrm{Other}} \quad (5)$$

Equation 5 can be written for each of the 12 combinations of V, X, and Y considered and the resulting system of 12 equations can be solved for four unique random measurement uncertainties $u_{\mathrm{R,M}}$ for each of the four variables (the sensitivity terms and $u_{\mathrm{Other},V}$ can be calculated using uncertainty propagation and averaged across the measurements in the GLODAPv2.2022_unadjusted product). However, in practice, these calculations are not well conditioned despite the 12 constraining equations and only 4 unknowns. This is because the impacts of uncertainties in A_T and C_T measurements are nearly colinear, and a similar statement can be made for the impacts of uncertainties in pH_T and fCO_2 measurements. Also, disagreements between pH_T measurements and pH_T calculated from C_T and A_T provide similar information to the disagreements between C_T measurements and C_T calculated from pH_T and A_T. Similar statements can be made for other combinations.

To compare the quality of the measurements of a value V to the desired quality, we introduce dimensionless ratios between the estimated ($u_{\mathrm{R,M},V}$) and the desired climate-quality ($u_{\mathrm{R,CQ,V}}$) uncertainties:

$$r_V = \frac{u_{\mathrm{R,M},V}}{u_{\mathrm{R,CQ,V}}} \quad (6)$$

We solve for the r_V ratios that minimize the apparent disagreement between the observed $u_{\mathrm{R},\Delta,\mathrm{IC}}$ using methods detailed in Supporting Information Text S8. Due to the difficulty of distinguishing the effects of uncertainty in C_T from uncertainty in A_T, we require the simplifying assumption that the r_{CT} and r_{AT} values approximately equal each other.

Table 2. Literature-estimated values of the combined standard uncertainty (or the total relative uncertainty expressed as a percentage) of the indicated constraints for the carbonate chemistry in oxygenated seawater. Combined standard uncertainty and other terms used herein are defined in the glossary. Where no estimate of combined uncertainty is yet available, random uncertainty has been given. Random uncertainty is given in addition to the uncertainty in several instances in which both have been assessed. These estimates are most appropriate for a salinity of ~ 35, a temperature of ~ 25°C, and a total pressure of 1 atm. For a comparison in this study, terms in bold are propagated to obtain systematic uncertainty estimates while the other quantities, as given in the random uncertainty column, are propagated for random uncertainty estimates. We consider all uncertainties in these constants as contributors to systematic uncertainty in our calculations.

Parameter	Combined standard uncertainty	Random uncertainty	Reference(s)
A_T	2 μmol kg^{-1}	2 μmol kg^{-1}	Orr et al. (2018); Newton et al. (2015)
C_T	2 μmol kg^{-1}	2 μmol kg^{-1}	Orr et al. (2018); Newton et al. (2015)
pH_T	0.01	0.003	Orr et al. (2018); Newton et al. (2015)
fCO_2	1%	0.5%	García-Ibáñez et al. (2022); Newton et al. (2015)
$Phosphate_T$	*	2%	Lauvset et al. (2022)
$Silicate_T$	*	2%	Lauvset et al. (2022)
S_T	0.16%	0.005 g kg^{-1}	Morris and Riley (1966); Lauvset et al. (2022)
B_T	**2%**	*	Orr et al. (2018)
F_T	**0.15%**	*	Warner (1971)
pK_W	**0.01**	*	Orr et al. (2018)
pK_{1C}	**0.0075**	0.0055	Orr et al. (2018)
pK_{2C}	**0.015**	0.010	Orr et al. (2018)
pK_B	**0.005**	*	Dickson et al. (1990)
pK_{1P}	**0.034**	*	Dickson and Riley (1979)
pK_{2P}	**0.015**	*	Dickson and Riley (1979)
pK_{3P}	**0.17**	*	Dickson and Riley (1979)
pK_{Si}	*	**0.02**	Millero (1995)
pK_S	*	**0.01**	Dickson et al. (1990)
pK_F	*	**0.02**	Dickson and Riley (1979)
pK_0	**0.002**	*	Orr et al. (2018)

*Value unknown.

Results and discussion

Random uncertainty in calculations ($u_{R,C}$)

Figure 2 shows distributions of residuals for GLODAPv2.2022_unadjusted data product (gray histograms) between measured and calculated values (i.e., ΔV) along with Gaussians representing random uncertainty distributions determined from inter-consistency ($\pm u_{R,\Delta,IC}$, blue lines) and from uncertainty propagation ($\pm u_{R,\Delta,UP}$, Eq. 3, orange lines), both centered on systematic uncertainties determined from inter-consistency ($u_{S,\Delta,IC}$). These figures are also provided for the GLODAPv2.2022 data product with adjustments (Supporting Information Fig. S2). The uncertainty propagation estimates (orange lines) imply a tighter distribution of residuals between measured and calculated values than is found in practice through inter-consistency comparisons (blue lines). It can also be seen that most of the residual distributions are not well-approximated as a normal distribution, implying that the data from the GLODAPv2.2022_unadjusted data product are heterogeneous with respect to random uncertainties.

Figure 3 shows "random uncertainty" estimates for calculations obtained from both inter-consistency (blue bars, $u_{R,\Delta,IC}$) and uncertainty propagation (orange bars, $u_{R,C,UP}$), equivalent to the 68th percentile widths of the Gaussians in Fig. 2. Figure 3 reaffirms that some combinations of measurements are better for calculating a given seawater carbonate chemistry variable (i.e., have shorter bars) than others. Calculations that use the pH_T and fCO_2 pair together are particularly problematic, and calculations that pair either A_T or C_T with either pH_T or fCO_2 are generally best. This is consistent with long-established principles regarding whether two measurements provide similar or independent information for constraining the seawater carbonate chemistry (e.g., Dickson and Riley 1979; Millero 2007; Cullison Gray et al. 2011; McLaughlin et al. 2015).

As noted, the $u_{R,\Delta,IC}$ estimates are always worse than the $u_{R,\Delta,UP}$ estimates, ranging from ~1.5 to ~2.3 times (average 2.0 times) as large across the various calculations (Fig. 3). This implies at least one of two unfortunate possibilities: (1) these measurements within GLODAPv2.2022, considered collectively, are not attaining the stringent climate-quality precision standard or (2) there are unknown gaps in our understanding of the seawater carbonate chemistry and other seawater acid–

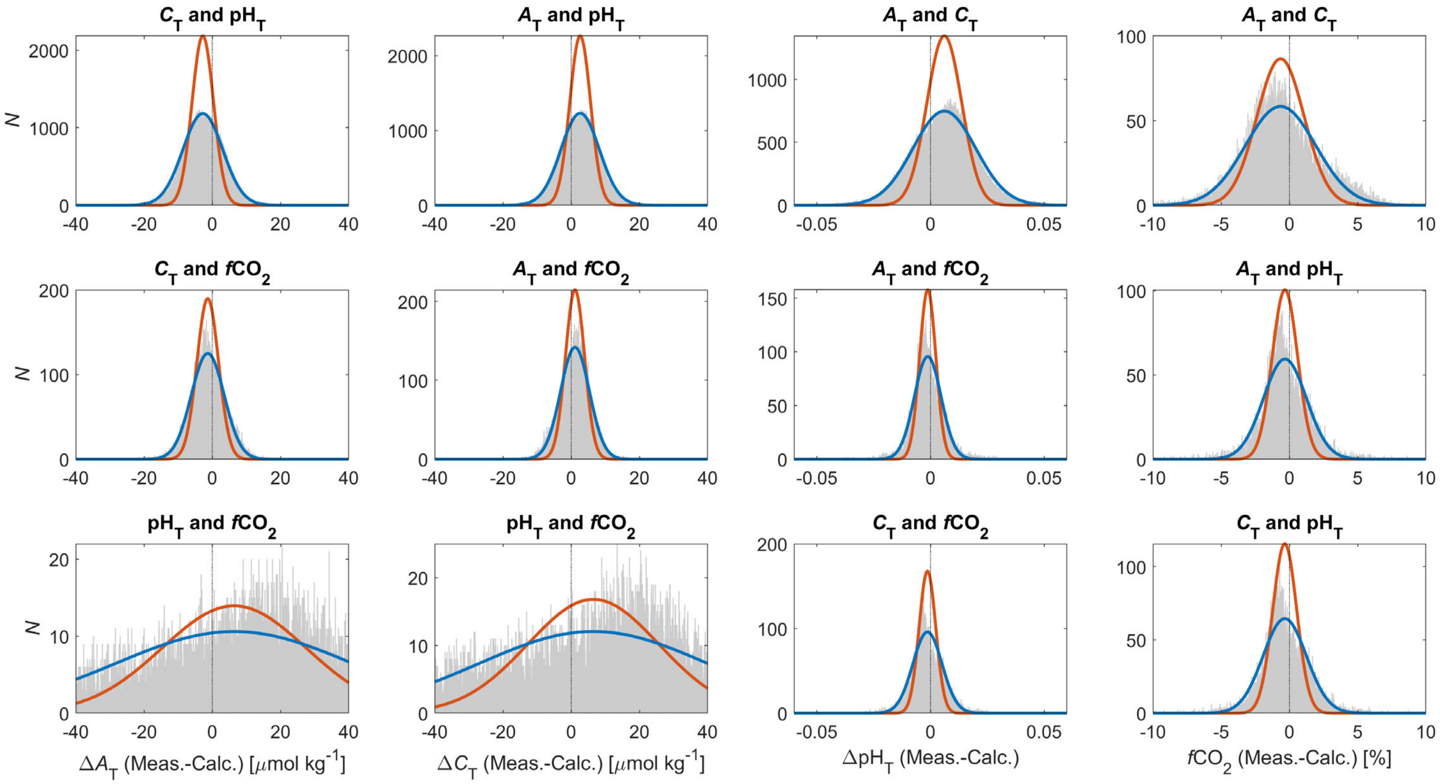

Fig. 2. Uncertainty estimates represented as equal area Gaussian distributions plotted over a histogram (also of equal area to the Gaussians) of residuals between the indicated measurements and calculations from the GLODAPv2.2022_unadjusted data product (i.e., after removing the GLODAPv2 inter-cruise consistency adjustments). The blue Gaussian lines are centered on the mean inter-consistency residual (i.e., the $u_{S,\Delta,IC}$) and have widths set to match the 68th percentiles of these residuals around the mean residuals (i.e., $u_{R,\Delta,IC}$ as in Fig. 1). The orange Gaussians are also centered on the mean residual and have a width set equal to the uncertainty propagation based random uncertainty estimates for calculations $u_{R,C,UP}$ (Eq. 3) after adding the climate-quality measurement uncertainty $u_{R,M,CQ}$ in quadrature to account for the added uncertainty in the measurement used to compute the residual. *See* the text for a discussion regarding the comparative widths of the blue and orange lines. A dotted vertical line at 0 is provided for reference. All comparisons are made at 1 atm and 25°C. Note the different number of measurements (*N*) for each pair.

base systems. One known gap in our understanding of carbonate chemistry relates to the presence of organic or other unidentified contributions to A_T (*see* Carter et al. (2024) for further discussion) and it is likely that these species contribute both to systematic and (through spatial or temporal variations in their amount contents in seawater) random uncertainties. When provided in cruise reports, most hydrographic cruise data QC metrics (e.g., from replicates and measurements of RMs) do not give clear evidence—nor do they typically give any reason to doubt—that the measurements provided are climate-quality, so we contend that both options are possible.

For discrete fCO_2, the comparison comes from only 14 cruises (with a 15th eligible cruise omitted, GLODAPv2.2022 cruise number 695, because it appears to be an outlier and including it nearly doubles the standard deviation of the measurement and calculation residuals). These comparisons therefore should be given lower confidence relative to comparisons that do not involve fCO_2, which are compiled from far more cruise datasets.

Like García-Ibáñez et al. (2022), we note that random uncertainty estimates from inter-consistency improve by an average of 22% for combinations of pH_T, C_T, and A_T when the GLODAPv2.2022 secondary QC adjustments are included (Supporting Information Fig. S3), while random uncertainty estimates involving both fCO_2 and pH_T worsen by an average of 58% when the adjustments are included. The adjustments improved inter-consistency between A_T, C_T, and fCO_2 by an average of 7%. While the fCO_2 data could not be adjusted during secondary QC due to sparse data, for some cruises, the data that they are compared to (i.e., A_T, C_T, and pH_T) are adjusted. Our results therefore reinforce the idea that GLODAP data product consistency adjustments are making the pH_T and fCO_2 measurements less inter-consistent (*see* García-Ibáñez et al. 2022 and Supporting Information Fig. S3).

The subset of pH_T measured on 47 cruises that used the spectrophotometric method with purified *m*-cresol purple indicator dyes has unusually strong inter-consistency, with 38% smaller average $u_{R,\Delta,IC,pHT}$ values than the data product at large (Supporting Information Fig. S4 compared to Fig. 3). The $u_{R,\Delta,IC,pHT}$ estimate from the C_T and A_T combination is only 20% greater than the $u_{R,\Delta,UP,pHT}$ estimate. In Supporting

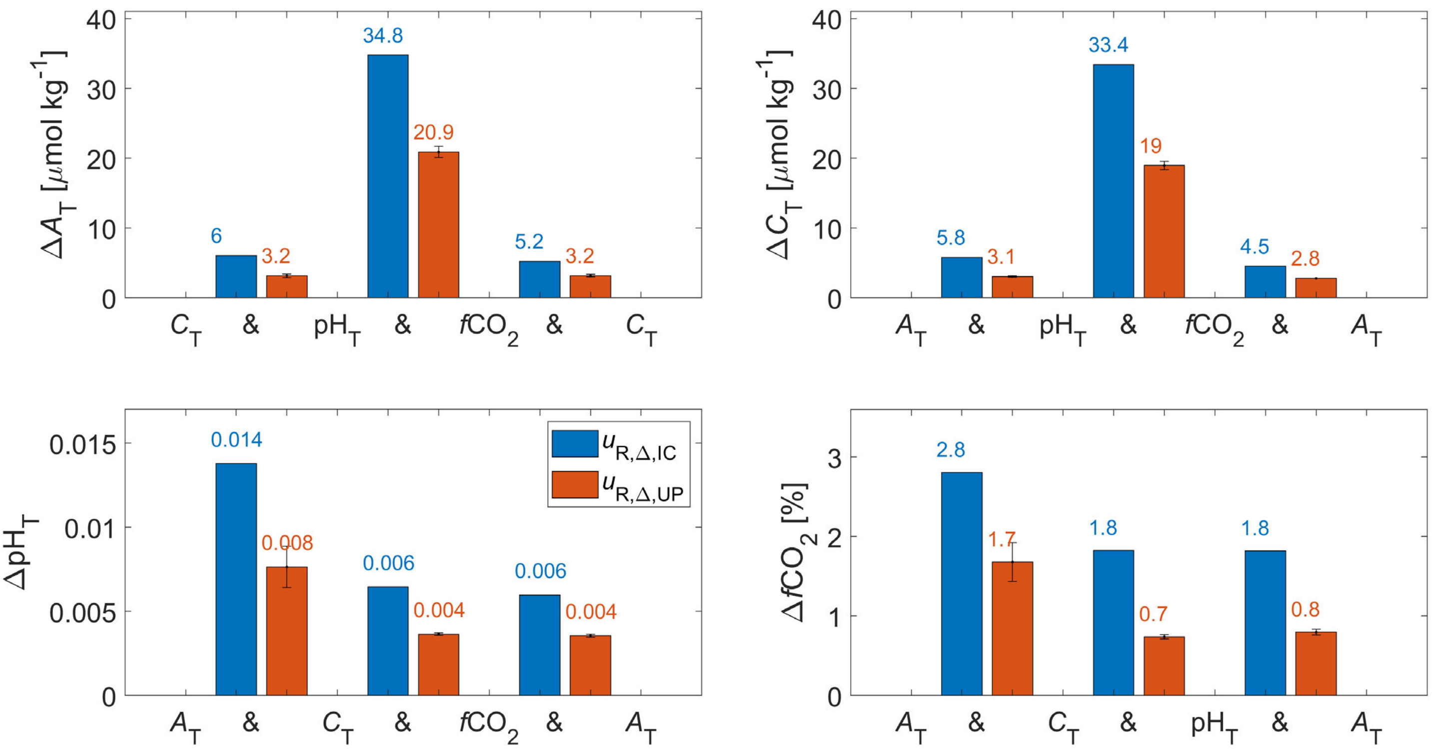

Fig. 3. Bar plots of estimated random standard uncertainties for residuals between measurements and calculations estimated from inter-consistency calculations ($u_{R,\Delta,IC}$, blue bars) and uncertainty propagation ($u_{R,\Delta,UP}$, orange bars) for the seawater carbonate chemistry variables listed on the *y*-axes (in the indicated units) when calculated using the pairs of measurements listed on the *x*-axes on either side of the bars. Of note, the blue bars are always greater than the associated orange bars. The whiskers on the orange bars indicate 1 standard deviation of the range of $u_{R,\Delta,UP}$ values obtained from the measurements in the GLODAPv2.2022_unadjusted data product, with the variability being attributable to the varying sensitivity of the calculations to assumed uncertainties in the input measurements as the seawater composition changes across the various measurements in the data product. In keeping with others (Newton et al. 2015; García-Ibáñez et al. 2022), fCO$_2$ is quantified as a relative precision (percentage of the measured value). Seawater pH$_T$ is seawater pH on the total hydrogen ion scale. All comparisons are made at 1 atm and 25°C. Calculations use the default carbonate chemistry constants for this analysis (Lueker et al. 2000; Lee et al. 2010).

Information Text S9, we discuss a subset of the GLODAPv2.2022 data product that was not subjected to the secondary QC with higher than average $u_{R,\Delta,IC}$ values.

Systematic uncertainty in calculations ($u_{S,C}$)

In our companion paper (Carter et al. 2024), we discuss many sources of uncertainty that collectively give rise to systematic offsets. However, in contrast to the estimates of random uncertainties, the inter-consistency based systematic discrepancies between measurements and calculations ($u_{S,\Delta,IC}$) for seawater carbonate chemistry variables in GLODAPv2.2022_unadjusted data product fall within the interval defined by the standard uncertainties that would be expected to result solely from assumed uncertainties in carbonate chemistry and other constants ($u_{S,\Delta,UP}$). Supporting this, the orange whiskers in all uncertainties reflecting $\pm u_{S,\Delta,UP}$ contain the $u_{S,\Delta,IC}$ estimates in all instances in Fig. 4. This is also true for many but not all the $u_{S,\Delta,IC}$ estimates obtained when using alternative constant sets (Supporting Information Text S10). Systematic uncertainty from carbonate chemistry constant sets therefore warrants further quantitative consideration.

Carbonate chemistry constant set comparisons

In recent years, a community consensus has emerged favoring the use of the carbonic acid dissociation constants measured by Mehrbach et al. (1973) and re-parameterized by Dickson and Millero (1987) on the seawater pH scale or by Lueker et al. (2000) on the total pH scale (Raimondi et al. 2019 and many references cited; Woosley 2021; Jiang et al. 2022), though the community consensus is less established for polar (Sulpis et al. 2020; Gattuso et al. 2023 and references therein) and coastal waters. Our calculations allow us to build upon this discussion by comparing the carbonate chemistry constants across multiple variables and calculations for the body of oceanographic measurements available in GLODAPv2.2022_unadjusted data product, as Woosley (2021) and Woosley and Moon (2023) did recently for A_T, C_T, and pH$_T$ measurements from their laboratory. We provide some insights from our comparisons in Supporting

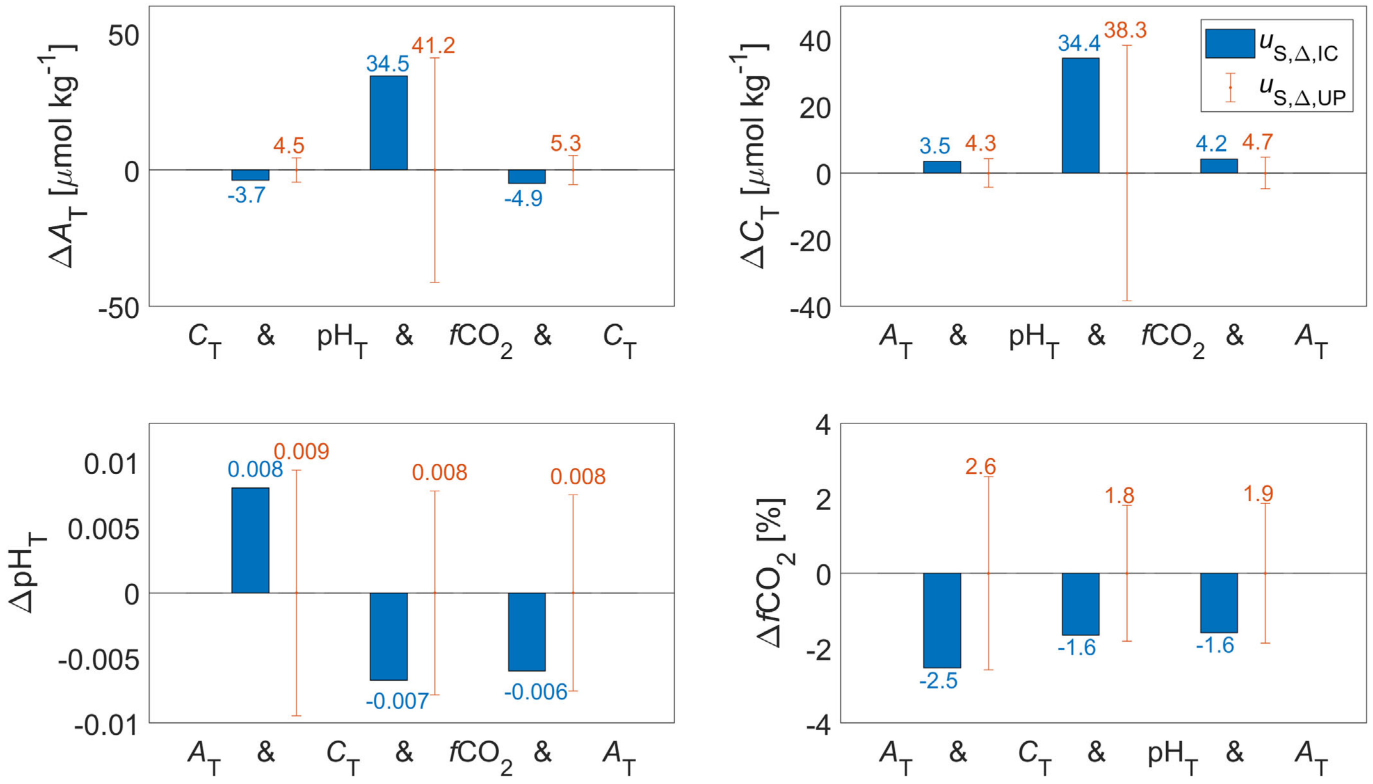

Fig. 4. $u_{S,\Delta,IC}$ Estimates obtained from the GLODAPv2.2022_unadjusted data product plotted as blue bars for the indicated combinations of variables used to calculate the variable of interest. The adjacent orange lines indicate the range of $\pm u_{S,\Delta,UP}$ (1σ) around the *x*-axis (black line). Seawater pH_T is seawater pH on the total hydrogen ion scale. All comparisons are made at 1 atm and 25°C. Calculations use the default carbonate chemistry constants for this analysis (Lueker et al. 2000; Lee et al. 2010).

Information Text S10. However, it is important to state that inter-consistency is not a measure of accuracy and having a small $u_{S,\Delta,IC}$ or $u_{R,\Delta,IC}$ does not guarantee a better value. This is because unidentified systematic uncertainties can significantly impact the results (Fong and Dickson 2019; Woosley and Moon 2023). For this reason—and because our findings vary with the data used, inter-consistency comparison variables considered, and the inclusion or omission of adjustments—we refrain from recommending a single carbonate chemistry constant set on the basis of this analysis and instead note that many community recommendations (by, e.g., Lee et al. 2000; Jiang et al. 2022; Sutton et al. 2022; Woosley and Moon 2023) advocate for constant sets that we find to be comparably inter-consistent (Supporting Information Text S10).

Random uncertainties in measurements in GLODAPv2.2022_unadjusted data product

Until now, we have only been able to say that the disagreements between measured and calculated values are larger than we would expect if all seawater carbonate chemistry measurements were being consistently measured at climate-quality by the community contributing to the GLODAPv2 data product. The various r_V ratio calculations allow us to go further and estimate the random uncertainties for measurements of individual variables ($u_{R,M}$), which can then be propagated to estimate the random uncertainties for calculated values ($u_{R,C}$), in the GLODAPv2.2022_unadjusted data product (Table 3). We recommend the $u_{R,M}$ values in Table 3 as more realistic community-level random measurement uncertainty estimates than the climate-quality values. Users can adopt these $u_{R,M}$ values for uncertainty propagation with their specific measurements and seawater compositions of interest.

If we solve for a single *r*—or ratio of the true random uncertainty to the climate-quality uncertainty that best aligns the uncertainty propagation and inter-consistency random uncertainties—for all equations and all variables then we would obtain a value of 1.7 for all data in the GLODAPv2.2022_unadjusted data product. This is a quantitative restatement that the apparent scatter in the differences between measurements and calculations is greater than expected from uncertainty propagation. The same is true when the *r* is derived after excluding the comparisons that use fCO_2 information ($r = 1.8$). This suggests that the issue is not limited to the comparatively few new-to-GLODAP cruises with discrete fCO_2 measurements. Interestingly, the GLODAPv2.2022 adjustments would improve the *r* calculated for the system of

Table 3. Random and systematic uncertainties in the GLODAPv2.2022_unadjusted data product and subsets thereof, estimated as indicated in the text. Values in italicized text are assumed whereas values in plain text are calculated. A dash indicates an estimation that is not attempted. $u_{R,M}$ stands for random measurement uncertainty and $u_{R,C}$ stands for random calculation uncertainty. $u_{R,C}$ varies with seawater composition and should be recalculated for seawater conditions of interest in an analysis. Units for A_T and C_T are μmol kg^{-1}, pH$_T$ is unitless, and fCO$_2$ uncertainties are expressed as a percentage.

Estimate for	$r_{AT} = r_{CT}$	r_{pHT}	r_{fCO2}	u	A_T	C_T	pH$_T$	fCO$_2$
Estimates of random uncertainty for measurements in GLODAPv2.2022								
Assumed climate-quality	*1*	*1*	*1*	$u_{R,M,CQ}$	*2*	*2*	*0.003*	*0.50*
All data	1.2	3.6	3.9	$u_{R,M}$	2.4	2.4	0.011	1.9
Subset: spec. pH$_T$	1.0	3.5	–	$u_{R,M}$	2.0	2.0	0.010	–
Subset: pure dye spec. pH$_T$	–	1.9	–	$u_{R,M}$	–	–	0.006	–
Subset: impure dye spec. pH$_T$	–	3.3	–	$u_{R,M}$	–	–	0.010	–
Subset: potentiometric pH$_T$	2.1	4.7	–	$u_{R,M}$	4.2	4.2	0.014	–
Estimates of random uncertainties in calculations from GLODAPv2.2022_unadjusted								
calcs. from A_T and C_T	*using values from the “all”*			$u_{R,C}$	–	–	0.008	1.8
calcs. from A_T and pH$_T$	*category above*			$u_{R,C}$	–	3.2	–	1.5
calcs. from A_T and fCO$_2$				$u_{R,C}$	–	2.2	0.003	–
calcs. from C_T and pH$_T$				$u_{R,C}$	3.3	–	–	1.4
calcs. from C_T and fCO$_2$				$u_{R,C}$	2.7	–	0.003	–
calcs. from pH$_T$ and fCO$_2$				$u_{R,C}$	38.8	34.6	-	–

equations relating A_T, C_T, and pH$_T$ (which decreases from 1.8 to 1.5 when adjustments are included) but worsen the r calculated for the system of equations that relates all variables (which increases from 1.7 to 2.0). This implies that the adjustments are indeed improving internal consistency between A_T, C_T, and pH$_T$, but they are also moving the carbonate chemistry data further from the constraint provided by the fCO$_2$ data.

The r_{AT}, r_{CT}, and r_{pHT} ratios are calculated to attribute the uncertainty to A_T and C_T collectively or to pH$_T$ separately. In Table 3, we see that the uncertainty propagation and interconsistency estimates are best aligned for all A_T, C_T, and pH$_T$ when pH$_T$ is assigned an uncertainty that is a larger multiple of the climate-quality uncertainty ($r_{pHT} = 3.6$) than the other two variables ($r_{CT} = r_{AT} = 1.2$). This is not surprising given that pH$_T$ has been measured in the absence of a seawater RM using varied techniques that have evolved methodologically over the decades covered by the GLODAPv2.2022 data product (*see* Supporting Information Text S1), and the development of a RM for seawater pH remains a noted priority (Capitaine et al. 2023). By contrast, A_T and C_T measurements have comparative methodological consistency and the benefit of a widely-used seawater RM. We can go further by examining the r_{pHT} values obtained for subsets of the seawater pH$_T$ measurements. The ratio is largest ($r_{pHT} = 4.7$) for non-spectrophotometric seawater pH$_T$ measurements (e.g., potentiometric measurements or where the method used is unspecified in the metadata data product of Carter et al. (2024)). Considering only spectrophotometric seawater pH$_T$ measurements made using unpurified indicator dyes improves the ratio ($r_{pHT} = 3.3$). Limiting the comparison to those made only using purified indicator dyes improves the ratio even more ($r_{pHT} = 1.9$). Interestingly, the combination of both spectrophotometric measurement types ($r_{pHT} = 3.5$) is worse than either individually, implying that unpurified indicator dye measurements, while more variable than purified indicator dye measurements, are more consistent with each other than they are with purified indicator dye measurements.

The r_{CT} and r_{AT} value (1.0 ± 0.2) is indistinguishable from 1 for the subset of measurements paired with spectrophotometric pH$_T$ measurements, implying that the associated A_T or C_T measurements have random uncertainties comparable to the climate-quality thresholds. By contrast, the r_{pHT} value of 1.9 implies a random uncertainty of ~ 0.006 for spectrophotometric seawater pH$_T$ measurements made with purified indicator dyes, and the r_{fCO2} value of 3.9 implies a random uncertainty of 2% in fCO$_2$ measurements, both of which are greater than the climate-quality thresholds (Table 3). However, we contend that values of the climate-quality thresholds for pH$_T$ and fCO$_2$ could be plausibly quantified as ~ 0.005 for pH$_T$ and $\sim 1.5\%$ for fCO$_2$ (instead of the nominal values of 0.003 for pH$_T$ and 0.5% for fCO$_2$). This is because changes of these magnitudes result in the $\approx 1\%$ carbonate ion content change used to define climate-quality measurement criteria *except* when pairing pH$_T$ with fCO$_2$ in the carbonate ion content calculation (which is not recommended when an alternative measurement combination is available). Unfortunately, for all subsets, the random uncertainties from any two measurements used in a calculation invariably result in random calculation uncertainties that are greater than the climate-quality thresholds (Table 3).

Random uncertainties and the adjusted GLODAPv2.2022 data product

Given the mixed impact of GLODAPv2.2022 adjustments on the inter-consistency of different carbonate chemistry variable pairings, we do not separately estimate random uncertainties for the adjusted data product and suggest using the uncertainty estimates in Table 3 regardless of whether adjustments are retained. We recommend that the GLODAPv2.2022 adjustments be retained when consistency between measurements of individual carbonate chemistry variables across multiple cruises (which is quantifiably improved by the adjustments; Lauvset et al. 2022) or *inter*-consistency between A_T, C_T, and pH_T are of primary interest.

We contend our random uncertainty estimates largely reflect semi-random cruise-specific offsets rather than simple random variations that are unique to each measurement (compare Supporting Information Fig. S5c,d). Despite this, our random uncertainty estimates are smaller than the adjustment limits used by the GLODAPv2 reference team for deciding when to apply adjustments to eliminate cruise-specific offsets and thereby improve internal data product consistency. This is because a cruise can only be said to be significantly different from the existing collection when it differs by a statistically significant amount. If we use the approximation of normally distributed variations between cruises, a cruise could be considered different with meaningful confidence when its mean offset is $> 2\sigma$, or $> 2u_{R,M,IC}$, from the mean of multiple other cruises. Thus, this study supports the current adjustment limits of $4\,\mu mol\,kg^{-1}$ in A_T and C_T, which are approximately double the $u_{R,M}$ estimates (Lauvset et al. 2022), but indicates that the pH_T adjustment limit of 0.01 is low relative to the $u_{R,M,pHT}$ ($2\sigma = 0.014$–0.030).

Determining the standard uncertainty to use for an analysis

As some systematic uncertainty contributions to the combined standard uncertainty can cancel for an analysis exploring temporal or spatial changes, it can be challenging to know when to include various uncertainty contributions in an analysis. In Table 4, we provide general guidelines for how these decisions might be made for several common types of analysis. An added complexity in this table, noted earlier, is that the $u_{R,M}$ reflects a combination of random variations that are specific to each measurement ($u_{Repeatability}$) and random variations that are specific to a cruise but uniform across the measurements from the cruise (u_{Cruise}). It can be important to distinguish between these contributions to $u_{R,M}$ for some analyses. Examples include observing system simulation tests of how well a given analysis can extract an exactly-known modeled signal from model output that has been perturbed with simulated measurement uncertainties, or any analysis that uses a Monte Carlo perturbation to assess the likely impacts of measurement uncertainties. An example calculation using this approach that also demonstrates the importance of considering cruise-wide offsets (i.e., u_{Cruise}) separately from measurement repeatability (i.e., $u_{Repeatability}$) is provided as Supporting Information Text S11.

Table 4. Guidelines for estimation of random uncertainty (u_R) for analyses that rely on various collections of measured and calculated values. Subscripts for u values are given in Table 1. Values for various terms are given in Figs. 3, 4 and Table 3. These uncertainties should be propagated through calculations used by the analyst. The $u_{S,\Delta,UP}$ should also be considered as a source of systematic uncertainty even when it is included below.

Information used	u_R
Measurements made on a single cruise	*See* cruise report for $u_{Repeatability}$ estimate or conservatively use $u_{R,M}$ (Table 3)
Measurements within a region measured by multiple cruises	$u_{R,M}$ (Table 3)
Calculations within a region measured by multiple cruises	$u_{R,C}$ (Table 3)
Combination of measurements and calculations	$\sqrt{u_{R,M}{}^2 + u_{R,C}{}^2 + u_{S,\Delta,UP}{}^2}$ (Fig. 4; Table 3)
Calculations of differences relative to a threshold value (e.g., air–sea $\Delta f CO_2$ or time spent with an $\Omega > 1$), here u_{Thresh} refers to uncertainty in the threshold value	$\sqrt{u_{R,M}{}^2 + u_{S,\Delta,UP}{}^2 + u_{Thresh}{}^2}$ (Fig. 4; Table 3)
Non-discrete measurements, or calculations from such measurements (e.g., autonomous sensors)	As in associated rows above, but with different u values than those estimated herein.

As we do not, from this analysis, have a reason to favor any set of carbonate chemistry constants to the exclusion of the others, and the choice of carbonate chemistry constant sets has a large impact on the calculated $u_{S,C,IC}$ (Supporting Information Text S10), we do not believe our $u_{S,C,IC}$ estimates are robust. We therefore recommend the use of $u_{S,C,UP}$ instead of $u_{S,C,IC}$ in Table 4.

Conclusions

Uncertainty quantification in seawater carbonate chemistry variables is complicated due to the large variety of variables that are measured, calculated, and estimated to constrain seawater carbonate chemistry; evolving measurement protocols for each carbonate chemistry variable; and analyses that are commonly applied to carbonate chemistry information. Here, we quantify the uncertainties that are theoretically expected for measurements and calculations based on literature uncertainty estimates and that appear in practice in a large compilation of open-ocean carbonate chemistry measurements. We argue for the importance of quantifying both the random and systematic uncertainty contributions that collectively give rise to the combined uncertainties.

We propagate random and systematic contributions to uncertainty through carbonate chemistry calculations for the collection of measurements in the GLODAPv2.2022 data product, after removing secondary QC adjustments, and compare these propagated uncertainties to similar estimates obtained by comparing measured values with and calculated values from this data product. This analysis reveals:

1. The variability in the residuals between measured and calculated carbonate chemistry variables is greater than we would expect from propagation of climate-quality measurement uncertainties. This implies that the measurements in the GLODAPv2.2022 data product (with or without adjustments) do not collectively meet the “climate-quality” standards.
2. Some of this greater than expected discrepancy between measured and calculated carbonate chemistry variables could be due to measurements of pH_T that predate the use of purified spectrophotometric indicator dyes, and we show that the subset of measurements of A_T and C_T in the GLODAPv.2022_unadjusted data product that are accompanied by spectrophotometric seawater pH_T measurements are consistent with the climate-quality criteria (i.e., have *r* values ≤ 1 in Table 3). We find the subset of measurements with pH_T obtained using purified spectrophotometric indicator dyes to have a particularly low random uncertainty estimate for pH_T.
3. The systematic offsets between measured and calculated carbonate chemistry values in this data product are consistent with systematic uncertainties inferred from literature estimates of uncertainties in both the thermodynamic carbonate chemistry constants and in the substance contents of non-carbonate molecules that contribute to seawater acid–base chemistry.
4. The large uncertainties in carbonate chemistry constant sets, at present, prevent us from using inter-consistency calculations to assert that any one set of carbonate chemistry constants is superior to all others. However, this analysis supports several earlier constant set recommendations. We suggest (in Supporting Information Text S5) that exploration of correlations between uncertainties in constant values could be a productive topic of future research that might reduce propagated calculation uncertainty estimates.
5. As García-Ibáñez et al. (2022) did before us, we show that the secondary QC adjustments applied to the GLODAPv2.2022 data product improve the consistency of the A_T, C_T, and pH_T, but at the cost of making the seawater pH_T measurements less consistent with the (relatively few) discrete measurements of seawater fCO_2.

We contend that the overall uncertainty applicable to an analysis depends upon the input data and type of analysis (e.g., the number of distinct cruises that provided the data considered; whether measured or calculated values or both were used; and whether the analysis is concerned with overall values vs. differences regionally, over time, or from a reference value). We provide some guidelines for how a study can determine an appropriate uncertainty estimate based on the particulars of its analysis.

Data availability statement

The GLODAPv2.2022 data product is available at https://glodap.info/index.php/merged-and-adjusted-data-product-v2-2022/, with the associated adjustment table located at https://glodapv2-2022.geomar.de/. The metadata data product used to distinguish between various types of pH measurements is available as Supporting Information Materials for the companion paper (Carter et al. 2024).

References

Álvarez, M., N. M. Fajar, B. R. Carter, E. F. Guallart, F. F. Pérez, R. J. Woosley, and A. Murata. 2020. Global Ocean spectrophotometric pH assessment: Consistent inconsistencies. Environ. Sci. Technol. **54**: 10977–10988. doi:10.1021/acs.est.9b06932

BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML. 2008. Evaluation of measurement data-guide to the expression of uncertainty in measurement. https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf/cb0ef43f-baa5-11cf-3f85-4dcd86f77bd6

Bockmon, E. E., and A. G. Dickson. 2015. An inter-laboratory comparison assessing the quality of seawater carbon dioxide measurements. Mar. Chem. **171**: 36–43. doi:10.1016/J.MARCHEM.2015.02.002

Bushinsky, S. M., and others. 2019. Reassessing Southern Ocean air-sea CO_2 flux estimates with the addition of biogeochemical float observations. Glob. Biogeochem. Cycles **33**: 1370–1388. doi:10.1029/2019GB006176

Capitaine, G., D. Stoica, T. Wagener, and P. Fisicaro. 2023. Production of a reference material for seawater pH_T measurements by a National Metrology Institute. Mar. Chem. **252**: 104244. doi:10.1016/j.marchem.2023.104244

Carter, B. R., and others. 2024. Uncertainty sources for measurable ocean carbonate chemistry variables. Limnol. Oceanogr. **69**: 1–21. doi:10.1002/lno.12477

Carter, B. R., R. A. Feely, N. L. Williams, A. G. Dickson, M. B. Fong, and Y. Takeshita. 2018. Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate. Limnol. Oceanogr.: Methods **16**: 119–131. doi:10.1002/lom3.10232

Chen, B., W. J. Cai, and L. Chen. 2015. The marine carbonate system of the Arctic Ocean: Assessment of internal consistency and sampling considerations, summer 2010. Mar. Chem. **176**: 174–188. doi:10.1016/j.marchem.2015.09.007

Cullison Gray, S. E., M. D. DeGrandpre, T. S. Moore, T. R. Martz, G. E. Friederich, and K. S. Johnson. 2011. Applications of in situ pH measurements for inorganic carbon calculations. Mar. Chem. **125**: 82–90. doi:10.1016/J.MARCHEM.2011.02.005

Dickson, A. G. 2010. The carbon dioxide system in seawater: Equilibrium chemistry and measurements, p. 17–40. *In* U. Riebesell, V. J. Fabry, L. Hansson, and J.-P. Gattuso [eds.], Guide to best practices for ocean acidification research and data reporting. Publications Office of the European Union.

Dickson, A. G., and C. Goyet. 1994. Handbook of methods for the analysis of the various parameters of the carbon dioxide system in sea water. Version 2. doi:10.2172/10107773

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 Res. I: Oceanogr. Res. Pap. **34**: 1733–1743. doi:10.1016/0198-0149(87)90021-5

Dickson, A. G., and J. P. Riley. 1979. The estimation of acid dissociation constants in seawater media from potentionmetric titrations with strong base. I. The ionic product of water—K_w. Mar. Chem. **7**: 89–99. doi:10.1016/0304-4203(79)90001-X

Dickson, A. G., C. L. Sabine, and J. R. Christian [eds.]. 2007. Guide to best practices for ocean CO_2 measurements. North Pacific Marine Science Organization.

Dickson, A. G., D. J. Wesolowski, D. A. Palmer, and R. E. Mesmer. 1990. Dissociation constant of bisulfate ion in aqueous sodium chloride solutions to 250°C. J. Phys. Chem. **94**: 7978–7985. doi:10.1021/J100383A042/ASSET/J100383A042.FP.PNG_V03

Dickson, A. G., J. D. Afghan, and G. C. Anderson. 2003. Reference materials for oceanic CO_2 analysis: A method for the certification of total alkalinity. Mar. Chem. **80**: 185–197. doi:10.1016/S0304-4203(02)00133-0

Doney, S. C., and others. 2022. A research strategy for ocean carbon dioxide removal and sequestration. National Academies of Science, Engineering, and Medicine. https://www.nationalacademies.org/our-work/a-research-strategy-for-ocean-carbon-dioxide-removal-and-sequestration

Doney, S. C., D. S. Busch, S. R. Cooley, and K. J. Kroeker. 2020. The impacts of ocean acidification on marine ecosystems and reliant human communities. Annu. Rev. Env. Resour. **45**: 83–112. doi:10.1146/annurev-environ-012320-083019

Doney, S. C., V. J. Fabry, R. A. Feely, and J. A. Kleypas. 2009. Ocean acidification: The other CO_2 problem. Ann. Rev. Mar. Sci. **1**: 169–192. doi:10.1146/annurev.marine.010908.163834

Ellison, S. L. R., and A. Williams [eds.]. 2012. Quantifying uncertainty in analytical measurement, 3rd ed. Eurachem/CITAC.

Fong, M. B., and A. G. Dickson. 2019. Insights from GO-SHIP hydrography data into the thermodynamic consistency of CO_2 system measurements in seawater. Mar. Chem. **211**: 52–63. doi:10.1016/j.marchem.2019.03.006

Friedlingstein, P., and others. 2023. Global carbon budget 2023. Earth Syst. Sci. Data **15**: 5301–5369. doi:10.5194/essd-15-5301-2023

García-Ibáñez, M. I., Y. Takeshita, E. F. Guallart, N. M. Fajar, D. Pierrot, F. F. Pérez, W. J. Cai, and M. Álvarez. 2022. Gaining insights into the seawater carbonate system using discrete fCO_2 measurements. Mar. Chem. **245**: 104150. doi:10.1016/J.MARCHEM.2022.104150

Gattuso, J.-P., S. Alliouane, and P. Fischer. 2023. High-frequency, year-round time series of the carbonate chemistry in a high-Arctic fjord (Svalbard). Earth Syst. Sci. Data **15**: 2809–2825. doi:10.5194/essd-15-2809-2023

Gruber, N., and others. 2019. The oceanic sink for anthropogenic CO_2 from 1994 to 2007. Science **363**: 1193–1199. doi:10.1126/SCIENCE.AAU5153

Guallart, E. F., and others. 2022. Spectrophotometric measurement of carbonate ion in seawater over a decade: Dealing with inconsistencies. Environ. Sci. Technol. **56**: 7381–7395. doi:10.1021/acs.est.1c06083

Humphreys, M. P., E. R. Lewis, J. D. Sharp, and D. Pierrot. 2022. PyCO2SYS v1.8: Marine carbonate system calculations in Python. Geosci. Model Dev. **15**: 15–43. doi:10.5194/GMD-15-15-2022

Jiang, L. Q., and others. 2022. Best practice data standards for discrete chemical oceanographic observations. Front. Mar. Sci **8**: 705638. doi:10.3389/fmars.2021.705638

Kuliński, K., B. Schneider, K. Hammer, U. Machulik, and D. Schulz-Bull. 2014. The influence of dissolved organic matter on the acid–base system of the Baltic Sea. J. Mar. Syst. **132**: 106–115. doi:10.1016/J.JMARSYS.2014.01.011

Lauvset, S. K., and others. 2022. GLODAPv2.2022: The latest version of the global interior ocean biogeochemical data product. Earth Syst. Sci. Data **14**: 5543–5572. doi:10.5194/ESSD-14-5543-2022

Lee, K., F. J. Millero, R. H. Byrne, R. A. Feely, and R. Wanninkhof. 2000. The recommended dissociation constants for carbonic acid in seawater. Geophys. Res. Lett. **27**: 229–232. doi:10.1029/1999GL002345

Lee, K., T.-W. Kim, R. H. Byrne, F. J. Millero, R. A. Feely, and Y.-M. Liu. 2010. The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans. Geochim. Cosmochim. Acta **74**: 1801–1811. doi:10.1016/j.gca.2009.12.027

Lueker, T. J., A. G. Dickson, and C. D. Keeling. 2000. Ocean pCO_2 calculated from dissolved inorganic carbon, alkalinity, and equations for K_1 and K_2: Validation based on laboratory measurements of CO_2 in gas and seawater at equilibrium. Mar. Chem. **70**: 105–119. doi:10.1016/S0304-4203(00)00022-0

McElligott, S., R. H. Byrne, K. Lee, R. Wanninkhof, F. J. Millero, and R. A. Feely. 1998. Discrete water column measurements of CO_2 fugacity and pH_T in seawater: A comparison of direct measurements and thermodynamic calculations. Mar. Chem. **60**: 63–73. doi:10.1016/S0304-4203(97)00080-7

McLaughlin, K., and others. 2015. Core principles of the California Current acidification network: Linking chemistry, physics, and ecological effects. Oceanography **28**: 160–169. doi:10.5670/oceanog.2015.39

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

Millero, F. J. 1995. Thermodynamics of the carbon dioxide system in the oceans. Geochim. Cosmochim. Acta. **59**: 661–677. doi:10.1016/0016-7037(94)00354-O

Millero, F. J. 2007. The marine inorganic carbon cycle. Chem. Rev. **107**: 308–341. doi:https://doi.org/10.1021/CR0503557/ASSET/CR0503557.FP.PNG_V03

Morris, A. W., and J. P. Riley 1966. The bromide/chlorinity and sulphate/chlorinity ratio in sea water. Deep Sea Res. Oceanogr. Abstr. **13**: 699–705. doi:10.1016/0011-7471(66)90601-2

Müller, J. D., and others. 2023. Decadal trends in the oceanic storage of anthropogenic carbon from 1994 to 2014. AGU Adv. **4**: e2023AV000875. doi:10.1029/2023AV000875

Newton, J. A., R. A. Feely, E. B. Jewett, P. Williamson, and J. Mathis. 2015. Global Ocean acidification observing network: Requirements and governance plan, 2nd ed. GOA-ON, www.iaea.org/ocean-acidification

Olsen, A., and others. 2019. GLODAPv2.2019—An update of GLODAPv2. Earth Syst. Sci. Data **11**: 1437–1461. doi:10.5194/essd-11-1437-2019

Orr, J. C., J.-M. Epitalon, A. G. Dickson, and J.-P. Gattuso. 2018. Routine uncertainty propagation for the marine carbon dioxide system. Mar. Chem. **207**: 84–107. doi:10.1016/J.MARCHEM.2018.10.006

Patsavas, M. C., R. H. Byrne, R. Wanninkhof, R. A. Feely, and W.-J. Cai. 2015. Internal consistency of marine carbonate system measurements and assessments of aragonite saturation state: Insights from two U.S. coastal cruises. Mar. Chem. **176**: 9–20. doi:10.1016/j.marchem.2015.06.022

Perez, F. F., and F. Fraga. 1987. Association constant of fluoride and hydrogen ions in seawater. Mar. Chem. **21**: 161–168. doi:10.1016/0304-4203(87)90036-3

Raimondi, L., J. B. R. Matthews, D. Atamanchuk, K. Azetsu-Scott, and D. W. R. Wallace. 2019. The internal consistency of the marine carbon dioxide system for high latitude shipboard and in situ monitoring. Mar. Chem. **213**: 49–70. doi:10.1016/J.MARCHEM.2019.03.001

Riebesell, U., V. Fabry, L. Hansson, and J. Gattuso. 2011. Guide to best practices for ocean acidification research and data reporting.

Sharp, J. D., D. Pierrot, M. P. Humphreys, J.-M. Epitalon, J. C. Orr, E. R. Lewis, and D. W. R. Wallace. 2021. CO2SYSv3 for MATLAB. doi:10.5281/ZENODO.4774718

Sulpis, O., S. K. Lauvset, and M. Hagens. 2020. Current estimates of K_1* and K_2* appear inconsistent with measured CO_2 system parameters in cold oceanic regions. Ocean Sci. **16**: 847–862. doi:10.5194/OS-16-847-2020

Sutton, A. J., and others. 2022. Advancing best practices for assessing trends of ocean acidification time series. Front. Mar. Sci. **9**: 1045667. doi:10.3389/fmars.2022.1045667

Takeshita, Y., and others. 2021. Consistency and stability of purified meta-cresol purple for spectrophotometric pH measurements in seawater. Mar. Chem. **236**: 104018. doi:10.1016/J.MARCHEM.2021.104018

Taylor, B. N., and C. E. Kuyatt. 1994. Guidelines for evaluating and expressing the uncertainty of NIST measurement results. National Institute for Standards and Technology.

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, CO2SYS. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory.

Warner, T. B. 1971. Normal fluoride content of seawater. Deep Sea Res. Oceanogr. Abstr. **18**: 1255–1263. doi:10.1016/0011-7471(71)90030-1

Williams, N. L., and others. 2017. Calculating surface ocean pCO_2 from biogeochemical Argo floats equipped with pH: An uncertainty analysis. Glob. Biogeochem. Cycles **31**: 591–604. doi:10.1002/2016GB005541

Woosley, R. J. 2021. Evaluation of the temperature dependence of dissociation constants for the marine carbon system using pH and certified reference materials. Mar. Chem. **229**: 103914. doi:10.1016/J.MARCHEM.2020.103914

Woosley, R. J., and J.-Y. Moon. 2023. Re-evaluation of carbonic acid dissociation constants across conditions and the implications for ocean acidification. Mar. Chem. **250**: 104247. doi:10.1016/j.marchem.2023.104247

Woosley, R. J., F. J. Millero, and T. Takahashi. 2017. Internal consistency of the inorganic carbon system in the Arctic Ocean. Limnol. Oceanogr.: Methods **15**: 887–896. doi:10.1002/lom3.10208

Acknowledgments

The Ocean Carbonate System Intercomparison Forum (OCSIF, https://www.us-ocb.org/ocean-carbonate-system-intercomparison-forum/) is a working group of subject matter experts that was established in 2019 with support from the Ocean Carbon and Biogeochemistry (OCB, us-ocb.org) Project Office, which receives funding from the National Science Foundation (NSF OCE-1850983) and the National Aeronautics and Space Administration (NASA NNX17AB17G). This paper represents the compiled outcomes and discussions of OCSIF members over the past 3 yr. OCSIF has sole responsibility for the contents of this report and the questions, findings, and recommendations within. The views expressed in this report do not necessarily represent the views of the U.S. Department of Commerce or the National Institute of Standards and Technology. BRC thanks the Global Ocean Monitoring and Observing program of NOAA for funding the Carbon Data Management and Synthesis Program (Fund Ref. 100007298, program officer: Kathy Tedesco) and thereby supporting his involvement in OCSIF activities. JDS thanks the Global Ocean Monitoring and Observing program of NOAA (Award NA21OAR4310251), NOAA PMEL, and the University of Washington CICOES. RJW acknowledges the National Science Foundation Division of Ocean Sciences (Awards OCE-1923312 and OCE-2148468) and the MIT mTerra Catalyst Fund. KMS thanks NOAA AOML and the University of Miami CIMAS. AJF was supported by NOAA PMEL. MÁ thanks the IEO-CSIC internal project RADPROF. WJC, RAE, and XL thank NOAA's Ocean Acidification Program via award: NA17OAR0170332. The CICOES and PMEL contributions to this effort are numbers 2024-1348 and 5577, respectively. We thank the other members of the OCSIF working group and the community at large for helpful contributions and insights to our discussions and analyses, and thank the generations of global ocean ship-based discrete carbonate chemistry analysts and investigators who have contributed to the many cruises that were incorporated into the GLODAPv2 data product.

Conflict of Interest

The authors declare no conflicts of interest.

Submitted 05 March 2024
Revised 25 July 2024
Accepted 03 August 2024

Associate editor: Christopher Cornwall