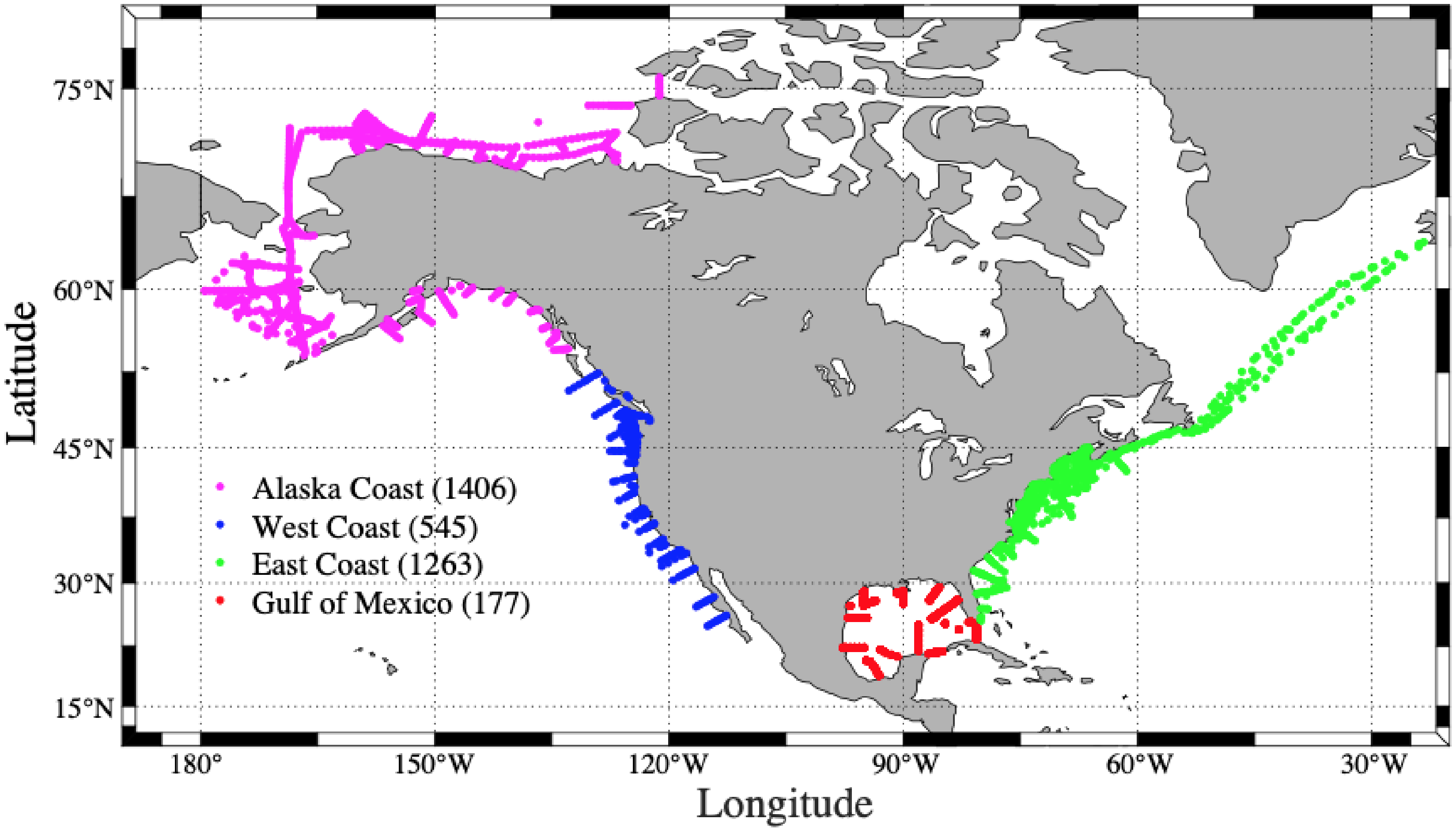

Figure 1. A map showing all the sampling profiles of the CODAP-NA data product (v2020, a total of 3,391 profiles). Magenta dots show the sampling profiles in the Alaska Coast (AC). Blue ones are for the West Coast (WC), green ones are for the East Coast (EC), and the red ones are for the Gulf of Mexico (GMx). Numbers within the parentheses indicate the total number of profiles in the region.

Steps for Primary and Secondary Quality Control of Discrete Cruise Data Sets

QC by Submitting PIs:

- Outlier detection
- Property to depth plots
- Property to property plots

Primary level QC conducted by the data product development community

Step 1. Data loading & calculation

- Calculate depth from pressure and vice versa.
- DIC + TALK to calculate pH, Carbonate ion, and fCO_2.
- Adjust pH, Carbonate ion and fCO_2 from report temperature to in-situ temperature.
- Calculate AOU and percent oxygen.
- Calculate algorithm derive values from CTDTMP, CTDSAL and CTDOXY.

Step 2. Property vs depth plots

(All variables against their sampling depths)

- CTDTEMP, CTDSAL
- CTDOXY, Oxygen, AOU, and Percent oxygen
- DIC, TALK
- Silicate, Phosphate, Nitrate
- Nitrite, Ammonium

Step 3. Property vs property

- CTDSAL vs CTDTMP
- TALK vs DIC
- TALK vs Salinity
- TALK vs Silicate
- DIC vs. fCO_2 (20°C)
- DIC vs. CTDOXY
- Phosphate vs Nitrate
- Nitrate vs Silicate
- Nutrients vs CTDOXY

Step 4. Internal consistency

(Property vs property, delta value vs depth, delta value vs property):

- CTDSAL vs Salinity (discrete)
- CTDOXY vs Oxygen (Winkler)
- Measured pH vs calculated pH
- Measured $[CO_3^{2-}]$ vs calculated $[CO_3^{2-}]$
- Measured fCO_2 vs calculated fCO_2.

Step 5 Property vs algorithm derived values

- DIC,
- TALK
- Silicate,
- Phosphate
- Nitrate
- Oxygen

Second level QC:

(focused on offsets instead of outliers)

- Crossover analyses
- Offset or factor adjustment

Step One was to ensure all of the cruise data files were ingested into NCEI's archives and documented with a rich metadata record (Jiang et al., 2015b). Maintaining a cruise data table allowing future users of the data product to access the original data files is an important component of any synthesis effort. For this study, a table with key metadata is available through this link: https://www.ncei.noaa.gov/access/ocean-acidification-data-stewardship-oads/synthesis/NAcruises.html. The following fields are listed in the table: A sequential number of the individual cruise data set (NO), expedition code (EXPOCODE), flags indicating the quality of the cruise (Cruise_flag, see Table 3), cruise identifier (Cruise_ID), Start_date, End_date, measured parameters, and links to NCEI's archive) .

Table 3. Cruise flags used for this product.

Flag value	Meaning
A	These were dedicated OA cruises that were executed following Best Practices for global ocean work as outlined in Hood et al. (2011) and other documents as can be found on GO-SHIP site*. Colloquially these are referred to as GO-SHIP quality. Traceable standards and certified reference materials were used, and deep stations (> 2500 m) were sampled to allow using near-constant deep-water concentrations as anchor points. A third inorganic carbon system parameter, such as pH or carbonate ion concentration were often measured, allowing consistency checks.
B	These are dedicated OA cruises that had onboard inorganic carbon measurements performed according to Best Practices (Dickson et al. 2007), and many other parameters to highest accuracy through use of standards and certified reference materials. However, the cruises did not necessarily have all other parameters analyzed to highest standards, such as freezing nutrients for shoreside analyses; not taking oxygen and nutrients samples on most Niskins; not normalizing CTD oxygen trace to Winkler oxygen values, insufficient metadata etc. There often are insufficient deep stations to compare data with open ocean data.
C	These were opportunistic cruises where OA parameters were measured in the water column. They include standard hydrographic, carbon, and OA parameters; T, S, O_2, nutrients, TALK, DIC, pH. Many parameters, including carbon and OA parameters were measured shoreside; CTD oxygen data were not adjusted to Winkler oxygen values. Generally, no dedicated OA personnel were onboard.
D	Underway samples only. These cruises had no CTD casts, and only had samples taken from the seawater supply line, with often a limited amount of other hydrographic parameters. T and S were obtained from thermosalinographs with limited or no salinity check samples.

(*https://www.go-ship.org/HydroMan.html)

Table 4. World Ocean Circulation Experiment (WOCE) World Hydrographic Program (WHP) (Joyce and Corry, 1994; Swift and Diggs, 2008) QC flags used for this product.

Flag value	Meaning
2	Acceptable
3	Questionable
6	Average of duplicates
9	Missing value

Step Two was to load the measurement values from the original cruise data files into MATLAB and conduct necessary calculations (Figure 2). All missing values were replaced with "-999" during this process. Variables without a QC flag from the original cruise data file were assigned an initial flag of 2 (good values, Table 4). Variables that were clearly out of range (e.g., a DIC value of < 0) were automatically assigned with a QC flag of "4" (bad values). The QC flags for all "-999" values or missing values were replaced with "9" (missing values).

Some surface samples from a few coastal cruises were collected from flow-through systems onboard research vessels, instead of Niskin bottles on sampling rosettes. In such cases, the temperature and salinity values were stored under the CTDTEMP and CTDSAL columns, respectively, although they were not measured from sensors mounted on a CTD rosette. Similarly, their sampling depth values were extracted from the metadata as the depth of the water inlet and stored under CTDPRES (Table 1). When water inlet depth information was not available, its sampling pressure was set to be 5 dbar. There is a column named "Observation_type" in the CODAP data product file to indicate whether a sample is from a "Flow-through" system or a "Niskin" bottle.

We calculated or assigned the below parameters:

(a) Sample_ID if not already included (Equation 1)

(b) depth from pressure and vice versa;

(c) recommended_Salinity_PSS78 (Table 1);

(d) conservative temperature, absolute salinity, sigma-theta;

(e) recommended_Oxygen

(f) apparent oxygen utilization (AOU);

(g) recommended_Nitrate_and_Nitrite;

(h) calculated pH, carbonate ion, and fCO_2 at *in-situ* conditions using CO2SYS from DIC and TALK, along with temperature, salinity, pressure, and nutrients; and

(i) *in-situ* pH, carbonate ion, and fCO_2 from their respective values at their measurement conditions.

Step Three was to identify outliers. Outliers were determined by visual inspection. Two types of outlier identification were used for this effort: (a) a broad-scale outlier identification by visually examining the plot of a variable against its sampling depth and other property-to-property plots, and (b) a fine-scale outlier identification based on consistency checks. Here, consistency checks refer to both the "internal consistency checks", i.e., the comparison of a measurement with its calculated value (e.g., spectrophotometrically-measured pH vs. pH calculated from other carbon parameters using CO2SYS), as well as validation checks, i.e., a measurement with one method against the same measurement made with a different method (e.g., oxygen measured from Winkler vs. a sensor, though in this case the oxygen profile is frequently adjusted to the Winkler titration values, so the measurements are not truly independent). For the broad-scale outlier identification we made plots of all variables against depth (or sigma-theta when only surface values are available), as well as these plots (Figure 2):

(a) CTDSAL against CTDTEMP
(b) TALK against Salinity
(c) TALK against Silicate
(d) TALK against DIC
(e) DIC vs. fCO_2 (20°C),
(f) DIC against CTDOXY,
(g) Phosphate vs Nitrate
(h) Nitrate vs Silicate
(i) all nutrients (silicate, phosphate, nitrate, nitrite, nitrate plus nitrite, ammonium) against CTDOXY.

Consistency check-based outlier identification was the primary way of finding outliers in this study. Consistency checks were conducted for these below variable pairs. This has been the most effective way of identifying outliers.

(a) CTDSAL vs. discrete salinity (discrete salinity as the reference value)
(b) CTDOXY vs. discrete oxygen measured from Winkler titration (Winkler oxygen as the reference value)
(c) pH measured with a spectrophotometer vs. pH calculated with CO2SYS from DIC, TALK and other parameters
(d) Carbonate ion ($[CO_3^{2-}]$) measured with a spectrophotometer vs. $[CO_3^{2-}]$ calculated with CO2SYS from DIC, TALK and other parameters

Table 7. Uncertainties of some variables based on an analysis that groups deep-water stations (>1000 m sampling depth) within 10 km radius, and 200 m depth difference. Stdev is the short for standard deviation. Numbers in parenthesis are expected errors based on propagating uncertainties in carbonate system calculations using the CO2SYS companion errors.m program. Refer to Table 1 for their full names and respective units.

Abbreviation	Units	Mean ± stdev	Percentage	Number of pairs
CTDTEMP_ITS90	°C	0.06 ± 0.07	-	54
CTDSAL_PSS78	-	0.007 ± 0.007	0.02%	53
Salinity_PSS78	-	0.003 ± 0.003	0.01%	33
CTDOXY	μmol kg^{-1}	3.4 ± 3.7	4%	51
Oxygen	μmol kg^{-1}	3.3 ± 5.9	3%	47
DIC	μmol kg^{-1}	2.4 ± 2.1	0.1%	48
TALK	μmol kg^{-1}	5.0 ± 3 .9	0.2%	45
pH_TS_insitu_measured	-	0.003 ± 0.005	-	6
pH_TS_insitu_calculated	-	0.01 ± 0.01 (0.02 ± 0.01)	-	44
Carbonate_insitu_measured	μmol kg^{-1}	1.4 ± 0.8	2%	12
Carbonate_insitu_calculated	μmol kg^{-1}	2.2 ± 2.0 (4.1 ± 0.7)	3%	44
fCO_2_insitu_measured	µatm	-	-	-
fCO_2_insitu_calculated	µatm	21 ± 22 (30 ± 16)	3%	44
Aragonite	-	0.02 ± 0.02	3%	44
Calcite	-	0.04 ± 0.04	3%	44
Revelle_Factor	-	0.14 ± 0.14	1%	44
Silicate	μmol kg^{-1}	5.3 ± 4.4	5%	50
Phosphate	μmol kg^{-1}	0.10 ± 0.13	5%	51
Nitrate	μmol kg^{-1}	0.6 ± 0.5	2%	29
Nitrite	μmol kg^{-1}	0.02 ± 0.02	69%	17
Ammonium	μmol kg^{-1}	0.06 ± 0.11	72%	29