NEON USER GUIDE TO DISSOLVED GASES IN SURFACE WATER (NEON.DP1.20097)

PREPARED BY	ORGANIZATION	DATE
Kaelin M. Cawley	AQU	12/15/2017
Keli Goodman	AQU	12/15/2017
Samantha Weintraub	TOS	12/15/2017
Stephanie Parker	AQU	12/15/2017

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	06/27/2017	Initial Release

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 DESCRIPTION

1.1 Purpose

This document provides an overview of the data included in this NEON Level 1 data product, the quality controlled product generated from raw Level 0 data, and associated metadata. In the NEON data products framework, the raw data collected in the field, for example the specific conductance of water, are considered the lowest level (Level 0). Raw data that have been quality checked via the steps detailed herein, as well as simple metrics that emerge from the raw data are considered Level 1 data products.

The text herein provides a discussion of measurement theory and implementation, data product provenance, quality assurance and control methods used, and approximations and/or assumptions made during L1 data creation.

1.2 Scope

This document describes the steps needed to generate the L1 data product Dissolved gases in surface water, the concentrations of carbon dioxide, nitrous oxide, and methane in surface water samples. This document also provides details relevant to the publication of the data products via the NEON data portal, with additional detail available in the file NEON Data Variables for Dissolved gases in surface water (NEON.DP1.20097) (AD[05]), provided in the download package for this data product.

This document describes the process for ingesting and performing automated quality assurance and control procedures on the data collected and analyzed following AOS Protocol and Procedure: Surface Water Dissolved Gas Sampling (AD[06]). The raw data that are processed in this document are detailed in the file NEON Raw Data Validation for Dissolved gases in surface water, Level 0 (NEON.DP0.20097) (AD[04]), provided in the download package for this data product. Please note that raw data products (denoted by 'DP0') may not always have the same numbers (e.g., '20093') as the corresponding L1 data product.

2   RELATED DOCUMENTS AND ACRONYMS

2.1   Associated Documents

AD[01]	NEON.DOC.000001	NEON Observatory Design (NOD) Requirements
AD[02]	NEON.DOC.001152	NEON Aquatic Sampling Strategy
AD[03]	NEON.DOC.002652	NEON Level 1, Level 2 and Level 3 Data Products Catalog
AD[04]	NEON.DP0.20097.001 _dataValidation.csv	NEON Raw Data Validation for Dissolved gases in surface water, Level 0 (NEON.DP0.20097)
AD[05]	NEON.DP1.20097.001 _variables.csv	NEON Data Variables for Dissolved gases in surface water (NEON.DP1.20097)
AD[06]	NEON.DOC.001199	AOS Protocol and Procedure: Surface Water Dissolved Gas Sampling
AD[07]	NEON.DOC.002905	AOS Protocol and Procedure: Water Chemistry Sampling in Surface Waters and Groundwater
AD[08]	NEON.DOC.000008	NEON Acronym List
AD[09]	NEON.DOC.000243	NEON Glossary of Terms
AD[10]	OS_Generic _Transitions.pdf	NEON Algorithm Theoretical Basis Document: OS Generic Transitions
AD[11]		NEON's Ingest Conversion Language (NICL) specifications

2.2   Acronyms

Acronym	Definition
MDR	mobile data recorder
ppmv	parts per million on a volumetric basis
CO_2	carbon dioxide
N_2O	nitrous oxide
CH_4	methane

3   DATA PRODUCT DESCRIPTION

The Dissolved gases in surface water (NEON.DP1.20097) data product provides carbon dioxide (CO_2), nitrous oxide (N_2O), and methane (CH_4) concentrations in reference air and water equilibrated air samples collected using AOS Protocol and Procedure: Surface Water Dissolved Gas Sampling (AD[06]). Field samples are collected from wadeable streams, non-wadeable rivers, and lakes, with sampling strategies specific to the type of water body and described briefly below. Barometric pressure from a handheld sensor, along with temperature and specific conductance of the water subsample for dissolved gas analysis, are also provided so that concentrations of dissolved gases in the original water body can be determined using Henry's law equilibrium constants. These procedures implement the guidelines and requirements described in the NEON Aquatic Sampling Strategy (AD[02]). Gas concentrations are reported at the spatial resolution of a single water sample, collected from a unique namedLocation within a sampled water body. The temporal resolution is that of a single collection date.

Measurments of surface water dissolved gases can provide valuable information for scientists, managers and decision makers regarding potential water quality responses to natural and anthropogenic changes. For example, supersaturation and undersaturation of pCO_2 in freshwaters may result from nutrient loading, point and non-point pollution sources, and groundwater inputs. Dissolved gas measurments can also provide insight into how changes in the surrounding landscape may influence aquatic ecosystem structure and function, enabling stakeholders to answer questions such as: How do increased nitrogen inputs influence stream denitrification? How are increases in CO_2 affecting the biological communities? Measuring long-term trends in concentrations of dissolved gases in surface water is part of the overall NEON biogeochemistry goal to understand how major nutrient and carbon fluxes within and across air, land and water systems change over 30 years.

3.1   Spatial Sampling Design

In wadeable streams, grab sampling in the thalweg is used to obtain water samples for dissolved gas analysis, assuming the stream channel is completely mixed and gradients in dissolved gases with water depth are minimal (Figure 1, top left). In non-wadeable rivers, grab sampling in the thalweg at 0.5 m depth is used (Figure 1, top right). For all stream types, samples are collected approximately 5 m downstream of sensor sets.

In lakes, up to three locations are designated for sampling: one in the deepest part of the lake, one near the most prominent inlet, and one near the outlet. Samples are taken from variable depths dependent on the degree of lake stratification, acknowledging that exchange and movement of dissolved gases may be limited due to density differences. Samples are collocated near lake buoy, inlet, and outlet sensor infrastructure (Figure 1, bottom left). Refer to AOS Protocol and Procedure: Water Chemistry Sampling in Surface Waters and Groundwater (AD[07]) for further details regarding field sampling.
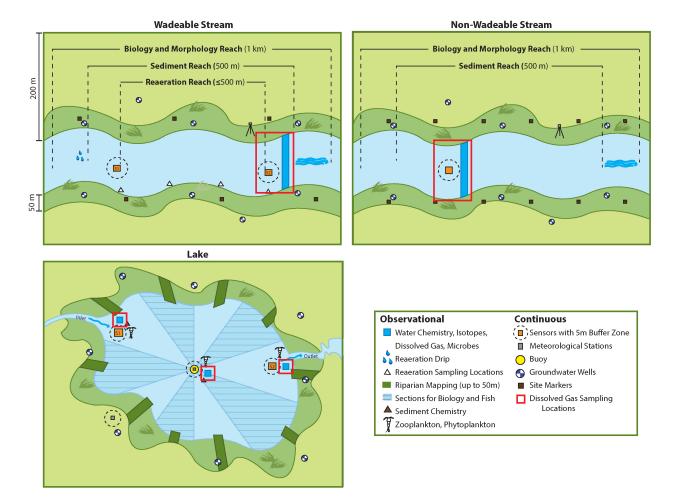
Figure 1: Summary of sampling locations for dissolved gas in water.

3.2   Temporal Sampling Design

For streams, sampling occurs up to 26 times per year, approximately every other week but guided by historical stream discharge data. Lakes are sampled 12 times per year, approximately monthly and during shoulder seasons to capture ice-on/ice-off and lake turnover events. The specific dates chosen for sampling depend upon environmental conditions at each NEON site and are outlined in site-specific sampling design documents for each Domain. Dissolved gas sampling will be conducted at the same time as other parameters such as surface water chemistry, stable isotopes, and surface water microbes.

Water subsamples for dissolved gas analysis are collected in syringes and are either immediately processed or stored in buckets of lake water while samples are collected from other stations at a given site. Once samples are collected from all stations at a site, a known volume of the water sample will be equilibrated with a known volume of ambient air by shaking the mixture for 5 minutes (Figure 2). A sample of ambient "reference" air will also be collected immediately prior to water and air equilibration. The reference air and water equilibrated air are

injected into evacuated vials which are stored upside down in water-filled centrifuge tubes for transport to an external laboratory facility for analysis of CO_2, CH_4, and N_2O concentration, which is reported as parts per million on a volumetric basis (ppmv).
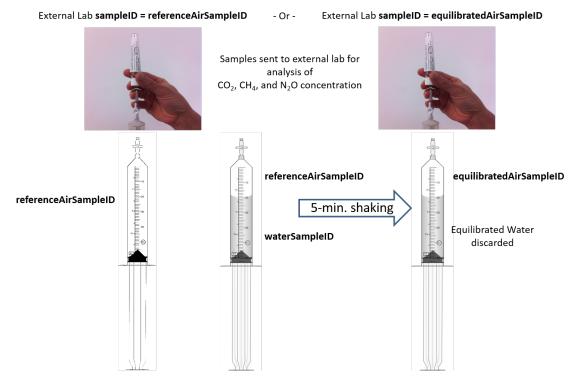

Figure 2: Diagram illustrating sample collection and field processing

3.3   Laboratory Quality Assurance and Uncertainty

NEON field procedures for collection of dissolved gas samples follow widely adopted community methods and all NEON technicians conducting this work receive proper training. For external laboratory analyses, facilities have been chosen for their use of analytical methods widely adopted by the aquatic chemistry community. Labs report the method detection limit, along with long-term analytical precison and uncertainty of standards analyzed as unknowns, for each gas species in a summary file. This allows users to interpret and model the dissolved gas data in the context of its uncertainty range. Contracted external facilities upload this summary file (sdg_externalLabSummaryData) when they begin work for NEON, then again once per year or whenever their information changes (for example, a new instrument is aquired or a change is detected in analytical precision). Additionally, NEON's Calibration/Validation department has regular procedures for auditing the quality assurance of external laboratories and their reports are available to data users.

3.4   Variables Reported

All variables reported from the field technician or laboratory (L0 data) are listed in the file NEON Raw Data Validation for Dissolved gases in surface water, Level 0 (NEON.DP0.20097) (AD[04]). All variables reported in the published data (L1 data) are also provided separately in the file, NEON Data Variables for Dissolved gases in surface water (NEON.DP1.20097) (AD[05]).

Field names have been standardized with Darwin Core terms (http://rs.tdwg.org/dwc/; accessed 16 February 2014), the Global Biodiversity Information Facility vocabularies (http://rs.gbif.org/vocabulary/gbif/; accessed 16 February 2014), the VegCore data dictionary (https://projects.nceas.ucsb.edu/nceas/projects/bien/wiki/VegCore; accessed 16 February 2014), where applicable. NEON AOS spatial data employs the World Geodetic System 1984 (WGS84) for its fundamental reference datum and Earth Gravitational Model 96 (EGM96) for its reference gravitational ellipsoid. Latitudes and longitudes are denoted in decimal notation to six decimal places, with longitudes indicated as negative west of the Greenwich meridian.

Some variables described in this document may be for NEON internal use only and will not appear in downloaded data.

3.5   Spatial Resolution and Extent

The finest resolution at which dissolved gas spatial data are reported is a single sample collected from a unique namedLocation. Overall, this results in a spatial hierarchy of:

waterSampleID (unique ID given to the individual water sample) → **namedLocation** (ID of the sampling location) → **siteID** (ID of NEON site) → **domainID** (ID of a NEON domain).

namedLocation in wadeable and non-wadeable (i.e. rivers) streams is indicated as 'ss', and namedLocations for lakes are designated as 'in', 'ot', 'c0', and 'c1','c2', and 'c3', as needed (if center is stratified), with 'c1' being the top layer.

The basic spatial data included in the data downloaded include the latitude, longitude, and elevation of the namedLocation where sampling occurred, plus associated uncertainty due to GPS error. Shapefiles of all NEON Aquatic Observation System sampling locations can be found in the Document Library: http://data.neonscience.org/documents. If for some reason samples cannot be taken at the designated stations, the coordinates for the center of the site will be returned. In this case, the actual coordinates of the sampling location are recorded in the **altLocation**, **altLatitude** and **altLongitude** fields.

3.6   Temporal Resolution and Extent

The finest resolution at which dissolved gas temporal data are reported is the **collectDate**, a single date on which water samples were collected. The total number of sampling events per year is expected to be 26 per wadeable and non-wadeable stream and 12 per lake.

The NEON Data Portal currently provides data in monthly files for query and download efficiency. Queries including any part of a month will return data from the entire month. All queries, regardless of the date range specified, will include a copy of sdg_externalLabSummaryData, which provides summary information from the external lab

about method detection limits, equipment used, and precision and accuracy. Code to stack files across months is available here: https://github.com/NEONScience/NEON-utilities

3.7 Associated Data Streams

The 'super parent' sample of the data from the Dissolved gases in surface water product is shared with several other related data products. **siteID**, **namedLocation** and **collectDate** (or **parentSampleID** of the xxx_fieldSuperParent table in each data product) are the linking variables that tie samples and their associated metadata to those found in the Chemical properties of surface water (NEON.DP1.20093), Stable isotope concentrations in surface waters (DP1.20206.001), and Surface water microbe cell count (DP1.20138.001) data products.

3.8 Product Instances

The NEON Observatory contains 34 aquatic sites, consisting of 24 wadeable streams, 3 non-wadeable streams (rivers), and 7 lakes.

Dissolved gas sampling yields one unique equilibrated air sample per namedLocation per sampling event. There will be up to 26 sampling events per year in wadeable and non-wadeable streams and up to 12 sampling events in lakes. Thus in wadable streams, there will be up to 26 unique equilibrated air sample records per site per year, non-wadeable streams will also yield up to 26 records per year, and lakes will yield 36-60 records per year, depending on degree of stratification. Observatory-wide, this will yield a total of 1226-1394 dissolved gas records per year.

NOTE: Replicate samples may be taken of a small percentage of our samples. If replicate samples are taken, there will be one unique sample per **replicateNumber** per namedLocation per sampling event, and the sample ID(s) of the replicate sample(s) will have the **replicateNumber** appended to the end.

3.9 Data Relationships

Each siteID x namedLocation combination is sampled at least once per event (one record expected per parentSampleID in sdg_fieldSuperParent). A record from sdg_fieldSuperParent may have zero or one child records in sdg_fieldData, depending on whether a water sample was sucessfully collected. In the event that a water sample cannot be taken, a record will still be created in sdg_fieldSuperParent, and sdg_fieldSuperParent.**samplingImpractical** will be something other than NULL, but there will be no corresponding record in sdg_fieldData. Each record from sdg_fieldData is expected to have one child record in sdg_fieldDataProc, which is a mixture of a water sample created in sdg_fieldData and reference air sample created in sdg_fieldDataAir. Each record from sdg_fieldDataProc and sdg_fieldDataAir is expected to have one child record in sdg_externalLabData. However, duplicates and/or missing data may exist where protocol and/or data entry abberations have occurred; *users should check data carefully for anomalies before joining tables*.

sdg_fieldSuperParent.csv - > One record expected per parentSampleID.

sdg_fieldData.csv - > One record expected per parentSampleID per collectDate (day of year, local time), generates a single waterSampleID to be used in dissolved gas equilibration procedure

sdg_fieldDataAir.csv - > One referenceAirSampleID record expected per namedLocation x collectDate combination

sdg_fieldDataProc.csv - > One record expected per waterSampleID, generates a single equilibratedAirSampleID

sdg_externalLabData.csv -> One record expected per equilibratedAirSampleID and referenceAirSampleID, associated with external laboratory gas species concentration measurments

sdg_externalLabSummaryData.csv - > One record expected per laboratoryName x analyte x method x labSpecificStartDate combination. Can use corresponding variables in externalLabData tables to associate sample data with relevant uncertainty values and method detection limits

sampleIDs and **sampleBarcodes** will be generated for each sampling event. After field processing and shipment to external labs are complete, any physical sample that remains will be discarded.

3.10   Special Considerations

For users interested in calculating dissolved gas concentrations with more precision, solubility levels at the time of sampling, or percent saturation for the dissolved gases, see Appendix and references cited therein. The information they provide is outside the scope of the basic QA/QC of L0 data used to create this L1 data product. However, the field and laboratory methods chosen were designed to enable these calculations by data users, thus a discussion of them is provided.

4   DATA QUALITY

4.1   Data Entry Constraint and Validation

Many quality control measures are implemented at the point of data entry within a mobile data entry application (field data) and web user interface (UI, lab data). For example, data formats are constrained and data values controlled through the provision of dropdown options, which reduces the number of processing steps necessary to prepare the raw data for publication (Figure 3). An additional set of constraints are implemented during the process of ingest into the NEON database. The product-specific data constraint and validation requirements built into data entry applications and database ingest are described in the document NEON Raw Data Validation for Dissolved gases in surface water, Level 0 (NEON.DP0.20097), provided with every download of this data product. Contained within this file is a field named 'entryValidationRulesForm', which describes syntactically the validation rules for each field built into the data entry application. Data entry constraints are described in Nicl syntax in the validation file provided with every data download, and the Nicl language is described in NEON's Ingest Conversion Language (NICL) specifications ([AD[11]).

Data collected prior to 2017 were processed using a paper-based workflow that did not implement the full suite of quality control features associated with the interactive digital workflow.
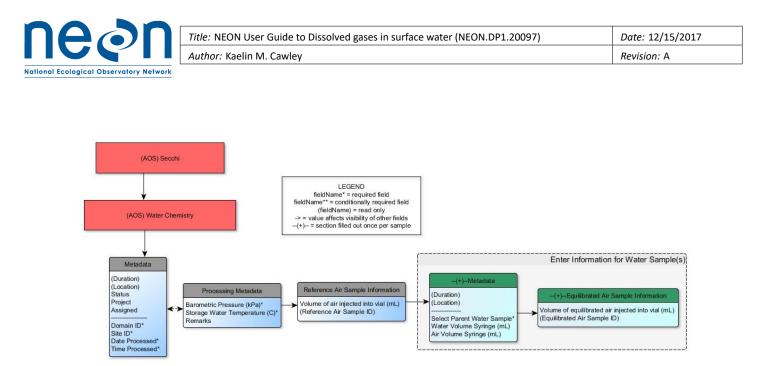
Figure 3: Diagram illustrating the mobile data entry application

4.2   Automated Data Processing Steps

Following data entry into a mobile application or web user interface, the steps used to process the data through to publication on the NEON Data Portal are detailed in the NEON Algorithm Theoretical Basis Document: OS Generic Transitions (AD[10]).

4.3   Data Revision

All data are provisional until a numbered version is released; the first release of a static version of NEON data, annotated with a globally unique identifier, is planned to take place in 2020. During the provisional period, QA/QC is an active process, as opposed to a discrete activity performed once, and records are updated on a rolling basis as a result of scheduled tests or feedback from data users. The Change Log section of the data product readme, provided with every data download, contains a history of major known errors and revisions.

4.4   Quality Flagging

The **dataQF** field in each data record is a quality flag for known errors applying to the record. Please see the *Special Considerations* section of this document for a list of known errors that may be present in the data, and below for an explanation of **dataQF** codes specific to this product.

fieldName	value	definition
dataQF	legacyData	Data recorded using a paper-based workflow that did not implement the full suite of quality control features associated with the interactive digital workflow

4.5 Analytical Facility Data Quality

Data analyses conducted for dissolved gas in surface water conform to the current data quality standards used by practitioners. On at least an annual basis a QAQC summary file is reported that contains a description of the method, method detection limit, precision, and measurement uncertainty associated with a specific time range and analyte. This is published in the **sdg_externalLabSummaryData** table in the expanded download package. In the **sdg_externalLabData** table, as part of the basic download package, a run detection limit and precision are also published. The run detection limit and precision may be different from the method detection limit and precision since they are based on the specific standards and settings used for a specific run rather than the method as a whole.

5 BIBLIOGRAPHY

1. Pawlowicz, R. (2008) Calculating the conductivity of natural waters. Limnology and Oceanography: Methods, vol. 6, pp. 489-501.
2. Sander (2015) Compilation of Henry's law constants (version 4.0) for water as a solvent. Atmospheric Chemistry and Physics, DOI: 10.5194/acp-15-4399-2015
3. Weiss, R.F. (1974) Carbon dioxide in water and seawater: The solubility of a non-ideal gas. Marine Chemistry, Vol. 2, pp.203-215.
4. Weiss, R.F. and Price, B.A. (1980) Nitrous oxide solubility in water and seawater. Marine Chemistry, vol. 8, pp.347-359.
5. Yamamato, S., Alcauskas, J.B., Crozier, T.E. (1976) Solubility of methane in distilled water and seawater. Journal of Chemical and Engineering Data, Vol. 21, pp. 78-80.

6 APPENDIX

Notes about calculating dissolved gas concentrations in surface water from reference air and equilibrated air gas concentrations. **ConcentrationCO2$_{gas}$, ConcentrationCH4$_{gas}$** and **ConcentrationCN2O$_{gas}$** are the external laboratory concentrations for the **equilibratedAir** sample and **ConcentrationCO2$_{air}$, ConcentrationCH4$_{air}$** and **ConcentrationN2O$_{air}$** are the external laboratory concentrations for the **referenceAir** sample.

1. For all gasses:
    a. The gas constant, R, is:
$$R = 8.3122598 \; L \; kPa \; K^{-1} \; mol^{-1}$$

    b. Temperature adjusted Henry's Law Constant is:
$$H(T) = H_0 e^{\frac{dln(H)}{d(\frac{1}{T})}(\frac{1}{T} - \frac{1}{T_0})}$$

    c. Dissolved gas concentration in water is:
$$C_w = \frac{mol_{gas}wat}{volH_20} = \frac{mol_{gas}aireq + mol_{gas}wateq - mol_{gas}air}{vol_{H_20}}$$

    d. where, Totalmol$_{gas}$

$$Totalmol_{gas} = mol_{gas}air + mol_{gas}wat = mol_{gas}aireq + mol_{gas}wateq$$

    e. Where, mol$_{gas}$aireq

$$mol_{gas}aireq = \frac{ppmv_{aireq} \cdot 10^{-6} \cdot BP \cdot vol_{aireq}}{RT}$$

    f. Where, mol$_{gas}$air

$$mol_{gas}air = \frac{ppmv_{air} \cdot 10^{-6} \cdot BP \cdot vol_{air}}{RT}$$

    g. Where, mol$_{gas}$wateq

$$mol_{gas}wateq = H(T) \cdot ppmv_{aireq} \cdot 10^{-6} \cdot BP \cdot vol_{H_2O}$$

2. Calculate dissolved carbon dioxide (CO2) concentration.
    a. The Henry's Law constant used is (Sander, 2015):

$$H_0 = 3.3 \cdot 10^{-4} \ mol \ m^{-3} \ Pa^{-1}$$

    b. The temperature relationship is:
$$\frac{dln(H)}{d(\frac{1}{T})} = 2400 \ K$$

    c. The overall equation is:

$$C_{CO_2} = \mathbf{ptBarometricPressure} \cdot 10^{-6} \cdot$$
$$\left[\frac{(\mathbf{gasVolumeSyringe})(\mathbf{concentrationCO2gas} - \mathbf{concentrationCO2air})}{R \cdot (\mathbf{storageWaterTemp} + 273.15) \cdot \mathbf{waterVolumeSyringe}} + \right.$$
$$\left. 3.3 \cdot 10^{-4} e^{\left(2400\left(\frac{1}{\mathbf{storageWaterTemp}+273.15} - \frac{1}{298.15}\right)\right)} \cdot \mathbf{concentrationCO2gas} \right]$$

3. Calculate dissolved nitrous oxide concentration.
    a. The Henry's Law constant used is (Sander, 2015):

$$H_0 = 2.4 \cdot 10^{-4} \ mol \ m^{-3} \ Pa^{-1}$$

    b. The temperature relationship is:
$$\frac{dln(H)}{d(\frac{1}{T})} = 2700 \ K$$

    c. The overall equation is:

$$C_{N_2O} = \mathbf{ptBarometricPressure} \cdot 10^{-6} \cdot$$
$$\left[\frac{(\mathbf{gasVolumeSyringe})(\mathbf{concentrationN2Ogas} - \mathbf{concentrationN2Oair})}{R \cdot (\mathbf{storageWaterTemp} + 273.15) \cdot \mathbf{waterVolumeSyringe}} + \right.$$
$$\left. 2.4 \cdot 10^{-4} e^{\left(2700\left(\frac{1}{\mathbf{storageWaterTemp}+273.15} - \frac{1}{298.15}\right)\right)} \cdot \mathbf{concentrationN2Ogas} \right]$$

4. Calculate dissolved methane concentration.
    a. The Henry's Law constant used is (Sander, 2015):

$$H_0 = 1.4 \cdot 10^{-4} \ mol \ m^{-3} \ Pa^{-1}$$

    b. The temperature relationship is:

$$\frac{dln(H)}{d(\frac{1}{T})} = 1900 \ K$$

    c. The overall equation is:

$$C_{CH_4} = \mathbf{ptBarometricPressure} \cdot 10^{-6} \cdot$$
$$\left[\frac{(\mathbf{gasVolumeSyringe})(\mathbf{concentrationCH4gas} - \mathbf{concentrationCH4air})}{R \cdot (\mathbf{storageWaterTemp} + 273.15) \cdot \mathbf{waterVolumeSyringe}} + \right.$$
$$\left. 1.4 \cdot 10^{-4} e^{\left(1900\left(\frac{1}{\mathbf{storageWaterTemp}+273.15} - \frac{1}{298.15}\right)\right)} \cdot \mathbf{concentrationCH4gas} \right]$$

6.1   Using sensor data for barometric pressure

Barometric pressure measured by the handheld devices published here are accurate +/- 0.0667 kPa (5 mmHg). Meteorological station data collected by sensors deployed at NEON Aquatic Sites often have reported uncertainty that is an order of magnitude lower. Therefore, if an application requires more precise estimates of dissolved gas concentrations, consider downloading barometric pressure sensor data from the NEON data portal (data.neonscience.org) and calculating dissolved gas concentrations using those values.

6.2   Accounting for salinity in dissolved gas concentrations

The NEON Aquatic Sites are comprised of freshwater lakes, yet some lakes occasionally have specific conductance values that indicate salinity could be significant when calculating dissolved gas concentrations. Based on data collected to date, it is estimated that these elevated specific conductance values could result in up to a 2% difference in calculated dissolved gas concentrations with and without salinity-based adjustments to Henry's Law or Bunsen coefficients.

Salinity for freshwater environments can be determined from the concentration of dissolved solutes and used along with Bunsen coefficients to calculate dissolved gas concentrations using reference air and equilibrated gas concentrations. The salinity of freshwater environments can vary depending on the composition of dissolved ions and therefore the salinity and conductivity need to be calculated for a given water body. The Chemical properties of surface water (NEON.DP1.20093) data product includes concentrations of dissolved ions for each NEON water body. Bunsen coefficients and their relationships with temperature and salinity for CO_2 can be fonud in Weiss, 1974, those for methane can be found in Yamamoto et al., 1976, and values for nitrous oxide can be found in Weiss and Price, 1980.