**Project Tile: Web Portal for Concentration and Flux of CO2 in Global Inland Waters**

**Phase 2 Project Development of the CO2 Web Portal**

**Two directions:**

**1. Refining and developing backend database models**

**2. Developing frontend views of the CO2 web portal**

See details below:

Note:

(1) In the following descriptions, measurement refers to column-wise or field-wise variable (e.g., CO2, CO2 flux, and DOC, etc.) in the database and sample refers to row-wise or case-wise serial measurements. One sample can have many measurements.

(2) Refer to docs/Date\_Template.xlsx when reading the following descriptions about the database model.

(3) Refer to docs/Web portal.pptx when reading about development of frontend views.

**1. Refining and developing backend database models:**

**Overall description**

The backend database contains the following **six** **database tables**: Site Location, Physical Properties, Greenhouse Gas Concentrations, Greenhouse Gas Fluxes, Gas Transfer Velocity, and Other Measurements.

The Site Location table contains fields such as *Site Type*, *Latitude*, *Longitude*, *Altitude* and the geometry field *Point*.

Physical properties refer to physical properties of the surveyed water body and/or the surrounding environment while taking field measurements of greenhouse gases. The Physical Properties table contains the following 8 fields: *Width*, *Depth*, *Surface Area*, *Discharge*, *Flow Velocity*, *Wind Speed*, *Water Temperature* and *Air Temperature*.

The Greenhous Gas Concentrations table contains concentration measurements of three types of greenhouse gases, namely, *CO2*, *CH4* and *N2O*, corresponding to the 3 fields in this database table.

The Greenhouse Gas Fluxes table contains measurements of greenhouse gas emission fluxes from the water body’s surface. It also contains 3 fields, *CO2 Flux*, *CH4 Flux* and *N2O Flux*, corresponding to the three greenhouse gases.

Gas transfer velocity (or, k) refers to the rate at which a gas is evaded from the water-air interface. It is usually calculated according to the Fick’s Law of diffusion (see <https://www.comsol.com/multiphysics/diffusion-equation>), or in our case, dividing a field measurement of gas flux by the gas concentration difference between the water body and the overlying atmosphere. It is a convention to normalize the measured gas transfer velocity to a Schmidt number of 600, called k600, in order to be comparable among different gases. Theoretically, k600 measured from different gases at the same location and time should be the same. In reality, however, k600 from different gases can vary significantly due to large uncertainty associated with this measurement. In addition, evasion of CH4 from water surfaces includes bubble fluxes (i.e., non-diffusive), making k600 calculated from CH4 fluxes deviates significantly from diffusive k600 in that case (read *docs/ Raymond\_et\_al-2012.pdf* if you want to know about gas transfer velocity). Considering the above factors, the Gas Transfer Velocity table therefore contains 6 six fields, two for each gas, namely, *k-CO2*, *k600-CO2*, *k-CH4*, *k600-CH4*, *k-N2O* and *k600-N2O*.

Other measurements refer to measurements that are not directly related to the greenhouse gas concentration or flux but are closely related to its production or consumption in freshwaters. The Other Measurements table contains 7 fields: *DOC*, *POC*, *TOC*, *pH*, *Alkalinity*, *DO* and *Chl a*.

Besides fields contained in each database table, two fields, *Date Time* and *User Note* are sample-specific and shared by all database tables expect the Site Location table.

In addition, the geometry field *Point* in the Site Location table serves as the foreign key and shared by the other database tables.

All fields listed above are optional for users to provide except the Latitude and Longitude fields.

**Fields description**

1.1 Shared Fields (among sample-specific database tables)

*1.1.1 Date Time*

The default format for *Date Time* is yyyy-mm-dd hh:mm:ss. We provide data template for users to fill in. However, users might still have various formats for *Date Time* and in some cases even not be a date or time (e.g., date or time range). Try read *Date Time* to see whether it can be read correctly and transformed to the correct format (i.e., yyyy-mm-dd hh:mm:ss). If not, prompt with a notice to remind us to manually modify the *Date Time* input. Time zone is not considered at this stage.

1.1.2 *User Note*

Text field for storing user’s notes for sampling and measurement, etc.

1.2 the Site Location table

1.2.1 *Site Type*

Change the Site Description field to Site Type in the original database model. We’ll just need to differentiate among different types of inland water bodies from the site description: Stream, River, Lake, Reservoir, Estuary, Pond, Wetland, Floodplain and Others. We therefore need to deploy a procedure for filtering through user site description, grabbing key words and assigning a correct type to the *Site Type* field in our database. See below:

1. If there is/are key word(s) reservoir, assign Reservoir to *Site Type* in the database;
2. Else if there is/are key word(s) estuary, assign Estuary to *Site Type* in the database;
3. Else if there is/are key word(s) lake, assign Lake to *Site Type* in the database;
4. Else if there is/are key word(s) pond, assign Pond to *Site Type* in the database;
5. Else if there is/are key word(s) stream, assign Stream to *Site Type* in the database;
6. Else If there is/are key word(s) river, assign River to *Site Type* in the database;
7. Else if there is/are key word(s) floodplain, assign Floodplain to *Site Type* in the database;
8. Else if there is/are key word(s) wetland, Marsh, swamp, bog, fen, meadow, assign Wetland to *Site Type* in the database;
9. Else if, assign Others.

1.2.2 *Latitude* and *Longitude*

The default unit for *Latitude* and *Longitude* in the database is decimal degree. We advise users to convert their coordinates to decimal degrees before submitting. In cases where the *Latitude* and *Longitude* inputs are not decimal degrees, prompt with a notification for manually converting the inputs.

1.2.3 *Altitude*

The default unit for altitude is meters above sea level (m). Users also have options of foot or kilometer (km). Detect and convert user units to m before storage in the database.

1 foot = 0.3048 meters, 1 kilometer = 1000 meters.

1.2.4 *Point*

The geometry field *Point* is a geometry representation of the site location in the database. The field also serves as the foreign key in other tables.

1.3 the Physical Properties table

1.3.1 *Width*

*Width* refers to the width of a water body like river, stream, estuary, or reservoir, etc. Default unit for *Width* is meter (m). Users also have options of km, foot, mile. Detect and convert user unit to m before storage into database.

1km = 1000 m, 1 foot = 0.3048 m, and 1 mile = 1609.34 m.

1.3.2 *Depth*

*Depth* refers to water depth of the surveyed water body. Default unit for *Depth* is meter (m). Users also have options of centimeter (cm), foot, inch. Detect and convert user unit to m before storage into database.

1 cm = 0.01 m, 1 foot = 0.3048 m, and 1 inch = 0.0254 m.

1.3.3 *Surface Area*

*Surface Area* refers to surface area covered by water of the surveyed water body. It applies to water bodies that are hard to describe with a width such as lake, wetland, and pond, etc. Default unit for *Surface Area* is square kilometer (m2). Users also have options of km2, ft2, and mile2. Detect and convert user unit to m2 before storage into database.

1 km2 = 1,000,000 m2, 1 ft2 = 0.0929 m2, and 1 mile2 = 2,589,990 m2.

1.3.4 *Discharge*

*Discharge* applies to water bodies that have a unidirectional net flow of water such as stream, river, estuary, and reservoir, etc. The default unit for *Discharge* is m3 s-1 but users also have options of L s-1, ft3 s-1, km3 yr-1. Detect and convert user unit to m3 s-1 before storage into database.

1 L s-1 = 0.001 m3 s-1, 1 ft3 s-1 = 0.028317 m3 s-1, and 1 km3 yr-1 = 31.71 m3 s-1.

1.3.5 *Flow Velocity*

*Flow Velocity* refers to water flow velocity of a flowing water body such as stream, river, estuary, and reservoir, etc. The default unit for *Flow Velocity* is m s-1 but users also have options of ft s-1, and cm s-1. Detect and convert user unit to m s-1 before storage into database.

1 ft s-1 = 0.3048 m s-1, and 1 cm s-1 = 0.01 m s-1.

1.3.6 *Wind Speed*

*Wind Speed* refers to wind speed measured at the location and time of the measurement. The default unit for *Wind Speed* is m s-1 but users also have options of ft s-1. Detect and convert user unit to m s-1 before storage into database.

1 ft s-1 = 0.3048 m s-1.

1.3.7 *Water Temperature*

*Water Temperature* refers to the temperature of water measured at the location and time of sampling. Default unit for *Water Temperature* is °C but users also have the option of °F. Detect and convert user unit to °C before storage into database.

(°F – 32) × 5 / 9 = °C.

1.3.8 *Air Temperature*

*Air Temperature* refers to the temperature of overlying air measured at the location and time of sampling. Default unit for *Air Temperature* is °C but users also have the option of °F. Detect and convert user unit to °C before storage into database.

(°F – 32) × 5 / 9 = °C.

1.4 the Greenhouse Gas Concentration table

1.4.1 *CO2*

*CO2* refers to dissolved CO2 concentration in the surveyed water body. Dissolved CO2 concentration in freshwater is often expressed as equivalent atmospheric partial pressure (µatm) or part per million (ppm). The default unit for *CO2* is µatm but users also have options of ppm, ppmv, µmol L-1, mmol L-1 and nmol L-1. Detect and convert user unit to µatm before storage into database. 1 ppm (or ppmv) = 1 µatm.

For conversion between µatm and mol L-1, we need to introduce a temperature sensitive Henry’s Law constant (H): - log10 (H) = - 700000 × T2 + 0.015 × T + 1.11, where T is *Water Temperature*.

µatm = µmol L-1 / H

µatm = mmol L-1 × 1000 / H

µatm = nmol L-1 / (H × 1000)

1.4.2 *CH4*

*CH4* refers to dissolved CH4 concentration in the surveyed water body. The default unit for *CH4* is nmol L-1 but users also have options of µmol L-1. Detect and convert user unit to µatm before storage into database.

1 µmol L-1= 1000 nmol L-1.

1.4.3 *N2O*

*N2O* refers to dissolved N2O concentration in the surveyed water body. The default unit for *N2O* is nmol L-1 but users also have options of µmol L-1. Detect and convert user unit to µatm before storage into database.

1 µmol L-1= 1000 nmol L-1.

1.5 the Greenhouse Gas Flux table

1.5.1 *CO2 Flux*

*CO2 Flux* refers to the amount of CO2 evaded from the water surface in a certain time. Default unit for *CO2 Flux* in the database is g C m-2 yr-1, but users also have options of mg C m-2 d-1, mg C m-2 hr-1, µg C m-2 s-1, mmol m-2 d-1, µmol m-2 hr-1, and µmol m-2 s-1. Detect and convert user unit to g C m-2 yr-1 before storage into database.

1 mg C m-2 d-1 = 0.365 g C m-2 yr-1, 1 mg C m-2 hr-1 = 8.76 g C m-2 yr-1, 1 µg C m-2 s-1 = 31.536 g C m-2 yr-1, 1 mmol m-2 d-1= 4.38 g C m-2 yr-1, 1 µmol m-2 hr-1 = 0.10512 g C m-2 yr-1, 1 µmol m-2 s-1 = 378.432 g C m-2 yr-1.

1.5.2 *CH4 Flux*

*CH4 Flux* refers to the amount of CH4 evaded from the water surface in a certain time. Default unit for *CH4 Flux* in the database is g C m-2 yr-1, but users also have options of mg C m-2 d-1, mg C m-2 hr-1, µg C m-2 s-1, mmol m-2 d-1, µmol m-2 hr-1, and µmol m-2 s-1. Detect and convert user unit to g C m-2 yr-1 before storage into database.

1 mg C m-2 d-1 = 0.365 g C m-2 yr-1, 1 mg C m-2 hr-1 = 8.76 g C m-2 yr-1, 1 µg C m-2 s-1 = 31.536 g C m-2 yr-1, 1 mmol m-2 d-1= 4.38 g C m-2 yr-1, 1 µmol m-2 hr-1 = 0.10512 g C m-2 yr-1, 1 µmol m-2 s-1 = 378.432 g C m-2 yr-1.

1.5.3 *N2O Flux*

*N2O Flux* refers to the amount of N2O evaded from the water surface in a certain time. Default unit for *N2O Flux* in the database is g N m-2 yr-1, but users also have options of mg N m-2 d-1, mg N m-2 hr-1, µg N m-2 s-1, mmol m-2 d-1, µmol m-2 hr-1, and µmol m-2 s-1. Detect and convert user unit to g N m-2 yr-1 before storage into database.

1 mg N m-2 d-1 = 0.365 g N m-2 yr-1, 1 mg N m-2 hr-1 = 8.76 g N m-2 yr-1, 1 µg N m-2 s-1 = 31.536 g N m-2 yr-1, 1 mmol m-2 d-1= 5.11 g N m-2 yr-1, 1 µmol m-2 hr-1 = 0.12264 g N m-2 yr-1, 1 µmol m-2 s-1 = 441.504 g N m-2 yr-1.

1.6 the Gas Transfer Velocity table

1.6.1 *k-CO2*

*k-CO2* refers to gas transfer velocity calculated from measured CO2 flux and concentration. *k-CO2* has the default unit of meter per day (m d-1) in the database but user has another option of cm hr-1. Convert the unit of *k-CO2* to m d-1 before storage into the database.

1 m d-1 = 25 / 6 cm hr-1.

1.6.2 *k600-CO2*

*k600-CO2* refers to gas transfer velocity calculated from measured CO2 flux and concentration which is further normalized to a Schmidt number of 600. *k600-CO2* has the default unit of meter per day (m d-1) in the database but user has options of cm hr-1. Convert the unit of *k600-CO2* to m d-1 before storage into the database.

1 m d-1 = 25 / 6 cm hr-1.

1.6.3 *k-CH4*

*k- CH4* refers to gas transfer velocity calculated from measured CO2 flux and concentration. *k- CH4* has the default unit of meter per day (m d-1) in the database but user has another option of cm hr-1. Convert the unit of *k- CH4* to m d-1 before storage into the database.

1 m d-1 = 25 / 6 cm hr-1.

1.6.4 *k600-CH4*

*k600-CH4* refers to gas transfer velocity calculated from measured CO2 flux and concentration which is further normalized to a Schmidt number of 600. *k600-CH4* has the default unit of meter per day (m d-1) in the database but user has options of cm hr-1. Convert the unit of *k600- CH4* to m d-1 before storage into the database.

1 m d-1 = 25 / 6 cm hr-1.

1.6.3 *k-N2O*

*k-N2O* refers to gas transfer velocity calculated from measured CO2 flux and concentration. *k-N2O* has the default unit of meter per day (m d-1) in the database but user has another option of cm hr-1. Convert the unit of *k-N2O* to m d-1 before storage into the database.

1 m d-1 = 25 / 6 cm hr-1.

1.6.7 *k600-N2O*

*k600-N2O* refers to gas transfer velocity calculated from measured CO2 flux and concentration which is further normalized to a Schmidt number of 600. *k600-N2O* has the default unit of meter per day (m d-1) in the database but user has options of cm hr-1. Convert the unit of *k600-N2O* to m d-1 before storage into the database.

1 m d-1 = 25 / 6 cm hr-1.

1.7 the Other Measurements table

1.7.1 *DOC*

*DOC* refers to dissolved organic carbon concentration in the surveyed water body. The default unit for *DOC* is mg L-1 in the database but user has another option of µmol L-1. Convert the unit to mg L-1 before storage in the database.

12 mg L-1 = 1000 µmol L-1.

1.7.2 *POC*

*POC* refers to particulate organic carbon concentration in the surveyed water body. The default unit for *POC* is mg L-1 in the database but user has another option of µmol L-1. Convert the unit to mg L-1 before storage in the database.

12 mg L-1 = 1000 µmol L-1.

1.7.3 *TOC*

*TOC* refers to total organic carbon concentration in the surveyed water body. The default unit for *TOC* is mg L-1 in the database but user has another option of µmol L-1. Convert the unit to mg L-1 before storage in the database.

12 mg L-1 = 1000 µmol L-1.

*TOC* = *POC* + *DOC*. Therefore, for storing the three variables in the database or upon a query on the database:

If two of the three exit, the third can be calculated from the exiting two.

If all of the three exit, respond with the user’s original inputs.

If one of the three exits, simply store them in the database.

1.7.4 *pH*

The *pH* value of water at the sampling location and time. pH has no unit.

1.7.5 *Alkalinity*

*Alkalinity* has the default unit of mg CaCO3 L-1 in the database table. Prompt with notification if user unit is not mg CaCO3 L-1.

1.7.6 *DO*

*DO* refers to dissolved oxygen concentration in the water body. The default unit for *DO* is mg L-1 in the database. Users have the options of mmol L-1 and µmol L-1. We ask users to convert the unit to mg L-1 or mmol L-1 if their record is %saturation. Convert the unit of *DO* to mg L-1 before storage in the database.

1 mmol L-1 = 32 mg L-1, 100 µmol L-1 = 3.2 mg L-1.

1.7.7 *Chl a*

*Chl a* refers to chlorophyll a concentration in the water body. The default unit for Chl a is µg L-1 in the database.

Note:

(1) Refer to docs/Web portal.pptx when reading about development of frontend views.

(2) Not mature thoughts. Let’s try out a few of these.

**2. Frontend Views to be developed**

2.1 View of home page

Contains a map interface showing background world map and site locations. It also contains links to views such as Upload Data, Download Data, etc.

2.2 View of Upload Data with Excel File

Contains links to downing the Data\_Template.xlsx and Metadata\_Template.xlsx files, and the link of Upload Data and Upload Metadata.

2.3 View of Uploading Data (Batch Upload)

Contains six parts of information from which to select parameters and units, corresponding fields in the database tables. Part 1: Site Information, Part 2: Water Physical Properties; Part 3: Greenhouse Gas Concentrations; Part 4, Greenhouse gas fluxes; Part 5, Gas transfer velocity; Part 6: Other measurements

Upon finishing selection of parameters and unit, ask users to specify a total number of observations and click on Generate Table.

View of Generated Spread Table: A spread table is generated upon user’s click on Generate Table with user-specified columns and unit. Small widget provided for user to adjust column sequence by giving sequence number of columns.

A submit button to send data to the backend.

2.4 View for displaying static site locations

(1) Using interactive world map as background;

(2) Having functionalities like changing symbol size or color, etc.

2.5 View for displaying site location marker clustering

2.6 View for displaying total measurements at each site location

2.7 View for displaying CO2 measurements at one site as time series

2.8 View for displaying all CO2 measurement

**Phase 1 Project Development of the CO2 Web Portal**

**The Problem:**

Observational data, especially *in situ* CO2 concentration and flux measurements, are essential for correctly modeling CO2 evasions from global inland waters. However, these measurements were collected and published separately by different research groups and there is a lack of a cohesive synthesis of direct measurements, hampering our ability to accurately model CO2 emissions from inland waters. We believe that a global synthesis of direct CO2 measurements would greatly enhance our understanding of the role that inland water plays in contributing CO2 to the atmosphere. A publically accessible, easy-to-use web portal for researchers to easily input, visualize and download data would be a favorable approach for the synthesis ahead.

**Application Example:**

An impressive web portal from Global Ocean Acidification Network for visualizing ocean field campaigns can be found via <http://portal.goa-on.org/Explorer>.

**Sample Data:**

We’ve compiled some 6000 individual CO2 concentration/flux measurements from the literature in MS Excel form, which can be used to jump start construction of the online web-portal. We expect, however, that most of the project time is spent on construction of the front- and backend of the web portal rather than using the data for actual analysis. We hope that size of dataset can be doubled or tripled over the next 1–2 years through involving the bigger community with a well-designed web portal.

**Specific Questions (breaking up of the project task):**

The project contains two main parts: designing the frontend web application and backend database of the web-portal.

*1. Designing frontend of the web-portal*

* ***The main page*** — the main page of the web portal includes three major functionalities for uploading, visualizing and downloading CO2 observational data in global inland waters. Included also is a map interface which displays background world map, site locations and CO2 data from a connected backend database.
* ***The map interface*** — choose an online map service which can be integrated into the main page as a background map. The map must also have capabilities to be connected to the backend database and to visualize georeferenced point data. Adjustable tools for simple visualizations of the data points (in symbol size, color ramp, etc.) will also be developed.
* ***Uploading data*** — an UPLOAD DATA button unfolds two possible data uploading options: Batch Upload and Single Upload, which lead users to web interfaces for batch and single uploading options, respectively.
* ***Batch Upload***
  + In the Batch Upload interface, users are asked to select a series (e.g., 10) of columns and units and specify the total number of observations (e.g., 15) to generate a spread table for data input. Selected columns, units and number of rows correspond to the structure and formats of user’s to-be-upload data.
  + A spread table is generated with user-specified columns, units and number of rows (15 rows × 10 columns with specified units in this case). The user can paste his/her data to the spread table. Column sequence is adjustable by giving sequence number to each column to be adjusted. A SUBMIT button sends data to the backend database.
* ***Single Upload***
  + In the Single Upload interface, users will type data directly into the data-submitting form and select units from dropdown menus. User has the option to save typed-in data at each intermediate step and submit data to the database by clicking on SUBMIT.
  + Note — considering most of users’ data are stored in spreadsheet formats (MS Excel sheet, csv file, etc.), the batch upload option which generates a spread table for users to paste and send data is probably recommended.
  + Note — allowing user to upload csv files to the database is probably not a good option for taking in user data considering very limited control on data format even with associating metadata.
* ***Viewing Data*** — on the main page, user can choose different data types (CO2 concentration, flux and other ancillary data) for visualization and viewing by checking on/off different boxes.
* ***Downloading Data*** — on the main page, for data downloading, user can filter through a series of data types and click on the DOWNLOAD DATA button (this option can be probably developed at a later stage of the web-portal development).

*2. Designing backend database of the web-portal*

* The backend database hosts data input through batch and/or single user upload and responds to user queries. The database should contain columns of the same data types and unit options as in the web interface. Columns in the frontend web interface and backend database are connected and work together to add, update and store user data and respond to user queries. The database can also perform computing (e.g., unit conversion to default) in the background.

*3. List of database/web-interface columns for user data taking-in*

* ***Part 1:*** ***Site Information*** —
  + ***Site Type***: stream, river, lake, reservoir, pond, estuary, wetland, estuary, floodplain, and others
  + ***Site Location***: longitude and latitude in decimal degree or “deg, min, sec“, and altitude in meters above sea level (masl.)
* ***Part 2:*** ***Physical Properties*** —
  + ***Sampling date:*** yyyy-mm-dd
  + ***Discharge:*** m s-1, L s-1, ft s-1
  + ***Water Temp:*** °C
  + ***Air Temp:*** °C
  + ***Width:*** m, km
  + ***Depth:*** cm, m, ft
  + ***Surface Area (for lakes, ponds, etc.):*** m2, ft2, km2
  + ***Flow Velocity:*** m s-1, ft s-1
  + ***Wind Speed:*** m s-1, ft s-1
* ***Part 3:*** ***Greenhouse Gas Concentrations*** —
  + ***CO2:*** ppm, µatm, µmol L-1, mg C L-1 mg L-1
  + ***CO2 method: …***
* ***Part 4:*** ***Greenhouse Gas Fluxes*** —
  + ***CO2 Flux:*** g C m-2 yr-1, mg C m-2 d-1, mg C m-2 hr-1, µg C m-2 d-1 g C m-2 d-1
  + ***CO2 Flux method:***
* ***Part 5:*** ***Gas Transfer Velocity*** —
  + ***k600-CO2:*** m d-1, cm s-1
  + ***k-CO2:*** m d-1, cm s-1
* ***Part 6: Other field measurements —***
  + ***DOC:*** mg L-1, µmol L-1
  + ***POC:*** mg L-1, µmol L-1
  + ***TOC:*** mg L-1, µmol L-1
  + ***pH***
  + ***Alkalinity:*** 100 mg CaCO3, µmol L-1
  + ***DO:*** mg L-1
  + ***Chl a:*** µg L-1

Note: listed selection options and units are to be added or changed.

**Existing methods**

Platform: AWS, MS AZURE, Google Cloud, Yale Server

Frontend web application: HTML, CSS, Python, etc.

Backend database: MySQL, Oracle, etc.

**Proposed methods/tools**

To be determined.

**Background reading**

Abril, G., S. Bouillon, F. Darchambeau, C. R. Teodoru, T. R. Marwick, F. Tamooh, F. Ochieng Omengo, N. Geeraert, L. Deirmendjian, and P. Polsenaere (2015), Technical Note: Large overestimation of pCO2 calculated from pH and alkalinity in acidic, organic-rich freshwaters, Biogeosciences, 12(1), 67-78.

Allen, G. H., and T. M. Pavelsky (2018), Global extent of rivers and streams, Science.

Raymond, P. A., J. Hartmann, R. Lauerwald, S. Sobek, C. McDonald, M. Hoover, D. Butman, R. Striegl, E. Mayorga, and C. Humborg (2013), Global carbon dioxide emissions from inland waters, Nature, 503(7476), 355-359.