

Water quality data for the manuscript Water column biogeochemistry in a tropical estuary

Autumn Oczkowski<sup>1</sup>, Ivelisse Cappielo-Cosme<sup>2</sup>, Morgan Schwartz<sup>1</sup>, Noemí Soto Nieves<sup>3</sup>, Marirosa Molina<sup>4</sup>, Eileen Villafane<sup>5</sup>, Evelyn Huertas<sup>6</sup>, Isabel Alayón Rivera<sup>2</sup>, David Katz<sup>1</sup>, Alana Hanson<sup>1</sup>

<sup>1</sup>*US Environmental Protection Agency, Atlantic Coastal Environmental Sciences Division, Narragansett, RI;* <sup>2</sup>*Oak Ridge Institute for Science and Education Participant with the US Environmental Protection Agency, Region 2 Caribbean Environmental Protection Division, Guaynabo, PR;* <sup>3</sup>*Universidad del Sagrado Corazón, School of Health & Sciences, San Juan, PR;* <sup>4</sup>*US Environmental Protection Agency, Watershed and Ecosystem Characterization Division, Research Triangle Park, NC;* <sup>5</sup>*Puerto Rico Department of Natural and Environmental Resources, Environmental Research Laboratory, San Juan, PR;* <sup>6</sup>*US Environmental Protection Agency, Region 2 Caribbean Environmental Protection Division, Guaynabo, PR*

## Overview

Here we provide data from water samples that were collected twice monthly, from April 26, 2021 and April 25, 2024, from sixteen stations throughout the San Juan Bay Estuary, Puerto Rico. The station locations are given in Fig. 1 and the descriptions of the data headers in the dataset are given in Table 1. An overview of methods follows. For methodological details, please see Oczkowski et al. (*Water column biogeochemistry in a tropical estuary*, In Review).

Fig. 1. Map of the San Juan Bay Estuary showing sampling locations and the delineations of the regions of the Estuary. Where possible, station locations (and names) were based on the San Juan Bay Estuary Program monitoring program (<https://estuario.org/landing-page-2021-en/>). CMP stands for the Caño Martín Peña, Suárez for Suárez Canal, and Piñones for Piñones Lagoon. Further, Torrecilla is commonly referred to as La Torrecilla Lagoon.

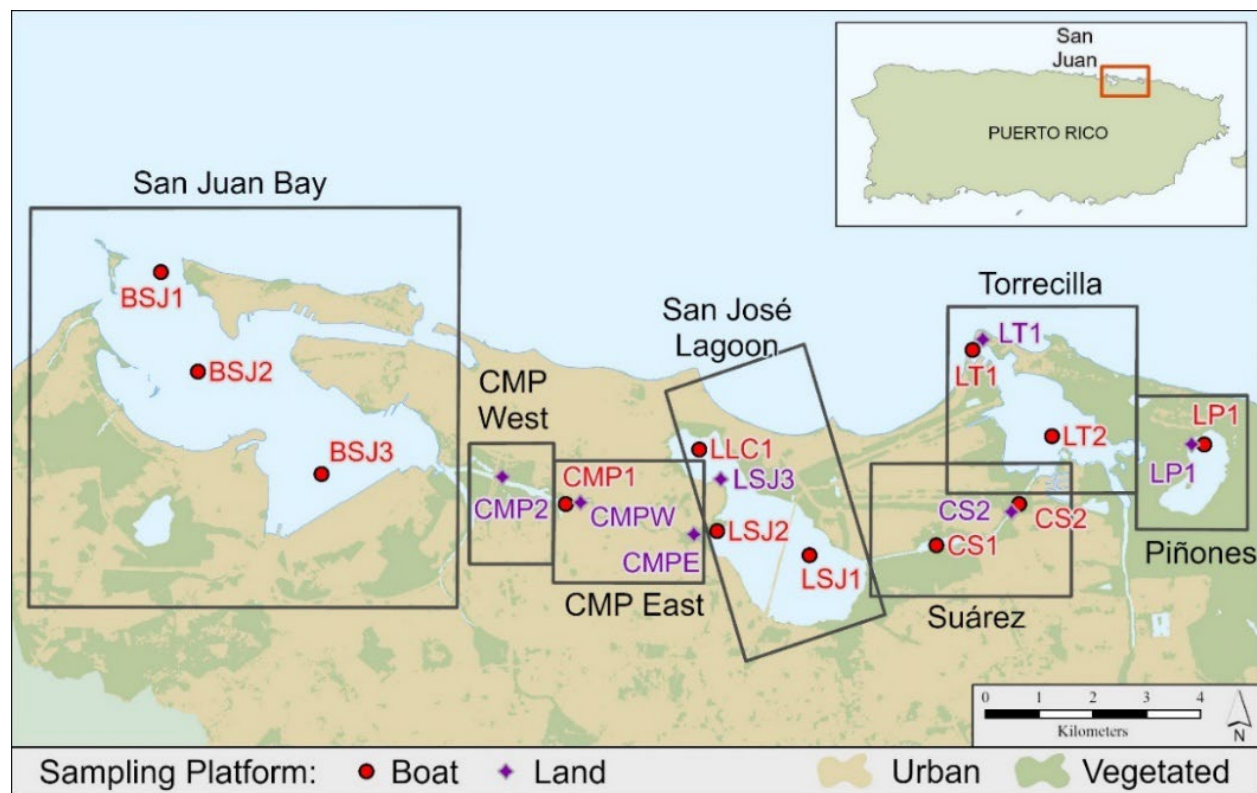


Table 1. Descriptions of the categories of data included in the attached datafile.

<b>Header</b>	<b>Description</b>
<b>number</b>	Unique identification number for each sample
<b>station</b>	Specific station identifier
<b>dupe</b>	This column indicates whether the sample is a duplicate. When dupe = True, the nutrient and particulate samples were collected, and subsequently analyzed, in duplicate
<b>region</b>	This is the generalized region associated with a particular station
<b>date</b>	The date of collection
<b>year</b>	The year of collection
<b>lnd/bt</b>	This column indicates whether the sample was collected as part of the land sampling or boat sampling event
<b>T.degC</b>	Water temperature in degrees Celsius
<b>DO.mgl</b>	Dissolved oxygen in milligrams per liter
<b>DO.%</b>	Dissolved oxygen as a percent
<b>salinity</b>	Salinity of the surface water in parts per thousand (ppt)
<b>ph.nbs</b>	pH measured on the scale maintained by the National Bureau of Standards (NBS)
<b>ph.mv</b>	pH reading in millivolts
<b>turbidity.ntu</b>	Turbidity measured in Nephelometric Turbidity Units (NTU)
<b>entero.mpn_per100ml</b>	Enterococci concentrations in most probable number (mpn) per 100 ml of water
<b>entero_numeric</b>	The same as above but greater and less than symbols are removed from measurements at the upper and lower limits of detection
<b>ecoli.mpn_per100ml</b>	E. coli concentrations in most probable number (mpn) per 100 ml of water
<b>po4.uM</b>	Phosphate concentrations in micromolar (micromoles per liter)
<b>nh4.uM</b>	Ammonium concentrations in micromolar
<b>n+n.uM</b>	Combined nitrate and nitrite concentrations in micromolar
<b>no2.uM</b>	Nitrite concentrations in micromolar
<b>DIN.uM</b>	Dissolved inorganic nitrogen, which is the sum of ammonium, nitrate, and nitrite, in micromolar
<b>NtoP</b>	Molar ratio of dissolved inorganic nitrogen to dissolved inorganic phosphorous
<b>d15n.permil</b>	The nitrogen stable isotope value of the suspended particulate matter in the surface water
<b>d13c.permil</b>	The carbon stable isotope value of the suspended particulate matter in the surface water
<b>PN.mg/l</b>	The particulate nitrogen concentration, in milligrams per liter, of the surface water
<b>PC.mg/l</b>	The particulate carbon concentration, in milligrams per liter, of the surface water

## Methods

Twice monthly water samples were collected from April 26, 2021 to April 25, 2024 from the San Juan Bay Estuary in Puerto Rico (Fig. 1). One monthly sampling was performed in conjunction with the San Juan Bay Estuary Program (boat based), while the other was conducted approximately two weeks later from land (land based). For specific methodological details see Oczkowski et al. (In Review). But briefly:

- At each station (Fig. 1), samples were collected via sterile autoclaved one-liter bottles from the surface water. Bottles were stored in a cooler, on ice, until processed in the laboratory. A calibrated YSI multi-parameter meter (YSI Inc., Yellow Springs, OH) was deployed at each station to measure conductivity, salinity, pH, dissolved oxygen, and temperature.
- In the laboratory, each water sample was filtered in duplicate through pre-combusted (450 °C, 6 hours) glass fiber filters (Whatman GFF, 0.7 µm pore size, 25 mm diameter) using filter flasks with funnels attached to an electric vacuum pump. After filtration, plastic, acid washed 20 mL vials were rinsed with a small portion of filtered water and then filled. Filters were carefully removed using forceps, folded, wrapped in foil, and placed in labeled bags. Both water and filters were stored at -20°C until analysis for dissolved and particulate nutrient content.
- To measure concentrations of *Enterococcus* and *E. coli*, we used the Defined Substrate Method via the IDEXX Quanti-Tray/2000 system. Samples were processed, in duplicate, within six hours of collection. For samples from locations with anticipated high bacterial concentrations, such as the CMP sites, a 1:100 dilution was performed to prevent results exceeding the upper detection threshold. Results were expressed as the most probable number (MPN) per 100 mL, with a detection limit ranging from 10 to 24,193 MPN/100 mL.
- We used a handheld fluorometer with a turbidity channel (AquaFluor, Turner Designs) up until February 28, 2023, after which turbidity was measured with a calibrated turbidity meter (2020i, LaMotte) in the lab.
- The dissolved inorganic nutrients, ammonium (NH<sub>4</sub><sup>+</sup>), nitrate+nitrite (NO<sub>3</sub><sup>-</sup>+NO<sub>2</sub><sup>-</sup>), nitrite (NO<sub>2</sub><sup>-</sup>), and orthophosphate (PO<sub>4</sub><sup>3-</sup>) were run simultaneously on two connected Astoria 2 segmented flow analyzers (Astoria-Pacific, Clackamas, Oregon, USA) following EPA methods (350.1, 353.2, and 365.1). Samples were calibrated against a minimum five-point standard curve for each analyte, blanks every 10 samples to correct for drift, and independent check standards were measured every 15 samples, with an allowable 20% difference.
- For particulate nitrogen (N) and carbon (C) content and stable isotope (δ<sup>15</sup>N and δ<sup>13</sup>C) analysis, filters were dried in a 60°C oven for at least 48 hours and then pelletized. Samples were analyzed using an Elementar Vario Select elemental analyzer connected to an Isoprime visION ratio mass spectrometer (Elementar Americas, Ronkonkoma, NY). Particulate N and C content was calculated by comparing sample peak areas to a standard curve of peak area plotted against elemental composition of standard reference material. Isotope values are expressed in δ notation where δX (‰) = [(R sample/R standard) - 1] × 1000 where X is <sup>13</sup>C or <sup>15</sup>N and R is <sup>13</sup>C/<sup>12</sup>C or <sup>15</sup>N/<sup>14</sup>N, respectively. USGS 40, USGS 41a, acetanilide, cystine, and internal working standards were included every 48 samples, as well as at the beginning and end of the run. Duplicate samples were analyzed in separate runs, with variance between duplicates generally better than 0.3‰. The precision of the laboratory standards was better than ±0.3‰ for isotope values, based on long term replication.