This preprint has not undergone peer review or any post-submission improvements or corrections. The Version of Record of this article is published in Limnology, and is available online at https://doi.org/10.1007/s10201-025-00804-w

# Biochemical oxygen demand as a proxy for dissolved organic carbon in Japanese rivers: Conservative estimates for ecological risk assessment

Yuichi Iwasaki and Wataru Naito

Research Institute of Science for Safety and Sustainability, National Institute of Advanced Industrial Science and Technology (AIST), 16-1 Onogawa, Tsukuba, Ibaraki 305-8569, Japan

Corresponding author: Yuichi Iwasaki, E-mail: yuichiwsk@gmail.com, yuichi-iwasaki@aist.go.jp

Number of words: 1271

Number of figures: 2

Number of tables: 0

## Abstract

Dissolved organic carbon (DOC) is a critical parameter for assessing metal bioavailability and toxicity in aquatic systems, but data from routine measurements in Japan are limited to specific sites. The goal of this study was to develop a statistical model to estimate DOC concentrations in Japanese rivers using biochemical oxygen demand (BOD) as a proxy. Because the relationship between BOD and DOC was expected to be highly variable, we focused on obtaining conservative (i.e., lower bound) rather than central tendency estimates of DOC concentrations to support "safe-side" screening-level ecological risk assessments. Based on BOD and DOC measurements from 30 river sites across Japan, we developed a quantile regression model at the 0.1 quantile to provide conservative estimates of DOC. Validation with additional monitoring datasets, including original field surveys in Kanagawa and Osaka Prefectures, demonstrated that the developed model provided reasonably conservative estimates of DOC and hence supported its use for "safe-side" screening-level ecological risk assessment. Because of the variability of the BOD-DOC relationship across sites, direct DOC measurements may be appropriate where screening-level assessments indicate potential ecological risks.

**Keywords**: bioavailability, dissolved organic matter, freshwater, biotic ligand model, quantile regression

This preprint has not undergone peer review or any post-submission improvements or corrections. The Version of Record of this article is published in Limnology, and is available online at https://doi.org/10.1007/s10201-025-00804-w

## INTRODUCTION

Knowledge of chemical speciation is essential for understanding and accurately predicting the bioavailability and toxicity of trace metals and cationic polymers in aquatic systems (Adams et al. 2020; Connors et al. 2023; Paquin et al. 2002). Dissolved organic matter (DOM) plays an important role in this speciation process because binding of trace metals and cationic polymers to DOM can reduce their toxicity. Dissolved organic carbon (DOC) is often used as a metric of DOM to predict metal bioavailability and toxicity (Farley et al. 2015; Tipping et al. 2008). In the case of Cu, for example, DOC has been the most influential parameter among several input parameters, including Ca, Mg, and alkalinity, in the derivation of predicted no-effect concentrations using a biotic ligand model (Peters et al. 2011). Despite its importance, routine monitoring of DOC is not done in many countries, including Japan (Iwasaki &Naito 2024; Peters et al. 2013). This lack of monitoring data has complicated efforts to assess the bioavailability and ecological risks of trace metals or cationic polymers in aquatic environments on a broader scale (e.g., at a country level). For example, Peters et al. (2013) have developed a statistical model to predict DOC from concentrations of dissolved iron to assess the ecological risks of nickel in the UK.

In Japan, 5-day biochemical oxygen demand (hereafter referred to as BOD) has been measured as an indicator of organic pollution in nationwide water quality monitoring programs, but there are no comprehensive data on DOC concentrations in rivers across the country, nor is there a model available to predict DOC concentrations. Our goal was thus to develop a statistical model based on 5-day BOD values to predict concentrations of DOC in Japanese rivers. Because the relationship between BOD and DOC was expected to be highly variable (see below), we focused on obtaining conservative estimates (i.e., very unlikely to be overestimates) of DOC concentrations to support a "safe-side" screening-level ecological risk assessment rather than central tendency estimates such as the arithmetic mean.

## MATERIALS AND METHODS

### Model development

To develop a statistical model for DOC prediction, we used two different types of DOC and BOD monitoring data: a nationwide water quality dataset for model development and three monitoring datasets for model validation. For the model development, we collected the relevant monitoring data from the Water Information System managed by the Ministry of Land, Infrastructure, Transport and Tourism (http://www1.river.go.jp/). By extracting measurement records where DOC and BOD were measured simultaneously at the same sites between 2011 and 2020, we obtained a total of 1583 records from 30 river sites across Japan (Fig. S1; Table S1 for the raw data). When selecting these sites, we excluded three sites that were likely located in brackish water based on their locations on a map and the fact that their measured electrical conductivities exceeded 100 mS/m. For a measurement below the reporting limit of quantification (BOD: 0.5 mg/L; DOC: 1 mg/L), we used half the limit of quantification in the later analysis.

Using the extracted data, we developed a quantile regression model at the 0.1 quantile to obtain conservative estimates of DOC from BOD. We modeled the BOD-DOC relationship using the following simple linear function:

$$\log(DOC) = a \times \log(BOD) + b \quad (1)$$

As a supplementary analysis, we also developed a quantile regression model at the 0.5 quantile to capture the median relationship. All the statistical analyses were performed using R version 4.4.0 (R Core Team 2024). The quantile regression models (Cade &Noon 2003; Koenker &Hallock 2001) were fitted using the function “rq” in the R package “quantreg” (version 5.97).

**Model validation**

To assess whether the developed quantile regression model (i.e., 0.1 quantile) could provide conservative estimates of DOC, we used a monitoring dataset obtained from Takeshita et al. (2019) and two additional datasets from original field surveys that we conducted in Kanagawa and Osaka Prefectures (Fig. S2). Takeshita et al. (2019) selected 50 sampling sites in Japanese rivers receiving nickel discharge, including reference sites (see Fig. S2 and Takeshita et al. (2019) for more details).

For our original surveys in Kanagawa and Osaka Prefectures, a total of 10 and 16 sampling sites, respectively, were selected from nationwide water quality monitoring sites in order to include both sites with relatively low BOD (≤ 1 mg/L) and those with relatively high BOD (> 2–3 mg/L) based on previously reported values (Iwasaki et al. 2022). The sampling was conducted in January and June 2024, respectively. For the Kanagawa Prefecture survey, we selected nine sampling sites from the Tsurumi River and one site from the Tama River; 16 sampling sites were set up in various rivers for the Osaka Prefecture survey (see Fig. S2 and Table S2 for more details, including latitudes and longitudes). The analyses of DOC and 5-day BOD were conducted in accordance with Japanese Industrial Standard (JIS) K0102 testing methods for industrial wastewater. Dissolved organic carbon was measured in water samples filtered through a 0.45-µm membrane filter using a total organic carbon analyzer (TOC-L CPH, Shimadzu).

## RESULTS AND DISCUSSION

Values of BOD at the 30 monitoring sites across Japan ranged from 0.1 to 14 mg/L. The DOC concentrations were highly variable at a given BOD (Fig. 1). The relationships between BOD and DOC were generally unclear, even when examined at individual sites (Fig. S1; Pearson's $r = -0.21$ to 0.71; median = 0.37). The magnitudes of ~90% of the correlation coefficients were less than 0.6.

Using these monitoring data for model development, the quantile regression model at the 0.1 quantile was estimated as follows (Fig. 1):

$$\log(DOC) = 0.5558 \times \log(BOD) - 0.2306 \quad (2)$$

Both the intercept ($b$) and slope ($a$) of the model were statistically significant ($p < 0.001$). This regression line was estimated so that 10% of the measured DOC values fell below the line. Compared with the model validation data, all but a few of the data points were above the 0.1 quantile regression line (Fig. 2). This result suggests that this quantile regression model provides conservative estimates of DOC from BOD.

We also estimated the quantile regression model at the 0.5 quantile:

$$\log(DOC) = 0.3155 \times \log(BOD) + 0.1364 \quad (3).$$

Both the intercept (0.1364) and slope (0.3155) of the model were statistically significant

This preprint has not undergone peer review or any post-submission improvements or corrections. The Version of Record of this article is published in Limnology, and is available online at https://doi.org/10.1007/s10201-025-00804-w

($p$ < 0.001). This 0.5-quantile model could facilitate obtaining moderate (i.e., median) estimates of DOC concentrations for environmental risk assessment or for examining how the results of a risk assessment depend on DOC values. However, note that given the variations in the BOD-DOC relationship among sites (and likely among seasons; Figs. 2 and S1), caution is required when using DOC values predicted from the 0.5 quantile regression model. Indeed, the observed BOD-DOC relationship from the Osaka Prefecture survey clearly differed from the relationships apparent in the other two datasets (Takeshita et al. (2019) and the Kanagawa Prefecture survey; see Fig. 2), although the underlying reasons for these differences are unclear.

In this study, largely due to limited data availability, we analyzed the relationship between BOD and DOC by pooling monitoring data from 30 sites in rivers across Japan and developed a quantile regression model that provided conservative estimates of DOC from BOD. However, the BOD-DOC relationships likely depended on site-specific characteristics, such as catchment land use, vegetative cover, and resident microorganisms. Further data accumulation might enable evaluations that take account of these characteristics. However, it is important to note an essential difference: BOD is an indicator of biodegradable organic matter, whereas DOC is the amount of dissolved organic carbon and includes both refractory and biodegradable organic matter. This

This preprint has not undergone peer review or any post-submission improvements or corrections. The Version of Record of this article is published in Limnology, and is available online at https://doi.org/10.1007/s10201-025-00804-w

important difference suggests that there may be inherent limitations to accurately estimating DOC from BOD. Direct measurements of DOC may be appropriate at sites where screening-level assessments based on the conservative estimates of DOC indicate ecological risks of concern.

**Author contributions**

Conceptualization: YI and WN; Data curation: YI; Formal analysis: YI; Funding acquisition: YI and WN; Methodology: YI; Writing—original draft preparation: YI; writing—review and editing: YI and WN.

**Funding**

This study was partly funded by the Environment Research and Technology Development Fund (JPMEERF20225005) of the Environmental Restoration and Conservation Agency of Japan provided by the Ministry of the Environment of Japan.

**Data and code availability**

All data are available in the Supplementary Materials.

**Acknowledgements**

This paper does not necessarily reflect the policies or views of any government agency. This study was performed in compliance with the current laws of Japan. During the preparation of this paper, the authors used ChatGPT to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

**Conflict of interest**

The authors declare no conflicts of interest.

**REFERENCES**

Adams W, Blust R, Dwyer R, Mount D, Nordheim E, Rodriguez PH, Spry D (2020) Bioavailability assessment of metals in freshwater environments: A historical review. Environ Toxicol Chem 39:48–59. doi:10.1002/etc.4558

Cade BS, Noon BR (2003) A gentle introduction to quantile regression for ecologists. Front Ecol Environ 1:412–420. doi:10.1890/1540-9295(2003)001[0412:AGITQR]2.0.CO;2

Connors KA, Arndt D, Rawlings JM, Brun Hansen AM, Lam MW, Sanderson H, Belanger SE (2023) Environmental hazard of cationic polymers relevant in personal and consumer care products: A critical review. Integr Environ Assess Manag 19:312–325. doi:10.1002/ieam.4642

Farley KJ, Balistrieri LS, De Schamphelaere KAC, Iwasaki Y, Janssen CR, Kamo M, Lofts S, Mebane CA, Naito W, Ryan A, Santore R, Tipping E, Meyer JS (2015) Metal Mixture Modeling Evaluation: 2. Comparative evaluation of four modeling approaches. Environ Toxicol Chem 34:741–753. doi:10.1002/etc.4560

Iwasaki Y, Kobayashi Y, Suemori T, Takeshita K, Ryo M (2022) Compiling physicochemical characteristics of water quality monitoring sites (environmental reference points) in Japanese rivers and site grouping. J Japan Soc Water Environ 45:231–237. doi:10.2965/jswe.45.231

Iwasaki Y, Naito W (2024) Metal exposure profiles at metal-contaminated sites in rivers across Japan. ChemRxiv10.26434/chemrxiv-2024-tdzt3

Koenker R, Hallock KF (2001) Quantile Regression. J Econ Perspect 15:143–156. doi:10.1257/jep.15.4.143

Paquin PR et al. (2002) The biotic ligand model: a historical overview. Comp Biochem Physiol C Toxicol Pharmacol 133:3–35. doi:10.1016/S1532-0456(02)00112-6

Peters A, Merrington G, de Schamphelaere K, Delbeke K (2011) Regulatory consideration of bioavailability for metals: Simplification of input parameters for the chronic copper biotic ligand model. Integr Environ Assess Manag 7:437–444. doi:10.1002/ieam.159

Peters A, Simpson P, Merrington G, Schlekat C, Rogevich-Garman E (2013) Assessment of the effects of nickel on benthic macroinvertebrates in the field. Environ Sci Poll Res10.1007/s11356-013-1851-2

R Core Team (2024): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Takeshita KM, Misaki T, Hayashi TI, Yokomizo H (2019) Associations of community structure and functions of benthic invertebrates with nickel concentrations: Analyses from field surveys. Environ Toxicol Chem 38:1728–1737. doi:10.1002/etc.4462

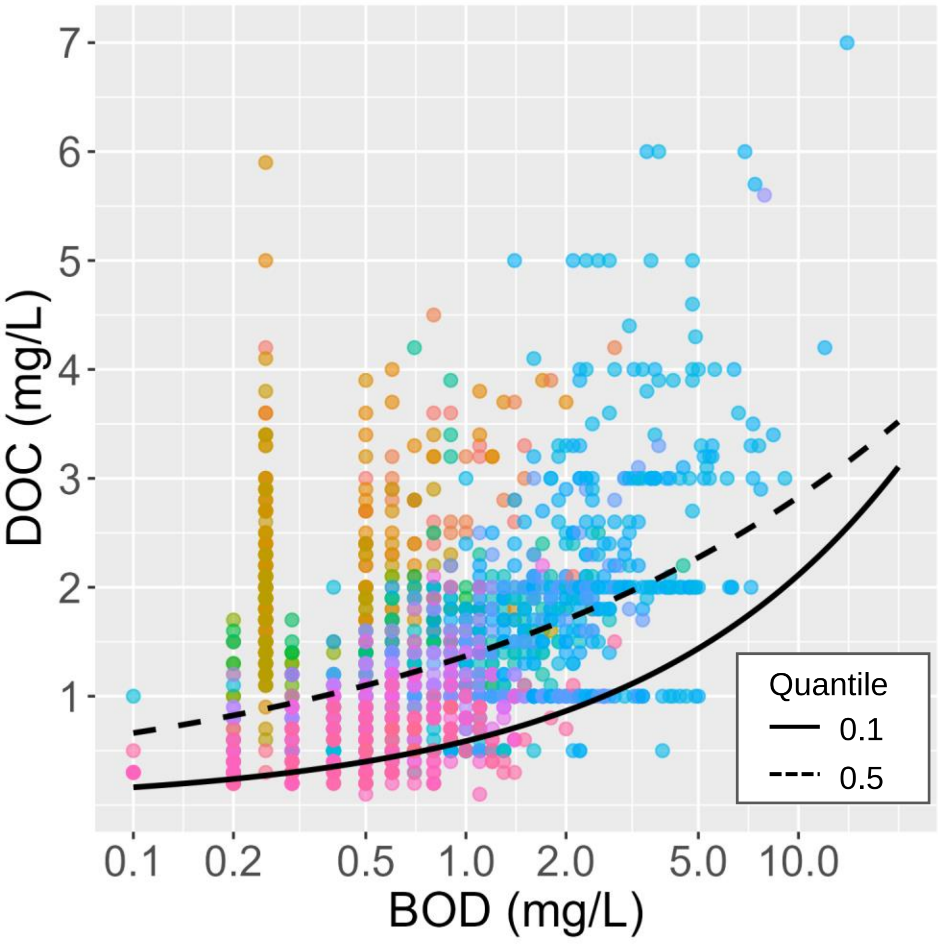
Tipping E, Vincent CD, Lawlor AJ, Lofts S (2008) Metal accumulation by stream bryophytes, related to chemical speciation. Environ Pollut 156:936–943. doi:10.1016/j.envpol.2008.05.010

[in Japanese with English abstract]

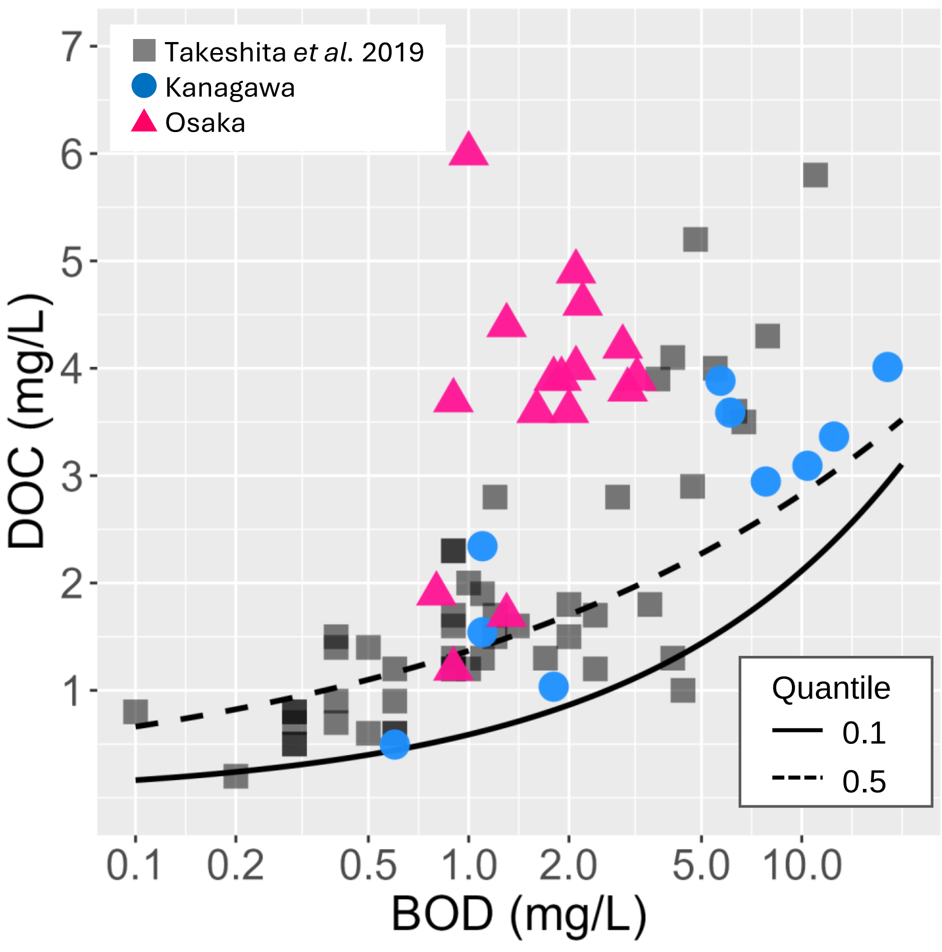
## Figure legends

**Fig. 1**. Relationship between biochemical oxygen demand (BOD) and dissolved organic carbon (DOC) in Japanese rivers. Different colors indicate 30 water-quality-monitoring sites selected from the Water Information System (see Fig. S1 for more details). The solid and dashed lines represent the results of quantile regression models at the 0.1 and 0.5 quantiles, respectively.

**Fig. 2**. Relationship between biochemical oxygen demand (BOD) and dissolved organic carbon (DOC) in three monitoring datasets (Takeshita *et al*., 2019 and original field surveys in Kanagawa and Osaka Prefectures). The solid and dashed lines represent the results of quantile regression models at the 0.1 and 0.5 quantiles, respectively, fitted to data from the Water Information System.



DOC (mg/L)
7
6
5
4
3
2
1
0.1
0.2
0.5
1.0
2.0
5.0
10.0
BOD (mg/L)
Quantile
0.1
0.5



Takeshita et al. 2019
Kanagawa
Osaka
Quantile
0.1
0.5
DOC (mg/L)
BOD (mg/L)
1
2
3
4
5
6
7
0.1
0.2
0.5
1.0
2.0
5.0
10.0