Science of the Total Environment 636 (2018) 12–19

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

# Profiling physicochemical and planktonic features from discretely/continuously sampled surface water

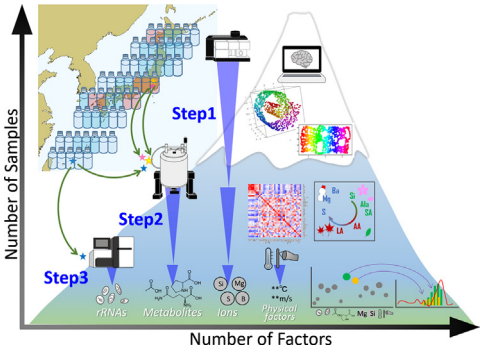
Azusa Oita [a], Yuuri Tsuboi [a], Yasuhiro Date [a,b], Takahiro Oshima [b], Kenji Sakata [a], Akiko Yokoyama [c,e], Shigeharu Moriya [a,b], Jun Kikuchi [a,b,d,*]

[a] RIKEN Center for Sustainable Resource Science, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[b] Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[c] Graduate School of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
[d] Graduate School of Bioagricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-0810, Japan
[e] Center for Regional Environmental Research, National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan

## HIGHLIGHTS

- Three analytical approaches for diverse -omics datasets are proposed.
- These approaches were used to explain the features of Odaiba in Tokyo Bay.
- Integrated use of the proposed approaches can be used for highlighting key factors.

## GRAPHICAL ABSTRACT




## ARTICLE INFO

*Article history:*
Received 18 January 2018
Received in revised form 2 April 2018
Accepted 10 April 2018
Available online xxxx

Editor: Daniel Wunderlin

*Keywords:*
Multi-omics
Metabolome
Microbiome
Multidisciplinary approaches
Machine learning
Vector autoregressive model

## ABSTRACT

There is an increasing need for assessing aquatic ecosystems that are globally endangered. Since aquatic ecosystems are complex, integrated consideration of multiple factors utilizing omics technologies can help us better understand aquatic ecosystems. An integrated strategy linking three analytical (machine learning, factor mapping, and forecast-error-variance decomposition) approaches for extracting the features of surface water from datasets comprising ions, metabolites, and microorganisms is proposed herein. The three developed approaches can be employed for diverse datasets of sample sizes and experimentally analyzed factors. The three approaches are applied to explore the features of bay water surrounding Odaiba, Tokyo, Japan, as a case study. Firstly, the machine learning approach separated 681 surface water samples within Japan into three clusters, categorizing Odaiba water into seawater with relatively low inorganic ions, including Mg, Ba, and B. Secondly, the factor mapping approach illustrated Odaiba water samples from the summer as rich in multiple amino acids and some other metabolites and poor in inorganic ions relative to other seasons based on their seasonal dynamics. Finally, forecast-error-variance decomposition using vector autoregressive models indicated that a type of microalgae (Raphidophyceae) grows in close correlation with alanine, succinic acid, and valine on filters and with isobutyric acid and 4-hydroxybenzoic acid in filtrate, Ba, and average wind speed. Our integrated strategy can be used to examine many biological, chemical, and environmental physical factors to analyze surface water.

© 2018 Published by Elsevier B.V.

*Abbreviations:* FEVD, forecast-error-variance decomposition; ICP-OES, inductively coupled plasma optical-emission spectroscopy; IBA, isobutyric acid; MDS, multidimensional scaling; NGS, next-generation sequencing; NMR, nuclear magnetic resonance; RF, Random Forest; rRNA, ribosomal RNA; SA, succinic acid; VAR, vector autoregression.

* Corresponding author.
*E-mail address:* jun.kikuchi@riken.jp. (J. Kikuchi).

https://doi.org/10.1016/j.scitotenv.2018.04.156
0048-9697/© 2018 Published by Elsevier B.V.

## 1. Introduction

Surface water is an essential part of aquatic ecosystems, including freshwater, marine, and estuarine. Aquatic ecosystems provide numerous benefits, such as a diverse biological production, safe freshwater, aesthetics, and recreation. However, aquatic ecosystems have been globally under threats from climate change and pollutants (Capon et al., 2015; Islam and Tanaka, 2004; Jambeck et al., 2015; Österblom et al., 2017; Seitzinger and Phillips, 2017; Worm, 2015). Both these factors can disrupt self-regulation of aquatic ecosystems that maintains structural and functional dynamics within normal ranges, thereby leading to profound changes in the ecosystem (Breitburg et al., 2018; Halpern et al., 2012). For the management of complex systems associated with aquatic ecosystems, it is necessary to consider multiple factors, including environmental physical factors, chemical factors, biological factors, and human factors, using diverse approaches and information (Burthe et al., 2016; Cajaraville et al., 2016; Cooper et al., 1994; Lynch et al., 2015). Based on advances in information and omics technology (Moran et al., 2016), of the ten dimensions for integrated assessment and modeling (Hamilton et al., 2015), methodological and disciplinary dimensions are becoming more important.

Although the omics revolution has contributed to environmental toxicology (Bahamonde et al., 2016; Martyniuk and Simmons, 2016), metaomics studies, including metagenomics, metatranscriptomics, metaproteomics, and metabolomics, were not incorporated in mathematical models to understand ecosystems until recently (Reed et al., 2014). A recently proposed modeling framework by Reed et al. (2014, 2015) links the functional genes of microbes and biogeochemical processes in water. The model was further developed by Louca et al. (2016) to integrate the sequence information of DNA, mRNA, and proteins with geochemical processes; their model describes the spatiotemporal dynamics of the concentrations of eight metabolites ($NH_4^+$, $O_2$, $NO_3^-$, $SO_4^{2-}$, $H_2S$, etc.) and six DNA profiles in the analyzed water. Another attempt to utilize multi-omics data is a data-driven approach. Ogawa et al. (2014) proposed biogeochemical typing using the 16S and 18S ribosomal RNA (rRNA) sequences via next-generation sequencing (NGS), ion data detected using inductively coupled plasma optical-emission spectroscopy (ICP-OES), metabolite data detected using Fourier transform infrared spectroscopy and nuclear magnetic resonance (NMR).

Among the target chemicals of omics studies, metabolites can describe the final products of transcription and translation, although there are some limitations because of their instability (Martyniuk and Simmons, 2016). Therefore, metabolites are closely related to phenotype and can reflect the surrounding environment of organisms. NMR was previously used for untargeted analysis of metabolite profiles for various environmental samples (Kikuchi and Yamada, 2017; Kikuchi et al., 2018; Simpson et al., 2018). Examples include Atlantic salmon (*Salmo salar L.*; Aursand et al., 2009), zebra mussel (*Dreissena polymorpha*; Lee et al., 2010), Araliaceae (*Panax ginseng*; Kang et al., 2008a; Kang et al., 2008b), grape varieties for wine (Godelmann et al., 2013), and green tea (*Camellia sinensis*; Watanabe et al., 2015). The studies of these profiles revealed the geographical origins of the products on a small scale. In addition, metabolic analysis was previously applied to provide strong discriminatory power to the evaluation of environmental variation and diversity in aquatic ecosystems (Asakura et al., 2014; Date and Kikuchi, 2018; Wei et al., 2018; Yoshida et al., 2014).

Recent work on early warning indicators emphasized the requirement for reliable tools to assess ecosystem resilience (Gsell et al., 2016), and additional efforts to integrate microbial ecology and geochemistry are required to better understand aquatic ecosystems. To this end, an integrated strategy linking multi-omics approach to utilize diverse types of information can screen for important features of surface water for use in modeling.

To investigate surface water features in a targeted area, we developed an integrated strategy linking three data-analytic approaches: step-1) machine learning, step-2) factor mapping, and step-3) forecast-error-variance decomposition (FEVD). These approaches utilize datasets that vary in sample size and analyzed factors. We applied our approaches to explore the features of bay water surrounding Odaiba, Tokyo, Japan, as a case study.

## 2. Materials and methods

### 2.1. Overview

We developed three data-analytic approaches and applied each to different datasets (Fig. 1). The machine learning approach was designed for datasets containing a relatively large number of samples with a small number of variables. Herein, this approach was applied to an ICP-OES data measuring 10 inorganic ions from 681 surface water samples collected from various places within Japan, including Lake Imba (Chiba Prefecture) and Iriomote Island (Okinawa Prefecture) (Table S1) and used to compare this data with those from other locations (see Section 2.3 for details). The factor mapping approach was designed to spatially or temporally analyze continuous data containing a relatively large number of variables. Herein, this method was applied to analyze three datasets describing ICP-OES-measured inorganic ions and NMR-measured metabolites in surface water from Odaiba and Lake Imba (Chiba Prefecture) sampled monthly and surface water sampled along the Kuira River on Iriomote Island (Okinawa Prefecture) (see Section 2.2 for detailed sampling locations) and compared seasonal dynamics between the locations (see Section 2.4 for details). The FEVD approach was designed to temporally analyze continuous datasets containing a large number of variables. This method was applied to a larger dataset from Odaiba that included NGS-analyzed small-subunit (18S and 16S) rRNA sequence data and physical information, as well as inorganic ion and metabolite data used in the factor mapping approach, for a brief investigation of biological correlations within Odaiba samples (see Section 2.5 for details). The R scripts used for our analysis can be found at http://dmar.riken.jp/Rscripts/. In the following subsections, we explain the methods used to obtain the data used for the Odaiba case study (Sections 2.2) and our three data-analytic approaches (Sections 2.3–2.5).

### 2.2. Sampling and processing

The 681 water samples (50-mL each) were collected for ICP-OES analysis (see Supplementary Methods Section 1 for details) from 15 prefectures in Japan, from Okinawa in the southwest to Hokkaido in the northeast, between 2010 and 2016 (Table S1 and Fig. 2a). Additional volumes of water were collected for NMR analysis (see Supplementary Methods Section 2, Table S2, and Fig. S1 for details) and NGS analyses (see Supplementary Methods Section 3 for details) using either a bucket (at a 1-m depth) or a Bandon water sampler (for samples from >1-m depth) at 14 selected locations. These samples were subsequently filtered using a 0.2-μm Durapore filter membrane (Millipore, Billerica, MA, USA) until the membrane clogged. The collected water, filters, and filtered water were stored in a 4 °C cooler box and transported to the laboratory, where they were stored at −80 °C.

The sampling locations included Odaiba [35.618180°N, 139.773110°E (WGS84; sample IDs of w549–w620 in Table S1)], Lake Imba (35.747494°N, 140.181877°E; w637–w648), and 12 locations on Iriomote Island [along the Kuira River, from the upper stream to its downstream areas, including Funauki Bay and offshore (Fig. 2b and c), including R1 (24.29058°N, 123.74693°E; w621), R2 (24.29123°N, 123.74757; w622), R3 (24.29481°N, 123.74757°E; w623), R3.5 (24.30095°N, 123.74678°E; w624), R4 (24.30251°N, 123.74746°E; w625), R5 (24.30951°N, 123.74916°E; w626), R6 (24.31363°N, 123.75034°E; w627), M1 (24.31071°N, 123.75259°E; w628), F1 (24.32843°N, 123.74301°E; w629, w630 and w633), F2 (24.33351°N,

Download English Version:

https://daneshyari.com/en/article/8859465

Download Persian Version:

https://daneshyari.com/article/8859465

Daneshyari.com