Environmental Modelling & Software 97 (2017) 72–85

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Detecting spatial patterns of rivermouth processes using a geostatistical framework for near-real-time analysis

Wenzhao Xu [a,*], Paris Collingsworth [b], Barbara Bailey [c], Martha Carlson Mazur [d], Jeffrey Schaeffer [e], Barbara Minsker [a]

[a] Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, United States
[b] Department of Forestry and Natural Resources and Illinois Indiana Sea Grant, Purdue University, West Lafayette, IN, United States
[c] Department of Mathematics and Statistics, San Diego State University, United States
[d] School of Environmental Studies, Bellarmine University, United States
[e] U.S. Geological Survey Great Lakes Science Center, United States

ARTICLE INFO

Article history:
Received 28 June 2016
Received in revised form
23 June 2017
Accepted 30 June 2017

Keywords:
Water quality
Rivermouth
Plumes
Dynamic
Decision support
Adaptive sampling

ABSTRACT

This paper proposes a geospatial analysis framework and software to interpret water-quality sampling data from towed undulating vehicles in near-real time. The framework includes data quality assurance and quality control processes, automated kriging interpolation along undulating paths, and local hotspot and cluster analyses. These methods are implemented in an interactive Web application developed using the Shiny package in the R programming environment to support near-real time analysis along with 2- and 3-D visualizations. The approach is demonstrated using historical sampling data from an undulating vehicle deployed at three rivermouth sites in Lake Michigan during 2011. The normalized root-mean-square error (NRMSE) of the interpolation averages approximately 10% in 3-fold cross validation. The results show that the framework can be used to track river plume dynamics and provide insights on mixing, which could be related to wind and seiche events.

Software availability

Product Title: Towed Undulating Vehicle Data Analysis Tool
Developer: Wenzhao Xu
Contact Address: Dept. of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign
Contact Email: xuwz.uiuc@gmail.com
Available Since: 2016
Programming Language: R
Source Code: http://stormxuwz.github.io/TUVTool/
Cost: Free

1. Introduction

Rivermouth ecosystems are dynamic transitional river and lake mixing zones that can extend many kilometers upstream of the river/lake confluence and a similar distance into the lake. Rivermouth ecosystems are not only economic centers for human populations (e.g., fish production and recreation) but also have significant influences on the lake ecosystem (Larson et al., 2013). River plumes affect nearshore water chemistry (Kaur et al., 2007; Makarewicz and Howell, 2012), bacteria transportation (Nekouee, 2012), and fish community composition (Janetski et al., 2013). However, the complexity of the rivermouth system impedes understanding of the river plume dynamics and their effects, which are controlled by many factors such as vertical/horizontal mixing, dispersion, density, and seiche effects (Rao and Schwab, 2007; Jackson et al., 2008). Seiche events, for example, are wind-induced water-level fluctuations that bring large volumes of lake water into rivermouths and can create backflow, which may affect the location of mixing zones (Pebbles et al., 2013). Moreover, phytoplankton distributions not only depend on temperature, bathymetry and hydrologic features such as watershed type and riverine input (Pavlac et al., 2012; Snow et al., 2000), but also are influenced by wind and the presence of older plumes (Hickey et al., 2005; Frame and Lessard, 2009). Therefore, it is important to understand plume dynamics to fully comprehend rivermouth systems.

* Corresponding author.
E-mail address: wxu15@illinois.edu (W. Xu).

http://dx.doi.org/10.1016/j.envsoft.2017.06.049
1364-8152/

In the Great Lakes, knowledge about rivermouth mixing patterns and especially plumes has become vital in understanding their role in maintaining nearshore and deepwater food webs (Hoffman et al., 2010; Larson et al., 2013). The recent (1990-present) invasion and proliferation of Dreissenid mussels have been implicated in the collapse of deepwater fish communities in Lakes Michigan and Huron (Riley et al., 2008; Madenjian et al., 2012). Mussels are thought to be sequestering energy and nutrients in nearshore areas that formerly supported fish in offshore and deepwater habitats (Hecky et al., 2004). Rivermouth ecosystems and their associated plumes may be one of the few areas where historical food webs are still intact, but food web assessments in such habitats have been limited due to the dynamic nature of plumes.

Understanding rivermouth dynamics requires comprehensive water quality data (Howell et al., 2012). Traditionally, rivermouth data are collected via fixed stations or buoys that continuously or periodically measure water chemistry. For example, the National Oceanographic and Atmospheric Administration (NOAA) have significant amounts of buoy data sampled at the coastline (https://coastwatch.glerl.noaa.gov). However, this approach provides data that are limited spatially by the existing buoy network of NOAA. Another approach is using a mobile sampling platform with a flow-through system that continuously pumps water from a fixed depth through a series of sensors to obtain water chemistry data (Pavlac et al., 2012; Twiss and Marshall, 2012; Jackson and Reneau, 2014). This extends the spatial range of data collection but fails to sample data throughout the water column. A promising approach to sample data at extensive three-dimensional (3-D) spatial scales is to use towed or autonomous undulating vehicles that carry multiple sensors. Such a vehicle may be autonomous or towed behind a ship that moves along different survey paths, undulating throughout the survey between the water surface and the near bed region of the water column. Such vehicles currently in operation include ScanFish (Ludsin et al., 2009), SARAGO (Marcelli et al., 2005), TRIAXUS (Jones et al., 2011) and V-Fin (Yurista et al., 2012), EcoMapper AUV (Jackson and Reneau, 2014) and various Gliders such as ROUGHIE (Page et al., 2017).

Monitoring with towed undulating vehicles requires expensive ship time so vehicles need to be deployed efficiently. Ships usually move along pre-defined transects or grid patterns and the towed vehicles collect data along each transect while undulating to sample at multiple depths. However, grids that are too small may fail to capture the river plume, while those large enough to capture the river plume fully also may expend excess time and effort to capture data outside of the plume that are not needed. In addition, analyzing data from gridded sampling assumes stationarity of the river plume, and the river plume state may change markedly during the time spent sampling a large grid, thus introducing temporal change into the spatial variability of the data. The adaptive sampling strategy, which involves adjusting collection strategies based on previously collected data to minimize effort while maximizing river plume coverage is one possible solution to this problem. However, adaptive sampling raises a second serious problem: the large amount of high-frequency data that are collected by towed undulating vehicles are difficult to analyze quickly enough to adjust sampling. This is especially true for tow-yo sampling, where kriging interpolation is used to provide direct visualization of sampling results. Existing commercial software (such as Surfer, Golden Software) requires researchers to manually fit a variogram (Ludsin et al., 2009; Yurista et al., 2009), which is time-consuming and such data are usually analyzed after collection, making adaptive sampling impossible. New and efficient methods are needed to analyze data onboard the vessel as it is being collected.

In this study, we propose an automated kriging method that interpolates raw data onto grid maps that allow users to visualize patterns and adjust sampling in near-real time. To highlight the spatial distribution of variables in a distinct and informative way, we use hotspot analysis with local G statistics (Ord and Getis, 1995). We then further cluster the water chemistry data to explore the mixing structure of the river and lake water. The analysis framework has been implemented in an interactive Web application developed with the Shiny package in the R programming development environment. This will allow researchers on research vessels to easily perform analysis in near-real time.

2. Study area and data description

We illustrate the utility of the methods developed in this work for illuminating details of the river plume dynamics using data collected by the TRIAXUS undulating vehicle at the Manitowoc, Muskegon and Pere Marquette rivermouth areas in Lake Michigan during the summer of 2011. TRIAXUS, developed by MacArtney Underwater Technology, was towed behind the research vessel, Lake Guardian (operated by the EPA-Great Lakes National Programs Office), along pre-defined transects parallel or perpendicular to the shoreline. Fig. 1 shows the transects located in nearshore areas outside of the Manitowoc River, Muskegon River, and Pere Marquette River in Lake Michigan that were sampled during summer 2011. At these three sites, the TRIAXUS vehicle was deployed in undulating trajectories to measure water chemistry at different depths as the ship moved along each transect. The sampling depth of all paths ranged from 3 to 34 m. Average wavelengths of the undulating cycles (i.e., the distance between two peak points or two valley points) ranged from 0.126 km to 0.6 km.

The TRIAXUS carried multiple sensors that measured specific conductance, temperature, turbidity (measured as beam attenuation coefficients (BAT)), dissolved oxygen (DO), indices of chlorophyll concentration and algal accessory pigments, and zooplankton biomass and density. Chlorophyll concentrations were measured by a Fluoroprobe sensor, which uses excitation light with varying wavelengths to distinguish algae fluorescence among different algal groups. The validation and potential cautions of using Fluoroprobe to estimate phytoplankton community are given by Catherine et al. (2012). Zooplankton biomass and density were derived from a laser optical plankton counter (LOPC), which counts the number of particles in different size bins (from 105 μm to 1920 μm with step size 15 μm). The methods for comparing LOPC output to zooplankton biomass and density derived from traditional sampling methods are described in Watkins et al. (2017). Other variables were measured by a SeaBird CTD (conductivity, temperature, and depth) sensor attached to the vehicle. As a result, multi-dimensional spatial data with longitude, latitude, and depth as coordinates were generated.

3. Methodology

3.1. General description

Fig. 2 shows the data analysis framework proposed and applied in this work. First, a data quality assurance/quality control (QA/QC) step removes outliers and anomalies in the data. Next, we use automated kriging interpolation to visualize water chemistry properties on grid maps from the sampling data. Based on the interpolations, two spatial statistical methods, local G statistics and k-means clustering algorithm, are implemented to identify patterns in the data. The proposed methods aim to extract the information from the raw data paths and require minimal human interaction. Such automated processes can extract information during the sampling activities, rather than as post-sampling analysis, enabling near-real-time adaptive observation.