Ecological Informatics 26 (2015) 27–35

Contents lists available at ScienceDirect

Ecological Informatics

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

Combining a locomotion indicator and data mining to analyze the interactive patterns between copepods and ciliates

Meng-Tsung Lee [a], Jiang-Shiou Hwang [b], Chih-Yung Hsu [c], Yang-Chi Chang [c,*]

[a] Department of Marine Leisure Management, National Kaohsiung Marine University, No. 142, Haijhuan Rd., Kaohsiung 811, Taiwan
[b] Institute of Marine Biology, National Taiwan Ocean University, No. 2, Pei-Ning Road, Keelung 202, Taiwan
[c] Department of Marine Environment and Engineering, National Sun Yat-sen University, No. 70, Lien-Hae Road, Kaohsiung 804, Taiwan

ARTICLE INFO

Article history:
Received 23 November 2013
Received in revised form 6 January 2015
Accepted 7 January 2015
Available online 14 January 2015

Keywords:
Data mining
Locomotion indicator
Qualitative motion description
Copepod and ciliate interaction

ABSTRACT

Interactions between zooplankton not only affect the world's carbon fixation but also have a direct impact on the yields of the fishing industry. Both copepods and ciliates have a crucial linkage role in constituting the marine food web. Analyzing the predator–prey interactions between these two species helps us to understand the productivity of oceans better. In this study, we explored the interactive patterns between copepods and ciliates and used the locomotion indicator net-to-gross displacement ratio (NGDR) to conduct quantitative analyses on the swimming patterns of copepods. We discovered that the movement trails of copepods are more distorted in undisturbed environments where the NGDR was significantly lower. In an environment where ciliates were present, the NGDR of copepods was significantly higher. This result indicated that the movement trails in the latter scenario were more linear and that the NGDR can clearly distinguish the swimming patterns of copepods. In addition, this study developed a qualitative motion description to be embedded in the interactive trail data of copepods and ciliates, which facilitates the data mining technologies needed to perform advanced analyses. The results of the association rules and decision tree analyses clearly demonstrated the interactive characteristics that could not be explored solely using locomotion indicators. The most obvious characteristic is that the swimming patterns employed by copepods to approach ciliates were either a downward vertical sinking or a horizontal movement. Upward movements for the copepods were not typically observed. More detailed swimming patterns of interactions between copepods and ciliates were revealed using the rule-based forms.

1. Introduction

Both copepods and ciliates play a critical role in the marine food web (Chen et al., 2012; Hwang and Martens, 2011; Verity and Smetacek, 1996). These species constitute the major zooplankton in the ocean (Hwang and Martens, 2011; Sanoamuang and Hwang, 2011) and affect the fish population, carbon cycle and energy flow in the marine ecosystem (Dahms et al., 2012; Kiørboe et al., 1998; Ohman and Hirche, 2001). The individual behaviors of copepods and ciliates have been the focus of numerous studies (Chang et al., 2011; Doall et al., 2002; Schmitt et al., 2006; Wu et al., 2010; Yen et al., 2008). Analyzing the interactions between these two organisms remains an important issue to be explored because the results of such studies would help in understanding the fundamental mechanisms of the marine food web (Hwang and Strickler, 2001).

Many studies of zooplankton behavioral patterns in small-scale environments use video recording to trace movement trails (Buskey et al., 1987; Dahms and Hwang, 2010; Strickler and Hwang, 1999; Vandromme et al., 2010; Yen and Fields, 1992). Analyses after acquiring trail data mostly rely on locomotion indicators to objectively illustrate the characteristics of the zooplankton movement (Chen et al., 2012; Fields and Yen, 1997; Lee et al., 2010; Mazzocchi and Paffenhöfer, 1999). There are several locomotion indicators, such as net-to-gross displacement ratio (NGDR), diffusion coefficient (Visser and Thygesen, 2003), fractal dimension (Uttieri et al., 2005, 2007; Cianelli et al., 2009), and multi-fractal dimension (Schmitt and Seuront, 2001; Seuront et al., 2004a,b). Among these indicators, the NGDR is simple and effective and has been frequently used in related studies (Buskey et al., 1983; Chang et al., 2011; Chen et al., 2012; Jakobsen et al., 2005; Mazzocchi and Paffenhöfer, 1999; Tseng et al., 2013; van Duren and Videler, 1995; Weissburg et al., 1998; Wu et al., 2011). Eventually, statistical analyses of the indicators summarize the specific movement patterns for certain species of zooplankton.

Traditional data-oriented analyses, such as statistical inference, are hypothetico-deductive processes (Galitsky et al., 2007). These processes always presume relevant null and alternative hypotheses based on the literature or experts' knowledge and then substantiate the assumptions through an appropriate test. For biological data analyses, statistical inference plays an important role in finding correlations between biological parameters, but it is incapable of interpreting causal links

* Corresponding author. Tel.: +886 7 5252000 5176.
E-mail addresses: masonlee@webmail.nkmu.edu.tw (M.-T. Lee), jshwang@ntou.edu.tw (J.-S. Hwang), changyc@mail.nsysu.edu.tw (Y.-C. Chang).

http://dx.doi.org/10.1016/j.ecoinf.2015.01.001
1574-9541/

between related parameters (Galitsky et al., 2007). To address this problem, an alternative data-oriented analysis called data mining (DM) has been applied to analyses biological data. Unlike statistical inference, DM is an inductive inference process that typically draws general rules or patterns from a large data set. DM is an effective tool to assist decision makers in exploring useful information hidden in the data and finding solutions to a given problem (Han and Kamber, 2000). The applications of DM have covered various domains, including ecology, environment, and medicine (Chang et al., 2011; Creighton and Hanash, 2003; Dixon et al., 2007; Ekasingh and Ngamsomsuke, 2009; Kanevski et al., 2004; Recknagel, 2001; Xia et al., 2005; Zhang et al., 2005), and with the advancement in DM many research domains, such as bioinformatics, medical and health informatics, and ecological and environmental informatics, were emerged.

DM characteristically reveals hidden rules or patterns in a large data set and works well with unstructured and highly non-linear data, making it suitable for studying foraging behavior (Chang et al., 2011). Although the NGDR indicator is able to describe the approximate changes of trails during the interactions of copepods and ciliates, the indicator is not sufficient to interpret more subtle interactive patterns spatially. Therefore, this study applied DM techniques for spatial analyses of interactive movements between copepods and ciliates. We adopted a qualitative motion description as proposed by Lattner et al. (2006) to represent the interactive trail data of two organisms, which facilitates the advanced analyses with DM technology. Lattner et al. (2006) applied sequential pattern mining to find frequent patterns in dynamic situations of 2D RoboCup simulation games. Researchers in various fields have used DM technology for spatial pattern analyses. Xia et al. (2005) studied the spatiotemporal movement patterns of tourists visiting Phillip Island. Using clustering and decision tree analyses, they were able to identify patterns between the tourists' profiles and their spatiotemporal movement. Chang et al. (2011) used the decision tree method and locomotion indicators, including NGDR, turning levels or degrees, and fractal dimensions to analyze swimming patterns of ciliates in various foraging environments. Overall, the objectives of this study can be summarized as follows:

1. Using NGDR indicators to perform measurements of the data obtained from the swimming trails of copepods and ciliates, we described the qualitative abstraction of swimming behaviors in various situations.
2. We restructured the swimming trails of copepods and ciliates using a qualitative motion description such that the interactive patterns can be better represented.
3. By using data mining technology to perform analyses based on the coded trail data, we further delineated the interactive patterns between these two organisms.

2. Materials and methods

2.1. Data collection and preprocessing

2.1.1. Laboratory experiment

In order to study the interactions between copepods (predator) and ciliates (prey), five experiments were conducted and video recording by Professor Hwang Jiang-Shiou, one of the coauthors. The code names of the five experiments are listed in Table 1, which "Co" stands for copepods, "Ci" for ciliates, and "A" for alga. The focus of the study was to analyze the interactive patterns of zooplankton, which made Co_Ci_A (1) and Co_Ci_A (2) the experimental group while Co, Ci_A, and Co_A the control group. The organisms used in the experiments were cultivated by the laboratory at University of Texas, Marine Science Institute (Texas, USA). The filming period lasted from August 1 to 2, 2004, and data were recorded in a dark room under the same environmental conditions at 22 ± 1 °C. The only source of light was an infrared LED lamp (peak wavelength 910 nm; 1.45 volts). 15 adult copepods of *Acartia tonsa* and 255 marine ciliates of *Strobilidium Sp.* were placed inside 4.5 × 1.2 × 2 (length, width, and height) mm laboratory vessels, containing 15 ml of sea water at a salinity of 30 ppt (parts per thousand) and a regular pH ranging from 7.4 to 8.5. The density of copepods was 1150 individual per liter which is common in the copepod aquaculture ponds or coral reef environments, and is also in the range of many habitats of coastal waters. The density of ciliates was 17,000 individuals per liter which is in the range of many marine habitats. The high density of copepods and ciliates allowed us to track and visualize the prey/predator interactions. The algae that were used consisted of *Isochrysis galbanas*, *Gymnodinium Sp.*, and *Rhodomonas Sp*. at a 1:1:1 ratio. Filming began approximately 15 min after placement of the copepods and ciliates to allow them to adjust to the environment, and filming time last approximately 1 h. The conditions and settings of the five experiments are summarized in Table 1.

2.1.2. Trail data collection

It is necessary to convert the videos into movement trail data for the subsequent analyses. Therefore, the data pre-processing work included digitizing the trail images and extracting the coordinates associated with the trails to generate the digital trail data. This study used LabTrack (Bioras, Kvistgård, Denmark) to digitize the trails in video recordings. LabTrack can be used to track moving objects that are observed in continuous images and to perform frame-by-frame continuous image analyses. This software is well-suited for tracking zooplankton movement trails (Kiørboe, 2008; Titelman, 2001). The original videos were recorded at a 29.97 fps (frame per second), and LabTrack extracted spatial coordinates from the trail images at a 0.033 second interval. For the experimental group, we recorded the trails of copepods and ciliates when both of them appeared in the filming screen throughout the videos. It is not necessary to sample too many trails from the control group whose function is to provide a basis of comparison with the experiment group (Broglio et al., 2001). The study took 10 s of videos every 5 min for tracking the trajectories of copepods and ciliates. Table 2 shows the total numbers of trails of the five experiments identified by the image analyses. DM used only the trails from the experimental group, while NGDR analyses used all the trails from the five experiments.

2.1.3. Qualitative motion description

This study used a qualitative motion description to replace the continuous coordinate data of the movement trails for DM. A numerical

Table 1
Conditions and settings of the five experiments.

Code name	Co	Co_A	Ci_A	Co_Ci_A (1)	Co_Ci_A (2)
Description	Only copepods exist	Both alga and copepods exist	Both ciliates and alga exist	Copepods, ciliates, and alga all exist	Repeated experiment
Date	August 1, 2004		August 2, 2004		
Light condition	Dark room with an infrared LED lamp (peak wavelength 910 nm; 1.45 volts)				
Volume/salinity/pH of water	15 ml of 30 ppt sea water and a regular pH from 7.4 to 8.5				
Room temperature	23 °C	22 °C	23 °C	23 °C	23 °C
Filming vessel	4.5 mm by 1.2 mm by 2 mm (length, width, and height) on the top with cover				
Ciliates	None		*Strobilidium Sp.* (255 individuals)		
Algae	None	Mixture of *Isochrysis galbana*, *Gymnodinium Sp.*, and *Rhodomonas Sp.* (about 1:1:1)			
Copepods	*Acartia tonsa* (15 individuals)		None	*Acartia tonsa* (15 individuals)	