Computers & Geosciences 84 (2015) 29–36

Contents lists available at ScienceDirect

Computers & Geosciences

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

Case study

Sensitivity of Self-Organizing Map surface current patterns to the use of radial vs. Cartesian input vectors measured by high-frequency radars

Hrvoje Kalinić [a,*], Hrvoje Mihanović [a], Simone Cosoli [b], Ivica Vilibić [a]

[a] Institute of Oceanography and Fisheries, Šetalište I. Meštrovića 63, 21000 Split, Croatia
[b] Istituto Nazionale di Oceanografia e di Geofisica Sperimentale, Borgo Grotta Gigante 42/c, 34010 Sgonico, Trieste, Italy

ARTICLE INFO

Article history:
Received 20 October 2014
Received in revised form
22 April 2015
Accepted 24 August 2015

ABSTRACT

In this paper, the Self-Organizing Map (SOM) method was applied to the surface currents data obtained between February and November 2008 by a network of high-frequency (HF) radars in the northern Adriatic. The sensitivity of the derived SOM solutions was tested in respect to the change of coordinate system of the data introduced to the SOM. In one experiment the original radial data measurements were used, and in the other experiment the Cartesian (total) current vectors derived from original radar data were analyzed. Although the computation of SOM solutions was not a demanding task, comparing both neural lattices yielded the nondeterministic polynomial time (NP) problem for which is difficult to propose a solution that will be globally optimal. Thus, we suggested utilizing the greedy algorithm with underlying assumption of 1-to-1 mapping between lattices. The results suggested that such solution could be local, but not global optimum and that the latter assumption could lower the obtained correlations between the patterns. However, without the assumption of 1-to-1 mapping between lattices, correlation between the derived SOM patterns was quite high, indicating that SOM mapping introduced to the radial current vectors and subsequent transformation into Cartesian coordinate system does not significantly affect obtained patterns in comparison to the SOM mapping done on the derived Cartesian current vectors. The documented similarity corroborates the use of total current vectors in various oceanographic studies, as being representative derivative of original radial measurements.

1. Introduction

Although being developed in the early 1980s, Self-Organizing Map (SOM) method or Kohonen maps (Kohonen, 1982, 2001), an artificial neural network based on an unsupervised learning, has been introduced to oceanography recently (Liu et al., 2006; Liu and Weisberg, 2011), when different remote and in situ techniques became robust for collection of significant and appropriate amounts of data (Le Traon, 2013). High-frequency (HF) radars, a remote sensing technique developed for mapping of surface current fields in an area (Essen et al., 2000; Paduan and Washburn, 2013), advanced in recent decades and became a standard observational technique in coastal oceanography, with numerous fields of application-from early detection of tsunamis (Gurgel et al., 2011) to being an input in assimilation for numerical models (Paduan and Shulman, 2004). The use of Kohonen maps in the mapping of surface current maps obtained by HF radars has been initiated by Liu et al. (2007), who objectively extracted different spatial structures and temporal evolution of physical phenomena over the West Florida shelf. Several studies followed, which mapped HF radar data only and extracted characteristic patterns in Long Island Sound outflow (Mau et al., 2007), which merged HF radar and operational meteorological products to the SOM and assessed their interconnectivity in the northern Adriatic area (Mihanović et al., 2011), and which introduced joined HF radar and altimetry data to the SOM along the West Florida shelf (Liu et al., 2012). Finally, Hisaki (2013) applied several mapping methods, including SOM, for proper classification of surface currents derived by HF radars close to Okinawa Island.

All these studies tested the sensitivity of SOM solutions to different setup parameters or to various combinations of input variables, trying to correlate them by an objective nonlinear method and even attempting to use them in an operational ocean forecasting system (Hales et al., 2012). However, the computation of total (Cartesian) surface current vectors from radial (raw) data measured by single HF radars (Lipa and Barrick, 1983; Liu et al., 2014) and setup of HF radar data processing parameters may introduce large errors in determination of final surface current maps (Cosoli et al., 2012a; Kirincich et al., 2012). These errors may

* Corresponding author.
E-mail address: hrvoje.kalinic@izor.hr (H. Kalinić).

http://dx.doi.org/10.1016/j.cageo.2015.08.005
0098-3004/

potentially affect the objective mapping and alter the characteristic patterns derived by the SOM. For this reason, we posed a question whether the direct introduction of radial vectors to SOM and subsequent computation of corresponding Cartesian current patterns alters the SOM solutions in respect to the patterns derived directly from Cartesian current vectors. The answer to this question will come out of the analysis performed in this paper, based on the HF radar data collected in the northern Adriatic during a 10-month period of measurements. Section 2 describes the HF data used in the study. Section 3 introduces the methodology for obtaining SOM patterns from radial and from total data vectors, including the correlation metrics between the patterns. Section 4 presents the results, while discussion and conclusions are presented in Section 5.

2. The HF radar data

A network of HF SeaSonde radar systems produced by Codar Ocean Sensors was installed and operated in the northeastern Adriatic Sea during the 2007–2008 period. Three stations, Rt Zub (ZUB), Savudrija (SAV) and Bibione (BIB) were operational between February and November 2008 (Fig. 1), providing radial currents data every hour. Working frequency was set to 25 MHz, bandwidth was 100 kHz providing radial data with 1.5 km radial resolution and maximum operating range up to 50 km, with a 5° angular resolution. The nominal accuracy of total vectors as specified by the producer is less than 7 cm/s in magnitude and less than 10° in direction (Kovačević et al., 2004).

The radial dataset used in the analyses is divided in two parts, the winter–spring period (1 February to 31 May 2008) and the summer–autumn period (1 June to 30 November 2008). For the purpose of the analysis, the measuring period was divided on the winter–spring and summer–autumn intervals. These intervals may significantly differ in surface current fields, due to changes in atmospheric forcing and hydrology in the wider area (Cosoli et al., 2013). Radials maps were extracted from the inversion of the sea-echo using a range-gating technique to determine distance, and a direction-finding algorithm to determine the azimuth of the scattering ocean surface. Radial maps were quality-controlled following Cosoli et al. (2012a) to remove Doppler lines that were either poorly constrained by signal-to-noise (SNR) or had insufficient spectral quality factor, and were then merged using a SNR-weighted average to provide more robustness to the currents. Total surface current maps were derived on a Cartesian grid with 2 km × 2 km horizontal resolution using a least square approach (Lipa and Barrick, 1983; Barrick and Lipa, 1986). They were also controlled for spikes following Kovačević et al. (2004) and had spatial gaps filled using a Gaussian-weighted spatial interpolation. The grid points with large geometrical dilution of precision caused by poor intersecting beam geometry and an insufficient number of radial velocities from each site were removed from the mapping procedure (Chapman and Graber, 1997). Only measuring points with more than 40% of the operational data in a particular interval were introduced to the SOM method. SOM patterns derived from original radial data were similarly mapped onto the same regular rectangular grid.

3. The methodology

The main application of the SOM method is to reduce the dimensionality of the problem, or to create the abstraction of the input data via clustering, where each cluster is represented by one reference (codebook) vector (Kohonen 2001; Gan et al., 2007). More details on the SOM method and its advantages to other relevant methods reducing data dimensionality are provided in Appendix. The dimensionality reduction is done in the following way: (1) the array of SOM nodes forms a lattice structure that can be visualized as an elastic-net of points, where each point represents a neuron, then (2) the neurons compete and adapt to fit the input data, i.e. to approximate its density function, and (3) finally, a set of neurons which best represents the input data vectors is obtained. These neurons are known as the winning neurons, Best Matching Units (BMU) or codebook vectors (Kohonen, 2001; Pölzlbauer, 2004). In our case, we prescribed 12 codebook vectors

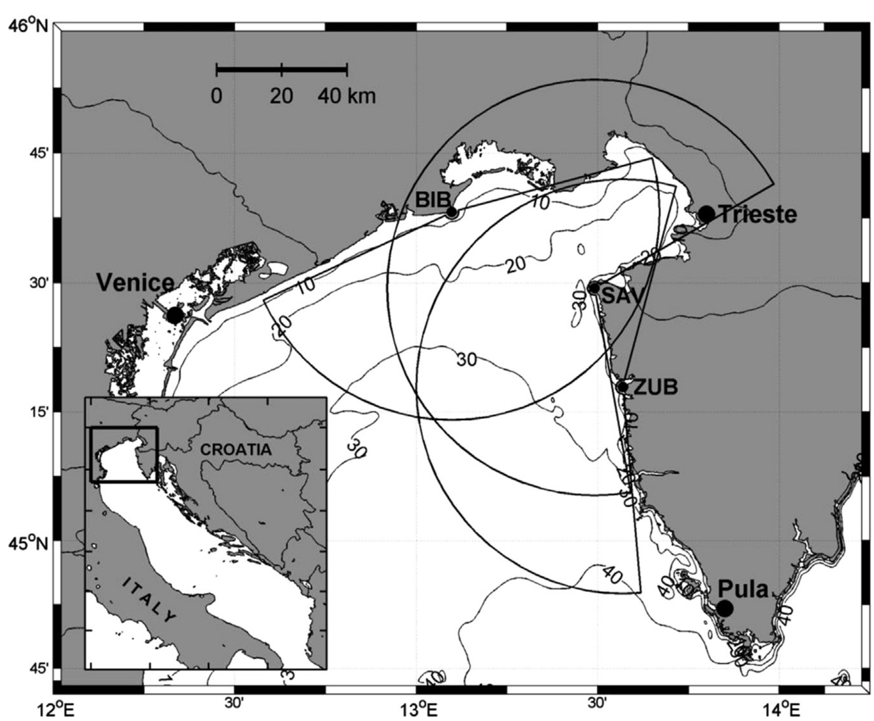

Fig. 1. Investigated area with bathymetry, positions of high-frequency radars (BIB – Bibione, SAV – Savudrija, ZUB – Zub) and their respective radial spatial coverage.