The pattern determination of sea surface temperature distribution and chlorophyll a in the Southern Caspian Sea using SOM Model

Bahmanzadegan A. R.[1*]; Lari K.[2]; Fatemi M. R.[3]; Azarsina F.[4]

Received: January 2012 Accepted: May 2012

Abstract

Remote sensing has changed modern oceanography by proving synoptic periodic data which can be processed. Since the satellite data are usually too much and nonlinear, in most cases, it is difficult to distinguish the patterns from these images. In fact, SOM (Self-Organizing Maps) model is a type of ANN (Artificial Neural Network) that has the ability to distinguish the efficient patterns from the vast complex of satellite data. In this study, the sea surface temperature data and chlorophyll a related to a part of south Caspian Sea were investigated weekly by NOAA satellite for three years (2003–2005) and the annual and seasonal patterns were extracted (elicited) with their relative frequency using the SOM model. In all patterns the Caspian Sea coast has the highest chl-a and when you go away from the shore the rate decreases and when you approach to the middle parts the chl-a is of the least proportion on the sea surface.

Keywords: Self-organizing maps, SST, Chlorophyll a, Caspian Sea

1- Department of physical oceanography, Faculty of marine science and technology, Islamic Azad University, Science and Research Branch, Tehran, Iran
2- Department of physical oceanography, Faculty of marine science and technology, Islamic Azad University, Tehran North Branch, Tehran, Iran
3- Departments of Marine Biology, Faculty of Marine Science and Technology, Islamic Azad University, Science and Technology Branch, Tehran
4- Department of marine Industries, Faculty of marine science and Technology, Islamic Azad University, Science and Technology Branch, Tehran
*Corresponding author's email: irali0641@gmail.com

[DOR: 20.1001.1.15622916.2013.12.1.9.2]

Introduction

SOM is an artificial intelligence method based on skills without control and it is also an important tool for eliciting the patterns(Kohonen, 1982). SOM is also a powerful technique to distinguish the pattern among complicated and vast satellite data. This is an attractive direct understanding method because it presents a simple image of the main pattern and uses different types of data.

If it is assumed that there are continuous patterns, above method especially can be useful for processing the oceanic dynamic data from satellite images. We examined some techniques promoting the comments concerning SOM to present a better comment about the patterns and annual and seasonal processing. By interpretation of the monthly relative frequency it is possible to explain the annual process; the importance of the situations with the same view is specified by synoptic maps. Special phenomena can be specified by following the time components of SOM. The time process for a defined point is stated by assisting parameters.

A problem in relation to use SOM is that the numbers of the chosen patterns are optional in a manner that the dimensions of the SOM are selected by the researcher. Essentially the data have a continuous pattern. The numbers of the chosen patterns depend on the comment concerning the details in the analysis. For instance, a very big SOM presents many patterns with detailed structure and a little one presents a general pattern. However, if the data have limited numbers of patterns SOM can distinguish the cases from the patterns with many relative frequencies to the patterns with zero relative frequency) Richardson, 2003).

One of the SOM method advantages is that the main patterns in a series of data can be spread into the format of the main data. So if the input data are SST images, finally our output data are SST patterns and this is an advantage in the multi-parameter techniques. If the output patterns are similar to the input ones, their interpretation is easier than the multi–parameter techniques outputs. Both SOM and empirical orthogonal functions methods are appropriate for the vast complex of data that makes it possible to reach to a facilitative interpretation of many images. Another advantage of SOM is that due to having an algorithm and strong calculations it can cover the lost data in interpretation.

Surface temperature and chlorophyll a determination in the South Caspian Sea is very important in different fields especially in the environmental cases. In this study by the use of self-organizing maps the main patterns are gained for temperature and chlorophyll a (chl-a) data with regard to the three years period (2003–2005) and the patterns are represented with regard to the relative frequency of different patterns and comparing temperature patterns with chl-a and with other assisting quantities such as topology and bed depth (Liu and Weisberg, 2004).

Oceanographic studies indicated that the measurements of Sea Surface temperature (SST) are from the main principles in determination of oceans treatments) Barton, 1995). SST is in contact with animal’s life, sea plants and the

[DOR: 20.1001.1.15622916.2013.12.1.9.2]

conception of the world's weather. The input SST data of boundary layers and similar complex of data are used in prediction models and atmospheric circulation (Emery et al., 2001).

Alexander and Scott (2002) described that the measurements of SST in large scales like El-Nino and possible processes with changes in universal weather are useful and practical. An important indicator for the presence of phytoplankton is Chl-a and this is the most common pigment in phytoplankton (Li et al., 2002). Bulter et al. (1998) reported that the existence of more than 0.2 mg/l of Chl-a shows that there would be enough food for fishes that is considerable in the case of commercial and fisheries.

Materials and methods

Self-organizing maps (SOM)

Self-organizing map is a new method to visualize the data in high dimensions. The method converts the nonlinear and complicated statistic relations of data with high dimensions to simple geometric relation with low dimensions.

Algorithm calculates the models of SOM in a manner that it explains the observations space in an optimal way. The models are arranged in a two dimensional meaningful sequence in a manner that the similar models are closer to each other than the dissimilar ones in a network. By this interpretation it is possible to consider SOM as a similarity graph or a cluster diagram.

The general principle of how SOM distinguishes the patterns from among the data is stated by an example of two dimensional data series. The data have two X, Y dimensions that in the Figure 1 were specified in gash-like circles, then these data were presented to the SOM network and the output network of SOM like filled rectangles entered to the data spaces.

These data have both linear and discrete situations. SOM puts the points between the data; here there are some prominent qualities. First nonlinear situation for $x < 3$ is presented well in the data. Secondly when there are more data the points are closer to each other rather than when there are less data and this causes to have images with good accuracy when we have enough data. It is noteworthy that the points out of the range (In fact, they are wrong) are not presented by SOM because they differ from other data. Finally SOM assumes that the data are continuous and put them in the center a point cut and it is useful when we have the lost data. So SOM distinguishes the empty places. In fact, the key power of SOM to distinguish the pattern is in its frequent algorithm. ANN works in frequent ways by which the data will be presented successfully.

The numbers of the points (output points) are defined by the user depending on the subject commentary details. Each point in the output layer relates to the input data through a weight link. Before beginning the frequent operations each link receives an amount by a random weight; that is each point assigns to one of the images randomly.

Then the first input data or line is compared with all the points of the network and a point which has the least difference with itself and the input data wins and forms the center of an up-to-date

[DOR: 20.1001.1.15622916.2013.12.1.9.2]

neighborhood. In an up–to-date neighborhood by getting away from the winner point the adaptability rate decreases.

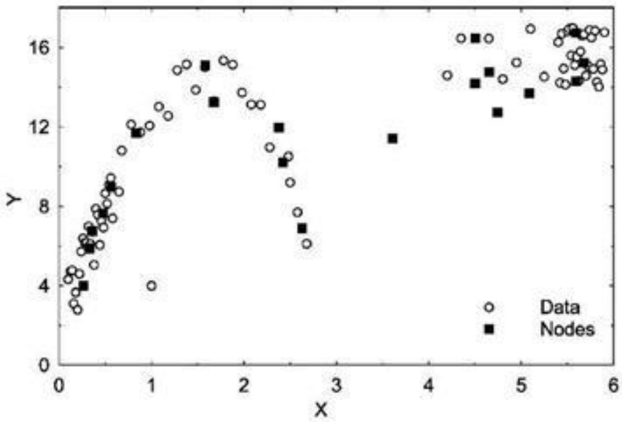


Figure 1: Showing SOM by an artificial two dimensional example (Richardson, 2003)

So the winner point becomes similar to the input data and the surrounding points go toward a direction showing similar but not exactly the same patterns. So in this method each point competes to adapt with the best input data. This process repeats for all input data as they pass through SOM serially. The input data circulate in SOM many times to have the patterns in SOM convergent toward the data. In this frequent process the speed or rate of the points becoming convergent toward the data is known as the ‘acquisition rate or speed’. The acquisition rate and neighborhood size (Radius) of up-to-dating decrease in frequent process in a manner that the primary, general pattern is distinguished. After frequent phase, SOM has many patterns of the data. Some patterns are similar and others are different.

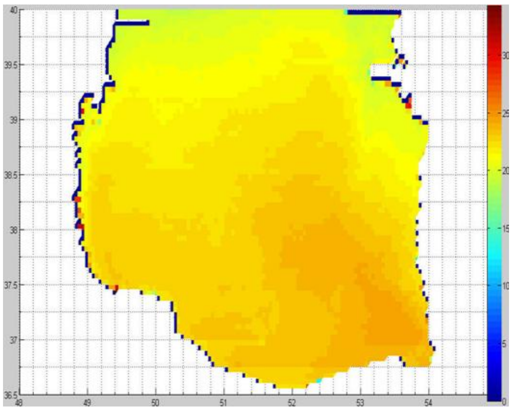
When essential patterns or the group of the patterns are distinguished there are some analyses promoting our comment. For instance, SOM can be used to classify the data and it is possible to define the patterns which are the best for each data.

The data and region

The region to be studied is between 36.50 N < lat < 40.00 N and 48.55 E < long < 54.30 E which is a region of the South Caspian Sea. The study method is an artificial intelligence way used for the field data concerning SST and surface chlorophyll a of this region for a three year period (2003–2005). The data were obtained from the NOAA site weekly. Finally 135 images were gained for each data. The data are related to the device MODIS and with the separation power of 0.05 grade for surface temperature and 0.041 grade for chl-a. Figure 2 shows the hydrography of the South Caspian region. During the steps before the processing, after extracting the data concerning SST and chl- a weekly belonging to years 2003-2005, 135 images gained for each one separately. The temperature images for each one had 8236 (76×116) pixel and the images concerning chl-a had 11815 pixel (85×135). The matrix of each image should be transformed to a column vector before entering them into the network.

[DOR: 20.1001.1.15622916.2013.12.1.9.2]

A



B

Figure 2: A) South Caspian hydrography, B) The patterns gained from self-organizing map

Then the input vectors are arranged next to each other and a primary matrix including 135 columns (same numbers of the input images) was obtained and the lost data were found by writing a program in the excel software and finally the input matrix became ready to be presented to the SOM network. By square location self organizing map appeared as 4×3 outputs with 10000 repetitions in the MATLAB software (Demuth, 2004).

The gained input matrixes were presented to the network in previous section and the network entered into repetition step and the weight matrix relating to the output technique was gained.

In the step after processing the separation column vectors and each vector transformed to matrix with regard to the latitude and longitude. Finally, the matrixes were drawn in the MATLAB software that has 12 main output patterns.

Results

Having prepared the main patterns by SOM it is possible to have different analyses. In first step it is possible to interperate each pattern separately. Relative frequency of different patterns is also specified in the presented interval, which is one of the self-organizing map advantages. With carrying out the model, twelve output patterns of Chl–a and SST with annual frequency percent and the patterns number were gained (Fig. 3& 4).

[DOR: 20.1001.1.15622916.2013.12.1.9.2]

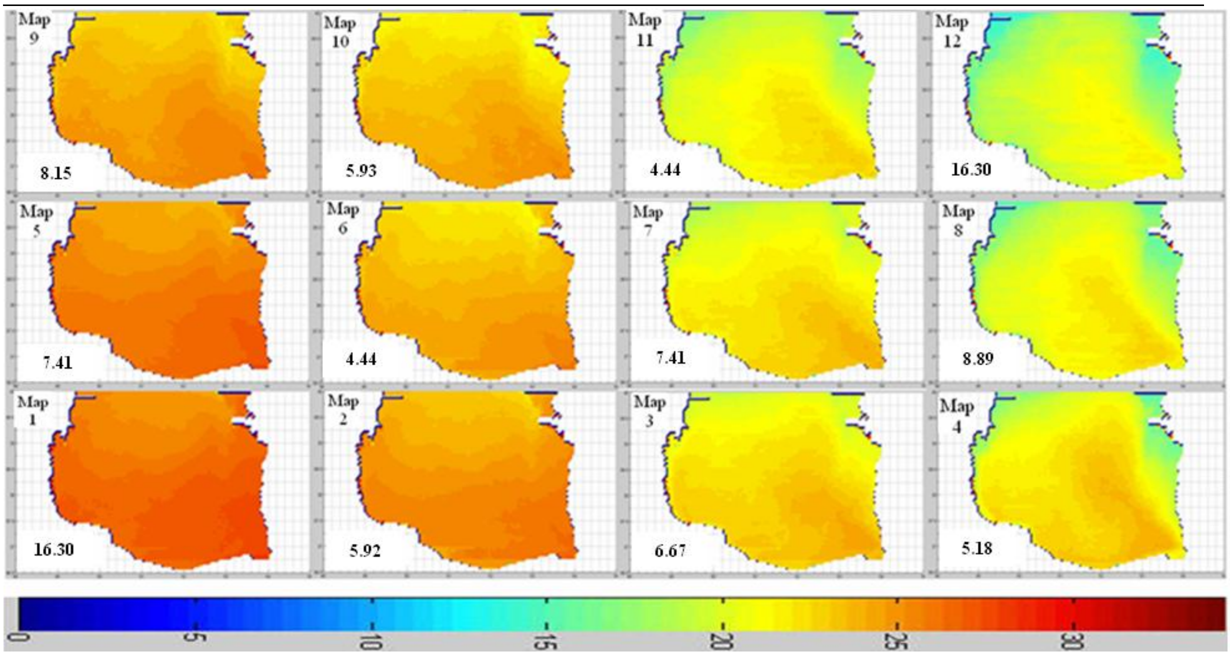


Figure 3: The occurrence frequency of twelve SST patterns in 2003 to 2005

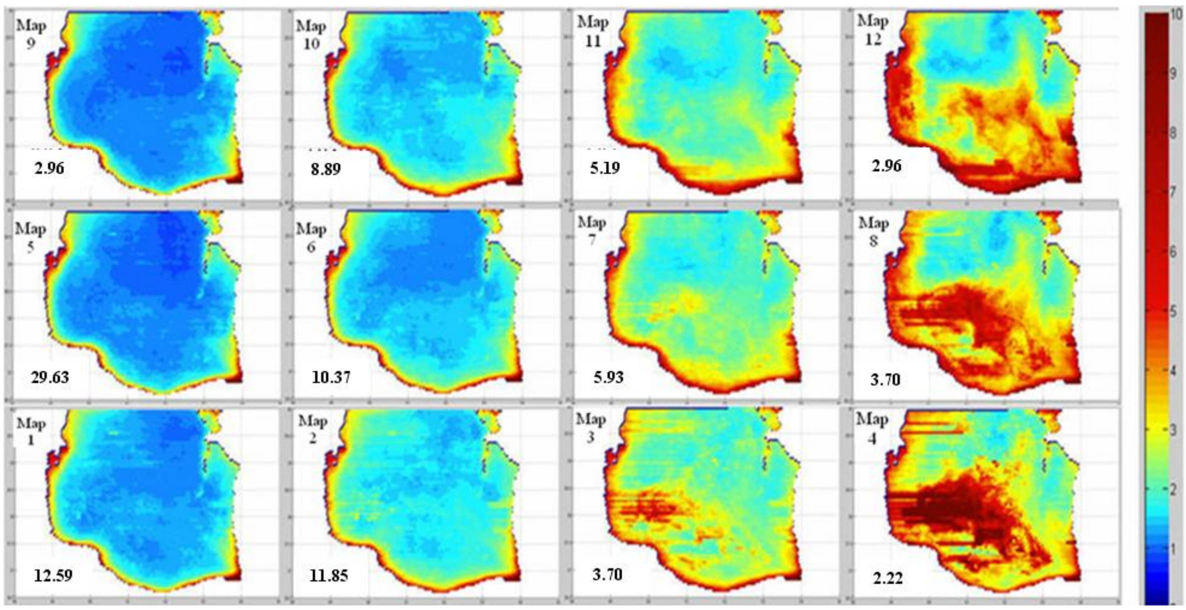


Figure 4: The occurrence frequency of twelve Chl–a patterns in 2003 to 2005

Also the minimum and maximum proportion of sea surface temperature and chlorophyll a in the output patterns (Table 1) and the most appropriate pattern for each month were obtained (Table 2).

Table 1: Minimum and Maximum value of SST and Chl-a associate with their coordinates in SOM patterns

Map Number		1	2	3	4	5	6	7	8	9	10	11	12
SST	Min (°C)	18.84	16.22	10.13	5.84	17.66	13.76	8.01	3.84	15.31	12.71	5.59	2.65
	Lat. (°N)	37.80	36.70	37.80	37.80	37.80	36.70	37.80	37.80	37.80	36.70	37.80	37.80
	Long.(°E)	53.85	52.40	53.85	53.85	53.85	52.40	53.85	53.85	53.85	52.40	53.85	53.85
	Max (°C)	34.36	33.00	27.58	22.97	33.42	30.53	24.67	20.39	32.00	29.19	22.07	19.10
	Lat. (°N)	37.50	37.50	37.50	37.50	37.50	37.50	37.50	37.50	37.50	37.50	37.50	37.50
	Long.(°E)	49.45	49.45	49.45	49.45	49.45	49.45	49.45	49.45	49.45	49.45	49.45	49.45
Chl-a	Min (°C)	1.00	1.24	1.55	1.43	0.88	1.08	1.45	1.34	0.90	1.17	1.37	1.33
	Lat. (°N)	36.62	39.58	38.96	38.49	38.96	39.00	39.00	39.82	38.96	38.92	39.08	39.82
	Long.(°E)	52.12	52.86	51.95	51.75	52.08	51.99	50.97	51.83	52.36	50.72	50.35	51.75
	Max (°C)	22.91	21.91	21.94	24.07	26.26	28.64	29.58	27.36	32.06	41.57	43.16	39.48
	Lat. (°N)	36.91	36.91	36.91	36.91	36.91	36.91	36.91	36.91	36.91	36.91	36.91	36.91
	Long.(°E)	53.55	53.55	53.55	53.55	53.55	53.55	53.55	53.55	53.55	53.55	53.55	53.55

Table 2: The best SST and Chl-a pattern for each month

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Agu	Sep	Oct	Nov	Dec
SST (° C)	8	12	12	11	3	9	5	1	1	9	3	4
Chl-a (mg/m3)	9	6	1	1	1	5	6	7	3	3	2	10

Seasonal relative frequency of sea surface temperature and chlorophyll a patterns are Shown in Figure 5, there is variety of the patterns in some seasons.

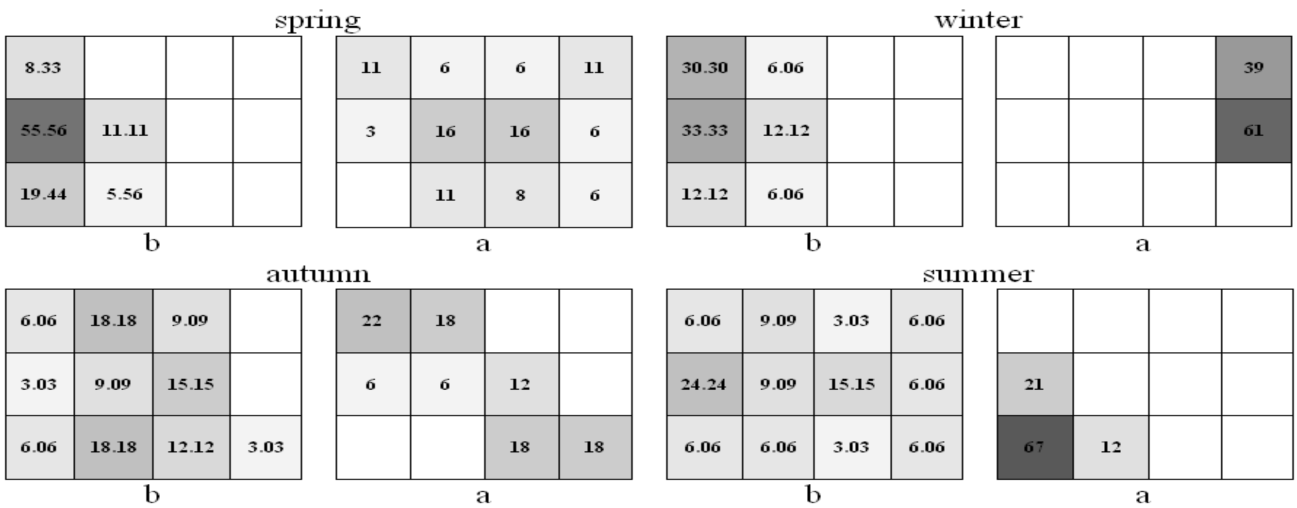


Figure 5: Seasonal relative frequency of SOM patterns, a) SST patterns, b) Chl-a patterns for years period of 2003-2005

In the studied years, the comparisons of changes in the relative frequency percent of the temperature warm patterns and relative frequency with high chl-a were conducted (Figure 6).

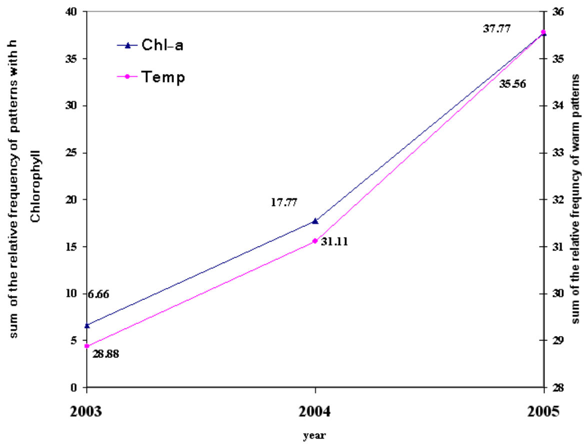


Figure 6: The changes of the temperature warm patterns and high chl-a patterns

Discussion

In a general view, by investigating the temperature patterns, it is possible to distinguish that the right patterns of map are cold and the left ones are warm patterns of the year (Fig3).

In all patterns there is a warm plum from the southeast side toward northwest and almost to the center of the map and the southeast part of the Caspian Sea is one of the warmest regions. It seems that there is a special thermal source in the southeast part. By attention of the Caspian Sea hydrography is specified that the coastal waters of the southeast of the Caspian Sea have less than 20 m depth and even less than 5m in many regions. Therefore receiving a defined value of thermal energy annually from the atmosphere and sun these regions are warmer than the deeper regions in the same latitude. In General when we move from the lower latitudes to the higher ones, the temperature is expected to be reduced that it is approved in the gained patterns, too.

The relative frequency of the thermal patterns in different seasons showed that there are varieties of the patterns in autumn and especially in spring in a manner that in spring almost all patterns occur; in such a way that it includes cold, mild and warm patterns. In autumn we have a mild situation as almost all mild situations are seen between the coldest and the warmest. These patterns show different situations that occurred with their relative frequency. To consider the monthly gained averages and patterns from SOM when winter begins the temperature decreases from January to February and a pattern appears in which is the coldest one in the map and during the year. There are very little changes from February to March as both of them have similar average patterns. By beginning of spring (from March to May) there is an increasing process in temperature and even it is more intensified from May to June and the temperature increase from June to July and August then we can see the warmest patterns in August. The average pattern does not change from August to September. The temperature decrease from September and is warmer in October then the warmness decreases a lot and the coldest patterns are from November to December when relatively the temperature increase.

Having investigated the patterns concerning chl-a in a general view it is possible to see that the right patterns have more chl-a than the left ones.

A general analysis concerning the gained patterns from chlorophyll a shows

[DOR: 20.1001.1.15622916.2013.12.1.9.2]

that in all patterns the Caspian Sea coast and seaside regions had the highest chl-a and when you go away from the shore the rate decreases and when you approach to the middle parts the chl-a is of the least proportion on the sea surface. So this paper shows the direct effect of the environmental factors and human activity on the sea and changes in the rate of chlorophyll a. The annual changes process of chl-a show that the least chl-a rate was in January and it increases from January to February relatively and decreases from February to March again. There is almost a similar average pattern with relative little chl-a rate in March, April and May and it decreases more in May to June in a manner that the process continues in July and August to September, but it stops from September to October and the chl-a decreases in little rate from October to November and these changes become more rapid in November to December and December to January.

Having investigated the seasonal changes of the chl-a with temperature changes it can be concluded that the patterns eight and twelve which are the coldest ones are in winter and all the chl-a patterns have little chl-a in the same season; especially pattern 9 with the least chl-a with relative frequency of 30.30 %.

The temperature patterns are drawn toward the middle ones in spring and this temperature change and increase has caused the relative frequency of pattern nine with very little chl-a decreases from 30.30 to 8.33% and this rate was distributed between other patterns with more chl-a specially the patterns 1 and 5. In summer the relative frequency concerning the patterns with less chlorophyll a decreases and the rate was distributed between the patterns with more chlorophyll a.

In autumn the temperature patterns distribute between the middle situations and the patterns related to chl-a are mostly distributed between the middle situations, too.

The relation between relative frequency and chl-a in the study period showed that the relative frequency of the warm patterns increased 28.88 to 35.56% in the studied years, respectively. Also the relative frequency of the chl-a patterns increased 6.66 to 37.77%, respectively.

References

Alexander, M. and Scott, J., 2002. The influence of ENSO on air–sea interaction in the Atlantic. Geophysical Research Letters, 29(14).

Barton, I. J., 1995. Satellite-derived sea surface temperatures: Current status. *Journal of Geophysical Research,* 100: 8777–8790.

Butler, M. J. A., Mouchot, M. C., Barale, V. and LeBlance, C., 1988. The application of remote sensing technology to marine fisheries:an introductory manual. FAO Fish.Tech.Pap.,(295):165 p.

Demuth, H. and Beale, M., 2004. Neural Network Toolbox for use with MATLAB, user’s guide, version 4.

Emery, W. J., Wick, G. A., Schluessel, P., and Donlon, C., 2001. Estimating Sea Surface Temperature from Infrared Satellite and in situ Temperature Data. *Bulletin of the American Meteorological Society* ,82: 2773-2785.

Kohonen, T., 1982. Self-organized information of topologically correct features maps Biological Cybernetics, 43, 59–69.

Li, H. P., Gong, G. C. and Hsiung, T. M., 2002. Phytoplankton pigment analysis by HPLC and its application in algal community investigations. *Botanical Bulletin of Academia Sinica*, 43: 283-290.

Liu, Y. and Weisberg, R. H., 2004. Sea Surface Temperature Patterns on the West Florida Shelf Using Growing Hierarchical Self-Organizing Maps. *Journal Atmospheric and Oceanic Technology*, 23, 325-338.

Richardson, A. J., 2003. Using self-organizing maps to identify patterns in satellite Imagery. Progress in Oceanography, 59, 223–239.

[DOR: 20.1001.1.15622916.2013.12.1.9.2]