Automated image processing for remote sensing data classification

Marco Reggiannini[1 [0000−0002−4872−9541]], Oscar Papini[1 [0000−0003−2069−5068]], and Gabriele Pieri[1 [0000−0001−5068−2861]]

Institute of Information Science and Technologies, National Research Council of Italy, Via G. Moruzzi 1, 56124 Pisa, Italy
{name.surname}@isti.cnr.it

Abstract Remote sensing technologies allow for continuous and valuable monitoring of the Earth's various environments. In particular, coastal and ocean monitoring presents an intrinsic complexity that makes such monitoring the main source of information available. Oceans, being the largest but least observed habitat, have many different factors affecting theirs faunal variations. Enhancing the capabilities to monitor and understand the changes occurring allows us to perform predictions and adopt proper decisions. This paper proposes an automated classification tool to recognise specific marine mesoscale events. Typically, human experts monitor and analyse these events visually through remote sensing imagery, specifically addressing Sea Surface Temperature data. The extended availability of this kind of remote sensing data transforms this activity into a time-consuming and subjective interpretation of the information. For this reason, there is an increased need for automated or at least semi-automated tools to perform this task. The results presented in this work have been obtained by applying the proposed approach to images captured over the southwestern region of the Iberian Peninsula.

Keywords: Image Processing · Remote Sensing · Mesoscale Patterns · Sea Surface Temperature · Machine Learning · Climate change

1 Introduction

To achieve a broader understanding and evaluation of the sea environment, an improvement in marine observation is required. Among all the relevant underlying processes in such a differentiated biological system, mesoscale events such as upwelling, countercurrents and filaments are of particular interest and constitute the subject of our analysis. These events, which transport deeper, colder and nutrient-rich waters to the surface, and affect the biological parameters of the habitat, enhancing the local biodiversity [7], can be observed by analysing Sea Surface Temperature (SST) recorded in remote sensing imagery.

Identifying and categorising upwelling regimes occurring in a marine ecosystem is an essential achievement for its characterisation. The main objective of this paper is to propose a method for performing an automatic classification of

images in place of the usual manual one completed by experts. When the number of images approaches the thousands, i.e. the typical order of magnitude having the goal to investigate long term and climate-related changes, the manual procedure is not manageable anymore. The method is applied to the Iberia/Canary Current System (ICCS), one of the least studied among the upwelling ecosystems [1]. Despite a general circulation similar to others, in ICCS we have diverse factors having a profound impact on the whole region.

The method proposed in this work is based on implementing an automatic procedure for classifying large datasets of images according to the different regimes of observable upwelling patterns. Such classification consists of several stages: starting from the extraction of quantitative features from a region of interest in the SST maps, proceeding to the characterisation of specific temperature patterns, which are correlated with the water flows between geographical points at different temperatures. The latter stage is performed by applying a set of rules to the computed features, which enable the assignment of a final class label to the considered region.

This paper follows a preliminary presentation given in [5] and represents a further extension of the work in [6]. It is arranged as follows: Section 2 provides a description of the employed dataset and the related ground truth classification; Section 3 reports on the pipeline used in our methods and describes a study case; Section 4 concludes the paper by discussing the outcomes of this work and providing a few considerations about future perspectives.

2 Materials

For the purposes of this work, SST data captured by Metop-A/B (EUMETSAT) and Aqua (NASA) have been collected and processed. Only data covering the region of interest were downloaded for each source (whose respective details are reported in Table 1). In particular, points with latitude between 35° and 40° N and longitude between 12° and 6° W were considered, resulting in 2–3 images per day at most.

Table 1. Data specifications

Satellite	Sensor Type	Spatial Resolution (km)	Temperature Resolution (°C)
Metop-A/B [3]	AVHRR	1	10^{-2}
Aqua [2]	MODIS	1	$5 \cdot 10^{-3}$

Expert oceanographers have preliminarily inspected the collected data to identify recurring SST patterns based on the detection of relevant mesoscale features (water filaments, upwelling jets and countercurrents). This way, it was possible to identify four prevailing patterns, named E1–E4 (see [5] for a detailed

description). Furthermore, each image was labelled according to the observed pattern, returning a ground truth dataset that could be used as a reference for the classifier implementation.

3 SST Analysis

In order to better analyse the different types of upwelling patterns, SST data are retrieved from the sources described in the previous section and arranged in a *spaghetti plot*, which is a simultaneous representation of the different SST trends for a given geographical area and a time interval. It is obtained by first dividing the considered area into a grid of small squares (whose size may be equal to or larger than the image spatial resolution). Then, for each square, the SST spatial average value is computed for each time sample in the dataset falling within the considered time window. Finally, the obtained ensemble of averaged SSTs is plotted versus time within the same diagram.

Figure 1 shows an example of an event classified as E4 in the ground truth and the spaghetti plots corresponding to the selected areas. Events of type E4 are characterised by the presence of a warm countercurrent originating in the Gulf of Cádiz and running along the southern Iberian coast, eventually reaching Cape St. Vincent (see Figure 1c). A cold water filament going westwards is also recognisable (see Figure 1b), which is a pattern typical for events of type E1. In this case, the squares' size and the time interval are 0.25° and 15 days respectively (notice that the ground truth event occurs at the end of the time window). After several tests, these specific values have been chosen since they return a better agreement between the results and the ground truth.

A spaghetti plot is then processed to extract statistical features, which depend on the SST signal in each square and its neighbourhood. These features are later used to classify the considered area, which is then associated with one of the four mesoscale patterns.

Let a be a square in the grid. As said, we have a temporal series of spatial SST averages in a, say μ_i, computed at times t_i, $i = 1, \ldots, n$. Notice that n may change from square to square, since it depends on the number of SST values captured by the sensor. In fact, the SST recording may fail for some parts of the area of interest (e.g. due to interfering clouds disturbances). Because of these considerations, the number of samples n can be considered as an index of reliability for the classification of the square a. The statistics features computed for a are:

1. the temporal mean $\mu(a)$, defined as the mean of the values μ_i;
2. the standard deviation $\sigma(a)$, defined as the standard deviation of the values μ_i;
3. the linear regression coefficient $\theta(a)$, defined as the slope of the straight line that better interpolates the values (t_i, μ_i).

The values μ, σ and θ are computed for every square in the grid. Figure 2 shows these values for our case study (event of 7 October 2017; see Figure 1).

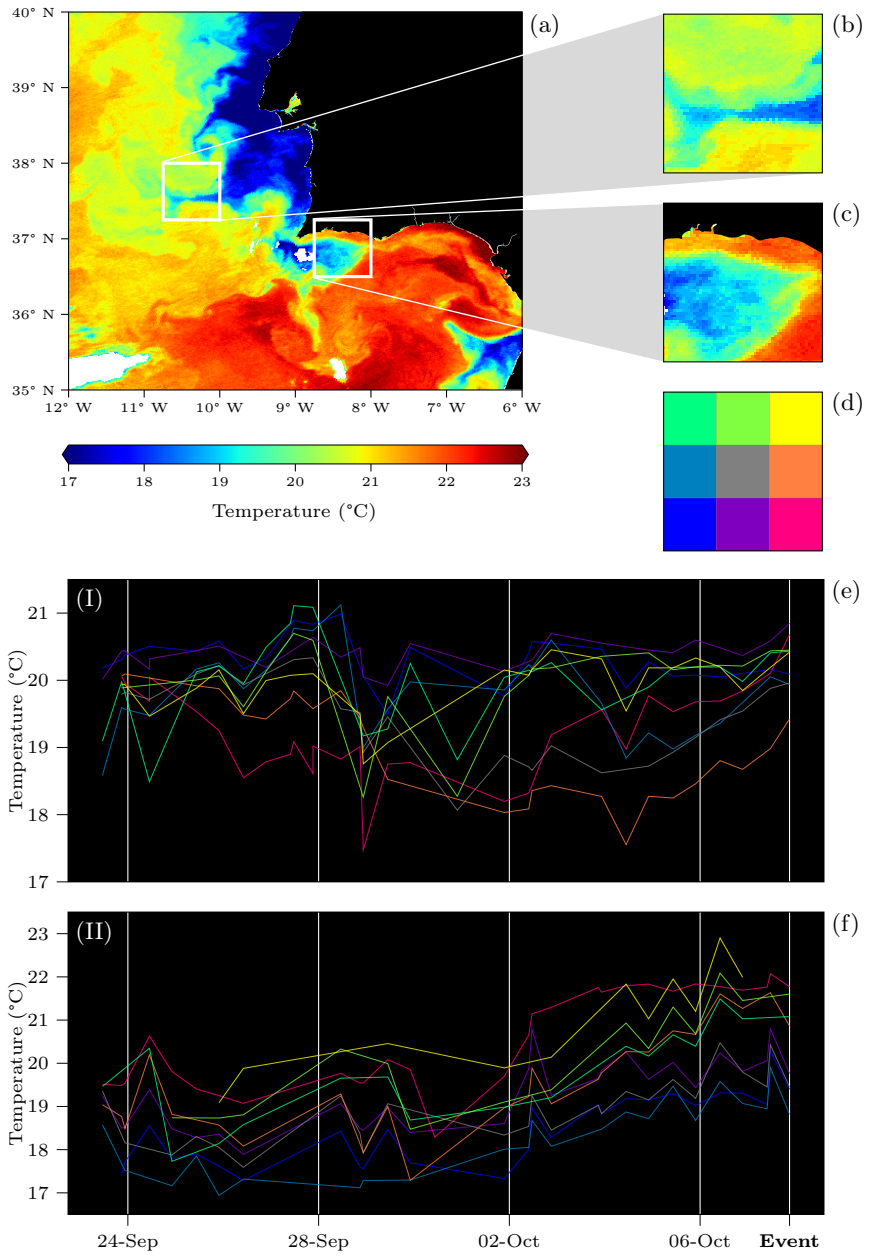

Figure 1. Event of 7 October 2017 at around 21:00 UTC. (a) SST map at the date of the event; (b) detail of the SST in the reference area for spaghetti plot I (latitude between 37.25° and 38° N, longitude between 10.75° and 10° W); (c) detail of the SST in the reference area for spaghetti plot II (latitude between 36.5° and 37.25° N, longitude between 8.75° and 8° W); (d) reference grid for both plots (dimension of squares 0.25°); (e,f) generated spaghetti plots.

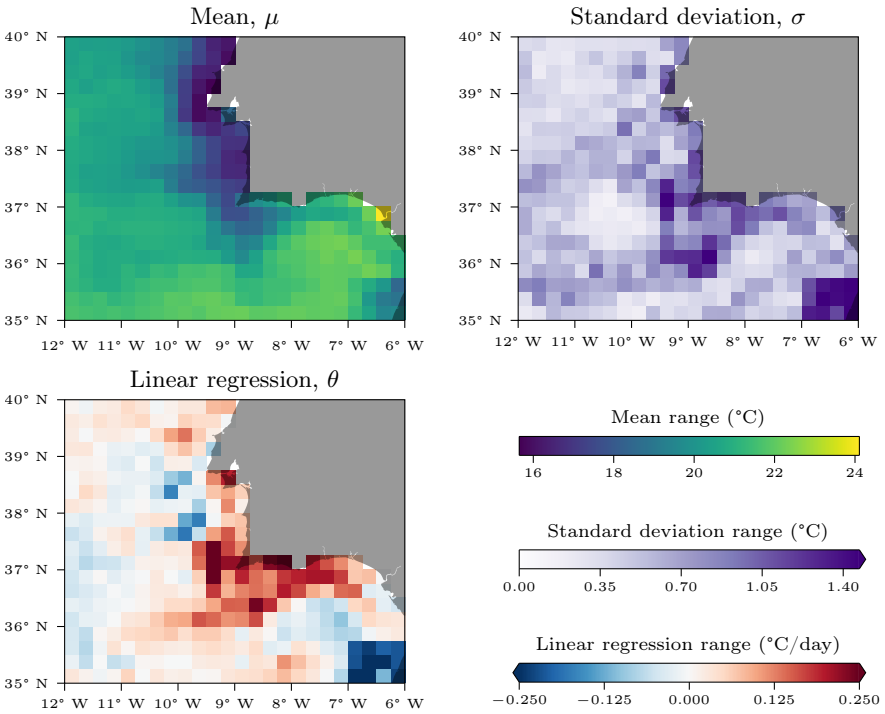

Figure 2. Maps representing the values of the statistics for each square, computed for the event of 7 October 2017 using data from the period between 23 September and 7 October.

The next step is to apply a set of rules to obtain, for each square a, an array of four scores (e_1, e_2, e_3, e_4), with $e_j \in [0, 1]$. The value e_j represents a belief index for the event of type Ej to have occurred inside a at the end of the considered time interval. The implementation of the rules is a crucial component for the classifier. Indeed, they are handcrafted so that the score e_j is boosted only if the behaviour of the features μ, σ and θ, inside and in the neighbourhood of the square a, matches the one observed in the case of an Ej pattern. Figure 3 shows the scores for each square of the grid for the event of Figure 1, computed using the values of the statistics depicted in Figure 2.

The classification of a square is finally completed by considering the maximum score $e_m = \max\{e_1, e_2, e_3, e_4\}$: if e_m is above a certain threshold, empirically defined, then the square is labelled "Em"; otherwise no label is assigned. Figure 4 represents a heatmap with the classification results applied to the event of Figure 1 using the scores of Figure 3, with each square coloured with the corresponding classification label. Also, each square is labelled with the numerical percentage of the related SST data, which is proportional to the n value as discussed above.

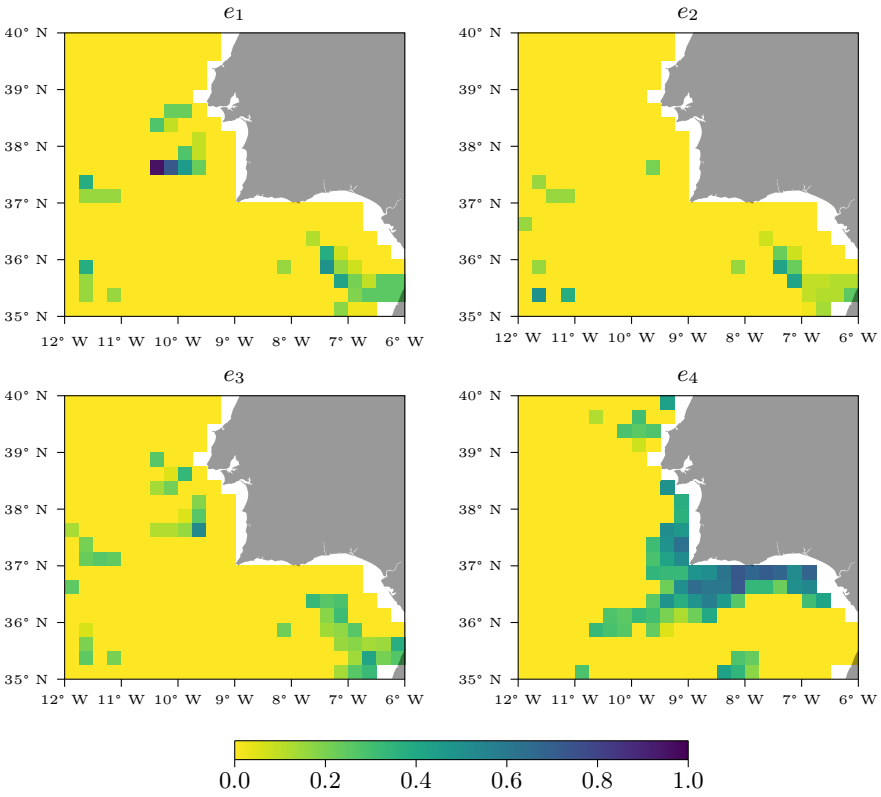

Figure 3. Array of the four maps with the scores of the squares relative to the event of 7 October 2017.

4 Discussion and Conclusion

In this work, a methodology for classifying upwelling events based on the analysis of SST time series has been proposed. Preliminary tests proved that the proposed method succeeds in classifying different mesoscale events. A few considerations can be pointed out concerning the presented case study (Figure 1). First, it is worth noticing that the labelling returned by the classifier agrees with the ground truth: among the squares located in the area where E4 events usually occur, those that fulfilled the previously mentioned data abundance constraints have been correctly labelled (Figure 4). Second, it is to remark that the occurrence of a mesoscale event is a phenomenon that depends on both space and time. In this work, the criterion adopted to formulate the classification rules is to estimate how much an SST pattern, observed within a given spatial and temporal neighbourhood, is close to a theoretical one. For the presented case study, these rules have been applied to the computed SST statistics (see Figure 2), eventually yielding the score maps in Figure 3. Each score map represents the

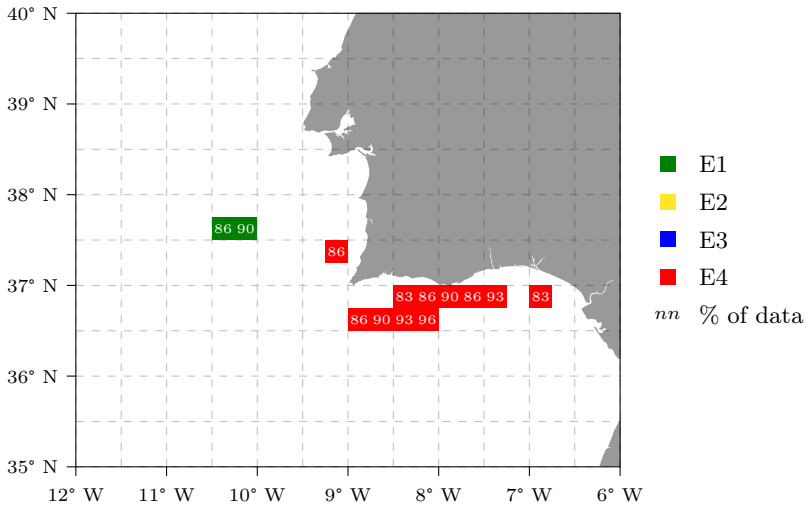

Figure 4. Labels given to each square of the grid, depending on their scores.

closeness between the SST signal and each possible mesoscale event. Moreover, the extracted features take into account not only the SST final observation, corresponding to the ground truth label, but also the SST variations captured in the preceding time window. This is the reason behind the presence of squares classified differently from E4, in apparent conflict with the ground truth. Since the proposed approach takes into consideration the SST signal over an extended range of time, it is reasonable that more than one label is assigned, in agreement with the multiple observed mesoscale events. It is even more so considering that inside the presented case study's dataset, different ground truth labels have been assigned to images captured very close in time. For example, on 6 October, two distinct events are observed: one classified as E1 in the ground truth approximately at 10:00 UTC, and a second one around 21:20 UTC classified as E4.

The test and validation of the proposed algorithm are carried out and will continue as part of the activities of the EU H2020 project NAUTILOS [4].

Acknowledgements

The authors express gratitude to Prof. Flávio Martins and Dr João Janeiro from the University of Algarve, Centre for Marine and Environmental Research, for their support.

Funding

This paper is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101000825 (NAUTILOS).

Conflict of Interests

The process of writing and the content of the article does not give grounds for raising the issue of a conflict of interest.

References

[1] Francisco P. Chavez and Monique Messié. 'A comparison of Eastern Boundary Upwelling Ecosystems'. In: *Progress in Oceanography* 83.1 (2009), pp. 80–96. DOI: 10.1016/j.pocean.2009.07.032.

[2] NASA/JPL. *GHRSST Level 2P Global Sea Surface Skin Temperature from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the NASA Aqua satellite (GDS2)*. 2020. DOI: 10.5067/GHMDA-2PJ19.

[3] OSI SAF. *Full resolution L2P AVHRR Sea Surface Temperature MetaGRanules (GHRSST) - Metop*. 2011. DOI: 10.15770/EUM_SAF_OSI_NRT_2013.

[4] Gabriele Pieri et al. 'New technology improves our understanding of changes in the marine environment'. In: *Proceedings of the 9th EuroGOOS International Conference*. EuroGOOS, 2021.

[5] Marco Reggiannini, João Janeiro, Flávio Martins, Oscar Papini and Gabriele Pieri. 'Mesoscale Patterns Identification Through SST Image Processing'. In: *Proceedings of the 2nd International Conference on Robotics, Computer Vision and Intelligent Systems — ROBOVIS*. SciTePress, 2021, pp. 165–172. DOI: 10.5220/0010714600003061.

[6] Marco Reggiannini, Oscar Papini and Gabriele Pieri. 'An Automated Analysis Tool for the Classification of Sea Surface Temperature Imagery'. In: *Pattern Recognition and Image Analysis* 32.3 (2022). DOI: 10.1134/S1054661822030336.

[7] Rubén Varela, Fernando P. Lima, Rui Seabra, Claudia Meneghesso and Moncho Gómez-Gesteira. 'Coastal warming and wind-driven upwelling: A global analysis'. In: *Science of The Total Environment* 639 (2018), pp. 1501–1511. DOI: 10.1016/j.scitotenv.2018.05.273.