


ORIGINAL ARTICLE

Piece-wise constant cluster modelling of dynamics of upwelling patterns

Susana Nascimento¹  | Alexandre Martins¹ | Paulo Relvas² | Joaquim F. Luís³ | Boris Mirkin^{4,5}

¹Department of Computer Science and NOVA Laboratory for Computer Science and Informatics (NOVA-LINCS), NOVA School of Science and Technology, Lisboa, Portugal

²Department of Earth Marine and Environmental Sciences, Universidade do Algarve / Centre of Marine Sciences (CCMAR), Faro, Portugal

³Department of Earth Marine and Environmental Sciences, Universidade do Algarve / Instituto Don Luis (IDL), Faro, Portugal

⁴Department of Data Analysis and Artificial Intelligence, National Research University Higher School of Economics, Moscow, Russian Federation, Moscow, Russian Federation

⁵Department of Computer Science, Birkbeck University of London, London, UK

Correspondence

Susana Nascimento, Department of Computer Science and NOVA Laboratory for Computer Science and Informatics (NOVA-LINCS), NOVA School of Science and Technology, Lisboa, Portugal.

Email: snt@fct.unl.pt

Funding information

NOVA LINCS, Grant/Award Numbers: LA/P/0101/2020, UIDP/04326/2020, UIDB/04326/2020, UIDB/04516/2020; Foundation for Science and Technology; Basic Research Program of the National Research University Higher School of Economics Moscow

Abstract

A comprehensive approach is presented to analyse season's coastal upwelling represented by weekly sea surface temperature (SST) image grids. Our three-stage data recovery clustering method assumes that the season's upwelling can be divided into shorter periods of stability, ranges, each to be represented by a constant core and variable shell parts. Corresponding clustering algorithms parameters are automatically derived by using the least-squares clustering criterion. The approach has been successfully applied to real-world SST data covering two distinct regions: Portuguese coast and Morocco coast, for 16 years each.

KEYWORDS

coastal upwelling, data recovery clustering, spatiotemporal clustering, SST images, time series segmentation

1 | INTRODUCTION

This article contributes to the field of spatiotemporal clustering focusing on the issue of modelling coastal upwelling patterns. In this introductory part we briefly discuss the subjects of spatiotemporal clustering, of upwelling modelling with clustering, and specifics of our approach to clustering.

1.1 | Spatio-temporal clustering

The richness of the field of spatiotemporal clustering systems stems from a variety of (a) applications, (b) data structures, and (c) mathematical models (see recent reviews in Alam et al. (2022), Ansari et al. (2019), Atluri et al. (2018), Kamenetsky et al. (2022), Mozdzen et al. (2022), Wazarkar and Keshavamurthy (2018) and Yu et al. (2020)).

Among applications, one should point to image data analysis including such areas as Medical imaging, 3D imaging, Oceanography, Industrial automation, Remote sensing, Mobile phones, Security, Face related applications, and so on (Wazarkar & Keshavamurthy, 2018). Other application areas are Earth Sciences (Donatelli et al., 2022; Martino et al., 2018; Ramachandra et al., 2019; Tonini et al., 2022; Yu et al., 2020), Health Informatics (Kamenetsky et al., 2022; Kulldorff & Nagarwalla, 1995; Mattera, 2022), as well as other societal phenomena (Mozdzen et al., 2022).

Concerning data structures, one should distinguish between at least four types of spatiotemporal (ST) data structures: '(i) event data, which comprises of discrete events occurring at point locations and times (e.g., incidences of crime events in the city), (ii) trajectory data, where trajectories of moving bodies are being measured (e.g., the patrol route of a police surveillance car), (iii) point reference data, where a continuous ST field is being measured at moving ST reference sites (e.g., measurements of surface temperature collected using weather balloons), and (iv) raster data, where observations of an ST field is being collected at fixed cells in an ST grid (e.g., fMRI scans of brain activity)' (Atluri et al., 2018, p. 6).

Accordingly, data types (i) and (iii) are represented by geographic points, data type (ii) by trajectories or time series, and data type (iv), by spatial maps. These representations to a large extent determine the types of mathematical and computational models applicable to them.

A set of statistical models for determining 'anomalous' events have been developed based on modelling the data by cylindrical bodies whose height reflected the temporal aspect while their circular part reflected the spatial aspect (Donatelli et al., 2022; Tonini et al., 2022). In Kamenetsky et al. (2022) log-linear models are considered for circular clusters. Bayesian framework, as well as conventional finite mixture of Gaussian distributions framework are scrutinized in Mozdzen et al. (2022). A review of spatiotemporal event modelling can be found in Yu et al. (2020). The work by Mattera (2022) separates spatial and temporal aspects into additively related parts, and Evers and Linsen (2022) proposes a model and method for partitioning the space in segments.

1.2 | Clustering in the analysis of oceanographic phenomena

Our concern is using clustering for modelling the phenomenon of upwelling which is characterized by the rising of nutrient rich cold waters from the deeper levels to the ocean, under the joint action of the wind stress and the earth rotation (Coriolis effect).

Long term spatiotemporal analysis of coastal upwelling is essential for the study of ocean dynamics, coastal resource managements and climate models. Coastal upwelling is responsible for the major highly productive coastal regions of the world ocean, with strong social-economic impact. Thus, the estimation of the long-term variability of the upwelling patterns is crucial to access the upwelling behaviour under a climatic change scenario.

Upwelling leaves a clear signal in the sea surface temperature (SST), that is truthfully recognized in the infrared satellite images. Regular and continuous operational global observations of the SST dates back to the eighties of the last century. Therefore, a climatic scale time-series of good resolution remote sensing SST records, at the global scale and collected on a daily basis, is available for analysis. This represents an enormous amount of data that cannot be examined by processes others than automatic ones.

Unfortunately, so far no general physical theory to explain the upwelling patterns has been developed. The literature presents discussions of local upwelling systems in various locations such as the off-coast of Peru, China, or the USA West coast. This latter subject is reviewed at length in Jacox et al. (2018). Even in the base operation of the identification of the upwelling front, the sudden transition between the cold upwelled waters and the warmer offshore waters, a crucial tool for coastal water management, is not fully automated yet. Therefore, this operation is currently carried out manually by experienced oceanographers.

In the relevant literature, one may distinguish between two streams of work: (1) methods for clustering at oceanographic phenomena and (2) computer-oriented studies of upwelling systems. Stream (1) includes works on

- Trajectory clustering: Specifically, a method from Liu et al. (2019) proceeds by using a similarity measure between trajectories, whereas the algorithm from Zhang et al. (2018) first partitions a trajectory into a set of line segments by characteristic points at which the object makes a sharp turn, and then groups similar line segments together into a cluster.
- Generative model clustering: The work in Ramachandra et al. (2019) models spatial clusters with Gaussian density functions, so that patterns that are highly different from the Gaussian are considered anomalous. The authors in Sambe and Suga (2022) put together temperature and salinity profiles while applying the Gaussian mixture model to cluster about 200 ocean depth layers; to discover that the resulting clusters form spatially contiguous regions.
- Hotspot discovery: The work in Martino et al. (2018) applies a version of fuzzy c – means method extended by using hyperspheres as cluster centers into a mix of spatial and temporal features. The resulting clusters apply to forecasting earthquake hotspots.
- Correlation study: Konstantaras (2020) performs spatiotemporal clustering upon seismic data; then uses the obtained clusters to investigate whether a potential correlation between seismic energy release rates and the time interval between large earthquakes existed. A feature of the study is that the correlation is analysed with a deep neural network.
- Time-stable clustering: Article by Chen et al. (2015) introduces the concept of stable spatiotemporal cluster as a set of 'core points' that never change cluster memberships at a given time window. To identify such clusters, the authors use a methodology extending the popular DBSCAN



method, adjusted to spatiotemporal specifics according to Birant and Kut (2007). The methodology inherits both advantages and shortcomings of the DBSCAN method. The main advantage is the ability to capture any spatiotemporal irregularity of the data in the shape of the density function being recovered by the method. The disadvantages relate to the presence of several ad-hoc user-specified parameters defining the sizes of neighbourhoods, the numbers of points and thresholds. The data include three sea surface characteristics: temperatures, surface heights and wave heights.

Stream (2), computer-oriented studies of upwelling systems, are mostly based on the so-called sea surface temperature (SST) records, grids of SST measured at some time moment. Some researchers combine this data with related features, first of all, winds. Article by Ramanantsoa et al. (2018) combines winds with ocean currents as upwelling drivers to propose a method for upwelling front detection in the waters near Madagascar. The work in Nowicki et al. (2019) goes much further to develop a working automated detection system for forecasting coastal upwelling in the Baltic Sea. To this end, they utilize a method from Lehmann et al. (2012) for upwelling modelling based on classifying winds in 'favourable' and 'not favourable' categories. Recent trends in SST, Chlorophyll-a, net primary production (NPP) and meridional wind stress in the Eastern North Atlantic Subtropical Gyre (NASE) are studied in Siemer et al. (2021).

A rather intense course within this stream is taken by articles using the SST data only. A great advantage of this approach is that at the temperature data only, there is no need in clustering algorithms looking for sophisticated spatial cluster shapes. Image pixels of the same or similar values of temperature can form real complex shapes indeed. Our articles (Nascimento et al., 2012, 2015) developed an approximation approach at which the spatial compactness of pixel clusters was guaranteed by a superficially technical trick—the user-imposed window of a pre-specified size, within which all the local operations are executed. This approach appears very close to the popular seeded region growing (SRG) image clustering by Adams and Bischof (1994), with two differences that are worth noticing: (a) the criterion involves the within-cluster average temperature rather than the temperature difference as in SRG; (b) main parameters are adjusted automatically according to the least-squares approximation criterion rather than heuristically. Similar constructions involving within-cluster average temperatures are being developed regarding different areas of the ocean in Aouni et al. (2021), Houghton and Wilson (2020), Huang and Wang (2019), Shi et al. (2021), and Tamim et al. (2015, 2019). As is well-known, the accuracy and efficiency of heuristic algorithms strongly depend on the appropriate selection of parameters, which is done manually in these articles (see also Ansari et al. (2019)). A rather interesting pre-processing option is developed in Aouni et al. (2021) considering the normalization of the SST data using perpendicular lines to the coastline, which tackles the so-called over-segmentation problem (see further in Section 2.3).

1.3 | Data recovery clustering for the analysis of upwelling phenomenon

Our work falls in the field of spatiotemporal data clustering and follows the type of approach introduced by Chen et al. (2015) where clusters may move and change their size, shape and location, but have 'core points' that never change cluster memberships for a given time window. However, unlike that approach which relies on the popular DBSCAN (Ester et al., 1996) to capture dense fragments of the data distribution, we follow a different methodology to explicitly distinguish between core points and boundary points, which is reflected in the concepts of cluster 'core' and cluster 'shell' in Rodin and Mirkin (2017)). Our data recovery clustering framework is to overcome shortcomings of the most popular spatiotemporal clustering algorithms like ST-DBSCAN (Birant & Kut, 2007) and ST-OPTICS (Agrawal et al., 2016) that demand parameter settings from the user, which heavily influence the quality of clustering results. Specifically, we maintain that the cluster structure to be found can be represented as an 'ideal data structure' in the format of the data under consideration. Such a representation enables us to compute the squared difference between the data and the structure to be found. We use this squared difference as a goodness-of-fit least-squares criterion: the smaller the difference, the better the fit. Monograph by Mirkin (2012) demonstrates that this view is consistent with the conventional clustering studies and, also, leads to some theoretically sound properties.

The additivity of the least-squares clustering criterion enables us to extract clusters one-by-one rather than simultaneously. Such a sequential extraction strategy has two advantages in the context of upwelling description: first, it is adequate to the nature of the upwelling patterns under consideration, because at each time instant, there is only one major upwelling region to occur at a coastal area (if any); second, it allows to apply the Self-Tuning Seed Expanding Cluster (STSEC) algorithm introduced by Nascimento et al. (2015), a version of the pioneer Seeded Region Growing (SRG) approach in image analysis (Adams & Bischof, 1994), as a part of the current core-shell methodology.

The region that concerns us is the Portuguese and Atlantic North African coastal region, that represents the major portion of the Canary Current Upwelling System. Good satellite SST records of the ocean surface for both regions are available for almost 40 years.

Rather surprisingly, our attempts to find a constant core in a temporal sequence of SST grids during an upwelling season (1 year) failed: the local upwelling here does change its location, however slightly, several times during that period, which made us to introduce an intermediate stage of finding relatively stable 'time windows', here called time 'ranges', for obtaining constant core clusters within the ranges. Once again we apply here the data recovery approach to automate the choice of the number of ranges for a seasonal time sequence of SST grids.

Therefore, we come up with a three-stage clustering approach for the spatiotemporal analysis of annual sequences of upwelling patterns. First, our STSEC algorithm is applied to each SST grid in the collection characterizing an upwelling season, resulting in the segmented upwelling patterns. Second, the Iterative Anomalous Pattern (IAP) algorithm from (Chiang & Mirkin, 2010; Mirkin, 2012) unsupervisedly finds stable time periods, here called 'time ranges', from the STSEC segmentations, to define periods of relative stability of upwelling within the upwelling season. At the third stage, the STSEC segmentations belonging to each time range are given as input to the core-shell clustering algorithm to further extract a constant core part in the range, together with corresponding shells and the intensities.

The remainder of the article is organized as follows. We describe our three-stage clustering approach in Section 2. Section 3 describes application of this approach to annual sequences of SST grids derived from SST remote sensing data. In Section 4, we describe conclusions drawn out of the computations. Section 4 concludes the article.

2 | THREE-STAGE SPATIO-TEMPORAL CLUSTERING ALGORITHM FOR PIECE-WISE CONSTANT MODELLING OF UPWELLING PATTERNS

2.1 | Three stages and corresponding cluster models

The input is a set of SST grids G_t taken in consecutive time moments $t = 1, 2, \dots, T$ related to an upwelling phenomenon. Each grid consists of nodes (i, j) , where $i \in I$ are latitudes and $j \in J$ are longitudes, at which the temperature values, g_{ij} , are specified. The output consists of the following:

U: A set of upwelling patterns $U_t \subset I \times J$ within grids G_t ($t = 1, 2, \dots, T$);

R: A partition $R = \{R_1, R_2, \dots, R_s\}$ of the time moment sequence $1, 2, \dots, T$ in a set of non-overlapping ranges R_s consisting of more-or-less similar upwelling patterns U_t , $t \in R_s$;

C: A set of core patterns C_s ($s = 1, 2, \dots, S$) which are parts of the upwelling patterns U_t with t within the range R_s , so that $C_s \subseteq U_t$ for all $t \in R_s$.

The set of core patterns above represents the piece-wise constant model of the upwelling according to our approach.

Let us describe the approach more precisely. Its three stages are:

1. Extracting upwelling patterns from individual grids: This stage applies to every individual SST grid G_t to produce an upwelling pattern $U_t \subset I \times J$ according to the model $g_{ij} = \lambda u_{ij} + e_{ij}$ where $\lambda > 0$ is a pattern intensity weight, u_{ij} is the pattern belongingness function so that $u_{ij} = 1$ for $(i, j) \in U_t$ and $u_{ij} = 0$ for $(i, j) \notin U_t$. Given temperature values g_{ij} , $(i, j) \in I \times J$, the model requires to find values λ and u_{ij} minimizing the least-squares criterion

$$L = \sum_{i \in I} \sum_{j \in J} (g_{ij} - \lambda u_{ij})^2. \quad (1)$$

It is not difficult to prove that, given upwelling pattern U , the optimal λ is the average temperature over the pixels $\lambda = g_U = \sum_{(i,j) \in U} g_{ij} / |U|$. By substituting this value into (1), one can easily obtain $L = \sum_{(i,j) \in I \times J} g_{ij}^2 - g_U^2 |U|$, so that minimizing L is equivalent to maximizing $g_U^2 |U|$ with regard to all possible patterns U .

2. Find time intervals where patterns remain stable: This stage takes upwelling patterns U_t , $t = 1, 2, \dots, T$, obtained at the previous stage, as its input, to produce a partition $R = \{R_1, R_2, \dots, R_s\}$ of the time moment sequence $1, 2, \dots, T$ into a set of non-overlapping ranges R_s consisting of more-or-less similar patterns U_t , $t \in R_s$. Rather than clustering the patterns as they are, we are going to simplify the task and characterize patterns by a few features so that the clustering process is run in the feature space. We have selected four features to characterize the upwelling patterns:
 - a. the total area of the upwelling pattern;
 - b. the average temperature over the upwelling pattern;
 - c. the maximum latitude of the upwelling region
 - d. the minimum latitude of the upwelling region

Therefore, the input data here is a $T \times 4$ matrix of values y_{tv} of the four features $v = 1, 2, 3, 4$ at the patterns U_t , $t = 1, \dots, T$. The data matrix is centered by subtracting the grand mean from every feature with the follow-up normalizing by the range. The output is a partition

$R = \{R_1, R_2, \dots, R_S\}$ of the time sequence in 'time ranges' R_s together with their centers $c_s = (c_{sv})$. This partition is sought so as to minimize the square error criterion $L = \sum_{s=1}^S \sum_{t \in R_s} \sum_{v=1}^4 (y_{tv} - c_{sv})^2$ with regard to unknown R_s and c_s , $s = 1, 2, \dots, S$, with S being adjusted according to the iK-Means version of k -means method (Mirkin, 2012).

3. Fine-tuning constant core: Given a range R_s of upwelling patterns U_t , $t \in R_s$, this stage extracts a constant core $C_s \subset I \times J$ of the range, which is part of all the range patterns, $C_s \subset U_t$, $t \in R_s$. The remaining part, $D_t = U_t - C_s$ is referred to as shell, so that the entire process can be dubbed as core-shell clustering. The concept of core-shell clustering has been introduced in Rodin and Mirkin (2017) as that of super-cluster. Here that is modified for using in the context of spatio-temporal clustering. The core-shell concept represents the time range R_s as the union of a constant set C_s and a variable set D^t of its complements in U^t ($t = 1, 2, \dots, T$). We assume that the core set C_s is composed of tightly related core pixels, whereas pixel temperatures in the shells may vary. Therefore, a core-shell cluster partition of pattern U^t , is represented by two non-overlapping sets $C_s \cup D^t$ of binary values c_{ij} , $(i, j) \in C_s$, the core, and d_{ij}^t , $(i, j) \in D^t$, the shell at moment t , such that $c_{ij} \times d_{ij}^t = 0$ always holds for $t \in R_s$. Assume that the shells D^t are characterized by their intensity values $\lambda^t > 0$. The intensity of the core should be greater than that for any $t \in R_s$, that is, the core's intensity is $\lambda^t + \mu^t$ with $\mu^t > 0$. Then, the model to define an upwelling SST, g_{ij}^t , at point (i, j) and moment t can be stated as:

$$g_{ij}^t = (\lambda^t + \mu^t)c_{ij} + \lambda^t d_{ij}^t + e_{ij}^t, \quad (2)$$

in which the residual values e_{ij}^t are to be minimized according to the least squares criterion:

$$\Delta = \sum_{t \in R_s} \sum_{i \in I} \sum_{j \in J} (g_{ij}^t - (\lambda^t + \mu^t)c_{ij} - \lambda^t d_{ij}^t)^2. \quad (3)$$

It is not difficult to derive, from the first order necessary conditions for minimization of function Δ over λ^t and μ^t , that

$$\lambda^t = \frac{\sum_{i,j} a_{ij}^t s_{ij}^t}{\sum_{i,j} s_{ij}^t}, \quad (4)$$

$$\lambda^t + \mu^t = \frac{\sum_{i,j} a_{ij}^t r_{ij}}{\sum_{i,j} r_{ij}}. \quad (5)$$

Substituting the intensity values λ^t and $\lambda^t + \mu^t$ derived in (4) and (5) into Equation (3) for Δ leads to: $\Delta = A - B$, where $A = \sum_t \sum_{i,j} (g_{ij}^t)^2$ defining the total data scatter and $B = \sum_t ((\lambda^t + \mu^t)^2 \times |C_s| + (\lambda^t)^2 \times |D^t|)$ being the core-shell cluster's contribution to that. Here $|C_s| = \sum_{i,j} c_{ij}$ is the number of data points in the core and $|D^t| = \sum_{i,j} d_{ij}^t$ is the number of points in the t -shell. Since A is constant, the minimization of the least squares criterion (3) is equivalent to maximization of the criterion B :

$$B = \sum_t ((\lambda^t + \mu^t)^2 \times |C_s| + (\lambda^t)^2 \times |D^t|). \quad (6)$$

The models described above lead to algorithms producing, in respect, upwelling patterns U_t (stage 1), partitioning of them in ranges with similar patterns R_s (stage 2), and extracted constant cores C_s of the ranges (Stage 3). All these stages are based on the data recovery approach to clustering (Mirkin, 2012) at which the data are approximated with a convenient cluster structure: anomalous single cluster, partition, and core-shell super-cluster.

2.2 | Algorithm PiWiCOMUP and its stages

Here we describe our three-stage algorithm for Piece-Wise Constant modelling of Upwelling Pattern PiWiCOMUP. The three stages have been described above together with corresponding clustering models. We present local search optimization algorithms for the criteria derived there.

Algorithm PiWiCOMUP

Stage 1. Extracting upwelling patterns from individual grids: Extraction of the coastal upwelling regions is performed by the unsupervised spatial clustering algorithm, Self-Tuning Seed Expanding Cluster (STSEC) (Nascimento et al., 2015). The STSEC method extends the popular Seeded Region Growing (SRG) within the framework of anomalous clustering which adds some features overcoming well recognized limitations of SRG algorithms. First, the clustering criterion derived above, $g_U^2 |U|$, takes the format of a product rather than the conventional difference between temperatures of a pixel and the mean of the region of interest, with a threshold adaptively optimized from that criterion. Second, the method involves a moving window which acts as one more regularizer of the cluster growing process. The STSEC algorithm involves preprocessing the temperatures by subtracting the average value. It grows pattern U by adding one pixel (i,j) at a time so that to locally maximize the increment of criterion $g_U^2 |U|$, which can be shown to have a format of inequality $c^*g(i,j) \geq \pi$ where c^* is the average temperature at the current pattern U and π is the similarity threshold chosen to be either expert-driven or equal to $c^*/2$. The STSEC algorithm starts by choosing the coldest pixel of a given SST grid as its initial seed. Pixels (i,j) are chosen from the front line of U within a window of pre-specified size ($n=8$). Whenever no more pixels satisfy the inequality above, the algorithm stops. The STSEC algorithm and its extension, S-STSEC, that sequentially retrieves more than one continuous upwelling region, have demonstrated promising results when applied to various geographical locations like the Portuguese and Morocco coasts (Nascimento et al., 2015, 2020).

Stage 2. Find time intervals where patterns remain stable: To accomplish this goal, we apply the Iterative Anomalous Patterns (IAP) algorithm (Mirkin, 2012) to a $T \times 4$ matrix of values y_{tv} of the four features $v=1,2,3,4$ at the patterns U_t , $t=1,\dots,T$, as described above. The IAP is a simple and effective clustering algorithm that sequentially extracts clusters one by one and, simultaneously, allows to derive the number of clusters to be found. It also serves as an effective initialization to fuzzy c -means with the IAP-FCM successfully applied to SST image segmentation (Nascimento et al., 2012). Here we use the IAP to unsupervisedly segment the time series of the extracted features. First, all the dataset is standardized by shifting the space origin into the gravity center, that is the point of grand mean at which all the features have their average values. After this all the values are divided by the corresponding feature range. The IAP finds an anomalous cluster in iterations akin to those of k -means except that only one, anomalous cluster, is sought here, and the origin never changes. It starts at an anomalous center c^* put into an entity farthest away from the origin according to the squared Euclidean distance. Then IAP updates the cluster by adding to it all those entities that are closer to c^* than to the origin. After this, the IAP updates the center by taking c^* to be equal to the mean of the cluster defined at the previous step. If the new c^* differs from that previous one, next iteration of cluster and center updates applies. The computation stops when the new center coincides with the previous one. The found anomalous cluster is removed from the dataset, and the process is repeated until all the entities are clustered. Then those clusters that consist of one or two elements only are removed. Elements of these clusters are reassigned to those larger clusters according to their distances to cluster centers. Each found group R_s , $s=1,2,\dots,S$ is designated as an upwelling time range. Figure 1 shows the time series of those (normalized) features obtained from the segmentations of the SST instants (for the upwelling season 2019 of Portuguese coast) with the IAP four-cluster partition marked by vertical dash lines.

Stage 3. Fine-tuning constant core: The core-shell clusters are constructed for each of the S range clusters R_1, R_2, \dots, R_S . Since the core cluster is, by assumption, constant and homogeneous, the initial core C_s is defined as the intersection of all R_s STSEC upwelling patterns, U_t , $t=1,2,\dots,R_s$. Each shell, D^t is defined as the set difference of pattern U_t and core C_s . Then, the initial intensity values λ^t and $\lambda^t + \mu^t$ are computed from Equations (4) and (5). To connect each initial core-shell cluster slice with the remaining upwelling region at instant t the set of the grid points forming a 4-neighbourhood, F^t , are merged with that. Thus, set $B = \{C_s \cup D^t \cup F^t\}_t$ define the initial core-shell cluster. After this, the algorithm iterates as follows. For each point (i,j) in $B = \cup_t B^t$ decide which of the alternatives: (A) to make (i,j) to belong to the core; or (B) (i,j) to belong to any of the shells; or, on the contrary, (C) to remove point (i,j) from any of them; leads to maximum increase of criterion δ_{ij} . This process requires $1+2^{|R_s|}$ tentative binary decisions. The process iterates until there is no improvement in criterion G , that is, until $\delta_{ij} \leq 0$. The derivation of formulas for values δ_{ij} is straightforward.

2.3 | Preprocessing SST grids

To improve the quality of SST grids, especially regarding the issue of over-segmentation faced by the SRG type algorithms (Aouni et al., 2021; Abidi et al., 2018; Nascimento et al., 2015), we apply a preprocessing method to every SST grid selected. The preprocessing method comprises three steps:

1. Removing North-South temperature gradient: This gradient results from the fact that higher latitudes receive less solar energy. To remove that, we apply the *grdtrend* module from the Generic Mapping Tools (GMT) software (Wessel et al., 2019).

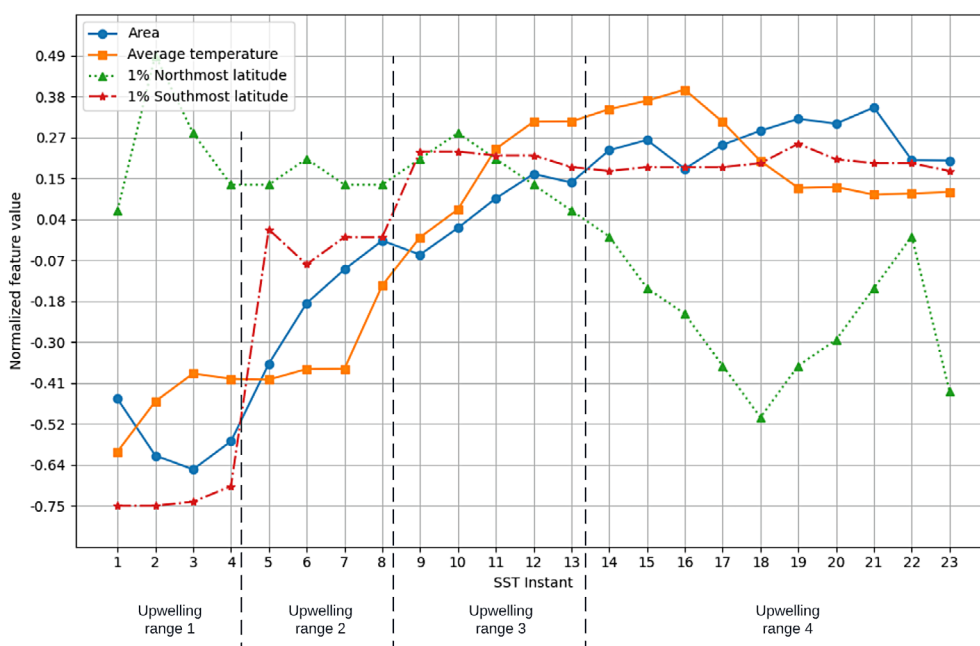


FIGURE 1 SST instant ranges obtained by IAP for time series extracted from STSEC segmentations.

2. The moving average filter: This involves the sliding window algorithm using a window size of 5 and a step size of 1, acting as the average filter. This removes white noise that might be present in the SST data.
3. Enhancement of East–West temperature gradient: This step normalizes temperature values of SST grids along the perpendicular lines to the coastline, adapting a method in (Aouni et al., 2021). Since the Portuguese/Moroccan coast has an almost constant longitude, the rows of each SST grid can be used as the lines perpendicular to the coastline. Then, each point (i,j) belonging to a line i is assigned with the difference between its temperature and the average temperature along the line if it is negative, or zero if it is positive. After this transformation all the temperatures on the grid become negative while the temperatures of the offshore ocean region are about zero. This transformation effectively prevents the over-segmentation because the automatic thresholding criterion of STSEC clustering algorithm above is a product of pixels's temperature and the average one. Thus, preprocessed SST grids are SST averaged grids.

Hereafter each SST averaged grid, result of the preprocessing stage, is designated as an *SST instant*.

3 | EXPERIMENTAL STUDY

Two annual collections of SST grids, for 16 years each, from the Portuguese coast (latitude from 36.0 N to 44.0 N and longitude from 13.0 W to 8.0 W) and from the Atlantic Ocean off North Africa were used in this study, covering the years from 2004 to 2019. Due to the large extension of the latter study area, the SST grids were divided into a northern branch from 30.0 N to 35.0 N (North Morocco) and a southern branch from 20.0 N to 27.0 N (South Morocco). The strip between 27.0 N and 30.0 N was left out due to the contorted coastline configuration and the presence of the Canary Islands, that is known to produce strong irregularities in the upwelling pattern (Barton & Arístegui, 2004).

We built SST grids, each grid having a size of 401×251 nodes with a spatial resolution of approximately 2×2 km, whose entries show temperature in degrees Celsius. Each SST grid represents the average of 8 days made with Level 2 data downloaded from the OceanColor site (<https://oceancolor.gsfc.nasa.gov/>), and filtered to use only the best quality data according to the products quality flags. Each yearly collection of SST grids consists of 46 weekly SST grids (after the moving average filter stage). Therefore, a total of 1472 SST grids were under analysis for North Africa (46 scenes for each subregion, 16 years). Since in the Portuguese coast the upwelling season is typically stronger from March to October, we considered SST data from approximately the 30th of March to the 30th of October, so that we had 27 SST grids per season. After applying a moving average filter with a window size 5 this number was reduced to 23. An example of an SST grid is given in Figure 2. The image represents the initial stage of the processing.

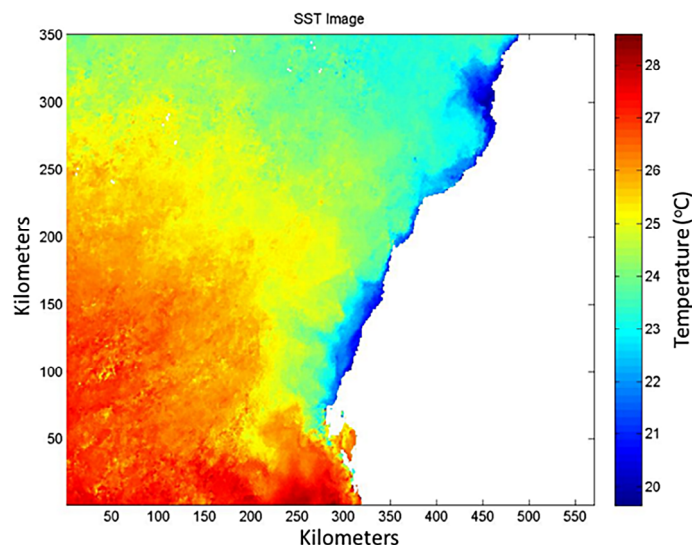


FIGURE 2 Example of an image constructed upon an SST 2×2 km grid representing the entire region under study. This grid was then divided in northern and southern parts for analysis. The continuous white region on the right side of each SST image corresponds to land surface (NW Africa), whereas white pixels in the ocean area correspond to missing values due to cloud cover.

3.1 | Description of PiWiCOMUP clustering results

Extensive analysis of SST images, like Figure 2, shows that the coastal upwelling regions are not that easy to distinguish on an SST grid. They have strong morphological variation, contorted protrusions, and transition zones with fuzzy thermal boundaries that correspond to quite irregular histograms, etc. Experienced oceanographers mostly rely in visual inspection to infer the upwelling patterns. However, that procedure is not suitable to analyse long time-series of SST grids.

The beginning and final steps of the PiWiCOMUP clustering process are illustrated in Figure 3 taking as example an SST grid for the South Morocco sector: (a) result of the moving average filter; (b) preprocessing to enhance the East–West temperature gradient; (c) segmentation of upwelling regions by STSEC algorithm; (d) core–shell cluster (highlight in orange–green) obtained from the upwelling regions segmentations.

From the 16 years long time-series of the Portuguese coast, we selected three most cloud-free years for a showcase analysis (2007, 2015 and 2019).

After the second stage of the PiWiCOMUP algorithm, each season was divided into a number of upwelling time ranges, automatically determined by the algorithm. Table 1 lists the obtained upwelling ranges for years 2007, 2015 and 2019.

It should be noted that the number of upwelling ranges for each year was between three and five for both collections of SST data from the Portuguese coast and the Atlantic off North Africa.

The graphic in Figure 4 shows the areas of upwelling cores (orange line), the areas of the corresponding shells (green line) forming each instant core–shell cluster, as well as the areas of the corresponding whole upwelling regions obtained by S-STSEC algorithm (blue line) for the SST collection of 2015 Portuguese coast, as an example.

The average temperature of upwelling cores (i.e., their intensities) against the average temperature of the offshore waters at each SST instant of the same collection (year 2015) are displayed in the graphic of Figure 5. The maximum average temperature of core upwelling with value 17.71°C occurs at instant 11, approximately at the middle of the upwelling season, and the maximum temperature difference occurs at instant 15. We found concordant results for the other years. In 2007 the core upwelling with maximum average temperature of 16.97°C occurred at SST instant 10 while the maximum temperature difference between core and offshore waters occurred at instant 16. A similar pattern is present for the season of 2019 with the maximum average temperature of 17.1°C for core upwelling occurring at instant 12, the middle of the upwelling season, whereas the maximum temperature difference is achieved at instant 16. These results show that the upwelling events tend to increase in strength around instants 10–12. In fact, these instants coincide with the summer season where the upwelling events tend to have the strongest activity (Leitão et al., 2019). In summary, the cores' intensities of core–shell clusters characterize very well the upwelling reference temperature.

3.2 | PiWiCOMUP clustering interannual time series analysis

The PiWiCOMUP algorithm has been independently applied to two collections of SST grids, from the Atlantic ocean off North Africa and from the Portuguese coast.

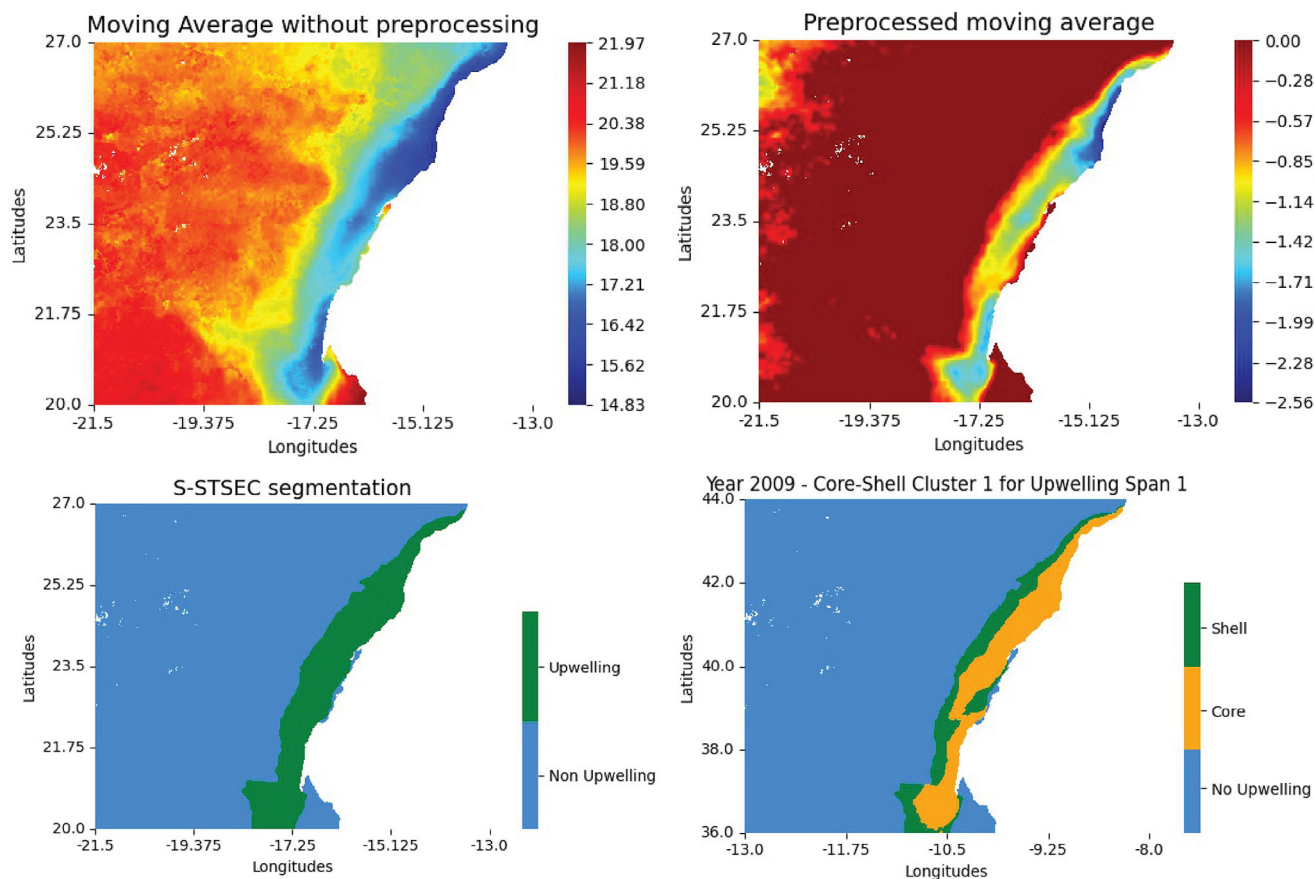


FIGURE 3 Beginning and final steps of the PiWiCOMUP clustering process: pre-processing stage (top figures); upwelling region segmentation and core-shell cluster (bottom figures).

TABLE 1 SST instants forming the upwelling *ranges* for each year.

Upwelling range	Year 2007	Year 2015	Year 2019
1	1–8	1–3	1–4
2	9–13	4–10	5–8
3	14–23	11–19	9–13
4		20–23	14–23

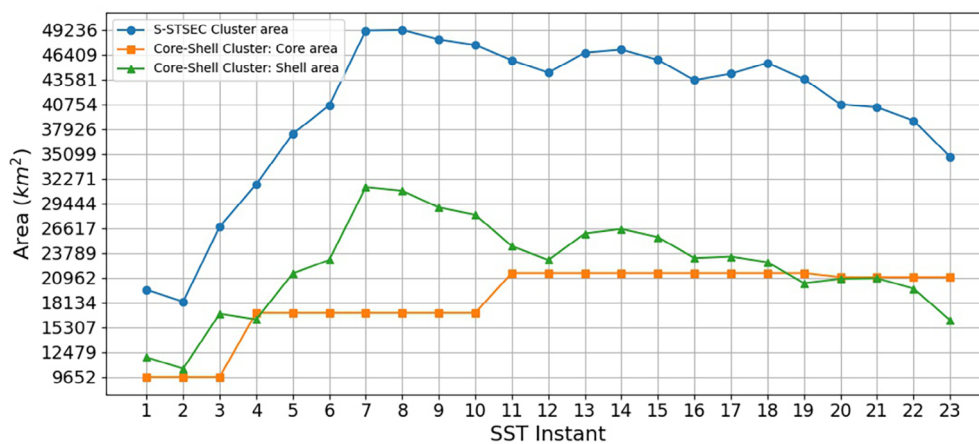


FIGURE 4 Areas of the cores, the shells of the core-shell clusters and the areas of upwelling regions by STSEC: year of 2015.

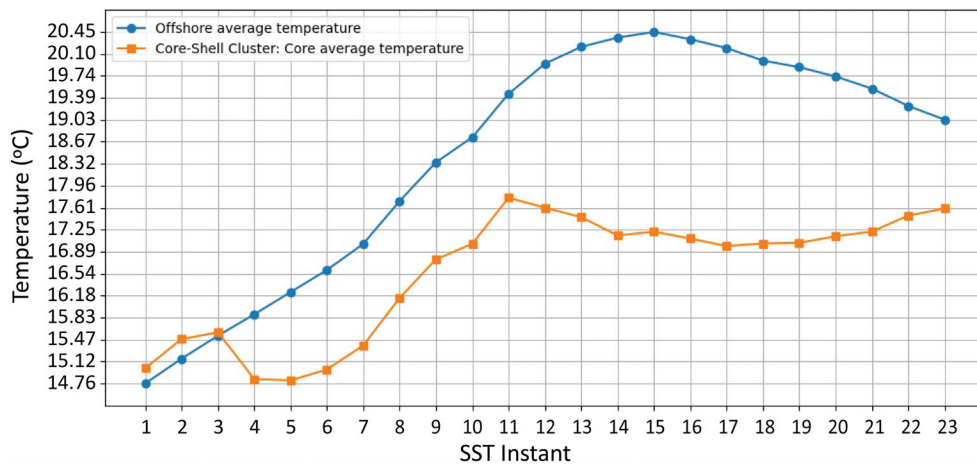


FIGURE 5 Comparison of average cores' temperatures against average temperatures of offshore ocean waters: year 2015.

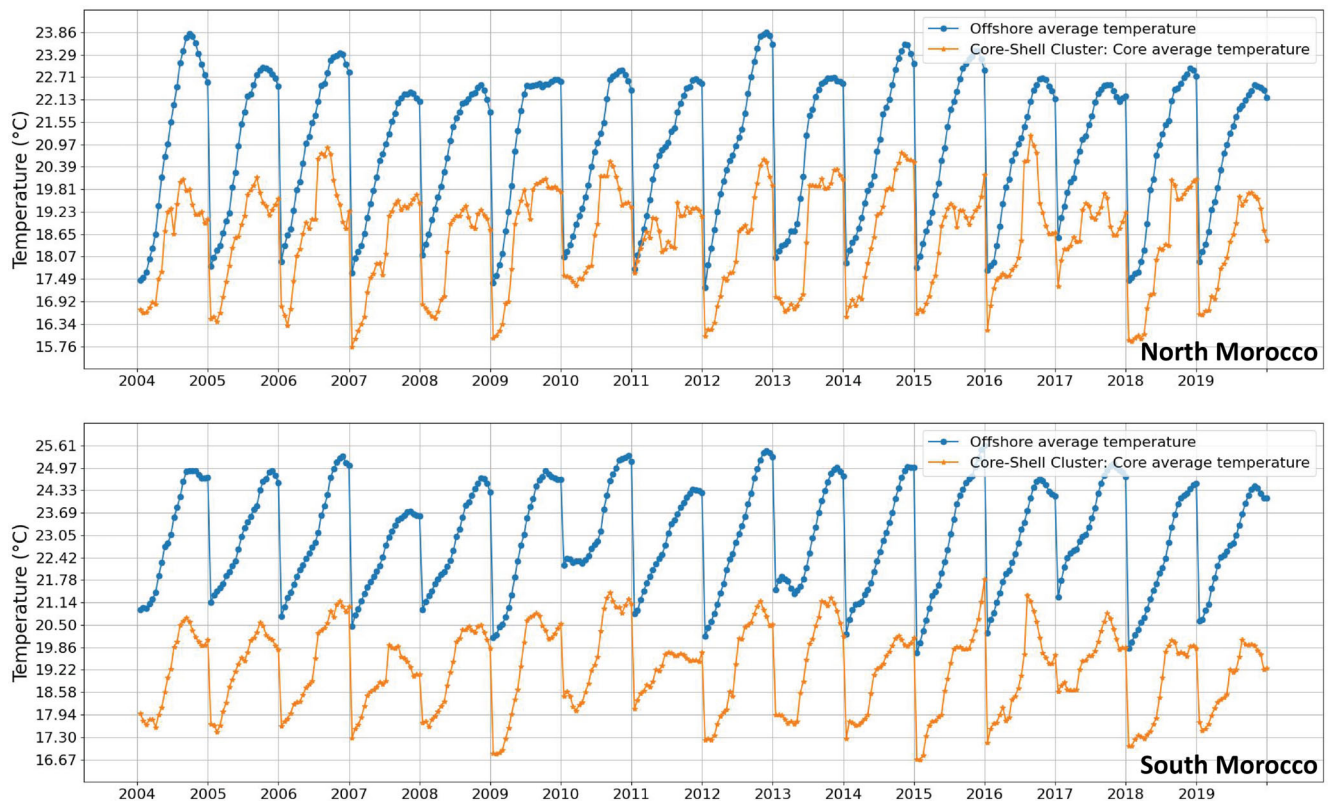


FIGURE 6 Interannual variability of the average cores' temperatures against average temperatures of offshore ocean waters: North Morocco (top) and South Morocco (bottom).

Time series were built from the core-shell clusters to analyse upwelling regularities. For that, the following features were taken: the average temperature of upwelling cores against average temperature of offshore ocean waters; the average temperature of each of these structures which are nothing else than the cores and shells' intensities defined by Equations (4) and (5); the areas of the cores and the areas of the shells forming the core-shell clusters characterizing their extent.

In what follows we discuss the interannual time series results off North Africa. As a first check to the oceanographic validity of the results, the temperatures of the upwelling cores were compared with the average contemporaneous temperatures of the remaining offshore areas, as shown in Figure 6 for North and South Morocco. The core temperatures are consistently lower than the offshore ones, reflecting the permanent upwelling regions close to the coast and captured by the core analysis. The annual cycle of the SST is observed both, in the core and offshore, but the differences between the core and the offshore temperatures are stronger during the summer peak upwelling months, consistent with the

upwelling regime. The differences are stronger in the South relatively to the North Morocco branch, reflecting the more intense upwelling in the southern region.

The time series of the cores and shells average temperatures displayed in Figure 7, for North and South Morocco, show that when the upwelling area under analysis is extended to the shell, the average temperature slightly increases. We are now capturing the fringe of the upwelled waters, relatively warmer than the core ones. However, in both regions North and South, the average temperatures are cold enough to have the confidence that we are capturing upwelling waters.

All the average temperature trends give $+0.01^{\circ}\text{C}$ for the North Morocco and -0.01°C to the core and 0.00°C to the shell and core+shell for the South Morocco. The obtained $+0.01^{\circ}\text{C}$ is slightly less that observed for the Iberian waters in Relvas et al. (2009). Going to the south, trend decays.

The interannual variability of the upwelling areas detected by the core-shell clustering are shown in Figures 8 and 9 for North and South Morocco, respectively. In each figure the top graphic displays the total area of the core-shells and the middle/bottom graphics represent the core/shells areas, respectively. Trend lines are drawn in orange. According to these results, the coastal areas occupied by cold upwelled waters have been increasing along the 16 years under analysis. An increase of 69.51 km^2 and 190.49 km^2 per year in the North and South Morocco respectively, was inferred from our analysis. Such increase was achieved by a negative contribution of the core areas (-13.05 km^2), compensated by a positive increase of the shell areas (82.56 km^2) in the North Morocco, and positive contributions of the core (29.64 km^2) and shell (160.85 km^2) in the South Morocco.

From an oceanographic point of view, the increased extent of the surface signal of the upwelling can be interpreted as an increase of the upwelling intensity or an increase of the positive (cyclonic) wind stress curl along the North Africa, that spreads the same amount of upwelled water further offshore.

The oceanographic interpretation of the obtained results is out of the scope of the present article. However, the results achieved by the proposed three-stage clustering approach are consistent with the oceanographic features of both regions.

All the computations were conducted in a 15-inch MacBook Pro 2018, with a 2.2 GHz 6-Core Intel Core i7 processor, 16 GB 2400 MHz DDR4 of memory and a Radeon Pro 555X 4 GB graphics card, on the macOS Monterey version 12.2.1 OS. The S-STSEC algorithm took on average 15.6 s to segment an SST grid. The computation times of the core-shell clustering algorithm depend on the cardinality T of the upwelling ranges being processed. The obtained times for $T = 3$ to $T = 6$ were: 8, 14, 26 and 54 s, respectively.

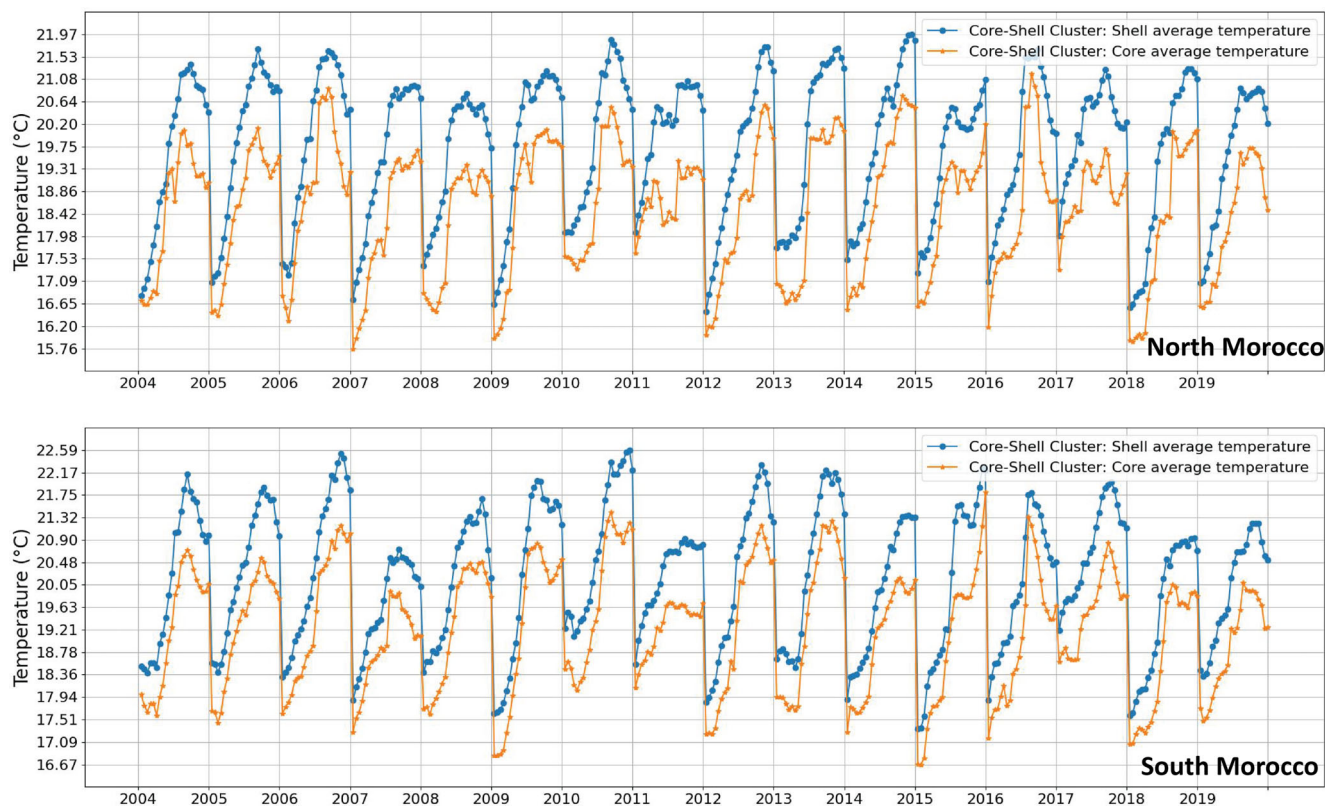


FIGURE 7 Interannual variability of the average cores' temperatures against shell's average temperatures: North Morocco (top) and South Morocco (bottom).

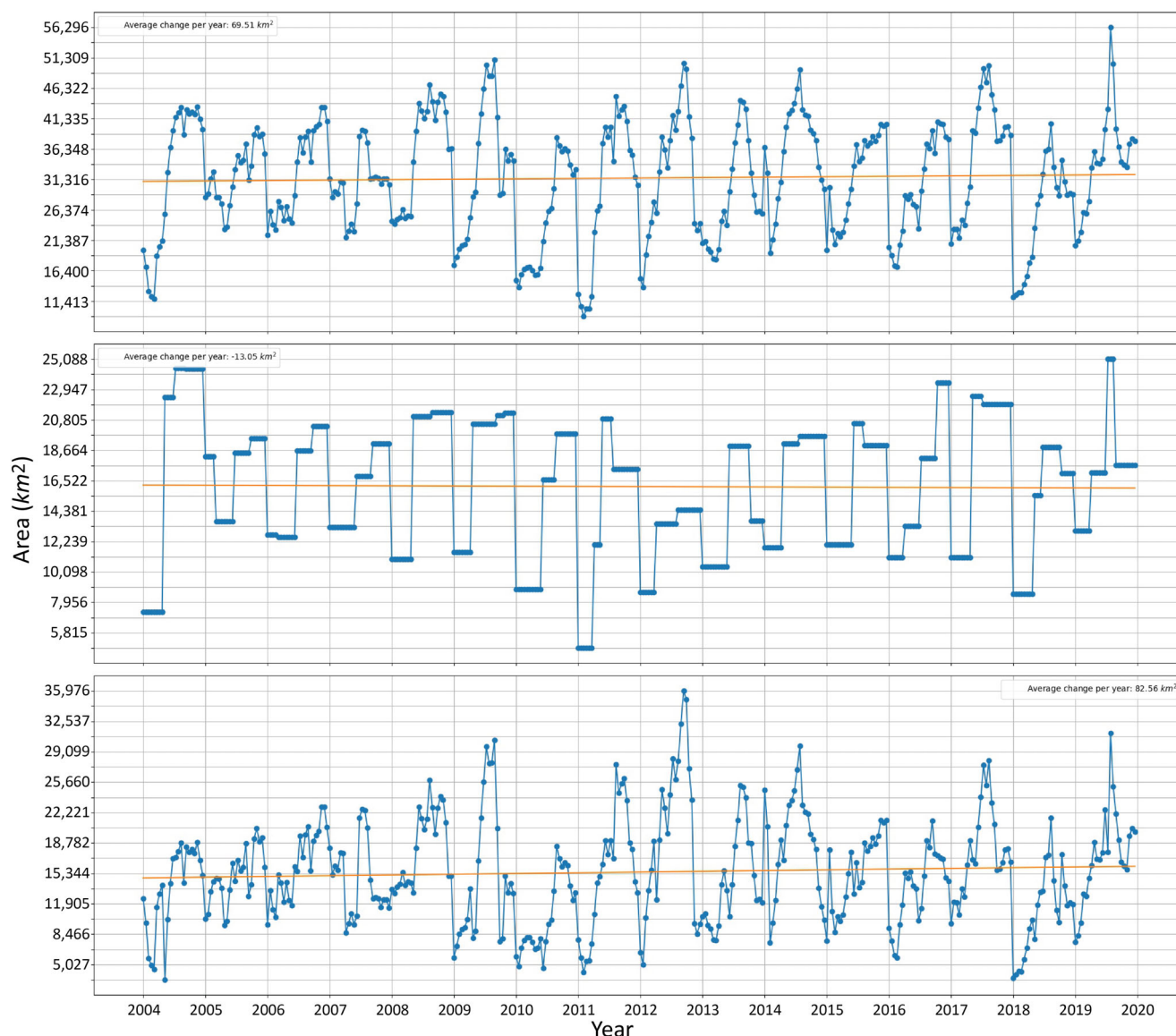


FIGURE 8 Interannual variability of the upwelling areas detected by the core-shell clustering, for North Morocco. Trend lines are drawn (orange lines). Top pictures represent the core and shell areas. The middle and bottom pictures represent the core areas and shell contributions. See text for details.

4 | CONCLUSION

In this article, we undertook an investigation of a spatiotemporal phenomenon of coastal upwelling. Such an investigation is useful from both theoretical and practical points of view because it helps finding regularities of oceanic systems that play an important role in many aspects of survival and prosperity of the humankind.

With the view of upwelling phenomenon as a realization of a stable soliton wave, we used our data recovery clustering methods to find some constant components in a season's real-world upwelling near Portugal and Morocco coasts. Our preliminary results (not reported) demonstrated to us that, during a season, upwelling is not necessarily stable. However, every season can be partitioned in periods of relative stability of coastal upwelling. Therefore, we developed a machinery based on data recovery clustering to

- Extract the patterns of coastal upwelling at every 'static' SST grid;
- Divide any time-determined sequence of upwelling patterns in periods of relative stability referred to here as time ranges; and then
- Extract a constant part, core, within every range.

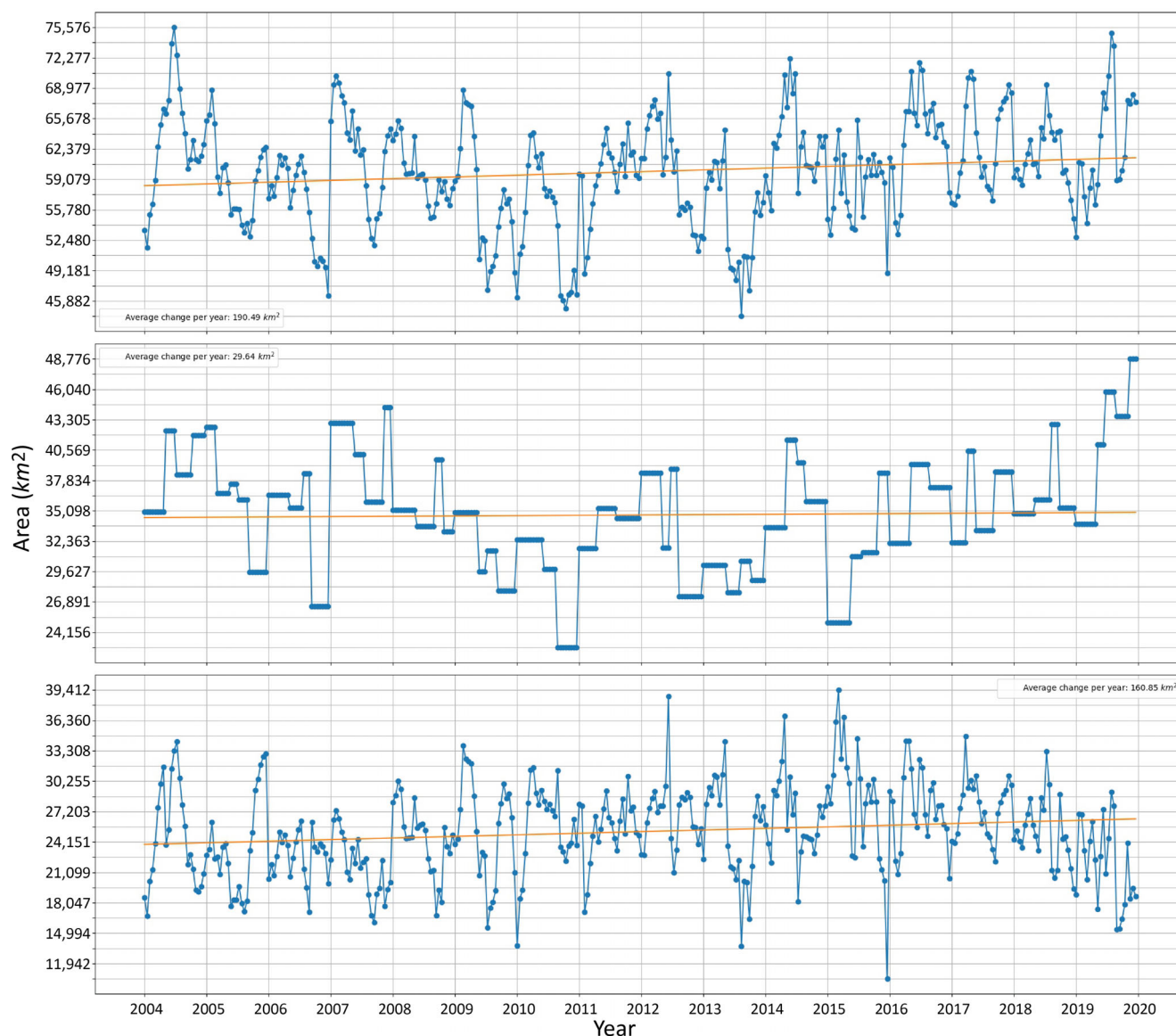


FIGURE 9 Interannual variability of the upwelling areas detected by the core-shell clustering, for South Morocco. Trend lines are drawn (orange lines). Top pictures represent the core and shell areas. The middle and bottom pictures represent the core areas and shell contributions.

We applied this machinery to real-world SST data for 16 years and found consistent portrayals of the upwelling as piece-wise constant core patterns, including observations of dynamics of areas and temperatures of them. We consider our results adequate because they are consistent (a) among themselves; (b) with existing expert domain knowledge, and (c) with oceanographers' judgement.

These results not only relate to the specific coastal areas, but also bring forward a number of novel issues before the science of oceanography. What makes upwelling to change several times within a season? Is there any regularity in the moments of change? We think that exploration of these and other issues, such as those related to specifics of upwelling area sizes and temperature differences, emerging in the light of our results, is of interest from both theoretical and practical points of view.

ACKNOWLEDGEMENTS

Susana Nascimento and Alexandre Martins acknowledge the support from NOVA LINCS (UIDB/04516/2020), Paulo Relvas acknowledges the support through projects UIDB/04326/2020, UIDP/04326/2020 and LA/P/0101/2020, Joaquim F. Luís acknowledges the support through project UIDB/50019/2020, all funded by Portuguese national funds from FCT-Foundation for Science and Technology. Boris Mirkin acknowledges support from the Russian Science Foundation under grant 22-11-00323 at the National Research University Higher School of Economics Moscow. The authors are grateful to the anonymous reviewers for their insightful and constructive comments that allowed us to improve the presentation.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Susana Nascimento  <https://orcid.org/0000-0001-5328-0577>

REFERENCES

- Abidi, Z. E., Minaoui, K., Tamim, A., & Laanaya, H. (2018). Delineation of moroccan coastal upwelling using the principal component analysis fusion algorithm on ssc and sst images. In *2018 9th international symposium on signal, image, video and communications—isivc 2018* (pp. 174–178). <https://doi.org/10.1109/ISIVC.2018.8709227>
- Adams, R., & Bischof, L. (1994). Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6), 641–647. <https://doi.org/10.1109/34.295913>
- Agrawal, K., Garg, S., Sharma, S., & Patel, P. (2016). Development and validation of optics based spatio-temporal clustering technique. *Information Sciences*, 369(C), 388–401. <https://doi.org/10.1016/j.ins.2016.06.048>
- Alam, M. M., Torgo, L., & Bifet, A. (2022). A survey on spatio-temporal data analytics systems. *ACM Computing Surveys*, 54(10s), 1–38. <https://doi.org/10.1145/3507904>
- Ansari, M. Y., Ahmad, A., Khan, S. S., Bhushan, G., & Mainuddin. (2019). Spatiotemporal clustering: A review. *Artificial Intelligence Review*, 53, 2381–2423. <https://doi.org/10.1007/s10462-019-09736-1>
- Aouni, A. E., Daoudi, K., Minaoui, K., & Yahia, H. (2021). Robust detection of the north-west african upwelling from sst images. *IEEE Geoscience and Remote Sensing Letters*, 18(4), 573–576. <https://doi.org/10.1109/LGRS.2020.2983826>
- Atluri, G., Karpayne, A., & Kumar, V. (2018). Spatio-temporal data mining: A survey of problems and methods. *ACM Computing Surveys*, 51(4), 1–41. <https://doi.org/10.1145/3161602>
- Barton, E., & Aristegui, J. (2004). The canary islands coastal transition zone: Upwelling, eddies and filaments. *Progress in Oceanography*, 62(2–4), 67–69. <https://doi.org/10.1016/j.pocean.2004.08.003>
- Birant, D., & Kut, R. A. (2007). St-dbscan: An algorithm for clustering spatial-temporal data. *Data & Knowledge Engineering*, 60, 208–221. <https://doi.org/10.1016/j.datak.2006.01.013>
- Chen, X. C., Faghmous, J. H., Khandelwal, A., & Kumar, V. (2015). Clustering dynamic spatio-temporal patterns in the presence of noise and missing data. In *Proceedings of the 24th international conference on artificial intelligence* (p. 2575–2581). AAAI Press. <https://doi.org/10.5555/2832581.2832609>
- Chiang, M. M. T., & Mirkin, B. (2010). Intelligent choice of the number of clusters in k-means clustering: An experimental study with different cluster spreads. *Journal of Classification*, 27(1), 3–40. <https://doi.org/10.1007/s00357-010-9049-5>
- Donatelli, C., Duran-Matute, M., Gräwe, U., & Gerkema, T. (2022). Statistical detection of spatio-temporal patterns in the salinity field within an inter-tidal basin. *Estuaries and Coasts*, 45, 2345–2361. <https://doi.org/10.1007/s12237-022-01089-3>
- Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd*, 96, 226–231. <https://doi.org/10.5555/3001460.3001507>
- Evers, M., & Linsen, L. (2022). Multi-dimensional parameter-space partitioning of spatio-temporal simulation ensembles. *Computers and Graphics*, 104, 140–151. <https://doi.org/10.48550/arXiv.2205.00980>
- Houghton, I. A., & Wilson, J. D. (2020). El niño detection via unsupervised clustering of Argo temperature profiles. *Journal of Geophysical Research: Oceans*, 125(9), e2019JC015947. <https://doi.org/10.1029/2019jc015947>
- Huang, Z., & Wang, X. H. (2019). Mapping the spatial and temporal variability of the upwelling systems of the australian south-eastern coast using 14-year of modis data. *Remote Sensing of Environment*, 227, 90–109. <https://doi.org/10.1016/j.rse.2019.04.002>
- Jacox, M. G., Edwards, C. A., Hazen, E. L., & Bograd, S. J. (2018). Coastal upwelling revisited: Ekman, bakun, and improved upwelling indices for the U.S. west coast. *Journal of Geophysical Research: Oceans*, 123(10), 7332–7350. <https://doi.org/10.1029/2018jc014187>
- Kamenetsky, M. E., Lee, J., Zhu, J., & Gangnon, R. E. (2022). Regularized spatial and spatio-temporal cluster detection. *Spatial and Spatio-Temporal Epidemiology*, 41, 100462. <https://doi.org/10.1016/j.sste.2021.100462>
- Konstantaras, A. (2020). Deep learning and parallel processing spatio-temporal clustering unveil new ionian distinct seismic zone. *Informatics*, 7(4), 39. <https://doi.org/10.3390/informatics7040039>
- Kulldorff, M., & Nagarwalla, N. (1995). Spatial disease clusters: Detection and inference. *Statistics in Medicine*, 14, 799–810. <https://doi.org/10.1002/sim.4780140809>
- Lehmann, A., Myrberg, K., & Höflisch, K. (2012). A statistical approach to coastal upwelling in the Baltic Sea based on the analysis of satellite data for 1990–2009. *Oceanologia*, 54(3), 369–393. <https://doi.org/10.5697/oc.54-3.369>
- Leitão, F., Baptista, V., Vieira, V., Silva, P. L., Relvas, P., & Teodósio, M. A. (2019). A 60-year time series analyses of the upwelling along the portuguese coast. *Water*, 11(6), 1285. <https://doi.org/10.3390/w11061285>
- Liu, J., Xue, C., Dong, Q., Wu, C., & Xu, Y. (2019). A process-oriented spatiotemporal clustering method for complex trajectories of dynamic geographic phenomena. *IEEE Access*, 7, 155951–155964. <https://doi.org/10.1109/access.2019.2949049>
- Martino, F. D., Pedrycz, W., & Sessa, S. (2018). Spatiotemporal extended fuzzy c-means clustering algorithm for hotspots detection and prediction. *Fuzzy Sets and Systems*, 340, 109–126. <https://doi.org/10.1016/j.fss.2017.11.011>
- Mattera, R. (2022). A weighted approach for spatio-temporal clustering of covid-19 spread in Italy. *Spatial and Spatio-Temporal Epidemiology*, 41, 100500. <https://doi.org/10.1016/j.sste.2022.100500>
- Mirkin, B. (2012). *Clustering: A data recovery approach* (2nd ed.). Chapman and Hall/CRC Press. <https://doi.org/10.1201/b13101>

- Mozden, A., Cremaschi, A., Cadonna, A., Guglielmi, A., & Kastner, G. (2022). Bayesian modeling and clustering for spatio-temporal areal data: An application to Italian unemployment. *Spatial Statistics*, 52, 100715. <https://doi.org/10.1016/j.spasta.2022.100715>
- Nascimento, S., Casca, S., & Mirkin, B. (2015). A seed expanding cluster algorithm for deriving upwelling areas on sea surface temperature images. *Computers & Geosciences*, 85, 74–85. <https://doi.org/10.1016/j.cageo.2015.06.002>
- Nascimento, S., Franco, P., Sousa, F., Dias, J., & Neves, F. (2012). Automated computational delimitation of SST upwelling areas using fuzzy clustering. *Computers & Geosciences*, 43, 207–216. <https://doi.org/10.1016/j.cageo.2011.10.025>
- Nascimento, S., Mateen, S., & Relvas, P. (2020). Sequential self-tuning clustering for automatic delimitation of coastal upwelling on SST images. In C. Analide, P. Novais, D. Camacho, & H. Yin (Eds.), *Intelligent data engineering and automated learning: Ideal 2020* (pp. 434–443). Springer International Publishing. https://doi.org/10.1007/978-3-030-62365-4_41
- Nowicki, A., Janecki, M., & Dzierzbicka-Głowacka, L. (2019). Operational system for automatic coastal upwelling detection in the Baltic Sea based on the 3d CEMBS model. *Journal of Operational Oceanography*, 12(2), 104–115. <https://doi.org/10.1080/1755876x.2019.1569748>
- Ramachandra, B., Dutton, B., & Vatsavai, R. R. (2019). Anomalous cluster detection in spatiotemporal meteorological fields. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 12(2), 88–100. <https://doi.org/10.1002/sam.11398>
- Ramanantsoa, J. D., Krug, M., Penven, P., Rouault, M., & Gula, J. (2018). Coastal upwelling south of Madagascar: Temporal and spatial variability. *Journal of Marine Systems*, 178, 29–37. <https://doi.org/10.1016/j.jmarsys.2017.10.005>
- Relvas, P., Luís, J., & Santos, A. M. P. (2009). Importance of the mesoscale in the decadal changes observed in the northern canary upwelling system. *Geophysical Research Letters*, 36(22), L22601. <https://doi.org/10.1029/2009gl040504>
- Rodin, I., & Mirkin, B. (2017). Supercluster in statics and dynamics: An approximate structure imitating a rough set. In L. Polkowski, et al. (Eds.), *Rough sets* (pp. 576–586). Springer International Publishing. https://doi.org/10.1007/978-3-319-60837-2_46
- Sambe, F., & Suga, T. (2022). Unsupervised clustering of argo temperature and salinity profiles in the mid-latitude northwest Pacific ocean and revealed influence of the Kuroshio extension variability on the vertical structure distribution. *Journal of Geophysical Research: Oceans*, 127(3), e2021JC018138. <https://doi.org/10.1029/2021JC018138>
- Shi, W., Huang, Z., & Hu, J. (2021). Using TPI to map spatial and temporal variations of significant coastal upwelling in the northern South China Sea. *Remote Sensing*, 13(6), 1065. <https://doi.org/10.3390/rs13061065>
- Siemer, J. P., Machín, F., González-Vega, A., Arrieta, J. M., Gutiérrez-Guerra, M. A., Pérez-Hernández, M. D., & Fraile-Nuez, E. (2021). Recent trends in SST, chl-a, productivity and wind stress in upwelling and open ocean areas in the upper eastern North Atlantic subtropical gyre. *Journal of Geophysical Research: Oceans*, 126(8), e2021JC017268. <https://doi.org/10.1029/2021JC017268>
- Tamim, A., Minaoui, K., Daoudi, K., Yahia, H., Atillah, A., & Aboutajdine, D. (2015). An efficient tool for automatic delimitation of Moroccan coastal upwelling using SST images. *IEEE Geoscience and Remote Sensing Letters*, 12(4), 875–879. <https://doi.org/10.1109/lgrs.2014.2365558>
- Tamim, A., Minaoui, K., Daoudi, K., Yahia, H., Atillah, A., Fella, S. E., & Ansari, M. E. (2019). Automatic detection of Moroccan coastal upwelling zones using sea surface temperature images. *International Journal of Remote Sensing*, 40(7), 2648–2666. <https://doi.org/10.1080/01431161.2018.1528513>
- Tonini, M., Pecoraro, G., Romaila, K., & Calvella, M. (2022). Spatio-temporal cluster analysis of recent Italian landslides. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 16, 536–554. <https://doi.org/10.1080/17499518.2020.1861634>
- Wazarkar, S., & Keshavamurthy, B. N. (2018). A survey on image data analysis through clustering techniques for real world applications. *Journal of Visual Communication and Image Representation*, 55, 596–626. <https://doi.org/10.1016/j.jvcir.2018.07.009>
- Wessel, P., Luis, J. F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H., & Tian, D. (2019). The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems*, 20, 5556–5564. <https://doi.org/10.1029/2019GC008515>
- Yu, M., Bambacus, M., Cervone, G., Clarke, K., Duffy, D., Huang, Q., & Yang, C. (2020). Spatiotemporal event detection: A review. *International Journal of Digital Earth*, 13(12), 1339–1365. <https://doi.org/10.1080/17538947.2020.1738569>
- Zhang, D., Lee, K., & Lee, I. (2018). Hierarchical trajectory clustering for spatio-temporal periodic pattern mining. *Expert Systems with Applications*, 92, 1–11. <https://doi.org/10.1016/j.eswa.2017.09.040>

AUTHOR BIOGRAPHIES

Susana Nascimento holds a PhD in Computer Science (2002) from Universidade NOVA de Lisboa and is Assistant Professor at the Department of Computer Science of NOVA School of Science and Technology since 2003 and researcher at NOVA LINCS. Her main research is on unsupervised learning and clustering oriented towards the development of clustering methods providing more flexible algorithms with special attention on the interpretability of clustering results in the framework of real-world applications.

Alexandre Martins Software engineer with a Masters in Computer Science at Universidade NOVA de Lisboa, 2022. Master thesis was developed on the topic of Unsupervised Clustering of Coastal Upwelling on Portugal and Northern Morocco coastal regions. Currently working as a Software Engineer at a software development company based in Slovakia.

Paulo Relvas Physicist with a PhD in Physical Oceanography at the University of Wales (U.K.), 1999. Professor at the University of Algarve and researcher at CCMAR. Director of the Marine and Coastal Systems MSc program. Large experience on students' supervision. Research experience on upwelling systems, mesoscale processes, coastal and shelf oceanography, inner shelf circulation, and applied remote sensing.

Joaquim Luis holds a PhD in Geophysics. He is currently an Assistant Professor at University of Algarve. His research interests include Marine Geophysics and scientific programming writing and development of visualization tools. He is a core developer of the Generic Mapping Tools package.

Professor Boris Mirkin holds a PhD in Computer Science and DSc in Systems Engineering degrees from Russian Universities. He published a dozen monographs and a hundred refereed papers. In 1991-2010 he extensively traveled taking visiting research appointments in France, USA, Germany, and a teaching appointment at Birkbeck University of London, UK. He develops methods for clustering and interpretation of complex data within the “data recovery” perspective.

How to cite this article: Nascimento, S., Martins, A., Relvas, P., Luís, J. F., & Mirkin, B. (2023). Piece-wise constant cluster modelling of dynamics of upwelling patterns. *Expert Systems*, 40(10), e13446. <https://doi.org/10.1111/exsy.13446>