Computers & Geosciences 179 (2023) 105421

Contents lists available at ScienceDirect

Computers and Geosciences

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

Research paper

Core–shell clustering approach for detection and analysis of coastal upwelling

Susana Nascimento [a,*], Alexandre Martins [a], Paulo Relvas [b], Joaquim F. Luís [c], Boris Mirkin [d,e]

[a] Department of Computer Science and NOVA Laboratory for Computer Science and Informatics (NOVA-LINCS) Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
[b] Centre of Marine Sciences (CCMAR), Universidade do Algarve, 8005-139 Faro, Portugal
[c] Instituto Dom Luís (IDL), Universidade do Algarve, 8005-139 Faro, Portugal
[d] Department of Data Analysis and Artificial Intelligence, National Research University Higher School of Economics, Moscow, Russian Federation
[e] Department of Computer Science, Birkbeck University of London, UK

ARTICLE INFO

Keywords:
Spatio-temporal clustering
Time series segmentation
SST images
Coastal upwelling
Core–shell cluster

ABSTRACT

A comprehensive approach is presented to analyze season's coastal upwelling represented by weekly sea surface temperature (SST) image grids. The proposed model, core–shell clustering, assumes that the season's upwelling can be divided into shorter periods of stability, time ranges, consisting of constant core and variable shell parts. A one-by-one core–shell clustering algorithm is provided. The algorithm parameters are automatically derived from the least-squares clustering criterion.

The approach applies to SST gridded data for sixteen successive years (2004–2019) of coastal upwelling in the western Iberian coast, the northernmost branch of the Canary Current Upwelling System. Our results show that at each season, there are 3 to 5 time intervals, the ranges, at which the upwelling presents stable core patterns of relatively cold water surrounded by somewhat larger shell areas of warmer waters. Based on other experimental computations performed by our team, we conclude that this pattern is not just a purely local phenomenon but has a more global meaning. Inter-annual time series analysis are consistent among themselves and with existing expert domain knowledge.

1. Introduction

This paper proposes a novel clustering framework for the spatio-temporal (ST) analysis of coastal upwelling, aligned with the increased interest in clustering for Oceanographic applications (e.g., Sambe and Suga, 2022; Wazarkar and Keshavamurthy, 2018, and references therein). The study area is the northernmost branch of the Canary Current Upwelling System, the western Iberian coast, where the upwelling regime prevails almost continuously, typically from April until October, driven by the northerly wind characteristic of this time of the year. The coastline is populated with some conspicuous protrusions that distort the upwelling front into a contorted line, which represents a challenge to the performance of any algorithm for automatic detection. The occurrence of a good percentage of days with clear sky, allowing good satellite SST records, along with the existence of oceanographic knowledge and expertise about the western Iberia upwelling, makes this region rather suitable to investigate and test methods for automatic upwelling detection and tracking.

This work belongs to the field of spatio-temporal data clustering (e.g. Shekhar et al., 2015; Atluri et al., 2018; Ansari et al., 2019), 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 DBSCAN methodology (Ester et al., 1996) to capture dense fragments of the data distribution, we apply a different approach to explicitly distinguish between core points and boundary points. This is reflected in the concepts of cluster 'core' and cluster 'shell' proposed in Rodin and Mirkin (2017). The Core–Shell clustering approach here is extended to overcome shortcomings of the popular spatio-temporal clustering algorithms like ST-DBSCAN (Birant and Kut, 2007) and ST-OPTICS (Agrawal et al., 2016) demanding several parameter settings from the user. It is well known that the accuracy and efficiency of those type of algorithms strongly depend on the appropriate selection of their parameters, which is of an issue in the absence of self-tuning parameter selection (Ansari et al., 2019; Shi and Pun-Cheng, 2019). Moreover, our approach can be easily adapted to cases at which the "constant core" is, in fact, variable, remaining constant only during relatively short time ranges. This is exactly the case under consideration.

* Corresponding author.
E-mail address: snt@fct.unl.pt (S. Nascimento).

https://doi.org/10.1016/j.cageo.2023.105421
Received 28 August 2022; Received in revised form 20 July 2023; Accepted 31 July 2023
Available online 4 August 2023

Long term spatio-temporal analysis of coastal upwelling is essential for the study of ocean dynamics, coastal resource managements and climate models (Baptista et al., 2018; Siemer et al., 2021). The pioneering installation of the Advanced Very High-Resolution Radiometer (AVHRR) sensor on board of the second generation National Oceanic and Atmospheric Administration (NOAA) satellites, together with the satellites that followed up, provide a regular and continuous operational global observations of the sea surface temperature (SST) for almost 40 years.

Traditionally, the identification of the upwelling front is carried out subjectively by experienced oceanographers. However, the large amount of available satellite SST records makes such assignment virtually impossible. In addition, real-time or short-delay monitoring of the upwelling patterns is a crucial tool for coastal water management and is not consistent with the intervention of an oceanographer. These issues motivate the need in developing of unsupervised computational methods for the identification of upwelling fronts.

Many studies have been conducted in this direction at various coastal regions of the world. Illustrative approaches include the works in Huang and Wang (2019), Shi et al. (2021) that explore the semi-automatic Topographic Position Index (TPI) algorithm, which is commonly used to recognize coastal upwelling regions; the obtained upwelling regions' characteristics along with wind data are used to study spatio-temporal variations of coastal upwelling for several years. Simulation models were used in Ramanantsoa et al. (2018) like the Coastal and Regional Ocean COmmunity (CROCO) to simulate ocean currents to help understanding how coastal upwelling events start and evolve through time. Other works (Marcello et al., 2011) explore additional features from SST data using bathymetric data to retrieve several upwelling features at a given sea depth to derive a coastal upwelling index (CUI). Rather than recognizing a coastal upwelling region, Oerder et al. (2018) focus was on the automatic detection of the upwelling fronts, by simply checking if the minimum cross-shore SST gradient was below a specified threshold between the coastal and the offshore regions. The work developed by Saldías et al. (2021) explores the tracking of coastal upwelling events through time by measuring coherence between several coastal upwelling regions using the Canny Edge-Detection algorithm, where chlorophyll fluorescence data was also used for the spatio-temporal variations study. However, most of these methodologies are either too complex and/or involve too many ad hoc parameters.

Using the concept of core–shell structure we assume a simplified model for the upwelling recognition and tracking from the SST grids as follows: consider a constant 'core' which forms quite fast and then gradually expands along the offshore waters up to some physical limits, and later shrinks to the core region. Then, we define a clustering criterion and apply the least-squares to approximate the SST data to the model, so that key parameters of the core–shell cluster model are found in an automated way as minimizers of the criterion. This aspect favorably compares our approach with the one in Chen et al. (2015) where key parameters like thresholds and the number of clusters to retrieve from data are to be taken ad hoc.

Since the least-squares clustering criterion is additive, this enables us to extract clusters one-by-one rather than simultaneously. Such a sequential extraction strategy has two advantages in the context of the issue of automating upwelling description: first, it is consistent with the nature of the upwelling phenomenon under consideration, because at each time instant, there is only one major upwelling region to occur at the coast off Portugal (if any); second, it allows to apply the Self-Tuning Seed Expanding Cluster (STSEC) algorithm (Nascimento et al., 2015), a version of the pioneer Seeded Region Growing (SRG) approach in image analysis (Adams and Bischof, 1994), as a prerequisite to the current core–shell methodology. The STSEC algorithm follows the same least-squares strategy, so that its application is, in fact, part of the core–shell method being developed here. Moreover, its application to SST images as temporal data puts the business of finding relatively stable "time windows", here called "ranges", for obtaining the core clusters onto an automated footing.

Therefore, we propose a three-stage clustering method for the spatio-temporal analysis of upwelling. First, the STSEC algorithm is applied to each collection of SST grids, characterizing an upwelling season, resulting in the segmented upwelling regions. Second, the Iterative Anomalous Pattern (IAP) algorithm (Mirkin, 2012) unsupervisedly finds those 'ranges', from the STSEC segmentations, to define stability periods within the upwelling season. At the third stage, the STSEC segmentations belonging to each *range* are given as input to the core–shell clustering algorithm to further transform the corresponding upwelling instants into a core–shell cluster and the corresponding intensities.

The remainder of the paper is organized as follows. We describe the core–shell clustering model in Section 2. Section 3 describes our methods for Core–Shell cluster analysis of a season's sequence of SST grids derived from SST remote sensing data. In Section 4, we describe the conducted experimental study and analyze the main results. Section 5 concludes the paper.

2. Core–shell clustering model

Let $A^t(I,J) = \left(a^t_{ij}\right)$ be a given preprocessed SST grid, with temperature values a^t_{ij} at points (i,j) where i $(i = 1,2,\dots,I)$ is the longitude, j $(j = 1,2,\dots,J)$, the latitude, and t $(t = 1,2,\dots,T)$, a time moment within a period T.

We are going to build a set of clusters U^t over the set of grids A^t to represent the phenomenon of upwelling and its dynamics. One may think that it can be easily done by finding individual clusters C^t representing the upwelling at each moment $t = 1,2,\dots,T$. Unfortunately, this is not the case because clusters C^t at distinct moments t are different, so that one needs to define the ways of how transitions from t to $t+1$ occur so that the cluster changes are more-or-less smooth. A common sense thinking suggests an expected pattern of the dynamics of an upwelling as a smooth process of steadily growing, in the beginning of the upwelling season, to then steadily declining in the end of the season. However, as our observations of the SST images show, the real world dynamics of the coastal upwelling phenomenon do not follow this pattern. Instead, coastal upwelling rather follows a 'piece-wise constant' pattern: at every year, the upwelling season can be divided into shorter periods at which the upwelling shape on the consecutive SST grids can be considered constant. To better model such dynamics we came up with the concept of core–shell cluster to cover each of the periods.

A core–shell cluster $U = \bigcup_{t=1}^{T} U^t$ is the union of a constant set R and a variable set S^t not overlapping R, so that $U^t = R \cup S^t$ $(t = 1,2,\dots,T)$ are the cluster slices. It is assumed that the core set R is composed of tightly related core pixels, whereas pixel temperatures in the shells may be more or less variable.

Therefore, a core–shell cluster slice U^t, is represented by two non-overlapping sets of binary values, $R \cup S^t$, $r_{ij} \in R$, the *core*, and $s^t_{ij} \in S^t$, the *shell* at moment t, such that $r_{ij} \times s^t_{ij} = 0$, for $t = 1,2,\dots,T$.

Assume that the shells S^t are characterized by their intensity values λ^t. The intensity of the core should be greater than that for any $t \in T$, that is, the core's intensity is $\lambda^t + \mu^t$ with $\mu^t > 0$. Then, the model to represent an upwelling sea surface temperature, a^t_{ij}, at point (i,j) and moment t can be stated as:

$$a^t_{ij} = \left(\lambda^t + \mu^t\right) r_{ij} + \lambda^t s^t_{ij} + e^t_{ij}, \tag{1}$$

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

$$\Delta = \sum_{t=1}^{T} \sum_{i=1}^{I} \sum_{j=1}^{J} \left(a^t_{ij} - (\lambda^t + \mu^t) r_{ij} - \lambda^t s^t_{ij}\right)^2. \tag{2}$$

Consider the first order minimality conditions for function Δ at moment t:

$$\frac{\partial \Delta}{\partial \lambda^t} = -2 \sum_{i,j} \left(a_{ij}^t - (\lambda^t + \mu^t) r_{ij} - \lambda^t s_{ij}^t \right) \left[-s_{ij}^t - r_{ij} \right] = 0, \tag{3}$$

$$\frac{\partial \Delta}{\partial \mu^t} = -2 \sum_{i,j} \left(a_{ij}^t - (\lambda^t + \mu^t) r_{ij} - \lambda^t s_{ij}^t \right) r_{ij} = 0. \tag{4}$$

Eqs. (3) and (4), after taking into account that $s_{ij}^t \times r_{ij} = 0$, lead to:

$$\lambda^t = \frac{\sum_{i,j} a_{ij}^t s_{ij}^t}{\sum_{i,j} s_{ij}^t}, \tag{5}$$

$$\lambda^t + \mu^t = \frac{\sum_{i,j} a_{ij}^t rij}{\sum_{i,j} rij}. \tag{6}$$

Substituting the derived intensity values λ^t and $\lambda^t + \mu^t$ in (5) and (6) into Eq. (2) leads to:

$$\Delta = \sum_{t=1}^{T} \sum_{i=1}^{I} \sum_{j=1}^{J} \left(a_{ij}^t \right)^2 - \sum_t \left(\left(\lambda^t + \mu^t \right)^2 \times |R| + \left(\lambda^t \right)^2 \times \left| S^t \right| \right), \tag{7}$$

where $|R| = \sum_{i,j} r_{ij}$ is the number of data points in the *core* and $\left|S^t\right| = \sum_{i,j} s_{ij}^t$ is the number of points in the t-shell.

Criterion (7) can be written as

$$\Delta = D - G, \tag{8}$$

with $D = \sum_t \sum_{i,j} \left(a_{ij}^t \right)^2$ defining the total data scatter and $G = \sum_t ((\lambda^t + \mu^t)^2 \times |R| + \left(\lambda^t \right)^2 \times \left| S^t \right|)$ being the core–shell cluster's contribution to that.

Since D is constant, to minimize the least squares criterion (7) is equivalent to maximize criterion G. An iterative algorithm, the Core–Shell Clustering, to optimize criterion G, is introduced in Section 3.

3. Core–shell cluster analysis of a sequence of SST grids

Our method takes a season's sequences of SST weekly grids derived from the remote sensed data as its input. Then the method proceeds in four stages: data preprocessing, segmentation of the sequence to separate the upwelling regions, obtaining upwelling stability periods (ranges), and core–shell clustering of SST's within the ranges. Fig. 1 shows the workflow of the procedure of the method. The following sections give a description of the key parts of the process.

3.1. Data preprocessing stage

In order to improve the quality of the SST grids, a preprocessing method has been developed. This allows to avoid the issue of over-segmentation faced by SRG type algorithms (Aouni et al., 2021).

The developed preprocessing stage comprises three steps and is illustrated in the workflow displayed in Fig. 2. The methods applied at each step, are summarized bellow:

1. **Removal of north-south temperature gradient**
   The goal of this step is to eliminate the north-south temperature gradient. This gradient emerges because higher latitudes receive less solar energy than those to the South. To do this, we use the *grdtrend* module from the Generic Mapping Tools (GMT) software (Wessel et al., 2019).
2. **Moving average filter**
   After removing the north-south gradient from the SST grids, we apply the sliding window averaging algorithm at the window size $W = 5$. This creates a new set of smoothed SST grids with noise removed to better expose the upwelling events.
3. **Enhancement of East-West temperature gradient**
   To further tackle the aforementioned over-segmentation problem, we added this step to the current preprocessing pipeline. The temperature values of each SST grid are normalized using perpendicular lines to the coastline, adapting a method from Aouni et al. (2021).
   Since the Portuguese coast has an almost constant longitude, the rows of each SST grid are used as the lines perpendicular to the coastline. Then, each point (i, j) belonging to each line l is assigned with the difference between its temperature and the average temperature of the points in line l, if this difference is negative, or zero if the difference is positive. After this transformation the temperatures of the points near the coast, i.e. the region of interest, become negative while the temperatures of the offshore ocean region are about zero. This transformation is effective in prevention the over-segmentation because the automatic thresholding criterion of STSEC clustering algorithm (described later) involves zero as a central value.
   This step marks the end of the preprocessing pipeline.

Each SST grid obtained in the end of the preprocessing stage, is referred to as an *SST instant*.

3.2. Unsupervised segmentation of SST grids to separate upwelling regions

The identification of the so-called upwelling front, that separates the cold upwelled waters at the coast from the warmer oceanic waters offshore, is critical to infer the coastal upwelling pattern. 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 within the framework of 'anomalous clustering' (Mirkin, 2012), and has proven to overcome well recognized limitations of SRG algorithms. First, the clustering criterion takes the format of a product rather than the conventional difference between 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 a regularizer of the cluster growing process.

After choosing the coldest pixel of a given SST grid as its initial seed, the STSEC algorithm sequentially expands the region while the pixels being evaluated satisfy a similarity condition, automatically obtained. Whenever no more pixels satisfy such condition, a bi-partition grid is obtained representing the upwelling/non-upwelling regions of the grid. Fig. 3 shows an SST image (on the left) and the corresponding STSEC upwelling region segmentation corresponding to the binary grid on the right.

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 Canary coasts (Nascimento et al., 2015, 2020).

3.3. Finding time ranges: segments of temporal stability

Given an upwelling season, we want to identify its temporal stability segments: groups of consecutive SST instants with similar upwelling characteristics. To accomplish this, we use the Iterative Anomalous Patterns (IAP) algorithm. The IAP is a simple and effective clustering algorithm that sequentially extracts clusters one by one in a manner similar to the Principal Component Analysis (PCA) and, simultaneously, allows to derive the number of clusters to be found. This algorithm was originally proposed to initialize K-means, leading to the so-called intelligent K-means (Mirkin, 2012). 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).

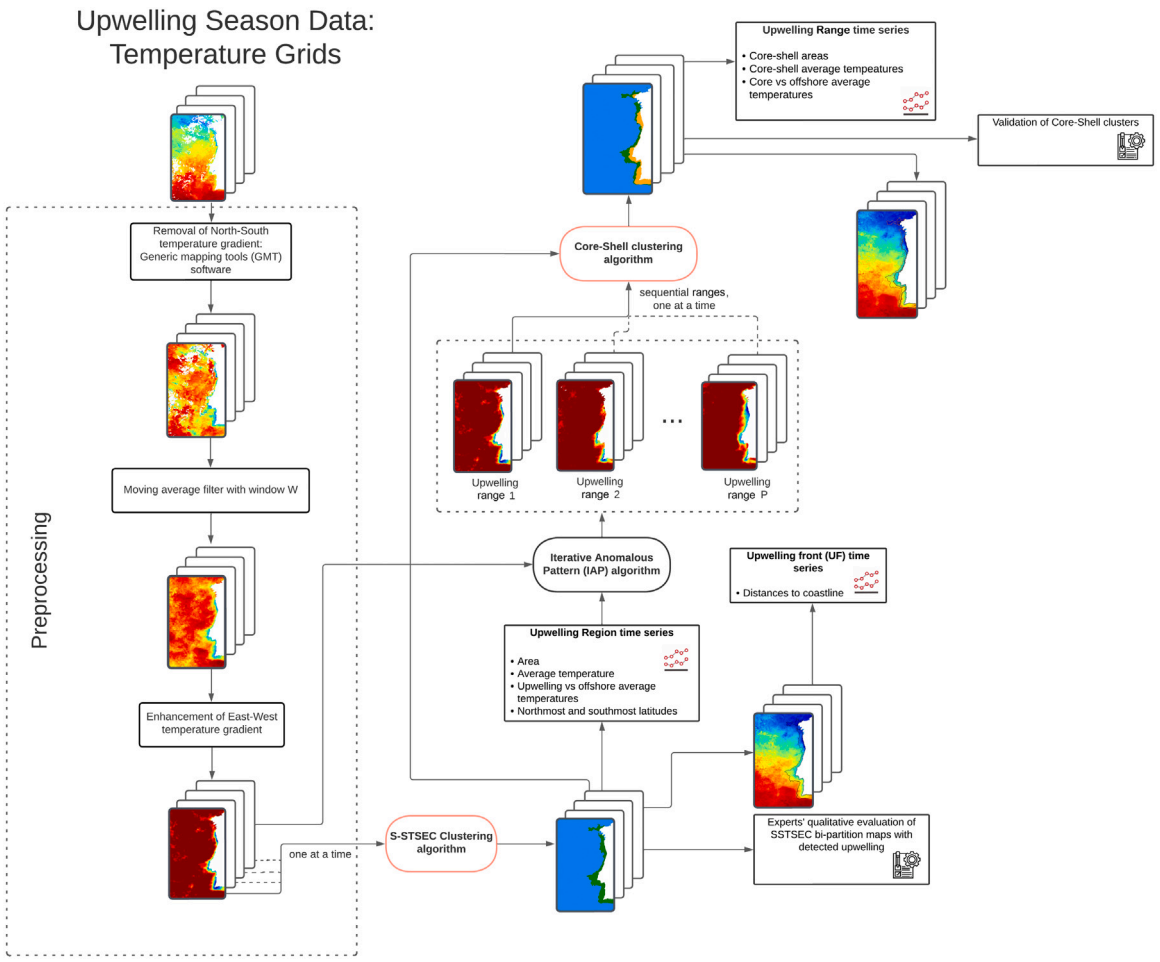

Fig. 1. Workflow of the core–shell clustering Process.

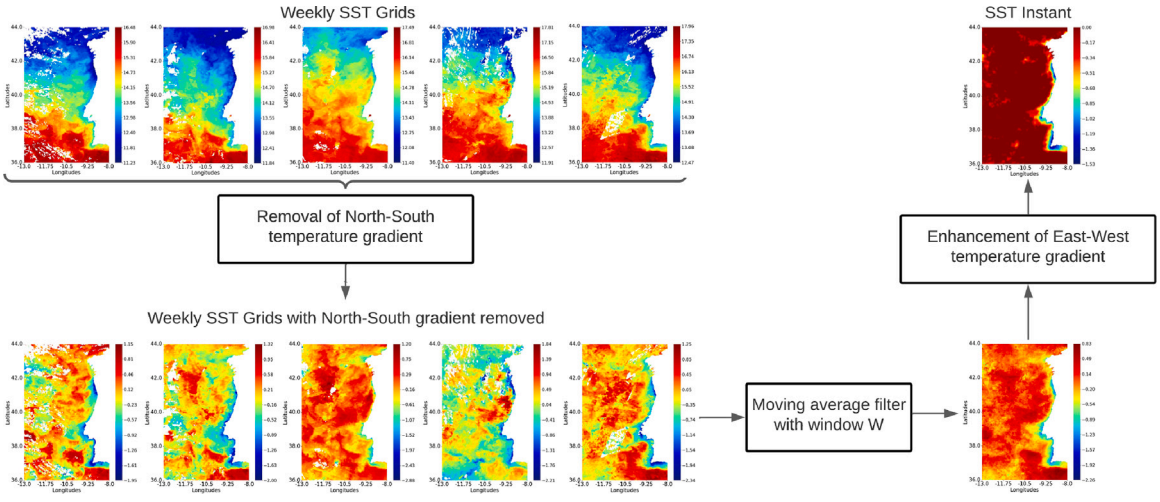

Fig. 2. Preprocessing pipeline of SST grids.

We take the consecutive upwelling regions segmented from the M SST instants ($M = 23$ for each upwelling season, which is 27 reduced by 4 because of the moving average filter with window size 5) to produce a partition of SST instants into disjoint sets of consecutive instants of more-or-less similar upwelling regions. To this end, four features are extracted from those segmented upwelling regions: the total area, average sea surface temperature, the maximum and the minimum latitudes. Then, a $M \times 4$ matrix is produced and given as input to the IAP. The data matrix is centered by subtracting the grand mean from every feature and then all the values are divided by the corresponding feature range. When running the IAP algorithm, imposing the cardinality of the clusters to be greater than two, it unsupervisedly finds at least three clusters, groups of consecutive SST instants, for each upwelling season. Each found group is designated as an upwelling *time range*. Fig. 4 shows the time series of those (normalized) features obtained

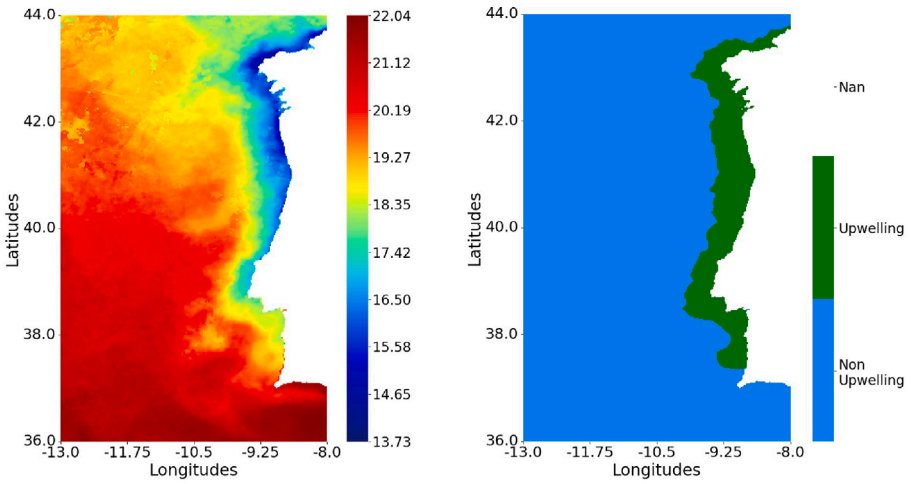

Fig. 3. SST image of Portugal with coastal upwelling (left); correspondent S-STSEC segmentation (right).

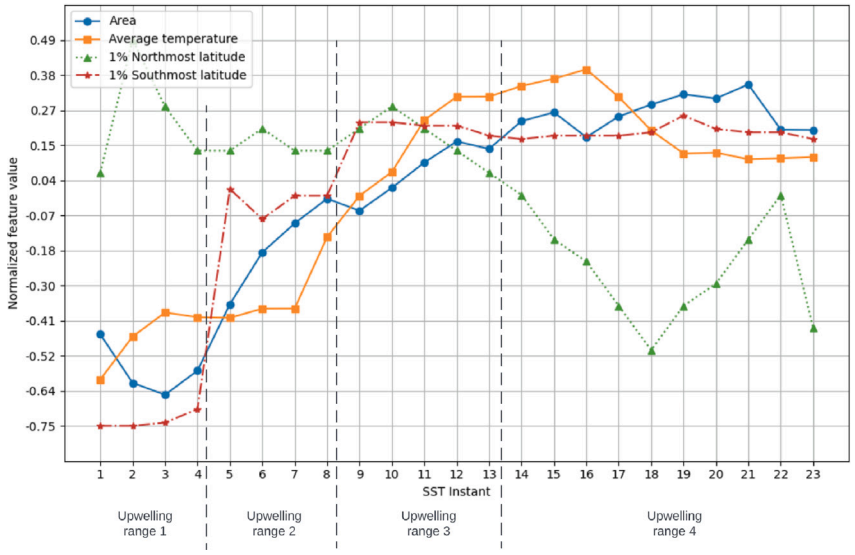

Fig. 4. Upwelling time ranges: consecutive SST instants clustered by IAP algorithm taking the four time series of upwelling extracted features.

from the segmentations of the SST instants (for the upwelling season 2019) with the IAP four-cluster partition marked by vertical dash lines.

3.4. Finding core–shell cluster within a time range

We propose an iterative algorithm, Core–Shell Clustering, to build a core–shell cluster by optimizing criterion (8). The algorithm receives as input a sequence of T SST instant grids defining a "range". Each grid is segmented by the STSEC algorithm, resulting in T clusters, $C^1, C^2, \ldots, C^T$, corresponding to the T sequential segmented upwelling regions. The output is a core–shell cluster, a sequence of T cluster slices $R \cup S^{t=1}, R \cup S^{t=2}, \ldots, R \cup S^{t=T}$ and corresponding intensity values λ^t and $\lambda^t + \mu^t$ $(t = 1, 2, \ldots, T)$.

The initial core–shell cluster slices are constructed from those T clusters $C^1, C^2, \ldots, C^T$. Since the core cluster is, by assumption, constant and homogeneous, the initial core is defined taking the intersection of T STSEC clusters, C^t, $t = 1, 2, \ldots, T$. Each shell, S^t is defined as the set difference of cluster C^t and core R, covering the upwelling region that may change at moment t along the period T. Then, the initial intensity values λ^t and $\lambda^t + \mu^t$ are calculated by Eqs. (5) and (6).

To connect each initial core–shell cluster slice with the remaining upwelling region at instant t, the set of the grid points forming a spatial 4-neighborhood, F^t, are merged with them. Thus, set $B = \left\{R \cup S^t \cup F^t\right\}_{t=1}^{T}$ defines the initial core–shell cluster.

After, the algorithm iterates as follows. For each point (i, j) in $B = \bigcup_{t=1}^{T} B^t$ decide which scenario to take: (A) to make (i, j) to belong to the core, or (B) to belong to any of the T shells, or, on contrary, (C) to remove point (i, j) from any of them, such that the increase of criterion G, δ_{ij} is maximum. This process requires $1 + 2^T$ tentative binary decisions. The process iterates until there is no improvement in criterion G, that is, until $\delta_{ij} \leq 0$. The derivation of values δ_{ij} (omitted due to lack of space) is straightforward.

The algorithm in pseudo-code is presented here.

4. Experimental study

4.1. Image data

Sixteen annual collections of SST grids from the Portuguese coast (latitudes ranging from 36°N to 44°N and longitudes ranging from 13°W to 8°W) were used in this study covering the years from 2004 to 2019. Each SST grid, with a size of 401 × 251 nodes and with a spatial resolution of 2 km × 2 km, represents the average of 8 daily SST grids made with Level 2 data downloaded from the OceanColor site (https://oceancolor.gsfc.nasa.gov/), and filtered to use only the

Algorithm 1 Core-Shell Clustering Algorithm

Input: Preprocessed T SST grids $A^t(I,J) = (a_{ij}^t)$, $i = 1,2,...,I$; $j = 1,2,...,J$, and $t = 1,2,....,T$.
Output: core-shell cluster $\{R \cup S^t\}_{t=1}^T$, corresponding intensities $\lambda^t + \mu^t$, λ^t $(t = 1,2,\cdots,T)$, and the final cluster contribution G.
Initialization: ▷ Construct the initial core R and the initial shells S^t from the STSEC clusters obtained from SST instant grids A^t at moments $t = 1,2,...,T$ as:
Run STSEC algorithm over SST instant grids A^t $(t = 1,2,....,T)$; let $C^1, C^2, \cdots, C^T$ be the obtained clusters
Calculate $R = \cap_{t=1}^T C^t$ and $S^t = C^t - R$ $(t = 1,\cdots,T)$
Let $B = \{R \cup S^t \cup F^t\}_{t=1}^T$ with F^t the set of the grid points forming a spatial 4-neighborhood with the region defined by $R \cup S^t$
Calculate intensity values λ^t and $\lambda^t + \mu^t$ by (5), (6)
Iteration: ▷ For each point (i,j) in B decide which of r_{ij}, s_{ij}^t should be 0 or 1, which requires $1 + 2^T$ binary decisions, to maximize the increase δ_{ij} of criterion G in (8)
repeat
  for each point $(i,j) \in B$ **do**
    if $r_{ij} \leftarrow 1$ **then**
      $s_{ij}^t \leftarrow 0$ for $t = 1,2,\cdots,T$
      calculate the change in criterion G, δ_{ij}
    else
      $r_{ij} \leftarrow 0$
      for $t = 1,2,\cdots,T$ **do**
        $s_{ij}^t \leftarrow 0$ **.or.** $s_{ij}^t \leftarrow 1$
        calculate the (two) change values in criterion G, δ_{ij}
      end for
    end if
  end for
  Let $\delta_{i^*j^*}^*$, be the maximum of those $1 + 2^T$ δ_{ij} for $i = 1,2,...,I$; $j = 1,2,...,J$
  if $\delta_{i^*j^*}^* > 0$ **then**
    assign the decision for point (i^*j^*) as $r_{i^*j^*}^*$

    for $t = 1,2,\cdots,T$ **do**
      assign the decision $s_{i^*j^*}^t$
      calculate the intensities λ^t and $\lambda^t + \mu^t$ by (5), (6);
    end for
  end if
until $\delta_{i^*j^*}^* < 0$

best quality data according to the products quality flags. The upwelling season along the Portuguese coast is typically stronger from March to October, thus we considered an upwelling season to be composed only by SST grids from approximately the 30th of March to the 30th of October. This gives a total of 27 8-day average SST grids per season and 432 SST grids for the whole sixteen annual collections (27×16).

Each SST grid is represented as a $I \times J$ matrix $A(a_{ij})$, where $i = 1,2,\ldots,I$ is the longitude, $j = 1,2,\ldots,J$ is the latitude, and a_{ij} represents the temperature Celsius value at node (i,j) (°C). Exceptions occur whenever an error is associated with the data collection due to meteorological events or values collected from a non-sea surface. In those cases the grid nodes are assigned a not-a-number (NaN) value. Fig. 5 shows an example of a SST grid. The continuous white region on the right side of each SST image corresponds to land surface, whereas white pixels in the ocean area correspond to missing values due to cloud cover.

4.2. SST grid data preprocessing

We illustrate the data preprocessing stages (described in Section 3.1) using data from the 2007 collection of SST grids, in Fig. 2.

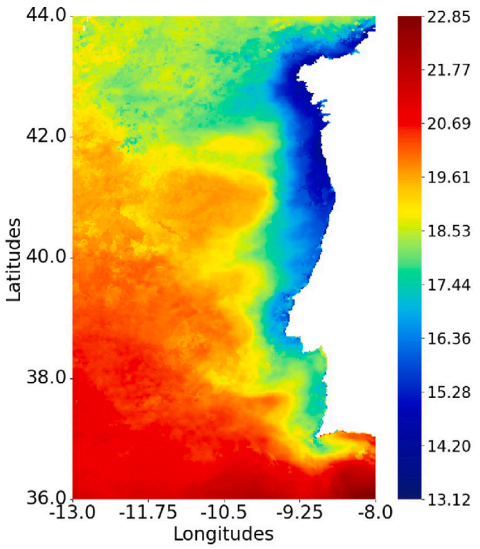

Fig. 5. Example of SST image of coastal upwelling in Portugal.

Fig. 6 shows a sample of five consecutive weekly noisy SST grids input of the preprocessing. Notice how the remote sensing derived data are imprecise because of intricate interactions controlling the process, as well as because of cloud cover. This is visualized by white pixels corresponding to NaN values. Each SST grid shows a strong north-south temperature gradient due to the latitudinal extent of these images: the further away the waters are from the Equator line, the colder they tend to be. As shown experimentally, the gradient negatively influences the quality of the SST segmentation leading to under-/over-segmentations. The role of the first preprocessing step is to remove that.

Fig. 7 shows results of removing the north-south temperature gradient from the SST grids in Fig. 6. Each SST grid has no variation of temperature over the latitudes, thus decreasing the temperature ranges with an average SST value of approximately 0 °C.

The transformed SST grids still contain noise and the upwelling patterns have low definition. Thus, at the second stage, a moving average filter applies. The window size w was experimentally set to $w = 5$ after an extensive experimental study with the sixteen collections of SST grids.

The left image in Fig. 8 has resulted from applying the moving average filter to SST grids in Fig. 7. This looks like a strong enhancement of the upwelling event pattern, along with the absence of noise.

Unfortunately, when running the S-STSEC algorithm with SST grids after the second preprocessing stage, we encountered over-segmentation in several cases, mainly in the north region. Therefore, we apply the third preprocessing stage by increasing the East-West temperature gradient at the SST grids to better separate the upwelling area from non-upwelling regions.

The image on the right of Fig. 8 shows the result of the third preprocessing stage to the SST grid of the image on the left of Fig. 8. Notice that the temperature values of the upwelling regions become negative while the temperatures of the offshore ocean are/tend to 0 °C. We refer to SST images after all the preprocessing stages as to SST instants.

The preprocessing pipeline was applied to each of the sixteen annual SST collections. Since the moving average filter has window size 5, each collection of 27 original SSTs is shrunk to 23 preprocessed SST *instants*. This totals to 368 SST instants for the sixteen years.

4.3. Upwelling regions retrieved from the SST instants

The S-STSEC algorithm was applied to each SST instant grid. To illustrate the impact of the preprocessing transformation in the obtained segmentation results, Fig. 9 shows the averaged SST grid on the

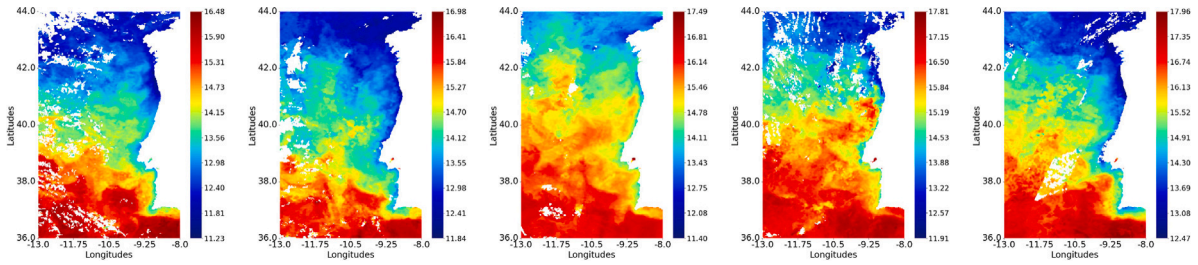

Fig. 6. Preprocessing input data: consecutive weekly SST grids from the year 2007.

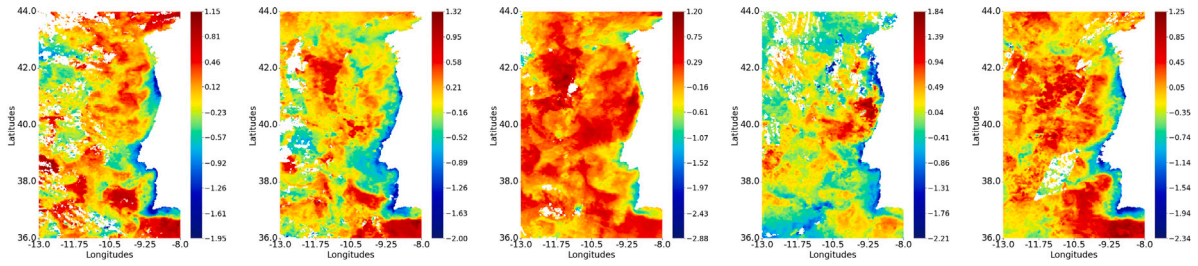

Fig. 7. Preprocessing output stage one: consecutive weekly SST grids after removal of the North-South temperature gradient from them.

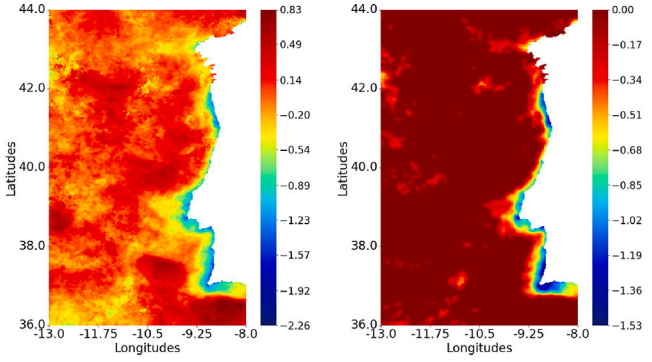

Fig. 8. Preprocessing second stage's output: SST grids after the moving average filter (left image); result of applying the East-West temperature gradient enhancement (right image).

left image, the result of S-STSEC segmentation without preprocessing at the middle image, and the result of S-STSEC segmentation after preprocessing (right image), with the segmented upwelling regions visualized in green.

The results of applying the S-STSEC to the sixteen collections of SST instant grids were evaluated by a team of expert oceanographers. Each S-STSEC SST instant's segmentation has received a grade in a 5-grade scale from 1 (bad) to 5 (excellent). The evaluations are rather concordant; their median is 4.5, and mode is between 4 and 5.

4.4. Main features of S-STSEC segmentations

We extracted four features to characterize the S-STSEC segmentations. Two main features were selected based on domain knowledge: the area of the upwelling regions, and the mean temperature of those regions against the mean temperature of the offshore waters. The offshore region was defined as a meridional strip with 0.4° of longitude width from 36°N to 44°N, parallel to the coastline and 320 km distant from shore. In this way, we guarantee that this region will not be affected by waters upwelled along the coast.

The upwelling areas time series for the years 2004 to 2019 are shown in Figs. 10. The upwelling areas tend to have the lowest value at the start of each upwelling season and reach a maximum value in the second half of such season. This pattern is present in every year with exception of 2013 collection where the lowest area was registered at SST instant 13. Looking at the time series as a whole, the lowest registered upwelling area occurred at the second SST instant of 2011 with an area of 9224 km^2 while the highest area value occurred at SST instant 18 of 2018 with a value of 55524 km^2. For the whole sixteen years the upwelling areas increased with an average of 320.81 km^2 per year. Such measurement was obtained by simple linear regression applied to the whole dataset.

Fig. 11 shows the time series of the average temperatures in the S-STSEC upwelling regions against the offshore waters average temperatures. The maximum average upwelling temperatures were obtained, in majority, at SST instant 15 of each year. The maximum average offshore temperatures were obtained, mostly, at instant 17 of each year. Overall, in each upwelling season the upwelling regions reached their maximum temperature, on average, two instants earlier than the offshore waters. A maximum average upwelling temperature of 20.26 °C was observed at the 22nd instant of 2014, and a minimum average upwelling temperature of 13.8 °C was observed at the 1st SST instant of 2004. Concerning the offshore average temperatures, the maximum offshore temperature of 21.78 °C occurred at the 16th SST instant of 2014 while the minimum offshore temperature of 14.0 °C occurred at the 1st instant of 2018. We conclude that, except for the year 2013, the coldest upwelling and offshore waters occurred at the first SST instant of each upwelling season.

Two other features were empirically selected to characterize morphological differences between the upwelling segmented SST instants. These features are the maximum and minimum latitudes of the detected upwelling regions.

4.5. Modeling upwelling ranges

Given an upwelling season, we want to cluster consecutive SST instants with similar upwelling characteristics, as reflected by the four extracted features. Each group is referred to as a *range*.

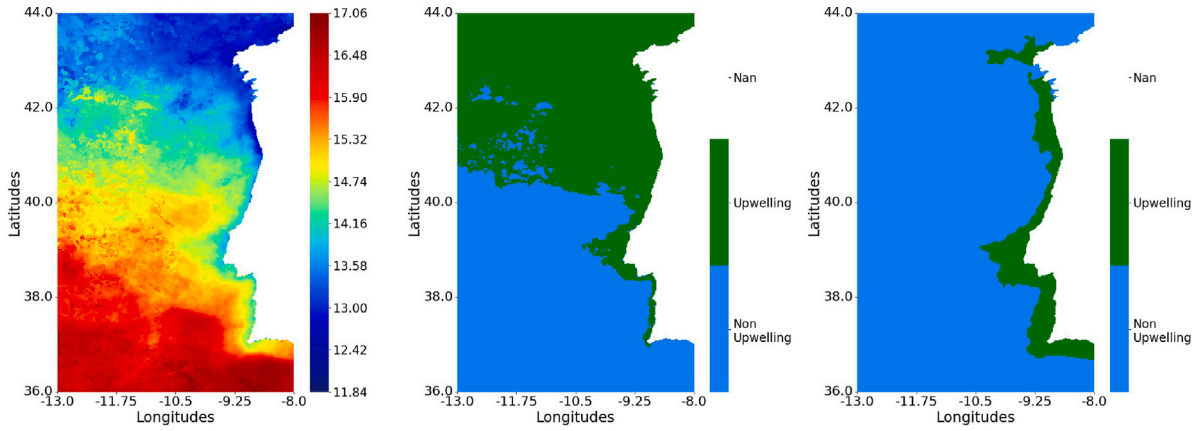

Fig. 9. Averaged SST grid (left image); the S-STSEC result without preprocessing, thus leading to over-segmentation (middle image); the S-STSEC segmentation after preprocessing (right image).

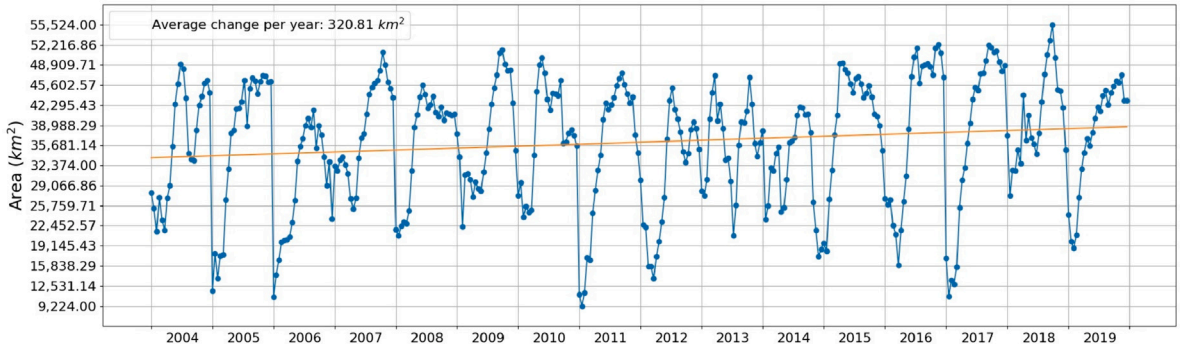

Fig. 10. Time series with the Areas (km^2) of S-STSEC segmented upwelling regions for the sixteen years.

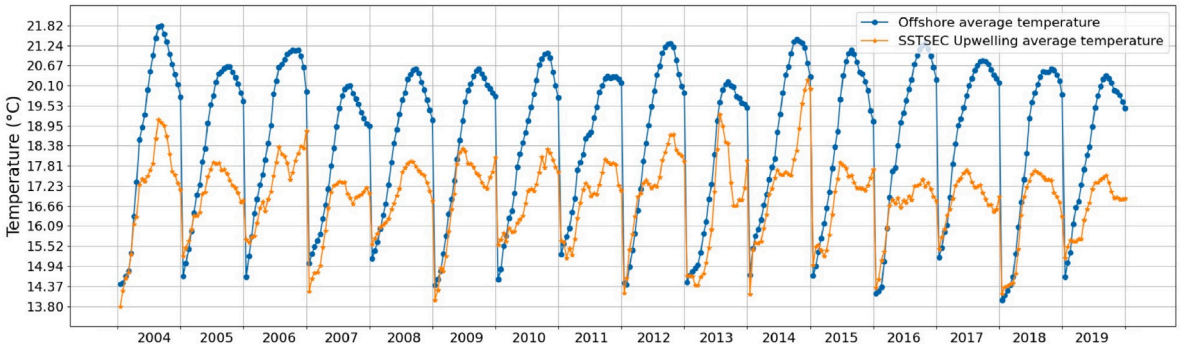

Fig. 11. Time series with the average temperatures (°C) of segmented upwelling regions against the offshore average temperatures for the sixteen years.

4.5.1. Unsupervised time series segmentation by anomalous clustering

For each annual collection of SST instants, a series of 23 SSTs resulting from the preprocessing stage, we take four time series defined by the four features, normalized by range, as input to the IAP algorithm. The IAP stop condition is specified at the clusters' cardinality higher or equal to three. The obtained groups of consecutive SST instants are listed at the columns of Table 1 as $a - b$ where a is the starting SST instant and b is the final SST instant. They define upwelling *ranges* for each year. Each group of consecutive SST grids defining a time range is the input to the core–shell clustering algorithm.

4.5.2. Illustration of core–shell clustering stages

Fig. 12 illustrates an example of the core–shell clustering algorithm applied to the second time range of SST instants in 2007 (instants 9 to 13) displayed at first row of the figure. At the initialization stage (line 4 of Algorithm 1), the S-STSEC algorithm is applied to each of these instant SST grids to produce the binary grid with its upwelling region visualized in green (second row of Fig. 12). The initial core–shell cluster slices are build from the clusters defining the upwelling regions (lines 5–7 of Algorithm 1). Finally, the core–shell cluster results (lines 9–31 of Algorithm 1) are visualized with core–shell cluster slices highlighted using orange-green colors (third row of Fig. 12). An additional result, shown at the bottom row of the figure, is the automatic front delineation of the core and shell structures in the original SST grids.

4.6. Validation of core–shell clustering results

Extensive analysis of SST images shows that the upwelling regions are not that easy to distinguish on a SST grid. They: have transition zones with fuzzy thermal boundaries; correspond to quite irregular

Table 1
Time series segmentations produced by the IAP algorithm. Each group of SST instants represents a time *range*; we label it as $a-b$, with a and b its boundary SST instants.

Upwelling time range	2004	2005	2006	2007	2008	2009	2010	2011
1	1–7	1–5	1–7	1–8	1–5	1–7	1–6	1–5
2	8–10	6–10	8–12	9–13	6–9	8–12	7–14	6–8
3	11–23	11–23	13–23	14–23	10–19	13–23	15–23	9–12
4					20–23			13–23

Upwelling time range	2012	2013	2014	2015	2016	2017	2018	2019
1	1–7	1–8	1–9	1–3	1–7	1–5	1–6	1–4
2	8–11	9–13	10–12	4–11	8–10	6–10	7–9	5–7
3	12–23	14–17	13–23	12–23	11–13	11–23	10–14	8–13
4		18–23			14–23		15–19	14–23
5							20–23	

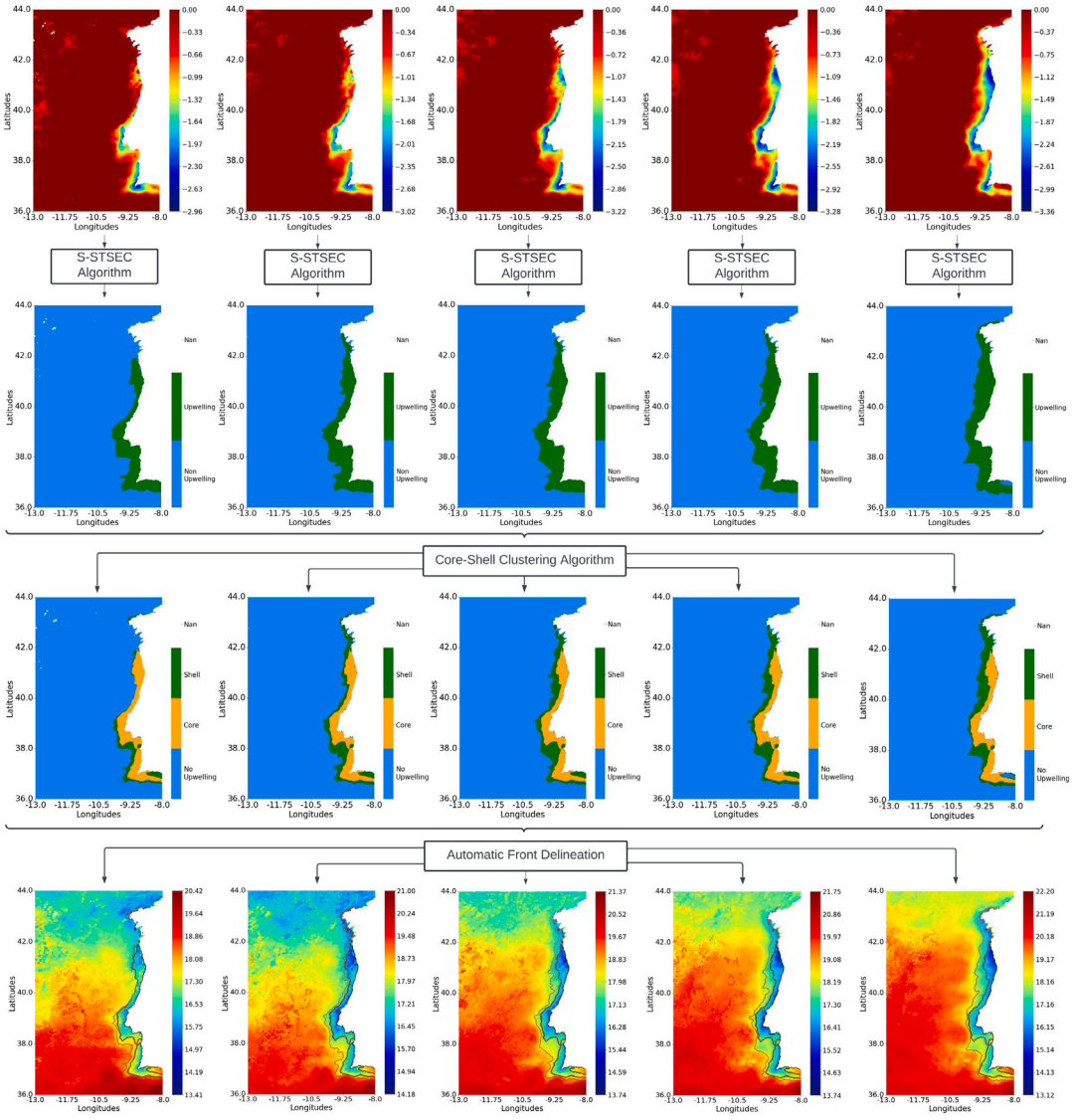

Fig. 12. Core–shell clustering algorithm example. Input: five SST instants corresponding to the 2nd time range of 2007 (first row); Initialization: S-STSEC segmentations results (second row); Final Iteration: core–shell cluster slices (third row); Additional result: original SST grids with core and shell fronts delineated (last row).

histograms; express strong morphological variation, etc. Due to the absence of a valid analytical model for the upwelling structures, one is confronted with a 'semantic gap' between the implicit Oceanographers knowledge and a systematic working definition of a 'gold standard' for upwelling images.

Thus, we are left with a qualitative assessment by an independent team of Oceanographers, which already proved rather successful at our previous publications related to S-STSEC algorithm (Nascimento et al., 2015). The obtained average validity score for S-STSEC upwellings at the current image collection was 4.5 (out of 5). Based on this, we took the S-STSEC upwelling results as ground-truth to assess the quality of the core–shell clustering segmentations with a more objective quality measures.

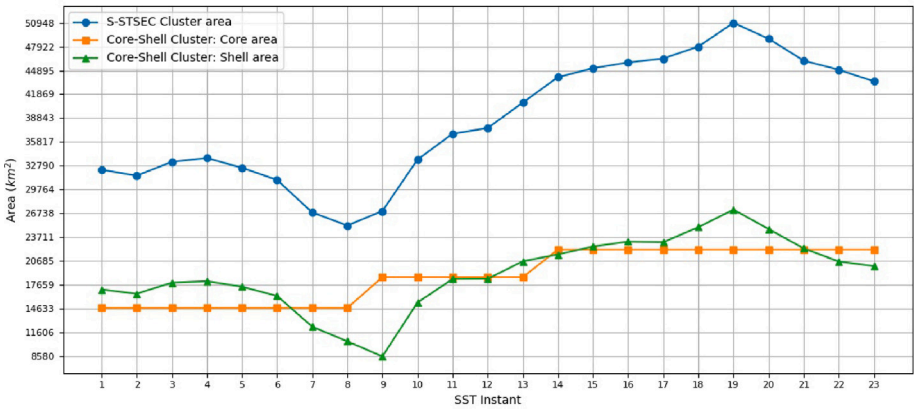

Fig. 13. Time series with the Areas (km^2) of 'cores'/'shells' of core–shell clusters, and the areas of upwelling regions by S-STSEC (year 2007).

We used the Kulczynski similarity index (Kulczynski, 1927; Zakani et al., 2016) to evaluate the quality of core–shell segmentations. This index, under the name "Maryland bridge coefficient", was proven to overcome drawbacks of the popular Jaccard coefficient in underestimating the similarity between two balanced sets with an equally balanced overlap or, contrastingly, between highly imbalanced sets (Mirkin and Koonin, 2003). We calculated the Kulczynski similarity between each core–shell segmentation and the corresponding S-STSEC ground-truth upwelling segmentation. The results show a real good matching, with an average Kulczynski similarity score of 0.978 ± 0.12 for the 368 images under study.

4.7. Analysis of core–shell clusters time series

To analyze upwelling using the obtained core–shell clusters, we defined the following features:

- the areas of the cores;
- the areas of the shells;
- the average temperature of the cores;
- the average temperatures of the shells (these are nothing else than the core and shell intensities, as defined by Eqs. (5) and (6));
- the average temperature of upwelling cores against average temperature of offshore ocean waters.

Fig. 13 illustrates, for the SST collection of 2007, the areas of upwelling cores (orange line), the areas of the corresponding shells (green line), as well as the areas of the whole upwelling regions obtained by S-STSEC algorithm (blue line). One can observe that the cores are constant along each time range (three ranges in 2007, as shown in Table 1), whereas the corresponding shells capture the evolving spatial pattern which is strongly concordant with the areas of the whole upwelling regions. This shows how the core represents the constant structure of the upwelling while the shells represent the evolving pattern associated with the spatial variability. Analyses of the corresponding time series for the other fifteen years were similar.

Two time series with the areas of cores and the areas of shells for the sixteen SST data collections are in Fig. 14. Since, by definition, the cores are constant, their areas are constant within each range. When comparing the shell areas to the S-STSEC segmentations for the same SST instants, it is possible to observe that the shells correctly capture the variability of the upwelling regions under analysis. Also, the shells generally correspond to larger regions when compared to the corresponding cores.

Fig. 15 shows the time series of core and shell average temperatures. The shell average temperatures are generally higher than the core average temperatures. Typically, this is not observed at the start of an upwelling season because the upwelling event is of a lower intensity at the time. An exception occurs in the year 2012 in which the cores have higher temperatures than the corresponding shells at the end of the upwelling season for that year. The highest average temperature of cores was recorded at the 21st SST instant of 2014 with a temperature of 19.88 °C while the maximum average temperature of the shells was observed at the 22nd instant of the same year with a temperature of 20.19 °C, 0.3 °C higher and one instant later than the highest average temperature of the core region. The lowest average core temperature was observed at the 1st SST instant of 2018 with a temperature of 13.98 °C while the lowest average temperature of the shells regions over the sixteen years was recorded at the 1st instant of 2004 with value 13.51 °C.

The time series of the average temperature of upwelling cores against the average temperature of the offshore waters at each SST instant of each year are displayed in Fig. 16. The results are aligned to the ones derived from S-STSEC segmentation leading to a concordant analysis. The upwelling core temperatures are considerably lower than the average temperature offshore. The temperature contrast increase towards the instants that correspond to the peak of the upwelling season. This result is expected given the nature of the coastal upwelling process.

The results achieved by the core–shell clustering are consistent with the oceanographic characteristics of the studied region. However, the oceanographic interpretation of these results is beyond the scope of the paper.

5. Conclusion

This paper proposes a description of the structure of a complex natural phenomenon, coastal upwelling, occurring in various locations within the world's ocean and being important for economy of countries nearby. Our case study concerns SST derived images of upwelling along the Portuguese coast. The results demonstrate that the approach is adequate to the case study. Based on other data available, we think our model may be adequate for other parts of the ocean as well. According to our computations, any specific upwelling pattern moves several times from one stable structure to another one during a year. These structures are captured here under the name of core–shell cluster, covering consecutive upwelling regions, within 'ranges' derived by the algorithm.

There can be 3 to 5 ranges during a season. Each core–shell cluster consists of two parts: a constant core and variable shell. Such a structure is not inconsistent with the oceanographic concept of wave soliton (Apel, 2003; Ibragimov and Lin, 2017). It requires, however, some further thinking into the nature of soliton shift, dissipation, and re-emergence.

We developed an automated machinery for discovery of core–shell cluster structures, which involves several stages: (a) obtaining and preprocessing yearly SST grid data; (b) finding time ranges and the

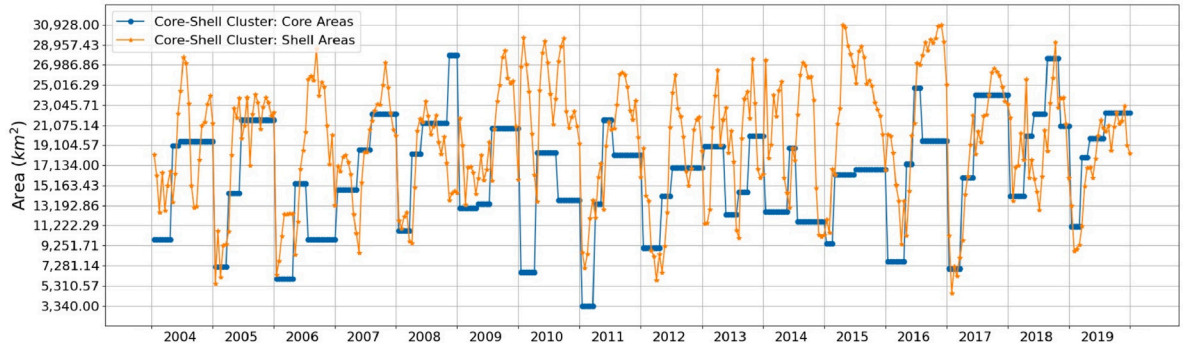

Fig. 14. Time series with the Areas (km^2) of 'cores'/'shells' of core–shell clusters for the sixteen years.

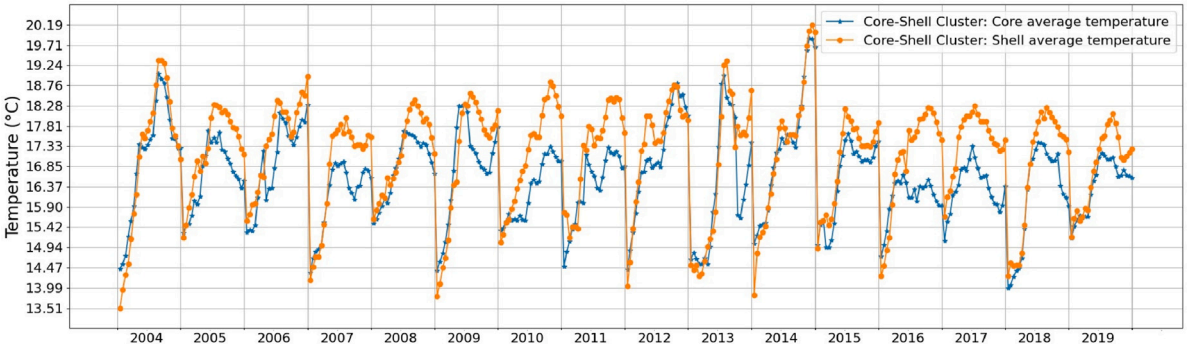

Fig. 15. Time series with the average temperatures (°C) of 'cores'/'shells' of core–shell clusters for the sixteen years.

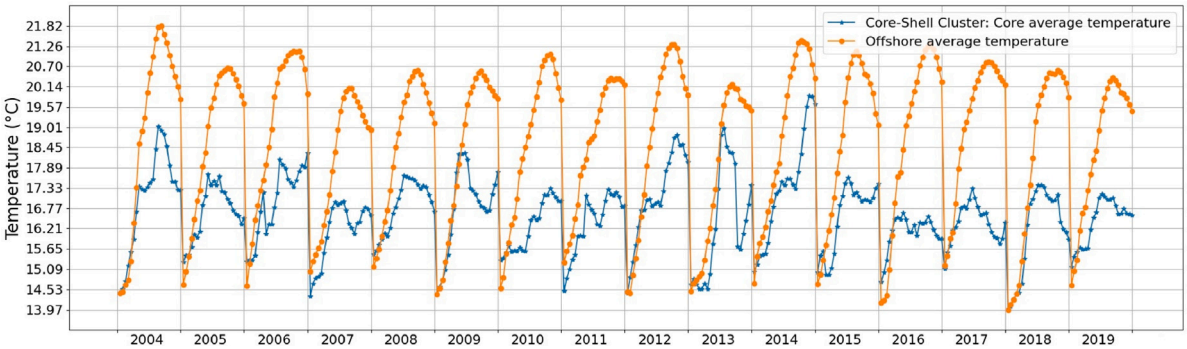

Fig. 16. Time series of cores average temperatures (° C) against the average temperatures of offshore waters for the sixteen years.

number of them; (c) determining core–shell structure within the ranges. According to our experimental results, we may safely claim that our method leads to automatically obtaining a full description of coastal upwelling as core–shell clusters. To the best of our knowledge, such a method is developed for the first time in the international literature.

Further work should include the application of our method to other upwelling regions such as Canary. More advanced extension would involve combining our approach with mathematical modeling of solitons as a theoretical construct.

CRediT authorship contribution statement

Susana Nascimento: Developed the conception, Methodology, and Design of the study, Participated in the analysis of results, Wrote the manuscript. **Alexandre Martins:** Implemented and tested the software code, Performed the experimental study, Participated in writing the experimental results, Produced the figures for the manuscript. **Paulo Relvas:** wrote parts of the introduction, Participated in the analysis of the results, Made the assessment of the segmentation results, Participated in oceanographic interpretation of the results. **Joaquim F. Luís:** acquired the SST data; provided access to pre-processing the *grdtrend* module from the Generic Mapping Tools (GMT), Participated in the analysis of results, Made the assessment of the segmentation results, Participated in oceanographic interpretation of the results. **Boris Mirkin:** Developed a mathematical model of 'time range' structure, Participated in formulation of the algorithms, Participated in the analysis of results, Participated in wrapping up the paper.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I provided a link where sample data (two years of SST grids derived from remote sensing SST images) are available for tests.

Acknowledgments

S.N. and A.M. acknowledge the support from NOVA LINCS, Portugal (UIDB/04516/2020), P.R. acknowledges the support through projects UIDB/04326/2020, UIDP/04326/2020 and LA/P/0101/2020, J. L. acknowledges the support through project UIDB/50019/2020, all funded by Portuguese national funds, Portugal from FCT- Foundation for Science and Technology, Portugal. B.M. acknowledges support from the Basic Research Program of the National Research University Higher School of Economics Moscow, UK. The authors are grateful to the anonymous reviewers for their insightful and constructive comments that allowed us to improve the presentation.

Computer code availability

- Name of the code: "CoreShellClustering"
- Contact: Alexandre Martins (ag.martins@campus.fct.unl.pt)
- Hardware requirements: No limits.
- Software requirements: To run the code at its current state, the user needs to have a Unix based system. The majority of the code produced is written in Python, except a small set of code lines that are written in Julia. As of software requirements, the user needs to have installed:
  - python: https://www.python.org/downloads/
  - julia: https://julialang.org/downloads/
  - GMT tools: https://github.com/GenericMappingTools/GMT.jl
- The source code is available for downloading at the link:
- https://github.com/thisDash/CoreShellClusteringAlgorithm

References

Adams, R., Bischof, L., 1994. Seeded region growing. IEEE Trans. Pattern Anal. Mach. Intell. 16 (6), 641–647. http://dx.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. Inform. Sci. 369 (C), 388–401. http://dx.doi.org/10.1016/j.ins.2016.06.048.

Ansari, M.Y., Ahmad, A., Khan, S.S., Bhushan, G., Mainuddin, 2019. Spatiotemporal clustering: A review. Artif. Intell. Rev. 53, 2381–2423.

Aouni, A.E., Daoudi, K., Minaoui, K., Yahia, H., 2021. Robust detection of the North-West African upwelling from SST images. IEEE Geosci. Remote Sens. Lett. 18 (4), 573–576. http://dx.doi.org/10.1109/LGRS.2020.2983826.

Apel, J.R., 2003. A new analytical model for internal solitons in the ocean. J. Phys. Oceanogr. 33 (11), 2247–2269.

Atluri, G., Karpatne, A., Kumar, V., 2018. Spatio-temporal data mining: A survey of problems and methods. ACM Comput. Surv. 51 (4), http://dx.doi.org/10.1145/3161602.

Baptista, V., Silva, P.L., Relvas, P., Teodósio, M.A., Leitão, F., 2018. Sea surface temperature variability along the Portuguese coast since 1950. Int. J. Climatol. 38 (3), 1145–1160. http://dx.doi.org/10.1002/joc.5231, URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.5231, arXiv:https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.5231.

Birant, D., Kut, R.A., 2007. ST-DBSCAN: An algorithm for clustering spatial-temp oral data. Data Knowl. Eng. 60, 208–221.

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. IJCAI '15, AAAI Press, pp. 2575–2581.

Ester, M., Kriegel, H.P., Sander, J., Xu, X., et al., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In: KDD, Vol. 96. pp. 226–231.

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 Sens. Environ. 227, 90–109. http://dx.doi.org/10.1016/j.rse.2019.04.002, URL https://www.sciencedirect.com/science/article/pii/S0034425719301397.

Ibragimov, R., Lin, G., 2017. Nonlinear analysis of perturbed rotating whirlpools in the ocean and atmosphere. Math. Model. Nat. Phenom. 12 (1), 94–114.

Kulczynski, S., 1927. Die pflanzenassoziationen der pieninen. Bull. Int. de l'Acad. Polonaise des Sci. et des Lett., Classe des Sci..

Marcello, J., Hernández-Guerra, A., Eugenio, F., Fonte, A., 2011. Seasonal and temporal study of the northwest African upwelling system. Int. J. Remote Sens. 32 (7), 1843–1859. http://dx.doi.org/10.1080/01431161003631576.

Mirkin, B., 2012. Clustering: A data recovery approach. Chapman and Hall/CRC Press (2nd Edition), p. 374. http://dx.doi.org/10.1201/b13101.

Mirkin, B., Koonin, E., 2003. A top-down method for building genome classification trees with linear binary hierarchies. In: Bioconsensus. DIMACS series, AMS Providence, pp. 97–112.

Nascimento, S., Casca, S., Mirkin, B., 2015. A seed expanding cluster algorithm for deriving upwelling areas on sea surface temperature images. Comput. Geosci. 85, 74–85. http://dx.doi.org/10.1016/j.cageo.2015.06.002, URL https://www.sciencedirect.com/science/article/pii/S0098300415001338, Statistical learning in geoscience modelling: Novel algorithms and challenging case studies.

Nascimento, S., Franco, P., Sousa, F., Dias, J., Neves, F., 2012. Automated computational delimitation of SST upwelling areas using fuzzy clustering. Comput. Geosci. 43, 207–216. http://dx.doi.org/10.1016/j.cageo.2011.10.025, URL https://www.sciencedirect.com/science/article/pii/S0098300411003608.

Nascimento, S., Mateen, S., Relvas, P., 2020. Sequential self-tuning clustering for automatic delimitation of coastal upwelling on SST images. In: Analide, C., Novais, P., Camacho, D., Yin, H. (Eds.), Intelligent Data Engineering and Automated Learning. IDEAL 2020, Springer International Publishing, Cham, pp. 434–443.

Oerder, V., Bento, J.P., Morales, C.E., Hormazabal, S., Pizarro, O., 2018. Coastal upwelling front detection off central Chile (36.5–37 °S) and spatio-temporal variability of frontal characteristics. Remote Sens. 10 (5), http://dx.doi.org/10.3390/rs10050690, URL https://www.mdpi.com/2072-4292/10/5/690.

Ramanantsoa, J.D., Krug, M., Penven, P., Rouault, M., Gula, J., 2018. Coastal upwelling south of Madagascar: Temporal and spatial variability. J. Mar. Syst. 178, 29–37. http://dx.doi.org/10.1016/j.jmarsys.2017.10.005, URL https://www.sciencedirect.com/science/article/pii/S092479631730249X.

Rodin, I., Mirkin, B., 2017. Supercluster in statics and dynamics: An approximate structure imitating a rough set. In: Polkowski, L., Yao, Y., Artiemjew, P., Ciucci, D., Liu, D., Ślęzak, D., Zielosko, B. (Eds.), Rough Sets. Springer International Publishing, Cham, pp. 576–586.

Saldías, G.S., Hernández, W., Lara, C., Muñoz, R., Rojas, C., Vásquez, S., Pérez-Santos, I., Soto-Mardones, L., 2021. Seasonal variability of SST fronts in the inner sea of Chiloé and its Adjacent Coastal ocean, Northern Patagonia. Remote Sens. 13 (2), http://dx.doi.org/10.3390/rs13020181, URL https://www.mdpi.com/2072-4292/13/2/181.

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. J. Geophys. Res.: Oceans 127 (3), e2021JC018138.

Shekhar, S., Jiang, Z., Ali, R.Y., Eftelioglu, E., Tang, X., Gunturi, V.M., Zhou, X., 2015. Spatiotemporal data mining: A computational perspective. ISPRS Int. J. Geo-Inf. 4 (4), 2306–2338.

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 Sens. 13 (6), http://dx.doi.org/10.3390/rs13061065, URL https://www.mdpi.com/2072-4292/13/6/1065.

Shi, Z., Pun-Cheng, L.S., 2019. Spatiotemporal data clustering: A survey of methods. ISPRS Int. J. Geo-Inf. 8 (3), http://dx.doi.org/10.3390/ijgi8030112, URL https://www.mdpi.com/2220-9964/8/3/112.

Siemer, J.P., Machín, F., González-Vega, A., Arrieta, J.M., Gutiérrez-Guerra, M.A., Pérez-Hernández, M.D., Vélez-Belchí, P., Hernández-Guerra, A., 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. J. Geophys. Res.: Oceans 126 (8), e2021JC017268. http://dx.doi.org/10.1029/2021JC017268, e2021JC017268 2021JC017268, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021JC017268, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021JC017268.

Wazarkar, S., Keshavamurthy, B.N., 2018. A survey on image data analysis through clustering techniques for real world applications. J. Vis. Commun. Image Represent. 55, 596–626.

Wessel, P., Luis, J.F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W.H., Tian, D., 2019. The generic mapping tools version 6. Geochem. Geophys. Geosyst. 20, 5556–5564. http://dx.doi.org/10.1029/2019GC008515.

Zakani, F.R., Arhid, K., Bouksim, M., Gadi, T., Aboulfatah, M., 2016. Kulczynski similarity index for objective evaluation of mesh segmentation algorithms. In: 2016 5th International Conference on Multimedia Computing and Systems. ICMCS, pp. 12–17. http://dx.doi.org/10.1109/ICMCS.2016.7905611.