Detecting and characterizing upwelling filaments in a numerical ocean model

Osvaldo Artal[1a], Héctor H. Sepúlveda[2b,*], Domingo Mery[3c], Christian Pieringer[4c]

[a] *Environmental Department, Aquaculture Research Division, Fisheries Development Institute (IFOP), Camino a Tenten S/N, Castro, Chile.*
[b] *University of Concepción, Geophysics Department, Barrio Universitario s/n, Casilla 160-C, Concepción, Chile.*
[c] *Pontifical Catholic University of Chile, Department of Computer Science, Av. Vicua Mackenna 4860 Macul, Santiago, Chile*

Abstract

Upwelling filaments are long (≈ 100's km) narrow (O ≈ 10 km) structures in the coastal ocean. They export nutrients and prevent the movement of larvae along the coast. Filaments can be observed in satellite images and in numerical models, but their manual identification and characterization is complex and time consuming. Here we present a Matlab code for a manual method to assist experts in this task, and a code for an automatic filament detection method (AFD) based on image processing and pattern recognition to identify and extract features in output files from a numerical ocean model. AFD was tested with a simulation of northern Chile. AFD had a similar performance in filament detection to that of human experts. AFD provides substantial time savings when analyzing a large number of images from a numerical ocean model. AFD is open source and freely available.

Keywords: Upwelling Filaments, Image Processing, Chile, Numerical Models, Coastal Ocean

1. Introduction

Ocean circulation patterns and their impact on marine ecosystems is a topic of interest to scientists (Bakun, 1996; Condie et al., 2005) that presents a technological challenge. This includes detecting features like eddies (Chaigneau et al., 2008), plumes (Chrysoulakis et al., 2005), and fronts and filaments (Cayula and Cornillon, 1992; Nieto et al., 2012) in results from remote sensors and numerical circulation models. These sources of information are

[1] Designed and wrote the code, analyzed data, and drafted the manuscript
[2] Designed the code, analyzed data, and drafted the manuscript
[3] Conception and design of code
[4] Conception and design of code

[*] Corresponding author
Email address: andres@dgeo.udec.cl (Héctor H. Sepúlveda)

often studied with the aid of image detection algorithms (Wang et al., 2010) or geoprocessing tools (Roberts et al., 2010).

Upwelling filaments are long (100-200 km) narrow (10-50 km) cold water structures that extend from the coast to the oceanic region. They have consistent characteristics of temperature and salinity, with a shallow vertical structure (100 m) (e.g. Sobarzo and Figueroa (2001)). They are strongly influenced by seasonal winds as a forcing factor, which associates the upwelling season with filament formation, with more and longer filaments during this period (Strub et al., 1991). Upwelling filaments are biologically important due to the volume of water with high concentrations of nutrients and chlorophyll-a (Chl-a) transported from the coastal zone to the open ocean (Jones et al., 1991; Álvarez-Salgado et al., 2001) due to their role in the advection of eggs and larvae of different marine species to the open ocean, preventing their transport along the coast (Rodríguez et al., 1999; Bécognée et al., 2009).

Because of their strong sea surface temperature (SST) and Chl-a signatures, filaments can be easily identified in remote sensing images Fonseca and Farias (1987); Cáceres (1992); Thomas (1999); Grob et al. (2003), or by direct measurement of water velocity (Marín and Delgado, 2007) and/or temperature and salinity profiles (Sobarzo and Figueroa, 2001). Satellite-based remote sensors are used to study large areas, generally with daily coverage, but clouds can limit the obtainable information about the surface structure of filaments (Thomas, 1999). Remotely sensed SST and Chl-a images only provide information about the surface characteristics of filaments, while in situ observations provide better understanding of the vertical structure of currents and density gradients affecting filaments (Flament et al., 1985; Navarro-Pérez and Barton, 1998; Barton et al., 2001). Observational studies provide detailed information about horizontal and vertical filament structure, but depend on the timing, duration, and frequency of hydrographic cruises, so generally describe only one or two filaments at a time (Álvarez-Salgado et al., 2001; Barton et al., 2001; Sobarzo and Figueroa, 2001).

There have been several studies of upwelling filaments in Chile (Sobarzo and Figueroa, 2001; Marín and Delgado, 2007; Morales et al., 2007; Letelier et al., 2009). Sobarzo and Figueroa (2001) presented the results of a research cruise in 1997 during which the vertical structure of an upwelling filament was registered near Mejillones Peninsula (23 °S) and described as ≈100 m deep and 165 km long. Marín and Delgado (2007) used Lagrangian drifters to study surface circulation between 23-30 °S and identified filaments over 295 km in length. In a study of 1,867 SST images recorded between 1987 and 1992 of northern Chile (18.3-24 °S), Barbieri et al. (1995) described upwelling filaments up to 222 km in length, with an average of 111 km. Sixty-three filaments were observed in this area between 1988 and 1990, mostly between the months of November and April. Most originated from specific coastal locations. Cloud cover seriously limits the study of major structures associated with coastal upwelling, with 54% of the images from coastal areas (defined as 71° W to the coast) and 83% from intermediate (defined as 71° W - 72° W) and oceanic areas (defined as 72° W - 73° W) were being classified as poor/bad for this purpose (Barbieri et al. (1995), Table 1).

Other studies have described the presence of filaments in numerical simulations of the Chilean coast (Escribano et al., 2004; Leth and Middleton, 2004). Parada et al. (2012) used

a climatological simulation of the region between 33 and 40 °S, with a spatial resolution of 5 km, combined with a module of numerical Lagrangian drifters. These studies describe a larger number of filaments in the austral summer, compared with austral winter, which is attributed to increased instability of coastal currents. In addition, the filaments were identified as presenting close and intense coastal currents that extended from the shore. The studies also concluded that these currents generally act as barriers to the latitudinal transport of passive particles like anchovies larvae. In some cases, the particles go back to the coast because of the 3D variability of the currents in the surface layers, including the change of orientation of the filaments.

In this context, numerical ocean simulations are useful for quantitatively characterizing the spatial dimensions of filaments (length, width, and depth). The use of numerical simulations requires proper validation of the results and a study of the sensitivity of models to changes in wind forcing, coastline or bathymetry. Satellite-based products like the Group for High Resolution Sea Surface Temperature (GHRSST) (Martin et al., 2012) and the results of numerical models (e.g. Oke et al. (2013)) are available to study the variability of upwelling filaments. However, manual analysis of large amounts of data is time-consuming and prone to error due to operators fatigue. There is also an inherent data interpretation error in which the expert changes the interpretation of what constitutes an upwelling filament. For this reason, it is useful to have computational tools to support the detection and characterization of filaments to reduce analysis time and provide consistency in the analysis.

This paper presents two computational tools developed in Matlab to identify and characterize upwelling filaments in output files from a numerical ocean model: a manual method and an automatic filament detection method (AFD). This work aims to complement the study of upwelling filaments using numerical schemes, automated identification (Eugenio and Marcello, 2009; Nieto et al., 2012; Cordeiro et al., 2015), and feature or particle tracking (Marín et al., 2003; Lett et al., 2008; Sayol et al., 2014; Otero et al., 2015) in ocean models. As an example of the application of the automatic method we present a characterization of filaments obtained from a numerical simulation in northern Chile (15-35 ° S). The article is divided as follows: Section 2 describes the two methods of filament identification, manual and AFD. Section 3 presents the validation of AFD against a panel of experts. Section 4 describes the main results regarding the variability of upwelling filaments on the Chilean coast. Finally, Section 5 discusses the main implications of this work.

2. Methodology

We developed two methods to identify filaments in output files from a numerical ocean model: one manual and the other automatic. The manual method consists of a graphic user interface (GUI) for manually labeling of filaments by clicking on a button to subsequently extract features automatically to a text file. The automatic method is based on image processing and pattern recognition techniques to automatically identify and characterize upwelling filaments. The user has to specify a series of criteria to define what constitutes an upwelling filament. Some of the factors to be considered in this decision are:

1. Sea surface current gradient.
2. Sea surface temperature gradient.
3. Filament length.
4. Persistence or duration of the filament.
5. Filament originating near the coast.

Since the numerical model used provided the results in the NetCDF format, we used Matlab's toolbox MexCDF (http://mexcdf.sourceforge.net/), to read the input file. Adapting this code to other programming languages like Octave should be fairly easy based on past experiences (Sepulveda et al., 2011), except for the GUI. This should be done with general purpose programming languages like Python.

2.1. Manual Method

The manual method is GUI that allows the user to label upwelling filaments manually by checking each image individually and clicking with the mouse at the start and end point of an upwelling filament. All the information relating to these clicks is stored in matrices that can then be processed efficiently. This method is adapted to read monthly outputs files in NetCDF from ROMS_AGRIF, with a daily temporal resolution. One can select a specific area to analyze, or study the entire model domain. The program reads the SST and surface speed magnitude for random time records, so that the filaments observed by the user in one image does not influence the users decision about the existence of filaments in the following image.

The GUI shows three images for every temporal record (Fig. 1): the horizontal temperature gradients considering 6 pixels, equivalent to 30 km, (Fig. 1a); the estimated surface temperature contours (Fig. 1b); and the contours of the magnitude of the surface current (Fig. 1c). Once an upwelling filament has been identified by the user by clicking on the start and end points, the algorithm calculates the start and end latitude and longitude, stores the year, month and day of the simulation, assigns an identification number to each filament number and calculates the length of each selection. The length is approximated using the Haversine formula (Robusto, 1957), which assumes a straight line between the start and end point considering the curvature of the Earth. To estimate the persistence of a filament, the coastal coordinates of its starting point (the closest pixel to the coast) are compared to those of the filaments selected the day before and the day after, within a spatial tolerance range (defined by the parameter rtol). If there is a coincidence, it is considered that the filament

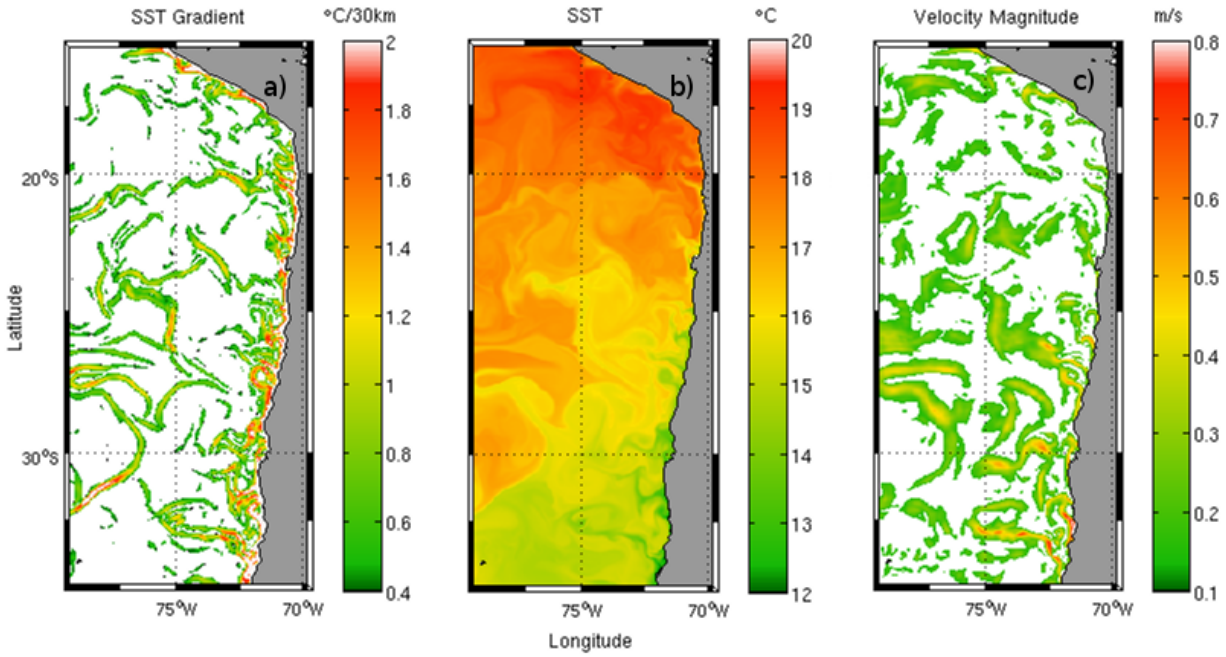

Figure 1: Example of the information used by the manual and automatic detection methods. Panel a) shows sea surface temperature (SST) gradients. Panel b) shows the SST field. Panel c) shows the magnitude of the surface current.

is the same. Once these values have been calculated, the selected filaments can be filtered by other user defined criteria (e.g. Table 1). Upwelling filaments in numerical models have been identified manually before (e.g. Cordeiro et al. (2015); Troupin et al. (2012)).

Magnitude of surface current	$\sqrt{U^2 + V^2} > 0.25 cm/s$
Minimum filament length	100 km
Minimum filament life (persistence)	3 days
Place of origin	First 50 km from the coast
rtol	1 pixel

Table 1: Criteria used to identify upwelling filaments. U and V are the east-west and north-south components of current velocity.

2.2. Automatic Method

The automatic filament detection method (AFD) is a processing algorithm to detect and characterize upwelling filaments in output files from a numerical ocean model. This algorithm applies techniques of digital image processing to highlight the object of study (filaments) and to calculate its main features: coordinate of origin, persistence, length, and direction. The algorithm is applied in this study to the results of a numerical simulation done with the ROMS_AGRIF ocean model which are stored in NetCDF. First, the input

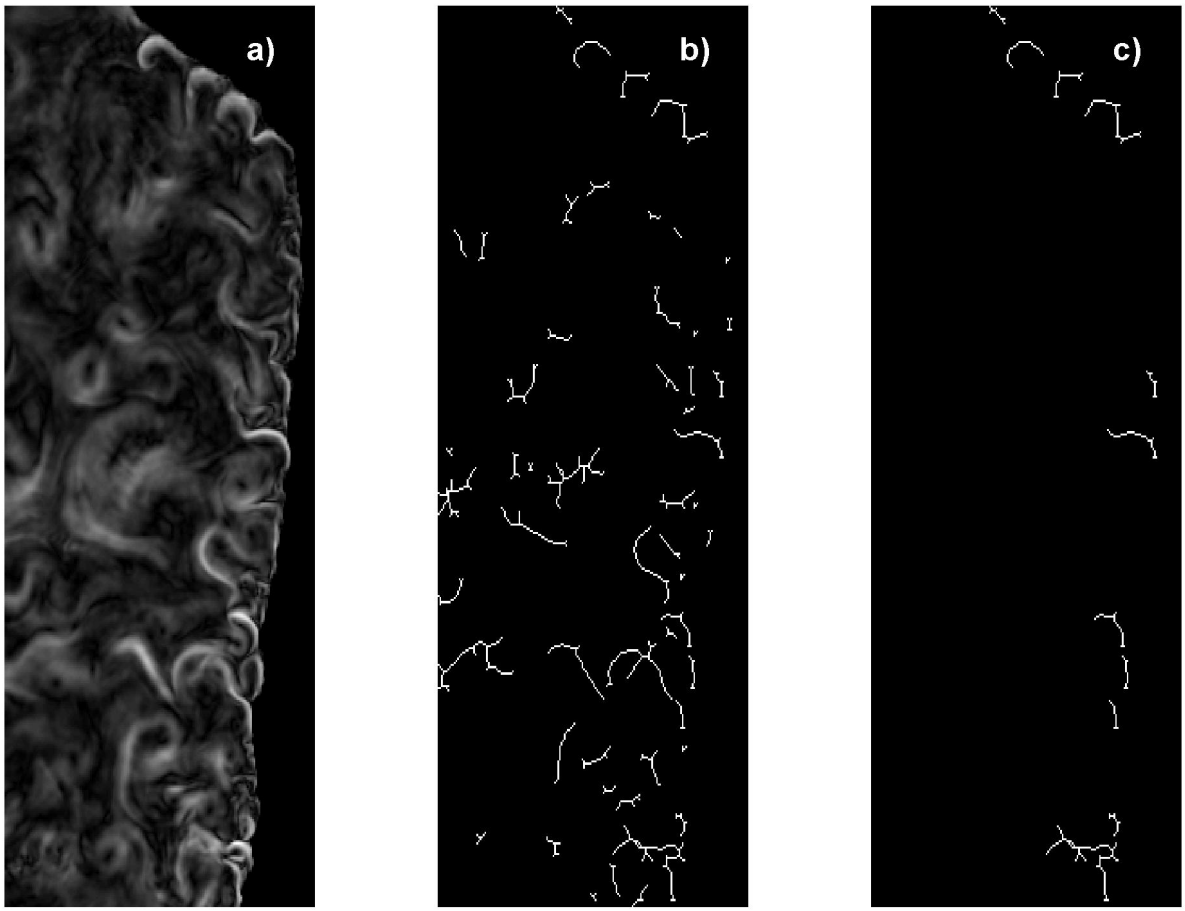

Figure 2: Example of automatic filament detection code working for an particular instant. Panel a) shows the magnitude of the surface velocity field in gray Panel b) shows the binary image (white/black) of the calculated morphological skeleton. Panel c) shows the segmentation of the skeleton, where we can observe the structures that were selected as filaments.

file is read and latitude, longitude, time, SST, and magnitude of the current, as well as the land/ocean mask are extracted. The later field is used to calculate the distance to the coast for each grid point. The code can be used to detect a filament using as a criteria SST gradient or a threshold value of surface velocity magnitude. The two methods will be combined in future versions of the code. In this study, the automatic method is applied to images of surface velocity magnitude, a dynamical indicator of the presence of an upwelling filament. Based on the classification proposed by Mery and Soto (2008), we summarize the organization of our algorithm in the following three steps.

1. **Image Acquisition**: A gray scale image is generated for the daily fields of surface current velocity, geographical coordinates, mask water/land and time. The algorithm applies the binary mask water/land use to ensure that only ocean data is represented. Grid points with current velocity magnitude values below the threshold are identified,

as defined in Table 1, transforming the grey scale image (Fig. 2a) into a binary image were each pixel has a 1 value if the threshold criteria is met, or a 0 value if it is not met. (image not shown).

2. **Pre-processing:** The binary image is modified to highlight filaments using elements of mathematical morphology. These morphological operations (Gonzalez et al., 2010) are used to retrieve the shape of the filaments. First a morphological opening is applied using a structuring element in the form of a 3x3 matrix that moves across the image to erode and then dilate the objects that compose it. A morphological skeleton is then generated by removing pixels at the boundaries of objects while avoiding breakage (Fig. 2b). A set of lines are obtained representing the full thinning of the region by maintaining the essential filament shape. The branches of the skeleton cut by only one pixel are joined by a "bridge" operation, that is, 1-0-1 pixel sequences are converted into 1-1-1 sequences. Finally, the option clean removes the pixels that have a single value of 1 and are surrounded by pixels of 0 value.
3. **Segmentation and Feature Extraction:** The skeletons resulting from the process above are analyzed individually. The result is a binary image where a pixel equal to 1 indicates membership in the region of interest and 0 indicates non-membership. Table 1 describes the criteria by which the algorithm evaluates the pixels belonging to the regions. If all the criteria are met, the algorithm classifies it as an upwelling filament (Fig. 2c) and obtains the main features from the objects: time, latitude and longitude of the origin of each filament, beginning and end time of the filament and its length. We use a matrix of distance to the coast to select the origin of each filament, which is taken as the pixel that is closest to the coastline. The length of a filament is calculated as the sum of all the pixels multiplied by the value of the spatial resolution of the model, this approach includes the length of all the branches of a filament in its total length. Once feature extraction was completed, we analyzed persistence over time in the same way as with the manual method: when two filaments have the same origin, defined as the closest pixel to the coast, within a spatial tolerance defined by rtol, they are considered as the same filament, regardless if the shape or size has changed.

A flow diagram of the automatic detection is shown in Fig. 3. To use it, the user has to configure the *ftd_param.m* file and call the *ftd_automatic.m* function to process the file defined in *ftd_param.m* without further intervention. A call to *ftd_assisted.m* launches the GUI for the manual method.

2.3. Ocean Model Configuration

The numerical ocean model ROMS_AGRIF, (Shchepetkin and McWilliams, 2005; Penven et al., 2005) was used to simulate the climatological ocean circulation in northern Chile, between 15-35 S and 69-80 W, with a horizontal spatial resolution of 1/20th of a degree (≈5 km), resulting in a matrix of 436x198 pixels. Several packages have been developed for this particular ocean model such as ROMSTOOLS, which is focused on pre- and post-processing of the model simulations (Penven et al., 2008), and LiveROMS, which facilitates installation,

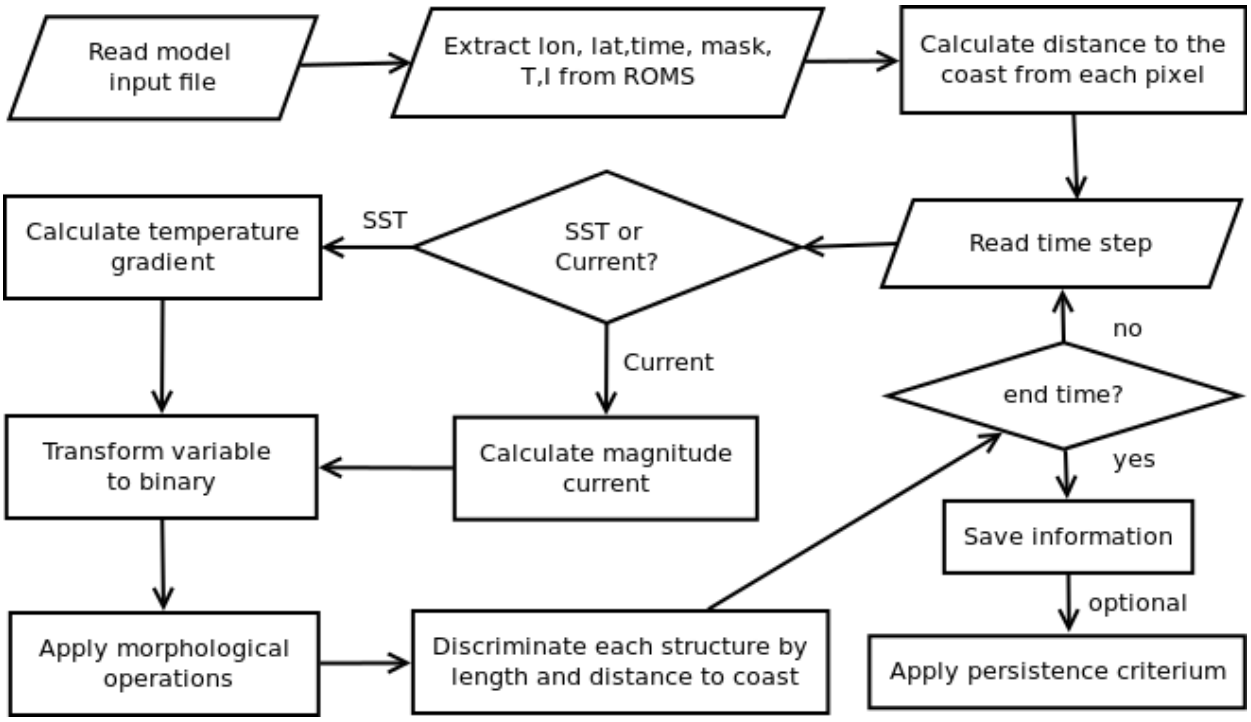

Figure 3: Flow diagram for the automatic detection code for each time step.

compilation and use (Sepulveda et al., 2011). The model was imposed on the surface with wind data, heat fluxes, and fresh water from COADS05 (Da Silva et al., 1994). Bathymetry was obtained from the ETOPO2 data base (Smith and Sandwell, 1997). The boundary and initial conditions were derived from the WOA2001 database (Conkright et al., 2001). In the vertical, we implemented 32 sigma layers, which are terrain-following divisions of the ocean depths, which were distributed so that there was a higher concentration of sigma layer near the top to properly resolve circulation. A sponge layer, a 50-km-wide region along the open boundaries, with artificially increased viscosity, was used to avoid the reflection of energy (waves) leaving the domain. We ran the simulation for 10 climatological years, saving the results in daily time steps. The last 4 years (7-10) of the simulation were analyzed, representing a total of 1440 images.

3. Automatic Method Validation

An important challenge in the classification of filaments is interpreting model results. This was evident when we asked experts with experience in ocean sciences to identify upwelling filaments. Some experts look at temperature gradients, others focus on the dynamic characteristic of surface currents in upwelling filaments, while still others consider both aspects important.

This generates different interpretations among experts on what constitutes an upwelling filament. For this study, we asked over 20 experts with experience in ocean sciences to identify upwelling filaments. A sample image similar to Figure 1 was sent to them. We

received replies from 8 experts, including an author of this article, and we sent 10 randomly selected images from the results provided by the numerical model (Table 2). The experts were asked to draw or circle the upwelling filaments in the images.

Image 1	Month 4, Day 4	Image 6	Month 7, Day 20
Image 2	Month 9, Day 6	Image 7	Month 2, Day 20
Image 3	Month 11, Day 19	Image 8	Month 5, Day 11
Image 4	Month 2, Day 8	Image 9	Month 9, Day 20
Image 5	Month 11, Day 10	Image 10	Month 10, Day 8

Table 2: Month and Day of the images studied by the experts.

We also applied the automatic detection algorithm to the same images evaluated by experts. The automatic method detected 117 filaments in the 10 images. The 8 experts detected between 32 (Expert 1) and 173 filaments (Expert 7) in the 10 images, with a median of 120 (Fig. 4). Most experts also identified short filaments since it was difficult to evaluate a filament's length due to the absence of a length scale in the images we provided. To have comparable results when the automatic method was applied, the filament's length was not considered as a filtering criteria. Thus, these results include filaments less than 100 km long. Hereafter we compare our results with those of the group of 5 experts that were more consistently in agreement with each other, thus excluding 3 experts (1, 6 and 7). A median of 116 filaments, with a standard deviation of 6.6, was obtained from the group of 5 experts, who identified between 8 and 17 filaments in each image, with a mode of 11 filaments being identified (not shown). Images 2 and 9 had the fewest filaments (average of 10.1 and 10.3 filaments among experts, respectively) and both were from the austral winter. Images 4 and 7 had the most filaments (both with an average of 13.8 filaments among experts). These images were from the austral summer. Some 90% of the analyzed images (10 images, 5 experts) had 11 to 13 filaments, with an average of $\approx$12 (12 $\pm$ 1 filaments with 90% certainty).

Following Chaigneau et al. (2008), we used three indicators to evaluate the results of the AFD method and those of the five selected experts. In the following, the AFD method is analyzed together with the experts, as if the method were another expert. The interpretations of every expert were defined as the truth, and the interpretations of the other experts were compared against this truth to calculate the following indicators: successful detection rate (SDR), undetected filaments (UDF) and excess number of detected filaments (ENDF).

$$SDR = Nc/Nt * 100 \quad (1)$$

$$UDF = Nt - Nc \quad (2)$$

$$ENDF = Ne - Nc \quad (3)$$

Nt corresponds to the total number of filaments detected in the true solution, Ne is the number of filaments detected by another expert, and Nc is the number of filaments in common between the truth and the experts. Nc was calculated by visually inspecting the

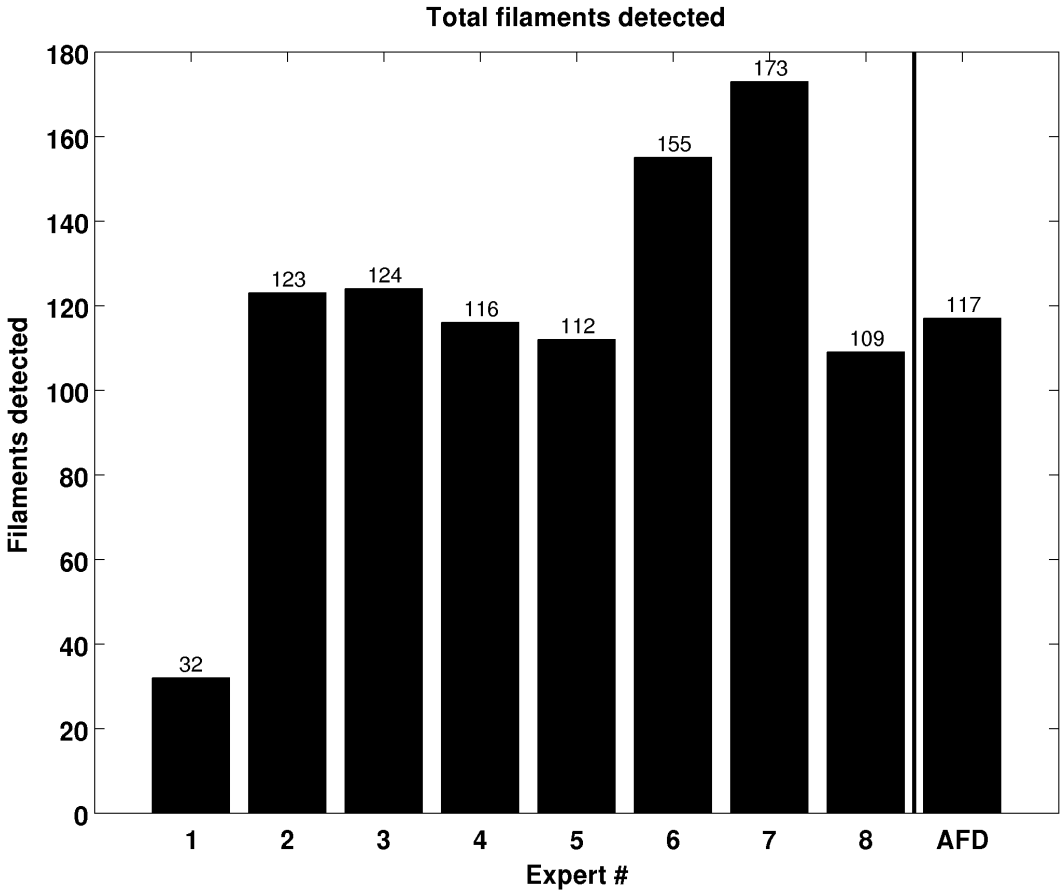

Figure 4: Histogram with the total number of upwelling filaments detected, manually by the experts, in 10 randomly selected images. AFD column shows the results for the automatic filament detection method.

filaments selected as true and the filaments selected by the other experts. Nc increased when the same object was selected in both images.

The results of SDR, UDF and ENDF obtained for the group of 5 experts and the AFD are presented as boxplots (Fig. 5) that were calculated considering 5 values. Although there are slight differences in the SDR of the 5 experts, these are not significant (Fig. 5a). SDR rates from experts 2, 3, 4, 5, 8 and AFD have a median value of 82.5%, 85%, 88%, 90%, 91% and 77%, respectively. Experts 5 and 8 obtained the highest median values of successful detection (Fig. 5a), indicating that the filaments detected by these experts were also observed by the other experts (including the AFD).

The AFD method is consistent with the experts, the difference being 1 or 2 undetected or over-detected filaments, which is equivalent to a 10-20% of the average number of filaments presented in the images. These results are confirmed by the median values for UDF and ENDF rates, which vary between 1 and 3 filaments (Fig. 5b, Fig. 5c). A study of SDR values for individual images used in the validation process (data not shown) reveals that Image 2 (Month 9, Day 6) presented more difficulties for the automatic method, since the median

value of its SDR was 63%. In contrast, results from Image 1 (Month 4, Day 4), presents more consistent results, with a median of 91%. Image 2 also presented major difficulties for the experts, where their SDR median was 79.5%.

The SDRs are low for Image 2 (Month 9, Day 6), but not for Image 9 (Month 9, Day 20). Conversely, the SDRs for Image 1 (Month 4, Day 4) are high, while they are not for Image 10 (Month 4, Day 8) (data not shown). This indicates that the difficulties in filament detection are not associated with the climatological month being studied. In a study using an eddy detection algorithm, Chaigneau et al. (2008) obtained a SDR of 92.7% and an excess of detection 18.7% for their proposed method, A widely used automatic method based on the Okubo-Weiss criteria applied to the same dataset obtained a SDR of 86.8% and an excess of detection of 63.3% (Chaigneau et al. (2008), Fig. 2). Other studies in automatic eddy detection (Nencioli et al. (2010), Table 1) present SDRs ranging from 85% to 100% when analyzing 10 images, and excess detection ranges between 0%-10%.

4. Temporal and spatial variability of upwelling filaments

Given the consistency between the AFD method and the 5 selected experts, we used this automatic approach to study 1,440 images from the climatological simulation of the Chilean coast described in Section 2.3. The criteria to define a filament are presented in Table 1. The AFD method required one minute per month to extract the characteristics of the filaments (30 images, daily output) on a machine with Ubuntu 10.04 64-bit OS (R) Intel (R) Pentium Dual CPU E2180 2.00 GHz and 4 GB of RAM, while it took an expert an average of an hour to do the same analysis, including writing down all the information. This could be reduced to 15 minutes per month by an expert if the GUI was used for manual selection.

First we describe here the characteristics of an arbitrarily selected a filament observed on November 8 (climatological month). The filament selected had a lifespan of 11 days. It was located between 71-76 W and 32-34 S. Observing the SST (Fig. 6a), we can see that the filament extended from the coast ($\approx$ 72 W) seaward past 74 W, with temperatures $\approx$ 2 °C colder than in surrounding waters. Surface currents along the filament are markedly larger than in the surrounding waters. These higher velocities and the offshore orientation of the filament suggests the offshore transport of nutrients and larvae by filaments (Fig. 6b). A vertical section along the filament (figure not shown) indicates that it has an impact down to a depth of 100 m, as seen in observational studies (Sobarzo and Figueroa, 2001).

The AFD analysis of the model results shows that filaments are generated across the coast of Chile, with at least one filament per degree of latitude. For our analyses we gridded the results by one degree of latitude. The areas with the largest number of filaments are 17 °S (45 filaments), 29 °S, and 30 °S (43 filaments each), with over 10 filaments per year (Fig. 7a). Barbieri et al. (1995) (Table 8) observed 69 filaments during a 3-year period (1988-1990) for the coastal region between 18.3-24 °S. For this same region 106 filaments were observed in a 4-year period, approximately 17% more per year. This difference could be due to a combination of three factors: a) the effect of cloud cover, which in the study of Barbieri et al. (1995) is mentioned as 33% for coastal regions; b) detection issues with the ADF method; or c) incorrect representation of upwelling filaments in the numerical model.

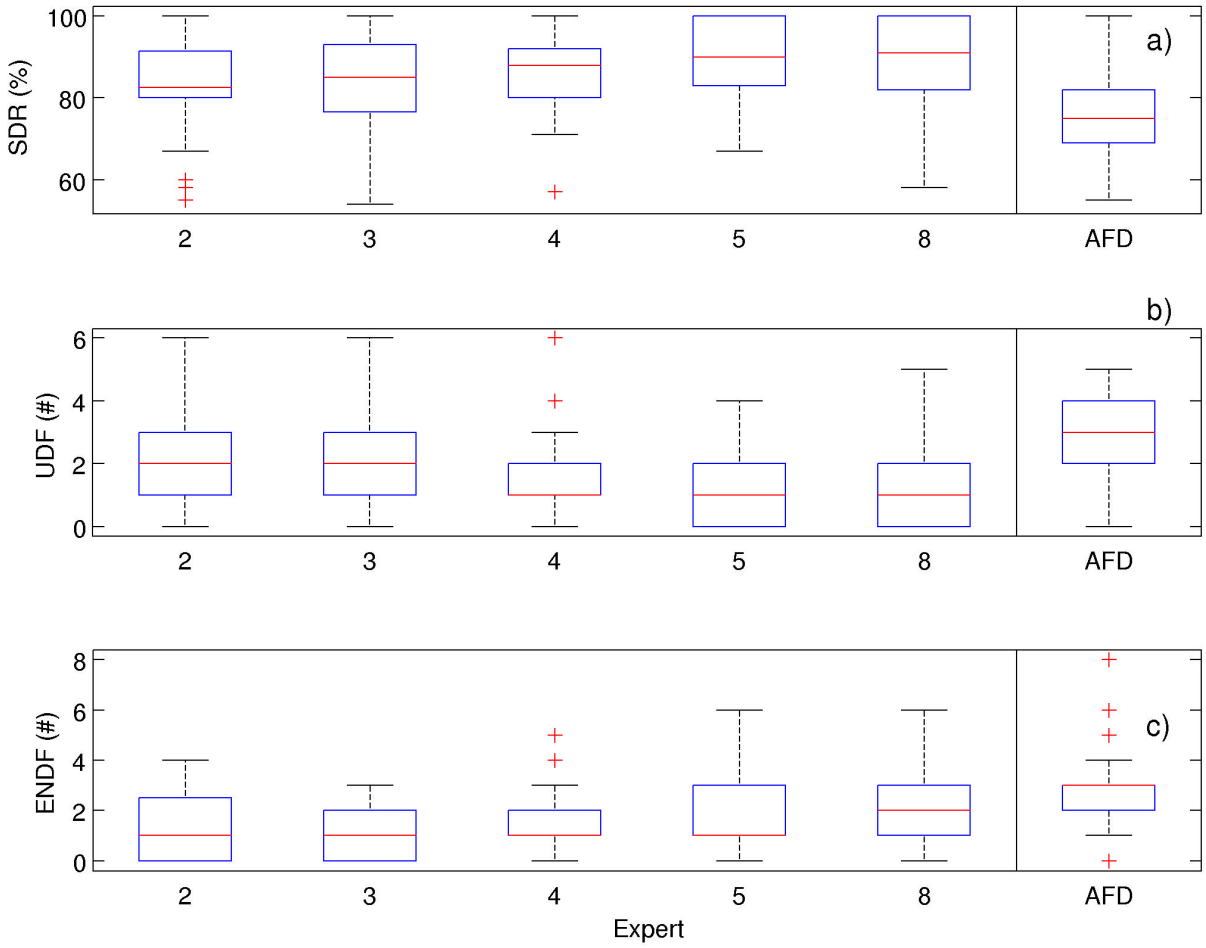

Figure 5: Variability of the identification indexes. The columns show the results when the interpretation of an expert was considered as the true value. The upper panel shows the Success Detection Rate (SDF), middle panel the number of Undetected Filaments (UDF) and the lower panel the Excess Number of Detected Filaments (ENDF). Each boxplot represents 50 evaluations of an experts criteria.

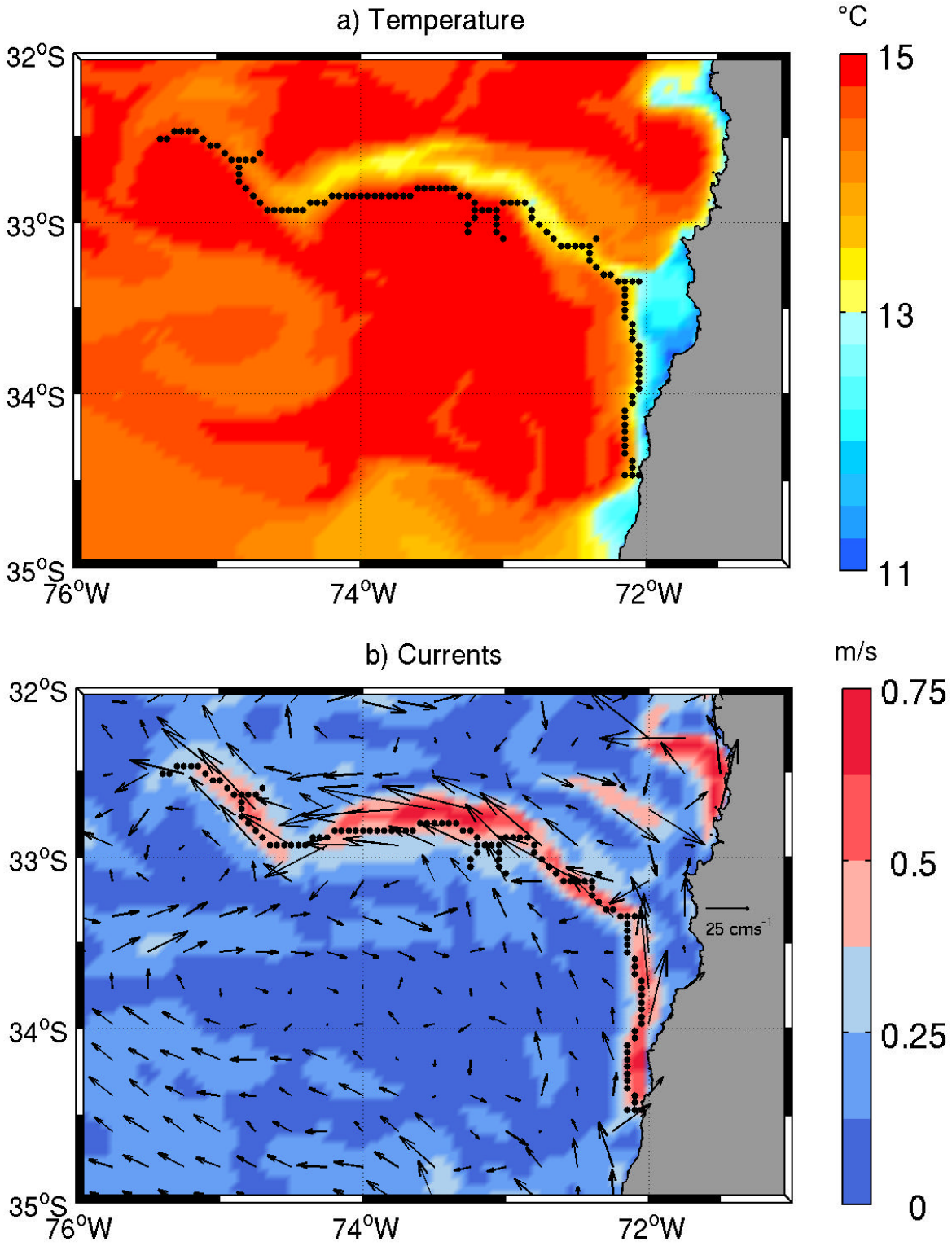

Figure 6: Example of an upwelling filament that was detected by the automatic method. Panel a) shows the sea surface temperature. Panel b) shows the magnitude of the surface current.

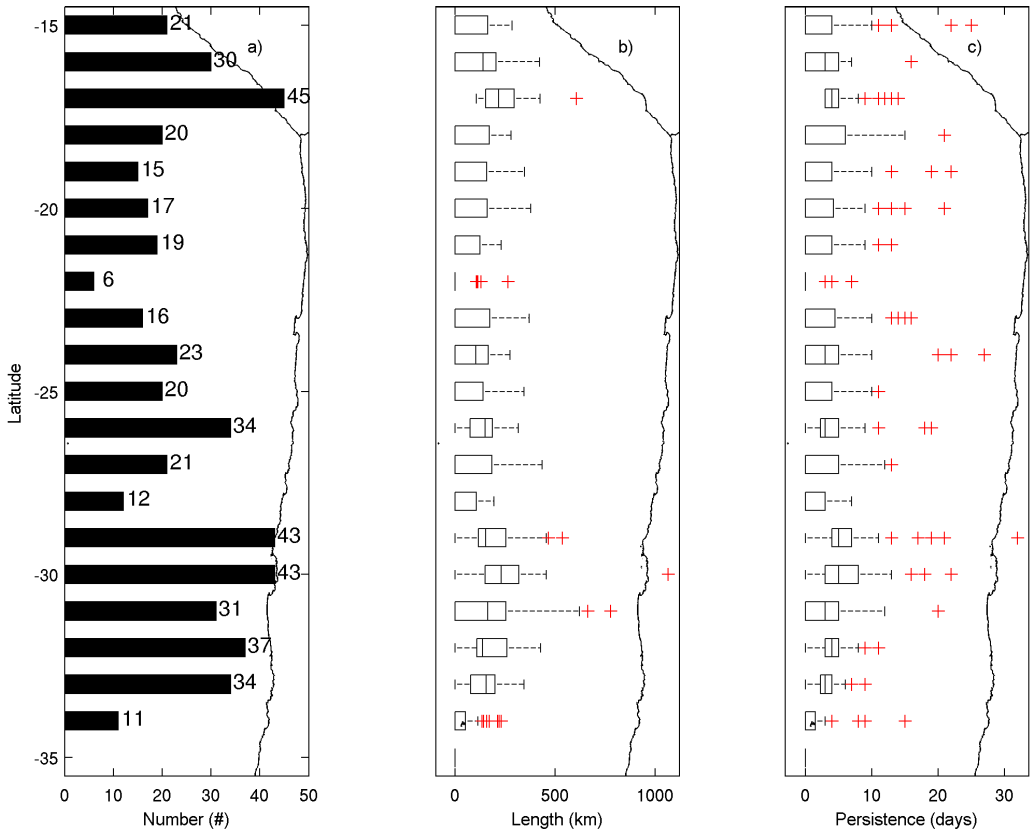

Figure 7: Spatial distribution of upwelling filaments, in bins of 1° of latitude detected using the automated method considering 4 years of a climatological simulation. Panel a) shows the total number of filaments, Panel b) shows a box plot of the length of the filaments (km). Panel c) shows a box plot of the persistence of the filaments (days).

A realistic hindcast of ocean circulation for this period could be used to further elucidate the origin of such differences. In terms of filament length, in this study, 77.1% (384) of the filaments were less than 250 km long and 34.7% (173) were less than 150 km long.

The median length of the filaments was 178 km, longer than the average of 111 km observed by Barbieri et al. (1995) in northern Chile. An overestimation of a filament's length is expected as it includes the length of all its branches. We observed longer filaments at 17 °S and 30 °S, with median lengths of 217 and 229 km, respectively (Fig. 7b). The longest recorded filament measured 1065 km, which was located at 30 S. The median persistence of a filament was 4 days (Fig. 7c), while 86% had lifespans of 10 days or less, comparable to the 6 days Barbieri et al. (1995) described for a filament in February 1989. The longest-lasting filament continued over 32 days, and was located at 29 S. Figure 8a shows information on the temporal variability of the filaments. We note that the number of filaments has a marked annual cycle, with peaks in spring-summer and autumn-winter lows. From December to

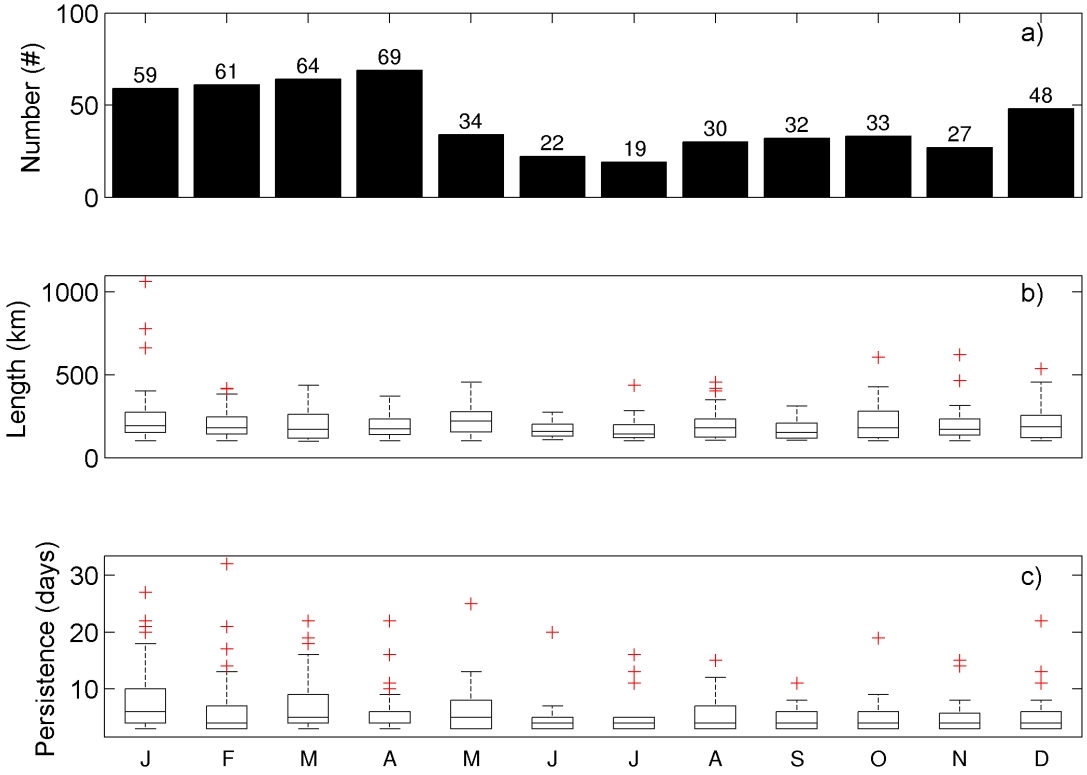

Figure 8: Temporal variability of the filaments detected using the automated method considering 4 years of a climatological simulation in all study area. Panel a) shows the number of filaments per month. Panel b) shows a box plot of the length of the filaments (km). Panel c) shows a box plot of the persistence at each month (days).

April of the four year of the study, there were more than 45 filaments every month, with a maximum of 69 in April, while between May and November, there were fewer than 35 filaments per month, with the fewest (19) in July. This annual cycle was expected by the association of filaments with coastal upwelling. In comparison, Barbieri et al. (1995) reports that upwelling filaments were numerous between November and April. Figures 8b and 8c show the temporal variability in length and persistence, respectively. In months with fewer filaments, there were some filaments with lengths and persistence above the average values. A non-parametric Kruskal-Wallis test was applied to compare the median values of length and persistence and indicated that persistence was comparable for all months ($p = 0.004$), while there was variability in median length ($p = 0.16$).

4.1. Sensitivity Study

We conducted a sensitivity study based on changes in surface velocity magnitude and the SST gradient (Table 3). The AFD method was applied in the study area to analyze four years of daily results (1440 images). In this case, a detected structure was considered a filament if it was longer than 100 km, the distance to the coast was less than 15 pixels (75 km), and if it persisted more than 3 days. The total number of detected filaments (N), the length (mean value and standard deviation), and persistence, in days (mean value and standard deviation) were calculated for changes in the threshold for the velocity magnitude [m/s] and for the SST gradient/30 km [°C] (Table 3). In comparison, Barbieri et al. (1995) reported SST gradients from images with a range of 1.6-9.6 °C/30 km, with the higher values obtained in summer and the lower ones in winter. A value of 2.25 °C/30 km (reported as 0.075 °C/km) was used by Cordeiro et al. (2015) to manually identify upwelling filaments from a numerical simulation of the Iberian Peninsula. These results show that the detection algorithm is highly sensitive to the choice of threshold values and that results should be validated before further analysis. In this study, the velocity magnitude was selected as the key parameter to determine whether a structure of interest was present, and the reference value was chosen after a bibliographic research of the area. As noted in Table 3, if we had used the SST gradient as a detection criteria, while length and persistence are similar, the total number of filaments would have been underestimated by almost half. Auxiliary plotting functions are included with this code in order to superimpose the detected structures over the SST, SST gradient, and velocity magnitude contour maps from the model results. These plots can be used to visually inspect a large number of images and quickly establish if the threshold parameters selected are capturing the structures of interest.

Parameter	**Value**	**N**	**Length [km]**	**Persistence [days]**
Velocity magnitude	0,25	74	213± 115	4.4 ± 2.8
Velocity magnitude	0,30	59	184± 76	4.4 ± 2.1
Velocity magnitude	0,50	1	231	3.0
SST gradient/30km	2,00	44	170± 67	4.6 ± 2.6
SST gradient/30km	2,50	17	169± 73	4.2 ± 1.8
SST gradient/30km	3,00	5	128± 11	3.2 ± 0.4

Table 3: Sensitivity study of filaments detected in the study area over 4 years (1,440 images). Results for total number (N), length (mean value and standard deviation), and persistence, in days (mean value and standard deviation) are presented for changes in the threshold value of the velocity magnitude [m/s] and the SST gradient/30 km [° C]

The automatic code is sensitive also to the rtol parameter. As mentioned in the Methodology section, to estimate the persistence of a filament, the coordinates of its starting point (the closest pixel to the coast) are compared to those of the filaments selected the day before and the day after, within a spatial tolerance range (defined by the parameter rtol). If there is a coincidence, it is considered that it is the same filament. By increasing rtol to 2 pixels, the total number of filaments increases by 110% in the months from May to October and by 49% in the remaining months (contrary to what one would expect from upwelling

dynamics). By considering a value of 0 for rtol, a $\approx 90\%$ decrease in the total number of filaments for all months was observed. As in the study of eddies in numerical models, the detection and the identification of two features as the same object are two different tasks, and improvements are needed in both areas.

It is recommended to select several random images from model results and compare what is obtained using the manual method. The initial set of parameters can be selected from the available literature in the area of interest and then compared with the results obtained by manual identification. Statistical indicators previously discussed can be used for guidance in the selection of new parameters. Once a satisfactory level of identification is achieved, use the automatic code to process several years of model results.

5. Conclusions

Filaments are oceanic structures along the coast in upwelling systems and play an important role in the export of nutrients to the open ocean and in the dispersion of larvae. Here we implemented a manual and an automatic detection algorithm of upwelling filaments in a numerical ocean model. The manual algorithm is a GUI that supports data extraction by an expert. It calculates properties like origin, linear length, and persistence. The automatic algorithm detects and characterizes upwelling filaments. Both methods were adapted to read results from the numerical ocean model known as ROMS_AGRIF, but are easily adaptable to other ocean models. The automatic method is based on pattern recognition techniques in images and presents significant savings in processing time over the manual detection using the GUI. The user can define different criteria of what constitutes an upwelling filament. The results obtained by the automatic method were compared with the analysis by 5 human experts. In the study of 10 images, the automatic method presented a median success detection rate of 77% compared to the filaments identified by the experts. The method over-detected a median of 3 filaments ($\approx$30%), while the experts over-detected 1-2 filaments. Detection algorithms applied to oceanic eddies have yielded comparable overdetection results (Chaigneau et al., 2008; Nencioli et al., 2010). While these results show the AFD method can obtain reasonable results, the method is highly sensitive to the threshold and selection parameter used (Table tab:sens), thus the results should be validated visually using the GUI. A plotting function to overlap the selected structures was developed and is available.

Analyzing a climatological simulation of the ocean in northern Chile, we found that the region between 29-30 °S had the largest number of detected filaments. Most filaments were less than 250 km long and lasted 4 to 6 days. The filaments presented an annual cycle, with peaks in the austral spring-summer and a lower presence in the austral autumn-winter. Our results indicate that the proposed method can be used to automatically detect upwelling filaments, and could be another tool in the complex task of evaluating the performance of ocean models.

6. Acknowledgments

This project was funded by Fondecyt 11080245 (HHS). We thank the experts who manually identified the filaments: Alexis Chaigneau, Danilo Calliari, Dante Figueroa, Elias Ovalle, Francesco Nencioli, James Pringle, Yazmina Olmos, Daniel Brieva, and Lorenzo Luengo. Comments by three anonymous reviewers were very helpful to improve our manuscript.

7. Computer Code Availability:

Name of code: FTD. **Developer:** Osvaldo Artal. **Contact address:** Environmental Department, Aquaculture Research Division, Fisheries Development Institute (IFOP), Camino a Tenten S/N, Castro, Chile. **Telephone number:** +56-33-3311369. **e-mail.** osvaldo.artal@ifop.cl. **Year first available:** 2016. **Hardware required:** Celeron CPU or better. **Software required:** Matlab, and image toolbox, or Octave ver 4.0.0 with toolboxes image (v2.4.1) and octcdf (v1.1.8) **Program language:** Matlab. **Program size:** 16 MB. **How to access the source code:** Available at:
https://github.com/oartal/FilamentDetection.

AFD	Automatic Filament Detection
SST	Sea Surface Temperature
Chl-a	Chlorophyll–a
GUI	Graphical User Interface
ROMS_AGRIF	A numerical ocean circulation model
NetCDF	Network Common Data Format
SDR	Successful Detection Rate
UDF	Undetected Filaments
ENDF	Excess Number of Detected Filaments

Table .4: Abbreviations used in this manuscript

8. References

References

Álvarez-Salgado, X. A., Doval, M. D., Borges, A. V., Joint, I., Frankignoulle, M., Woodward, E. M. S., Figueiras, F. G., nov 2001. Off-shelf fluxes of labile materials by an upwelling filament in the nw iberian upwelling system. Prog. Oceanogr. 51 (2-4), 321–337.

Bakun, A., dec 1996. Patterns in the ocean: Ocean processes and marine population dynamics. Calif. Sea Grant, La Jolla, CA, Cent. Investig. Biol. del Noroeste, La Paz, BCS, Mex. 17 (3), 1945–1946.

Barbieri, M., Bravo, M., Farías, M., González, A., Pizarro, O., Yáñez, E., 1995. Fenómenos asociados a la estructura térmica superficial del mar observados a través de imágenes satelitales en la zona norte de Chile. Investig. Mar. 23, 99–122.

Barton, E. D., Inall, M. E., Sherwin, T. J., Torres, R., nov 2001. Vertical structure, turbulent mixing and fluxes during Lagrangian observations of an upwelling filament system off Northwest Iberia. Prog. Oceanogr. 51 (2-4), 249–267.

Bécognée, P., Moyano, M., Almeida, C., Rodríguez, J. M., Fraile-Nuez, E., Hernández-Guerra, A., Hernández-León, S., mar 2009. Mesoscale distribution of clupeoid larvae in an upwelling filament trapped by a quasi-permanent cyclonic eddy off Northwest Africa. Deep. Res. Part I Oceanogr. Res. Pap. 56 (3), 330–343.

Cáceres, M. A., 1992. Vortices y filamenos observados en imagenes de satelite frente al area de surgencia de Talcahuano, Chile central. Investig. Pesq. 37, 55–66.

Cayula, J.-F., Cornillon, P., feb 1992. Edge Detection Algorithm for SST Images. J. Atmos. Ocean. Technol. 9 (1), 67–80.

Chaigneau, A., Gizolme, A., Grados, C., oct 2008. Mesoscale eddies off Peru in altimeter records: Identification algorithms and eddy spatio-temporal patterns. Prog. Oceanogr. 79 (2-4), 106–119.

Chrysoulakis, N., Adaktylou, N., Cartalis, C., dec 2005. Detecting and monitoring plumes caused by major industrial accidents with JPLUME, a new software tool for low-resolution image analysis. Environ. Model. Softw. 20 (12), 1486–1494.

Condie, S. A., Waring, J., Mansbridge, J. V., Cahill, M. L., sep 2005. Marine connectivity patterns around the Australian continent. Environ. Model. Softw. 20 (9), 1149–1157.

Conkright, M., Locarnini, H., Garcia, T., Brien, O., Boyer, T., Stephens, C., Antonov, J., 2001. World Ocean Atlas 2001: Objective analyses, data statistics, and figures, cd-rom documentation. Tech. rep., National Oceanographic Data Center, Silver Spring, MD.

Cordeiro, N. G. F., Nolasco, R., Cordeiro-Pires, A., Barton, E. D., Dubert, J., aug 2015. Filaments on the Western Iberian Margin: A modeling study. J. Geophys. Res. Ocean. 120 (8), 5400–5416.

Da Silva, A., Young, C., Levitus, S., 1994. Atlas of surface marine data. vol. 1, algorithms and procedures. Tech. rep., U.S. Department of Commerce, NOAA Department of Commerce, NOAA, Silver Spring, MA.

Escribano, R., Rosales, S. A., Blanco, J. L., jan 2004. Understanding upwelling circulation off Antofagasta (northern Chile): A three-dimensional numerical-modeling approach. Cont. Shelf Res. 24 (1), 37–53.

Eugenio, F., Marcello, J., sep 2009. Featured-based algorithm for the automated registration of multisensorial/multitemporal oceanographic satellite imagery. Algorithms 2 (3), 1087–1104.

Flament, P., Armi, L., Washburn, L., 1985. The Evolving Structure of an Upwelling Filament. J. Geophys. Res. 90 (C6), 11765–11778.

Fonseca, T., Farias, M., 1987. Estudio del proceso de surgencia en la costa chilena utilizando percepción remota. Investig. Pesq. 34, 33–46.

Gonzalez, R., Woods, R., Eddins, S., 2010. Digital image processing using MATLAB. Tata McGraw-Hill Education.

Grob, C., Quiñones, R. A., Figueroa, D., 2003. Cuantificacion del transporte de agua costa-oceano a traves de filamentos y remolinos ricos en clorofila a, en la zona centro-sur de Chile (37.5-37.5 S). Gayana 67 (1), 55–67.

Jones, B. H., Mooers, C. N. K., Rienecker, M. M., Stanton, T., Washburn, L., dec 1991. Chemical and biological structure and transport of a cool filament associated with a jet-eddy system off northern California in July 1986 (OPTOMA21). J. Geophys. Res. 96 (C12), 22207–22225.

Letelier, J., Pizarro, O., Nuñez, S., dec 2009. Seasonal variability of coastal upwelling and the upwelling front off central Chile. J. Geophys. Res. Ocean. 114 (12), C12009.

Leth, O., Middleton, J. F., dec 2004. A mechanism for enhanced upwelling off central Chile: Eddy advection. J. Geophys. Res. C Ocean. 109 (12), 1–17.

Lett, C., Verley, P., Mullon, C., Parada, C., Brochier, T., Penven, P., Blanke, B., sep 2008. A Lagrangian tool for modelling ichthyoplankton dynamics. Environ. Model. Softw. 23 (9), 1210–1214.

Marín, V. H., Delgado, L., Luna-Jorquera, G., dec 2003. S-chlorophyll squirts at 30S off the Chilean coast (eastern South Pacific): Feature-tracking analysis. J. Geophys. Res. 108 (C12), 3378.

Marín, V. H., Delgado, L. E., mar 2007. Lagrangian observations of surface coastal flows North of 30 S in the Humboldt Current system. Cont. Shelf Res. 27 (6), 731–743.

Martin, M., Dash, P., Ignatov, A., Banzon, V., Beggs, H., Brasnett, B., Cayula, J.-F., Cummings, J., Donlon, C., Gentemann, C., Grumbine, R., Ishizaki, S., Maturi, E., Reynolds, R. W., Roberts-Jones, J., nov 2012. Group for High Resolution Sea Surface temperature (GHRSST) analysis fields inter-comparisons. Part 1: A GHRSST multi-product ensemble (GMPE). Deep Sea Res. Part II Top. Stud. Oceanogr. 77-80, 21–30.

Mery, D., Soto, A., 2008. Features: The More The Better. In: 7th WSEAS Int. Conf. Signal Process. Comput. Geom. Artif. Vis. pp. 20–22.

Morales, C., González, H., Hormazabal, S., Yuras, G., Letelier, J., Castro, L., nov 2007. The distribution of chlorophyll-a and dominant planktonic components in the coastal transition zone off Concepción, central Chile, during different oceanographic conditions. Prog. Oceanogr. 75 (3), 452–469.

Navarro-Pérez, E., Barton, E. D., jun 1998. The physical structure of an upwelling filament off the North-West African coast during August 1993. South African J. Mar. Sci. 19 (1), 61–73.

Nencioli, F., Dong, C., Dickey, T., Washburn, L., McWilliams, J. C., Nencioli, F., Dong, C., Dickey, T., Washburn, L., McWilliams, J. C., mar 2010. A Vector GeometryBased Eddy Detection Algorithm and Its Application to a High-Resolution Numerical Model Product and High-Frequency Radar Surface Velocities in the Southern California Bight. J. Atmos. Ocean. Technol. 27 (3), 564–579.

Nieto, K., Demarcq, H., McClatchie, S., aug 2012. Mesoscale frontal structures in the Canary Upwelling System: New front and filament detection algorithms applied to spatial and temporal patterns. Remote Sens. Environ. 123, 339–346.

Oke, P. R., Griffin, D. A., Schiller, A., Matear, R. J., Fiedler, R., Mansbridge, J., Lenton, A., Cahill, M., Chamberlain, M. A., Ridgway, K., 2013. Evaluation of a near-global eddy-resolving ocean model. Geosci. Model Dev 6, 591–615.

Otero, P., Banas, N., Ruiz-Villarreal, M., apr 2015. A surface ocean trajectories visualization tool and its initial application to the Galician coast. Environ. Model. Softw. 66, 12–16.

Parada, C., Colas, F., Soto-Mendoza, S., Castro, L., jan 2012. Effects of seasonal variability in across- and alongshore transport of anchoveta (Engraulis ringens) larvae on model-based pre-recruitment indices off

481 central Chile. Prog. Oceanogr. 92, 192–205.
482 Penven, P., Echevin, V., Pasapera, J., Colas, F., Tam, J., oct 2005. Average circulation, seasonal cycle,
483 and mesoscale dynamics of the Peru Current System: A modeling approach. J. Geophys. Res. 110 (C10),
484 C10021.
485 Penven, P., Marchesiello, P., Debreu, L., Lefevre, J., may 2008. Software tools for pre- and post-processing
486 of oceanic regional simulations. Environ. Model. Softw. 23 (5), 660–662.
487 Roberts, J., Best, B., Dunn, D., Treml, E. A., Halpin, P., oct 2010. Marine Geospatial Ecology Tools:
488 An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++.
489 Environ. Model. Softw. 25 (10), 1197–1207.
490 Robusto, C. C., jan 1957. The Cosine-Haversine Formula. Am. Math. Mon. 64 (1), 38.
491 Rodríguez, J., Hernández-León, S., Barton, E., nov 1999. Mesoscale distribution of fish larvae in relation to
492 an upwelling filament off Northwest Africa. Deep Sea Res. Part I Oceanogr. Res. Pap. 46 (11), 1969–1984.
493 Sayol, J., Orfila, A., Simarro, G., Conti, D., Renault, L., Molcard, A., feb 2014. A Lagrangian model for
494 tracking surface spills and SaR operations in the ocean. Environ. Model. Softw. 52, 74–82.
495 Sepulveda, H. H., Artal, O. E., Torregrosa, C., nov 2011. LiveROMS: A virtual environment for ocean
496 numerical simulations. Environ. Model. Softw. 26 (11), 1372–1373.
497 Shchepetkin, A. F., McWilliams, J. C., jan 2005. The regional oceanic modeling system (ROMS): a split-
498 explicit, free-surface, topography-following-coordinate oceanic model. Ocean Model. 9 (4), 347–404.
499 Smith, W. H., Sandwell, D. T., sep 1997. Global Sea Floor Topography from Satellite Altimetry and Ship
500 Depth Soundings. Science (80-.). 277 (5334), 1956–1962.
501 Sobarzo, M., Figueroa, D., dec 2001. The physical structure of a cold filament in a Chilean upwelling zone
502 (Peninsula de Mejillones, Chile, 23 S). Deep Sea Res. Part I Oceanogr. Res. Pap. 48 (12), 2699–2726.
503 Strub, P. T., Kosro, P. M., Huyer, A., aug 1991. The nature of the cold filaments in the California Current
504 system. J. Geophys. Res. 96 (C8), 14743.
505 Thomas, A. C., nov 1999. Seasonal distributions of satellite-measured phytoplankton pigment concentration
506 along the Chilean coast. J. Geophys. Res. Ocean. 104 (C11), 25877–25890.
507 Troupin, C., Mason, E., Beckers, J., Sangrà, P., jan 2012. Generation of the Cape Ghir upwelling filament:
508 A numerical study. Ocean Model. 41, 1–15.
509 Wang, Z., Jensen, J. R., Im, J., oct 2010. An automatic region-based image segmentation algorithm for
510 remote sensing applications. Environ. Model. Softw. 25 (10), 1149–1165.