scientific **data**

Check for updates

DATA DESCRIPTOR

An ocean front dataset for the Mediterranean sea and southwest Indian ocean

Floriane Sudre [1 ✉], Ismael Hernández-Carrasco[2], Camille Mazoyer[1], Joel Sudre[3], Boris Dewitte[4,5,6,7], Véronique Garçon [8] & Vincent Rossi [1 ✉]

Fronts are ubiquitous discrete features of the global ocean often associated with enhanced vertical velocities, in turn boosting primary production. Fronts thus form dynamical and ephemeral ecosystems where numerous species meet across all trophic levels. Fronts are also targeted by fisheries. Capturing ocean fronts and studying their long-term variability in relation with climate change is thus key for marine resource management and spatial planning. The Mediterranean Sea and the Southwest Indian Ocean are natural laboratories to study front-marine life interactions due to their energetic flow at sub-to-mesoscales, high biodiversity (including endemic and endangered species) and numerous conservation initiatives. Based on remotely-sensed Sea Surface Temperature and Height, we compute thermal fronts (2003–2020) and attracting Lagrangian coherent structures (1994–2020), in both regions over several decades. We advocate for the combined use of both thermal fronts and attracting Lagrangian coherent structures to study front-marine life interactions. The resulting front dataset differs from other alternatives by its high spatio-temporal resolution, long time coverage, and relevant thresholds defined for ecological provinces.

Background & Summary

Fronts are ubiquitous dynamical features of both coastal and open ocean. This is where water masses of different properties converge, thereby enhancing nutrient availability and local primary productivity[1–3]. Ocean fronts thus form ephemeral ecosystems where numerous species gather[4–6]: from phytoplankton communities[7,8], all the way up the food chain to top predators such as sea birds[9,10], tuna[11,12] and marine mammals[13,14].

Numerous techniques have been developed to define and identify ocean fronts. Front detection methods most commonly used by the scientific community rely on the horizontal distribution of oceanic tracers (temperature, salinity, nutrients, etc.) and are either gradient-based[15,16] or histogram-based[17,18], while others make use of entropy[19], neural networks[20] or singularity exponents[21,22]. While remotely-sensed chlorophyll-a and salinity could be studied, satellite observations of sea surface temperature (SST) provide the most easily-accessible and longest input dataset for front detection[23,24]. Additionally thermal fronts are usually highly spatially correlated to density and chlorophyll-a fronts as well[11,25,26]. Subsequently, SST observations are well-suited to Eulerian front diagnostics and to study front-marine life interactions. Histogram-based front detection, such as the Cayulla and Cornillon algorithm[17] which was further improved by Nieto *et al.*[18], is the most widely used method to detect persistent front edges on large spatial scales from satellite SST observations[11,12,23]. However, gradient-based methods present significant advantages to the study of regional fronts dynamics compared to histogram-based methods[6]. They detect local fronts when the gradient of a specific property (temperature, salinity, nutrient, sea surface height, etc.) exceeds a pre-defined threshold. Gradient-based methods, such as the Belkin and O'Reilly Algorithm (BOA)[16] used in this study, thus provide a gradient magnitude instead of a simple binary edge,

[1]Aix Marseille Université, Université de Toulon, CNRS, IRD, MIO, Marseille, France. [2]Mediterranean Institute for Advanced Studies (UIB-CSIC), Miquel Marques, 21, Esporles, 07190, Balearic Islands, Spain. [3]UAR 2013 CPST, IR DATA TERRA, Z.P. de Brégaillon - CS 20330, 83507, Montpellier, La Seyne Sur Mer, France. [4]Centro de Estudios Avanzados en Zonas Aridas, Facultad de Ciencias del Mar, Universidad Catolica del Norte, Coquimbo, Chile. [5]Departamento de Biología Marina, Universidad Católica del Norte, Coquimbo, Chile. [6]Center for Ecology and Sustainable Management of Oceanic Islands (ESMOI), Faculty of Marine Sciences, Catholic University of the North, Coquimbo, Chile. [7]UMR5318 Climat, Environnement, Couplages et Incertitudes (CECI), Toulouse, France. [8]University of Liège, Liège, Belgium. ✉e-mail: floriane.sudre@mio.osupytheas.fr; vincent.rossi@mio.osupytheas.fr
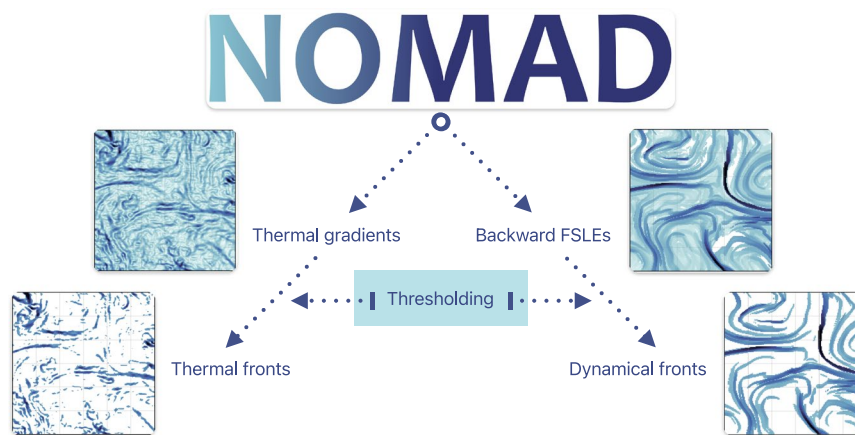
Fig. 1 Illustration showing how NOMAD[51] users can go from thermal gradients (resp. backward FSLEs) to thermal (resp. dynamical) fronts via region-specific thresholding, with example of fields before and after thresholding. All fields were taken on July 1st, 2020. The colormap is from cmocean[80] and with the same range as in Fig. 2.

require less parameter fine-tuning, and thus better reveal region-specific dynamics and variability[6,27]. The BOA improves on the original Canny Algorithm[15] by reducing gradient-generated noise while preserving the shape of frontal structures[16] and it has been successfully applied to the South China Sea[28], the East China Sea[29], the Kuroshio current[30] and the Mozambique Channel[27].

While the Eulerian perspective brought by thermal gradients describes well the spatial features of the flow at a given time, the Lagrangian dynamical approach has the added benefit of integrating both spatial and temporal current variability and brings additional information about transport and mixing properties of the fluid. In particular, Finite-size Lyapunov exponents (FSLEs)[31] used in this study are well-suited to evaluate horizontal mixing and transport in geophysical flows[7,32,33]. High FSLE values reveal lines acting as transport barriers, also expressed as hyperbolic Lagrangian coherent structures (LCS), which help to identify filaments, fronts or eddy boundaries. Ridges of backward-in-time FSLE (also referred to as "backward FSLEs") fields characterize regions of locally high convergence of particle trajectories (attracting LCS) while forward-in-time FSLEs identify regions of locally high divergence (repelling LCS). These dynamical structures greatly organize the fluid motion around them, providing a flow mapping with the main routes of transport[34,35]. Here we provide backward FSLEs to unveil attracting LCS. Tracers (such as temperature, nutrients, pollutants, etc.) converge and spread along attracting LCS, creating filament-like structures of high tracer concentration. This property offers a direct and clear physical interpretation of the attracting LCS.

FSLEs are complementary to thermal gradients because they are not vulnerable to cloud cover, they are able to reveal dynamics at a smaller scale than the resolution of the velocity field[34], and they provide useful transport diagnosis of tracers[7,36]. Additionally, FSLEs are especially useful to study marine bio-physical interactions. Indeed, the LCS revealed by the FSLE field separate regions with different dynamics and often coincide with sharp tracer gradients[37] so that they can be interpreted as dynamical fronts. FSLEs have been used to understand the dynamics of tracers like SST and chlorophyll-a[7,8,38,39], which have direct implications on the behavior of marine life across all trophic levels[5,10,13,14,40–42].

The present dataset offers the synergy of both Eulerian and Lagrangian perspectives by providing a multi-decadal record of both thermal gradient and backward FSLE fields at high resolution. Thermal gradients (resp. backward FSLEs) fields are computed on a 1 km (resp. 1.56 km) grid, covering the period 2003–2020 (resp. 1994–2020), for two oceanic regions characterized by high biodiversity under conservation concerns: the Mediterranean Sea (MedSea) and the Southwest Indian Ocean (SWIO). Thermal gradients (resp. backward FSLEs) lead to thermal (resp. dynamical) fronts after region-appropriate thresholding, as illustrated in Fig. 1.

Other data products informing on ocean fronts have been developed. Belkin *et al.*[23] produced a 9 km-resolution persistent front database for major Large Marine Ecosystems[43]. Additionally, daily backward FSLEs[44] (accessed on Jan 20th 2023) have been computed from altimetry derived daily velocities on a 4 km global grid since 1994 following the method of d'Ovidio *et al.*[32]. Finally, the BOA has been used by the National Oceanic and Atmospheric Administration (NOAA) since 2013 to produce near real-time chlorophyll-a frontal maps (accessed on Jan 20th, 2023) for the U.S. coastal areas only. Those existing datasets target analyses of much larger spatial scales, their resolution being 2.5 to 9 time coarser than the dataset presented here, and/or do not cover the present geographical domains of interest. Our front dataset differs to those existing by being specifically designed for multidisciplinary research, especially topics linking ocean weather and marine ecology[45]. It is dedicated to researchers who investigate front dynamics at space and time scales that relate to the highly patchy distribution of biological production and the rapidly evolving behavior (for foraging, mating, socializing, migrating, etc.) of marine species, especially top marine predators. Additionally, the greater length of the provided time series allows for analyzing long-term variability and trends. Our ocean front dataset for the Mediterranean Sea and the Southwest Indian Ocean aims to

bridge the gap between advances in the identification of physical oceanic features and the search of general rules governing how marine life across trophic levels exploits the highly heterogeneous seascape.

Methods

Thermal gradients.   We use the BOA[16], and more specifically Benjamin Galuardi's pseudo-code written in R (available on https://github.com/galuardi/boaR, accessed on May 19th, 2022) to compute thermal gradient magnitudes. We apply the BOA on daily satellite Sea Surface Temperatures (SST) from the Multi-scale Ultra-high Resolution sea-surface temperature analysis (MUR SST, accessed on Jan 10th, 2023)[46], the resulting field is a thermal gradient expressed in °C/km. The resulting Southwestern Indian Ocean (SWIO) thermal gradient dataset covers latitudes between 5°S and 35°S and longitudes between 30°E to 80°E, and the Mediterranean Sea (MedSea) dataset covers latitudes between 30°N and 46°N and longitudes between −6°E to 37°E. The outputs for both geographical domains are produced daily from 2003 to 2020 and at the same spatial resolution of 1/100° which is about 1 km.

Each daily thermal gradient map includes a flag variable indicating whether a given pixel is located on land (flag = 2), or if it is "suitable" (flag = 0) or "unsuitable" (flag = 1) for thermal front analysis. Pixels are deemed "unsuitable" based on MUR SST standard deviation of the formal estimation error, which is considered to be a good estimate on MUR analysis uncertainty[46]. When the MUR SST standard deviation of the estimation error is high, the thermal gradient field is smoothed out and possible front or eddy structures are obfuscated. We thus recommend applying the flag before any further analysis of the thermal gradient field. We set the MUR SST standard deviation of the estimation error threshold, i.e. threshold above which a pixel is deemed "unsuitable" for front analysis, at 0.0384 °C. The choice of threshold is based on a preliminary comparison between MUR SST standard deviation and our thermal gradient dataset.

Finite-size Lyapunov exponents.   In this work, backward-in-time FSLEs in the SWIO are computed from remotely-sensed Sea Surface Height (SSH) at 1/4° spatial resolution from Global Total Surface and 15 m Current (CLS, 2018, COPERNICUS-GLOBCURRENT, accessed on Jan 10th, 2023) from Altimetric absolute geostrophic velocities and Modeled Ekman Current Reprocessing[47]. In the MedSea, backward FSLEs are computed from daily absolute geostrophic surface currents derived from Sea Level Anomalies (SLA) at 1/8° spatial resolution from a SSALTO/DUACS multimission altimeter regional L4 product released in 2016 by AVISO + and based on regional mean dynamic topography[47] (https://data.marine.copernicus.eu, accessed on Feb 10th, 2023).

Backward FSLEs were obtained following the algorithm described by Hernandez-Carrasco et al.[34]. Essentially, the algorithm computes backward FSLEs by integrating backward in time two neighbouring particle trajectories advected in a two-dimensional flow using a fourth order Runge-Kutta scheme with a bilinear interpolation in space and linear in time. The total period of integration is 90 days. The Runge Kutta time step was set at 6 hours to reduce numerical diffusion. As we are interested in simulating fluid particles, we computed trajectories of infinitesimal passive particles. The FSLE value at a given position and time (x, y, t) can be expressed as:

$$\text{FSLE}(x, y, t, \delta_f, \delta_0) = \frac{1}{\tau}\log\left(\frac{\delta_f}{\delta_0}\right).$$

δ_0 is the initial distance between a particle at (x, y, t) and its 4 closest neighbors. δ_0 corresponds to the spatial resolution of the FSLE grid, which is here 1/64° or approximately 1.56 km. δ_f is the final distance between particles, here it is set so that $\delta_f = 10\delta_0$. τ is the minimum time (among the 4 particle pairs) that it takes for the particles to reach the distance δ_f. FSLEs are thus expressed in day^{-1}. Similarly to the thermal front dataset, the SWIO FSLE dataset covers latitudes between 5°S and 35°S and longitude between 30°E to 80°E, and the MedSea dataset covers latitudes between 30°N and 46°N and longitudes between −6°E to 37°E. Both geographical domains have a spatial resolution of 1/64° which is about 1.56 km, and the output is produced daily from 1994 to 2020. Each daily FSLE dataset includes a flag variable indicating whether pixels are located on land (flag = 2), or if they are "suitable" (flag = 0) or "unsuitable" (flag = 1) for front analysis. Pixels are deemed "unsuitable" based on their FSLE value. Pixels whose FSLE value is 0 day^{-1} correspond to particles which have not reached the final distance δ_f after the integration time, i.e. after 90 days. Pixels whose FSLE value is above 900 day^{-1} correspond to beached particles. We thus recommend applying the flag before any further analysis of the FSLE field. Figure 2 is an example snapshot of the thermal gradient and backward FSLE fields in the MedSea and the SWIO.

Thresholds.   One of the advantages of using the BOA and backward FSLEs to detect fronts is that both outputs are continuous fields. Users can then determine the threshold above which thermal gradients or FSLE values belong to a meaningful front, as illustrated in Fig. 1. Since different regions present different ranges of front intensity, that are somehow non-trivially linked to contrasting levels of Eddy-Kinetic Energy[27], it is relevant to define region-specific thresholds. Moreover, it is very likely that marine organisms aggregate actively along the most adequate fronts for foraging not based on their absolute characteristics but rather on their relative properties (i.e. targeting the most advantageous fronts as compared to all those found in the surroundings). As such, and to adapt our dataset to users interested in the interaction between fronts and marine life at both basin and eco-region scales, we provide thresholds for the MedSea and the SWIO basins but also for each Longhurst's provinces[48,49] and Spalding's ecoregions[50] falling into each geographical domain (Fig. 3). For both levels of analysis, we set the regional and subregional thresholds at the value of the 70th percentile computed over all the "suitable" pixels in the multidecadal dataset.
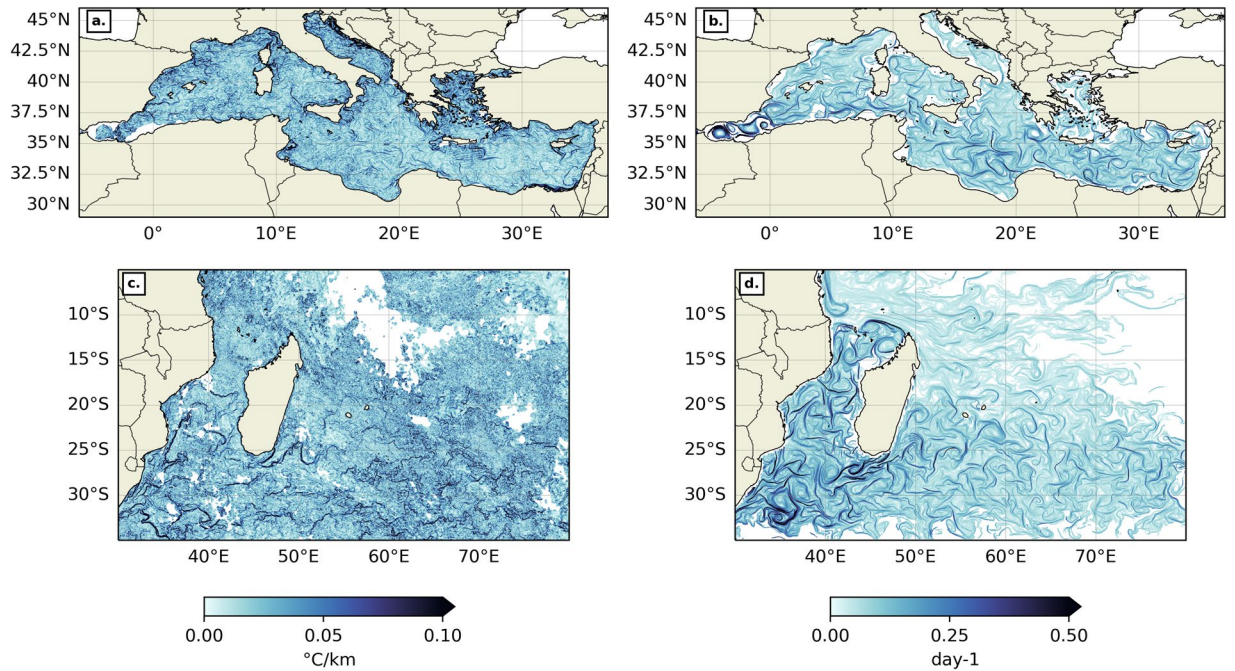
Fig. 2 Example snapshots of thermal gradient (**a,c**) and FSLE (**b,d**) fields in the MedSea (**a,b**) and SWIO (**c,d**), for "suitable" pixels only, on July 1st, 2020. The colormap is from cmocean[80].
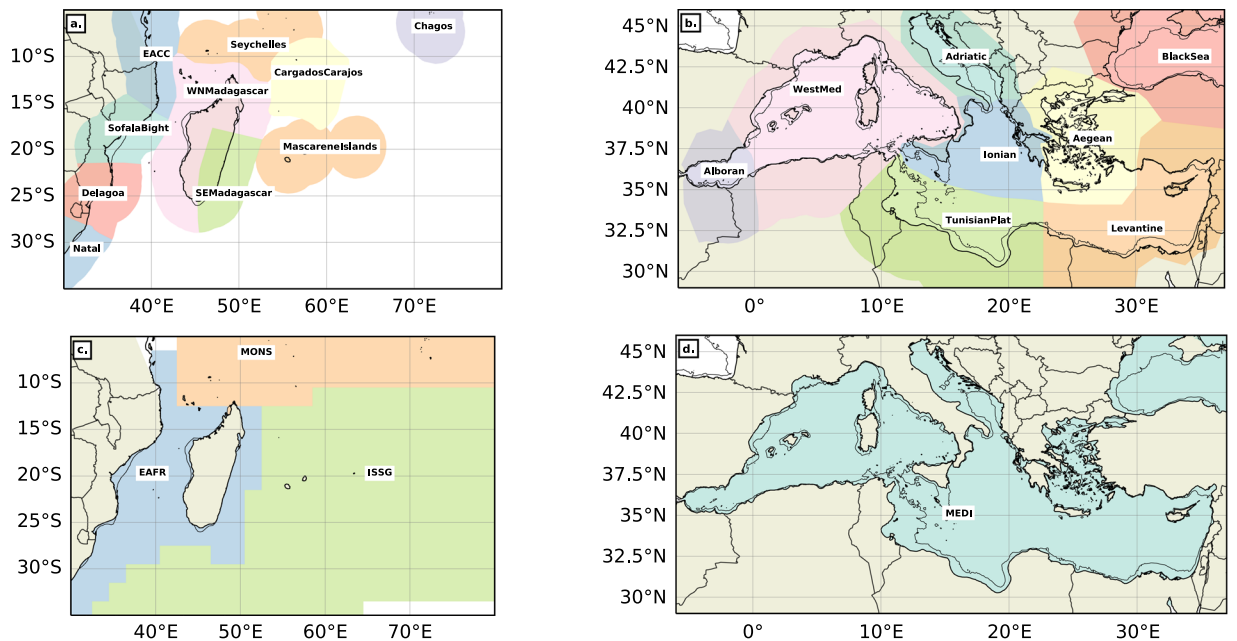

Fig. 3 Spalding's ecoregions[50] (**a,b**) and Longhurst's provinces[48,49] (**c,d**) in the SWIO (**a,c**) and in the MedSea (**b,d**).

Percentage of "suitable" pixels. The percentage of "suitable" pixels in each region of interest (SPP) is given by:

$$SPP(region) = \frac{SPC(region)}{OPC(region)} * 100.$$

Where SPC is the number of "suitable" pixels in the region of interest and OPC is the total number of ocean pixels in the same region. Thresholds and SPP are provided for each region and subregion in Table 2 and as attributes in each netcdf file.

Folder name	Region	Variable	size	time period	temporal resolution	spatial resolution
TG_MedSea	MedSea	thermal gradient	128 Gb	2003–2020	daily	1/100° (approx. 1 km)
FSLE_MedSea	MedSea	backward FSLE	61 Gb	1994–2020	daily	1/64° (approx. 1.56 km)
TG_SWIO	SWIO	thermal gradient	507 Gb	2003–2020	daily	1/100° (approx. 1 km)
FSLE_SWIO	SWIO	backward FSLE	244 Gb	1994–2020	daily	1/64° (approx. 1.56 km)

Table 1. Dataset subfolders details.

Region	Classification	Subregion	Thermal Gradients (°C/km)		FSLEs (day^{-1})	
			Threshold	SPP	Threshold	SPP
MedSea	WHOLE	n/a	0.033	83.8%	0.123	57.3%
MedSea	LH	MEDI	0.033	83.6%	0.123	59.8%
MedSea	MEOW	Adriatic Sea	0.039	78.0%	0.086	44.0%
MedSea	MEOW	Aegean Sea	0.039	78.9%	0.135	49.0%
MedSea	MEOW	Alboran Sea	0.039	75.9%	0.254	41.3%
MedSea	MEOW	Ionian Sea	0.031	84.4%	0.113	73.3%
MedSea	MEOW	Levantine Sea	0.032	86.4%	0.142	75.6%
MedSea	MEOW	Tunisian Plateau	0.030	85.9%	0.111	76.4%
MedSea	MEOW	Western Mediterranean	0.033	85.6%	0.119	69.3%
SWIO	WHOLE	n/a	0.035	74.1%	0.105	73.3%
SWIO	LH	EAFR	0.036	78.5%	0.166	89.3%
SWIO	LH	ISSG	0.036	77.1%	0.094	74.7%
SWIO	LH	MONS	0.034	60.3%	0.083	56.3%
SWIO	MEOW	Bight of Sofala	0.033	79.2%	0.190	90.4%
SWIO	MEOW	Cargados Carajos	0.034	68.3%	0.077	77.0%
SWIO	MEOW	Chagos	0.036	49.7%	0.075	42.7%
SWIO	MEOW	Delagoa	0.038	83.3%	0.179	95.2%
SWIO	MEOW	East African Coral Coast	0.032	71.7%	0.128	79.0%
SWIO	MEOW	Mascarene Islands	0.033	81.1%	0.089	88.0%
SWIO	MEOW	Natal	0.041	76.5%	0.210	89.0%
SWIO	MEOW	Seychelles	0.034	67.5%	0.083	66.3%
SWIO	MEOW	Southeast Madagascar	0.037	76.6%	0.142	93.7%
SWIO	MEOW	Western and Northern Madagascar	0.034	75.8%	0.146	86.6%

Table 2. Thresholds and SPP for each region, Longhurst's provinces (LH)[48,49] and Spalding's Marine Ecoregions of the World (MEOW)[50].

Data Records

The oceaN frOnt dataset for the Mediterranean seA and southwest inDian ocean (NOMAD)[51] can be accessed through Ifremer online data repository. Figure 2 is a sample snapshot of the thermal gradient and backward-in-time FSLE field in the MedSea and the SWIO. The front dataset is 940 Gb in total, and is separated into 4 folders as detailed in Table 1. Each folder contains yearly subfolders, consisting of daily netcdf files. Each netcdf file provides: thermal gradient or backward FSLE values, pixel quality flags, but also front threshold values specific to each Longhurst's province and Spalding's ecoregion found in the MedSea or SWIO, as well as the corresponding percentage of "suitable" pixels found in those same ecologically-relevant subregions of the MedSea and the SWIO. Each netcdf file contains variables as shown below:

Thermal gradient netcdf files (TG_*region*_*date*.nc)

(1) time: date in seconds since 1950-01-01 in the proleptic gregorian calendar.
(2) lat: latitude, in decimal degree north.
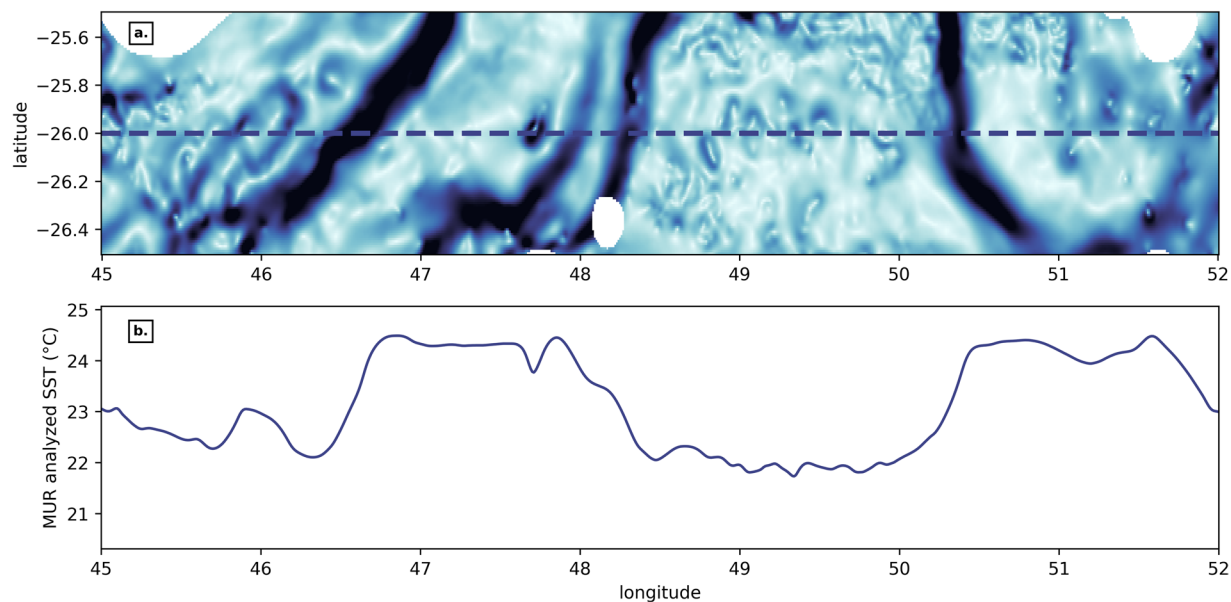(3) lon: longitude, in decimal degree east.
(4) tg: thermal gradient, in °C/km.
(5) flag: quality flag, which options are 0 = suitable, 1 = unsuitable, 2 = land.
(6) threshold_*subregion*: value of the 70th percentile of the thermal gradient field in a specific subregion, in °C/km.
(7) SPP_*subregion*: percentage of pixels deemed suitable in a specific subregion, in %.

Backward FSLE netcdf files (**FSLE_*region*_*date*.nc**)

(1) time: date in seconds since 1950-01-01 in the proleptic gregorian calendar.
(2) lat: latitude, in decimal degree north.

Fig. 4 Example of thermal gradient magnitude from the BOA (**a**) and corresponding MUR analysed SST value (**b**) for a section at 26°S in the SWIO, for July 1st, 2020. The colormap of Fig. 4a is from cmocean[80] and with the same range in °C/km as in Fig. 2.

- -

(3) lon: longitude, in decimal degree east.
(4) fsle: backward finite-size Lyapunov exponents, in day^{-1}.
(5) flag: quality flag, which options are $0 =$ suitable, $1 =$ unsuitable, $2 =$ land.
(6) threshold_*subregion*: value of the 70th percentile of the FSLE field in a specific subregion, in day^{-1}.
(7) SPP_*subregion*: percentage of pixels deemed suitable in a specific subregion, in %.

Technical Validation

About thermal fronts. NOMAD[51] provides a multidecadal record of thermal gradients, which can be translated into thermal fronts via region-specific thresholding (Fig. 1). The BOA[16], which we used to compute NOMAD[51] thermal gradients, has been successfully employed multiple times: in physical oceanography[27,28], marine ecology[30,52], as well as by the National Oceanic and Atmospheric Administration (NOAA) since 2013[24]; and can thus be considered as an established method recognized by the research community. Figure 4 illustrates the BOA performance by putting side by side the horizontal thermal gradient issued from the BOA (top panel) and its corresponding MUR SST variation on a section at 26°S. Sharp changes of temperature (ranging approximately from 0.25 to 2°C in this particular example) coincide with high values of thermal gradients (occurring at spatial scales ranging 10 to 50 km, in this example). High thermal gradient values thus are of the order of several degrees Celsius per 100 km.

Choice of SST data. When choosing a suitable SST product, we compromised spatial and temporal resolutions, cloud-contamination, time coverage and public availability. Cloud contamination is a real issue, especially when high spatial resolution is needed: infra-red radiometers cannot sense through cloud cover but micro-wave sensors do not have a sufficient spatial resolution yet to access small-scales dynamics. We thus opted for the Multi-scale Ultra-high Resolution (MUR) SST analysis[46], which combines 1 km resolution Moderate Resolution Imaging Spectroradiometer (MODIS) images, to 5–9 km resolution Advanced Very High Resolution Radiometer (AVHRR) products, 25 km resolution micro-wave data, and point-wise *in-situ* SST data. The integration and interpolation of those different satellite products is done by the Multi-Resolution Variational Analysis (MRVA) method, which uses scale-dependent time windows to adjust for various spatial resolutions[46]. MUR analysis thus provides improved spatial coverage and resolution of its SST product compared to conventional L3 data. However, such interpolation procedure also tends to blur oceanic features of interest such as eddies and fronts. It is thus necessary to find a way to identify pixels which have been blurred beyond a certain limit. The MUR SST product comes with a standard deviation of the estimation error (in °C) at each grid point[46]. We use this estimate of MUR analysis uncertainty to identify a suitable threshold above which the multidecadal dataset is considered too blurred for front analysis. The MUR analysis product also provides the time lag to most recent 1 km data, which would help the filtering, but the variable is only included after July 2016. Figure 8 of Sudre *et al.* work[27] further illustrates the relationship between thermal front detection and the two quality flags available in the MUR dataset. We finally opted for using only the standard deviation of the estimation error to provide a uniform and consistent multidecadal database. Visual comparisons between MUR standard deviation of the estimation error and our thermal gradient dataset were performed to find an arbitrary threshold that would return well-defined frontal structures without too much gaps. A too high estimation error limit would
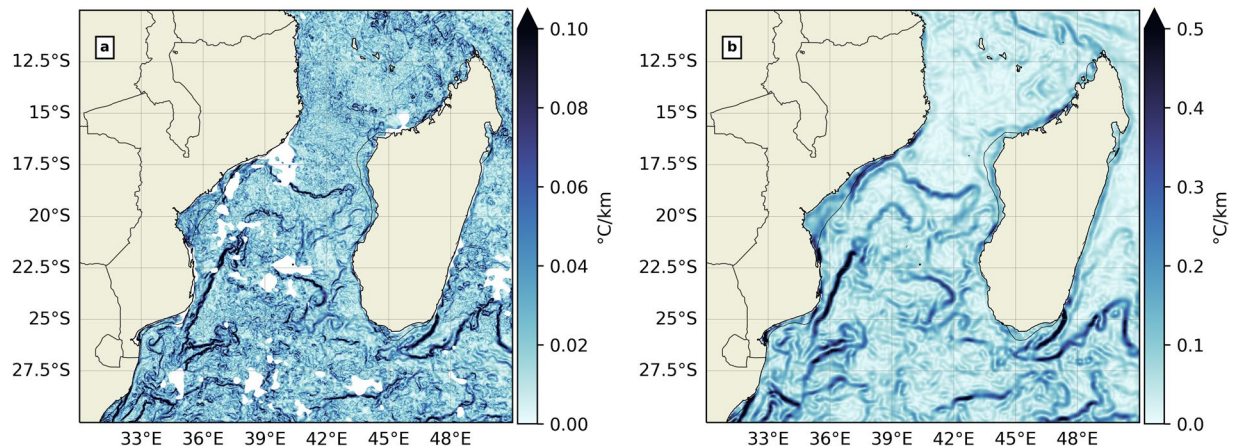
Fig. 5 (**a**) thermal gradient output from MUR SST and (**b**) thermal gradient output from SST CCI analysis, in the Mozambique Channel for July 1st, 2020. The colormap is from cmocean[80].

not filter out pixels where the MUR fusion and interpolation procedure blurred fronts and eddies. Yet a too low estimation error limit would lead to many pixels being filtered out, and some front features being hidden despite looking sharp enough to be exploitable. An additional difficulty resides in the fact that the standard deviation of the estimation error is a discrete variable in the MUR SST netcdf files. We found a good compromise for a standard deviation of the estimation error equal to 0.0384 °C.

The European Space Agency Sea Surface Temperature Climate Change Initiative: Analysis product[53] (SST CCI analysis) could be an attractive alternative to the MUR SST due to its climate-readiness. However, the SST CCI analysis is offered on a 1/20° grid (about 5 km), which is 5 times coarser than that of MUR SST. When pairing with biological observations such as GPS tracks, aerial surveys or sightings, the SST CCI may return spurious and artificial correlations. In designing NOMAD[51], the authors favored the grid resolution over the climate readiness of the final product to better fit typical scales of marine animal trajectories in the SWIO[54–56], hence the final choice of SST MUR over the SST CCI analysis. The comparison between both MUR SST and SST CCI analysis is illustrated in Fig. 5 over the front-rich Mozambique Channel[27]. The larger thermal gradient magnitudes from SST CCI could be explained by the contextual median filtering step in the BOA[16]. We hope future ocean front datasets will be able to combine the advantage of both fine-grid resolution and climate-readiness without much compromise.

About dynamical fronts. NOMAD[51] provides a multidecadal record of backward-in-time FSLEs, which can be interpreted as dynamical fronts through region-specific thresholding (Fig. 1). In the past 20 years, FSLEs have been recognized by the ocean science community as a reliable way to identify Lagrangian coherent structures[32,34,57–59]. Indeed, FSLEs tend to be better-suited to transport barrier detection than other diagnostics such as the Okubo-Weiss parameter and finite-time Lyapunov exponents[33,59,60]. Additionally, FSLEs perform well in the presence of noise and unresolved scales[34]. Finally FSLEs pairs well with tracer-based front indicators, such as sea-surface temperature[37,61] and phytoplankton communities[8,42].

Choice of velocity data. The choice of suitable velocity products for the MedSea and the SWIO was motivated by their spatial and temporal resolutions but also the extent of the datasets. The velocity products chosen for the MedSea and SWIO are different because we opted to select the best available product for each domain. Globcurrent 1/4° total velocities[47] are optimal for the SWIO because of the integration of the Ekman component to the velocity field, its spatial and temporal resolutions and coverage. To the authors knowledge, there is no available operational product providing total velocities (geostrophic + Ekman components) for the MedSea. While the ageostrophic components of the circulation may become relevant in a few rare instances, such as during short-living storms in the Northwestern MedSea[62], geostrophic currents resolved well the main surface dynamics. We thus opted for the next best: 1/8° geostrophic velocities derived from a SSALTO/DUACS multimission altimeter regional L4 product[47] for the MedSea. Although both velocity datasets perform well at the mesoscale (typical horizontal scales of 100 km), it is worth noting that they are not climate-ready. The choice of a uniform global climate-ready alternative, such as DUACS geostrophy-only currents distributed by the Copernicus Climate Change Service (C3S) (https://data.marine.copernicus.eu, accessed on Feb 10th, 2023), would have been uniform between both domains but would also have resulted in the loss of the Ekman component in the SWIO and led to a spatial resolution downgrade from 1/8° to 1/4° in the MedSea. The comparison between FSLEs derived from Globcurrent total velocities[47] and FSLEs derived from geostrophy-only currents from the C3S[63] in the SWIO is illustrated in Fig. 6. The difference between FSLEs computed from total and geostrophy-only currents reveals small changes in their magnitude and overall pattern, but also show some variations in the positions of the LCS (in the approximate range of 0.5–1°or 50–100 km). These displacements are especially strong in front-rich regions such as the Mozambique Channel and South East of Madagascar. We thus chose to let go of uniformity and climate-readiness to ensure the best rendering of front structures with available
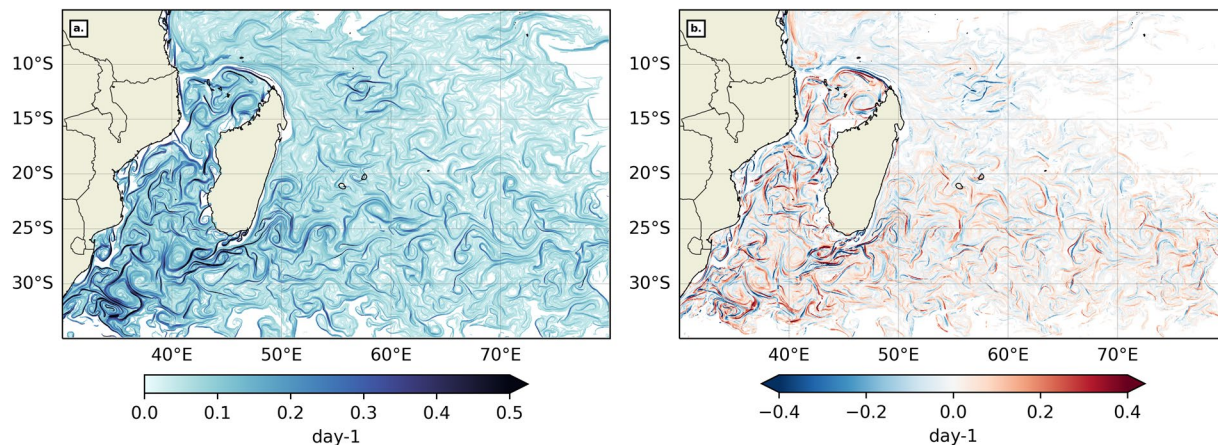
Fig. 6 (**a**) Snapshot of FSLEs computed from geostrophy-only currents from the Copernicus Climate Change Service and (**b**) its difference from FSLEs computed from total currents (shown in Fig. 2d), in the SWIO for July 1st, 2020. The colormaps are from cmocean[80].

altimetry products. Should climate-readiness be of interest to end-users, future front data products could use newer climate-ready products with our proposed methodology.

Added-value of spatial resolution. The NOMAD[51] dataset distinguishes itself from existing front databases by its higher spatial and temporal resolutions. A great advantage of FSLEs is that they are robust to small changes in the velocity field, and they also exhibit typical multifractal properties[34]. These properties ensure that FSLEs produced on a finer-grid than the original velocity grid are indeed meaningful and not artificial. They cannot resolve all the small scales processes, but they render well the dynamics driven by the coarser velocity field. We argue that spatial resolution is critical when investigating fronts-marine life interactions. To illustrate this statement, we compare NOMAD[51] backward FSLE dataset to the AVISO FSLE product[44] in the SWIO. The AVISO FSLE field is computed from 1/4° SSALTO/Duacs delayed-time global ocean absolute geostrophic currents (DUACS2018 DT MADT UV products) on a 4 km grid. The AVISO FSLE product[44] differs from NOMAD[51] FSLEs in several parameters including velocity field, spatial resolution, times of integration, δ_0 and δ_f. We have already contrasted the difference between total and geostrophy-only currents in FSLE computations in the previous paragraph. Additionally, we assume that the AVISO FSLE product time of integration, δ_0 and δ_f have been optimized for the flow field. We will thus only focus here on why we think the higher spatial resolution of the NOMAD[51] FSLE field is a significant added-value for biological use. We generated 1000 series of 100 to 200 points. Each point is separated from the next by a random direction and a distance of the order of 50–100 km, consistent with typical scales of marine animal presence/absence data (including sightings, GPS tracks, aerial survey, etc.), like those of whale sharks[54], sea turtles[55] or the great frigatebirds[56] in the SWIO. The next step is to collocate each series of points with both NOMAD and AVISO FSLE fields[44], taken on the same date. For each series the date was picked randomly in the year 2020 to avoid any bias brought on by seasonality. Figure 7 is an example snapshot of NOMAD[51] and AVISO[44] FSLE fields overlapped by a sample series of randomly-generated points in the South Mozambique Channel. Lastly, for each of the 1000 series, and relating to both finer NOMAD[51] FSLE (1.56 km resolution) and coarser AVISO[44] FSLE (4 km resolution) fields, we extracted 3 metrics:

- the mean FSLE value located under each randomly-generated point,
- the percentage of points collocated with FSLE front pixels (defined by subregional thresholds as in Table 2),
- and the mean distance between each point and its nearest FSLE front (defined by subregional thresholds as in Table 2)

We thus obtain, for each series of randomly-generated points, a difference (between NOMAD[51] and AVISO[44] FSLE) for each metrics described above. These analyses were performed for both the EAFR and ISSG Longhurst subregions[48,49] (Fig. 3) to illustrate the region-specific sensitivity to spatial resolution. The mean difference in FSLE values located under each randomly-generated point in the EAFR (resp. ISSG) subregion is of the order of 0.028 day^{-1} (resp. 0.016 day^{-1}), and reaching up to 0.202 day^{-1} (resp. 0.096 day^{-1}) (Fig. 8). The difference in collocation percentage is on average 8.03% in the EAFR and 6.5% in the ISSG subregions (Fig. 8), and going up to 35.55% and 24.04% respectively. Finally, the mean difference in distance to nearest front is smaller in EAFR at 4.3 km compared to 16.3 km in ISSG, likely due to the higher concentration of fronts in the Mozambique Channel[27]. Overall, the mean and maximum differences between fine and coarse FSLEs for those three metrics are non-negligible, especially in front-rich regions such as the Mozambique Channel and southeast of Madagascar, which will then reflect into any statistical comparison between biological observations and front occurrence. NOMAD's[51] finer spatial resolution is thus a significant added-value for high-resolution tracking of marine life-front interactions.
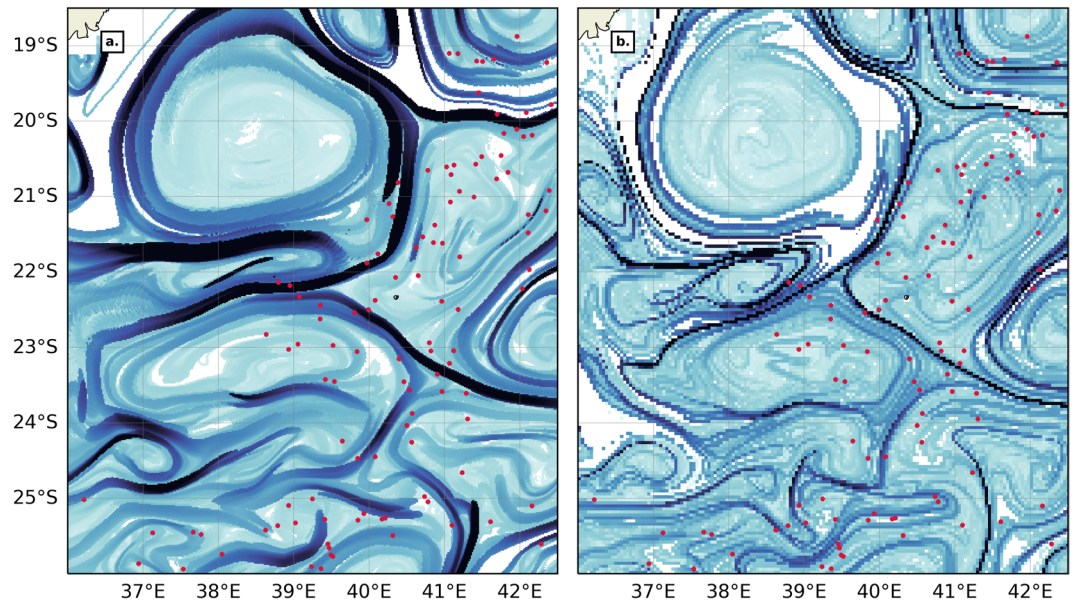
Fig. 7 Example of the collocation between randomly-generated points (in red) and FSLE fields for (**a**) NOMAD[51] backward FSLEs on a 1.5 km grid and for (**b**) AVISO backward near-real-time FSLEs[44] on a 4 km grid, for Jan 15th, 2020. The colormap is from cmocean[80] and with the same range as in Fig. 2.
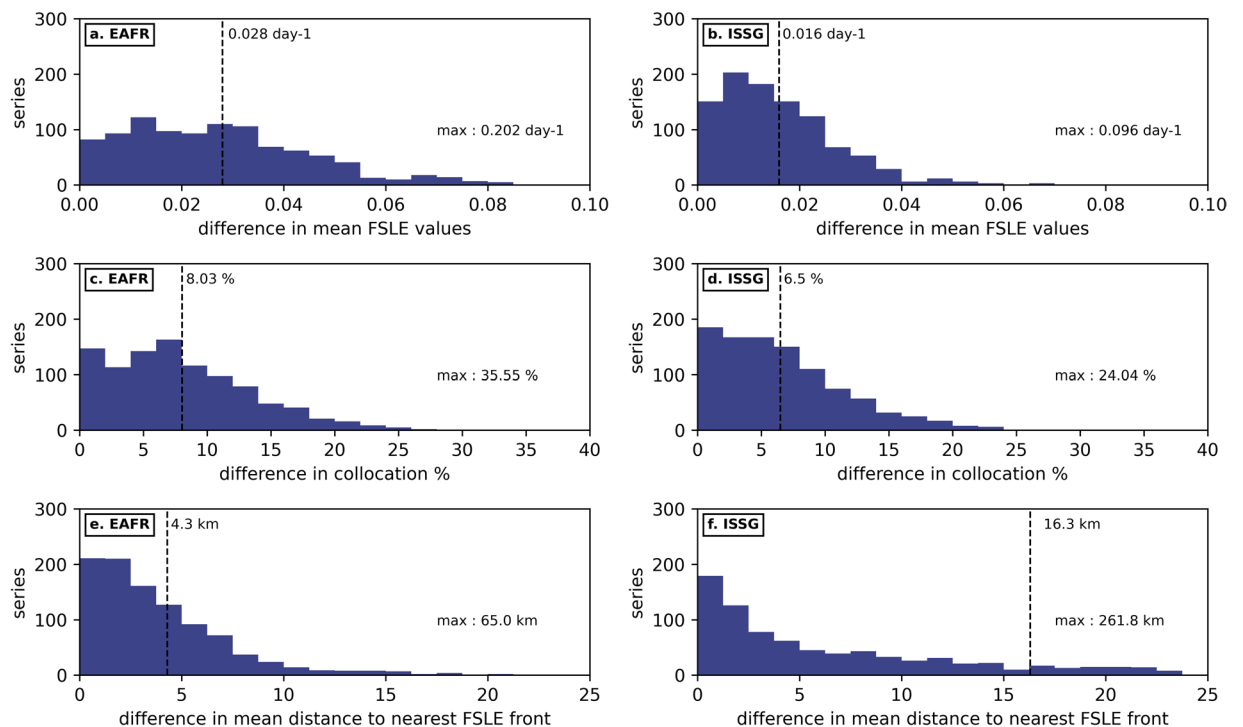

Fig. 8 Distributions of the differences in mean FSLE values (**a, b**), collocation percentage (**c, d**) and distance to nearest front (**e, f**) between NOMAD[51] and AVISO[44] FSLE front fields for 1000 series of 100–200 randomly-generated points. We chose 2 Longhurst[48,49] subregions to illustrate the region-specific sensitivity to spatial resolution: EAFR (**b**), ISSG (**c**). The vertical dotted lines mark the mean difference for all 1000 trajectories, weighted on the number of points per trajectory.

Finally, the backward FSLEs described in this paper were previously studied to analyze transport processes in the MedSea and were validated with trajectories of real drifters, as shown in Figs. 2b, 3. of Hernández-Carrasco and Orfila's work[64]. Similar validation has been performed in the SWIO and is available upon request.
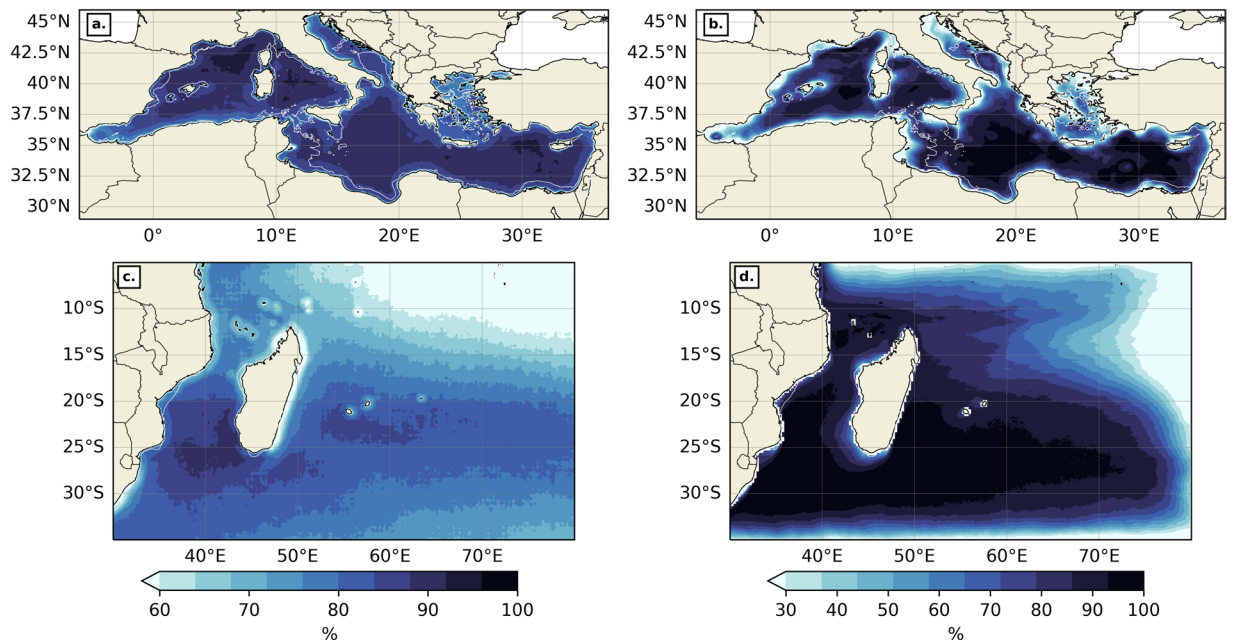
Fig. 9 Frequency of suitable days for thermal gradients (**a, c**) and FSLEs (**b, d**) in the MedSea (**a, b**) and the SWIO (**c, d**). The background white contour shows the 200 m isobath. The colormap is from cmocean[80].

Numerical diagnostics. *Suitable day frequency.* For each position in space (x, y) the frequency of "suitable" days (FSD) has been computed as follow:

$$FSD(x, y) = \frac{NSD(x, y)}{TND} * 100.$$

Where NSD (x,y) is the number of "suitable" days at (x, y), i.e. when the flag is equal to 0, and TND is the total number of days in the whole dataset. The FSD for thermal gradients reveals which regions are more prone to cloudiness between 2003 and 2020 (Fig. 9a,c). In the MedSea, most of the open ocean presents a FSD above 85% thanks to the dominant clear sky typical of Mediterranean climate. The Mediterranean regions most covered by clouds are the western Alboran Sea, the northern Aegean Sea and more generally coastal areas. In the SWIO, most of the open ocean in the Mozambique Channel and east of Madagascar has a FSD above 80%. Continental shelves, Islands (the Comoros archipelago, Mayotte, La Réunion and Mauricius islands) and northeast of the South Equatorial Current are marked by a reduced FSD, but rarely fall below 50%. The FSD for the backward FSLE dataset reveals regions most accessible to advected ocean particles between 1994 and 2020 (Fig. 9b,d). In the MedSea, the FSD is lowest on continental shelves and western Alboran Sea but above 90% for most of the open ocean. In the SWIO, the FSD is above 90% in the Mozambique Channel and most of the western Indian Ocean. Lesser FSD can be found along the path of the South Equatorial Current, along the west Madagascar continental shelves and more generally at the border of the geographical domain. Overall both datasets and both geographical domains present a very good ratio of suitable pixel for front analysis.

Mean fields. Time average maps have been produced for both thermal gradients (Fig. 10a,c) and backward FSLE fields (Fig. 10b,d) in both geographical domains over their respective time period (2003–2020 for thermal gradients and 1994–2020 for backward FSLEs). Only the "suitable" pixels were counted when time-averaging. In the MedSea, thermal gradients are stronger on average along continental shelves and highlight specific regional features, such as the quasi-permanent eddies of the Alboran Sea, the Almeria-Oran front[65], the Cartagena-Tenes Front[64], and the Liguro-Catalan current. Time-averaged backward FSLE values are higher in the Alboran Sea, South of Sardinia, along the Liguro-Catalan current and in the Levantine Sea. In the SWIO, mean thermal gradients are most intense on continental shelves as well, but also in the wake of both Northeast and Southeast Madagascar currents (NEMC and SEMC) and more generally south of 25°S. These results were also shown from a 1/36° resolution realistic CROCO simulation of the Mozambique Channel[27] and from Level 3 National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) SST data[11]. Time-averaged backward FSLE values are also stronger in the path of the NEMC and SEMC, but also at the side edges of the eddy train that characterizes the Mozambique Channel circulation[66,67], and towards the Agulhas current[68]. These results are consistent with previous works on time-averaged thermal gradients[58] and FSLEs in the MedSea[32,69], but also in the SWIO[11,27], thus validating our dataset for large time scales.

Front frequency. Another helpful metric to assess our dataset is the front frequency (FF), i.e. frequency of front days (amongst suitable days only) at a position (x, y), computed as follow:
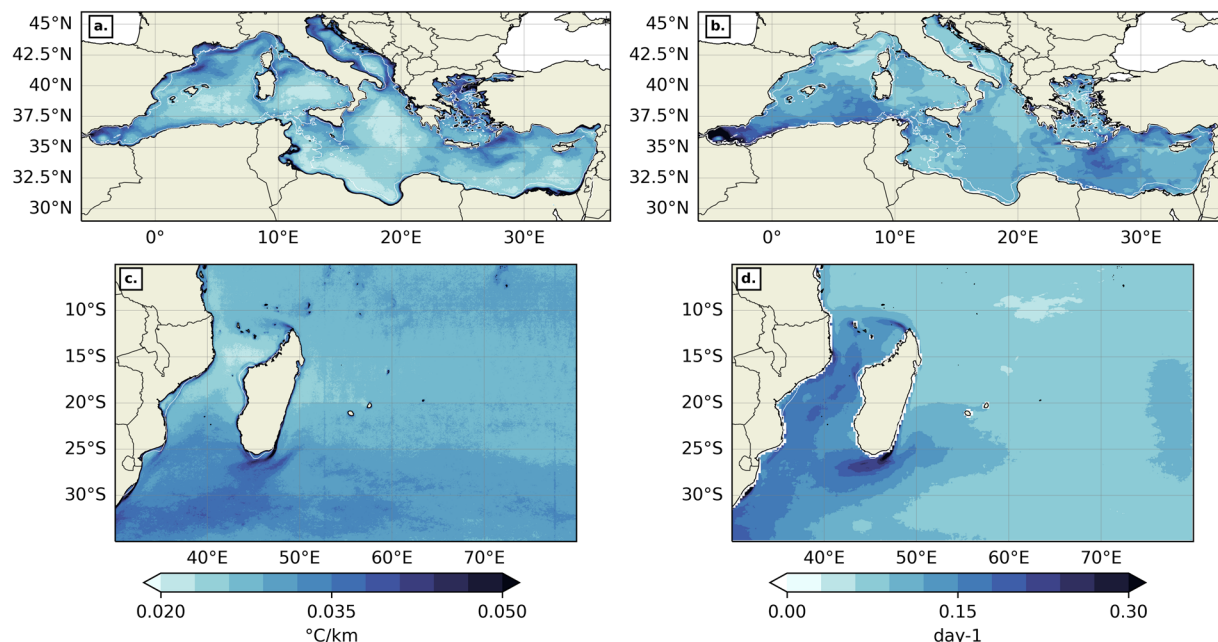
Fig. 10 Time average for thermal gradients (**a, c**) and FSLEs (**b, d**) in the MedSea (**a, b**) and the SWIO (**c, d**). The background white contour shows the 200 m isobath. The colormap is from cmocean[80].
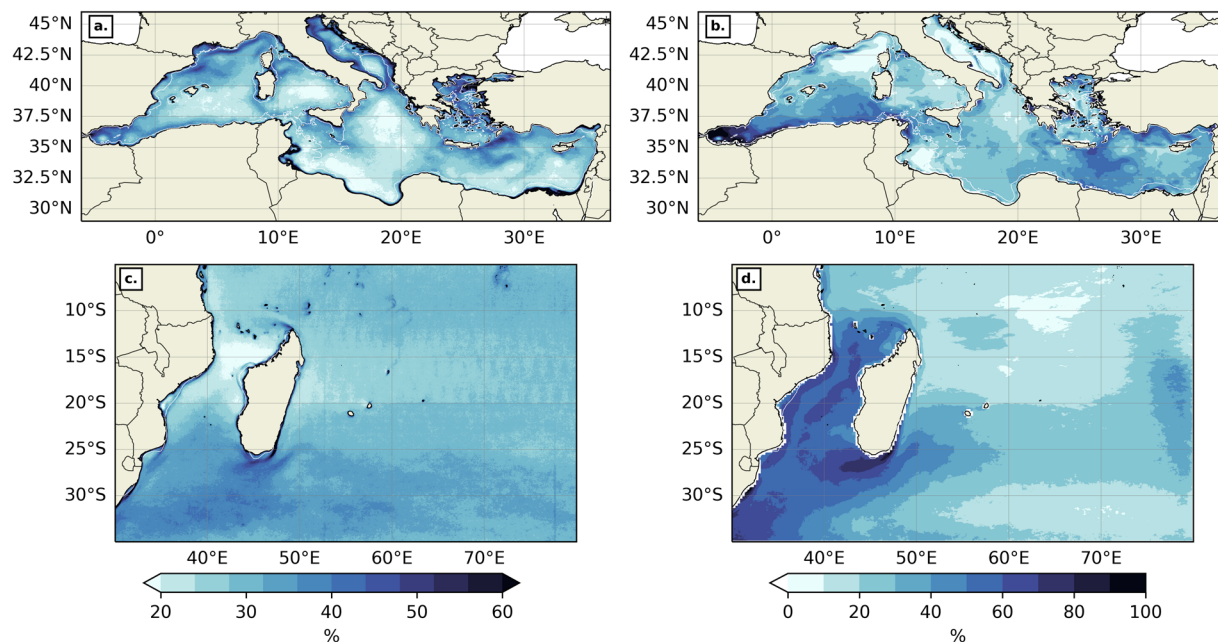

Fig. 11 Front frequency for thermal gradients (**a, c**) and FSLEs (**b, d**) in the MedSea (**a, b**) and the SWIO (**c, d**). The background white contour shows the 200 m isobath. The colormap is from cmocean[80].

$$FF(x, y) = \frac{SNFD(x, y)}{TNSD(x, y)} * 100.$$

Where SNFD (x, y) is the number of front days (amongst "suitable" days only) at (x, y), and TNSD (x, y) is the total number of "suitable" days at the same position. "Front days" correspond to days when thermal gradient of FSLE values at (x, y) are above a pre-determined 70th percentile threshold. We chose to take whole region thresholds for both geographic domains. In the MedSea (resp. SWIO), the thermal gradient threshold was taken at 0.033 °C/km (resp. 0.035 °C/km), and the backward FSLE threshold at 0.123 day^{-1} (resp. 0.105 day^{-1}). In both the MedSea and the SWIO, the spatial patterns of thermal (Fig. 11a,c) and FSLE-derived (Fig. 11b,d) front frequencies tend to overlap with the mean intensity maps of the same variables (Fig. 10). These results are consistent with the works of

Nieblas *et al.*[11] which used a different front detection method and a different SST product, and Sudre *et al.* when applying the BOA to numerically simulated temperature at different depths in the Mozambique Channel[27].

About MUR SST analysis artefacts.   Some artefacts are visible on thermal gradient mean and frequency figures (Figs. 9a,c, 10a,c, 11a,c). They are most visible in the SWIO, but can also be found in the MedSea. These criss-cross and square-like patterns are characteristics of satellite fingerprints and enhanced by MUR MRVA interpolations[46]. Similar satellite-SST-derived artefacts are also seen in Nieblas *et al.* Figure 8.D[11]. And indeed, the artefacts do not appear when the BOA[16] is applied on numerical simulations instead of remotely-sensed SST[27]. More importantly, they are not visible in NOMAD[51] thermal gradient snapshots (Fig. 2a,c) and therefore will not affect front-marine life interaction analysis. To conclude, since the artefacts are only produced after aggregating thermal gradients on long time-scales, the authors are confident they should not alter the reliability of the NOMAD[51] thermal gradient product for biological use.

Usage Notes

The oceaN frOnt for the Mediterranean seA and southwest inDian ocean dataset (NOMAD[51]) is a multi-decadal record of thermal gradients and backward FSLEs in the Southwest Indian Ocean and the Mediterranean Sea. The NOMAD[51] distinguishes itself from existing alternatives[23,44] by its higher temporal and spatial resolution and extent. NOMAD[51] is suitable for a broad variety of applications where the detection of thermal fronts and attracting Lagrangian coherent structures in the MedSea and the SWIO is needed. These include the identification of spatial and temporal dynamics of thermal and dynamic fronts, along with multidecadal variability analyses. Additionally, the release of our front dataset makes the investigation of the relationship between fronts and marine life possible thanks to a high temporal and spatial resolution.

The provided dataset is divided into 4 folders as described in Table 1. Each folder contains yearly subfolders with daily netcdf files. The netcdf files can be read as is, as they directly provide direct thermal gradient or backward-in-time FSLE fields. We suggest to use Python language software with xarray and dask packages, which are best suited to handle large datasets and High Power Computing. NOMAD[51] will be updated on a yearly basis.

NOMAD visualization.   The authors suggest using the quality flag provided in each dataset to select "suitable" pixels (i.e. flag = 0) before further front analysis. Then front detection is a simple thresholding of the field of interest as illustrated in Fig. 1. Subregion-specific thresholds are added as attributes of each netcdf file, but custom thresholding is also possible. Python xarray package makes filtering and thresholding of the data very intuitive with the function xarray.where. An example Jupyter Notebook for visualization of NOMAD[51] data and the reproduction of Fig. 2 is available on Github: https://github.com/FlorianeSudre/NOMAD_notebooks/.

Front collocation with biological data.   NOMAD[51] is primarily designed to study front-marine life interactions. To this end, the authors recommend using both thermal gradients and FSLE fields, and to interpolate them on the same grid, whose temporal and spatial scale is congruent with the behavior of the specie of interest. Temporal degradation of the dataset is easily performed by xarray.Dataset.resample while spatial regridding can be done with tools such as the Climate Data Operator (CDO) or with the Python package xgcm used in tandem with xarray). The authors recommend comparing quantitatively and qualitatively BOA outputs solely between same-grid products. For example, one can compare thermal gradient magnitudes and patterns within the entire NOMAD[51] thermal gradient dataset, as the SST field comes from a single product. However, while qualitative comparison (i.e. spatial patterns) is possible among thermal gradients originating from different SST fields, a quantitative comparison would require to determine a suitable threshold for each product since BOA algorithm is sensitive to the grid resolution of the input data. Attention should also be brought to the fact that thermal gradients and FSLEs do not have NaNs (where the pixel is deemed "unsuitable", i.e. where the quality flag is not equal to 0) at the same time and place. NaNs appear in FSLE fields near domain boundaries and in coastal areas, while NaNs in thermal gradient fields depend on cloud cover (Fig. 9). The difference in NaN coverage in both dataset should be taken into account when pairing with biological data to avoid any additional bias due to suitable day frequency.

Other possible uses for NOMAD.   Beside the study of front-marine life interactions[10], investigating trophic relationships[70] and the accumulation of fish larvae[71], NOMAD[51] could also be employed to monitor marine pollutants such as microplastics[72,73] and oil spills[74] along fronts. One could also examine relationships between surface thermal fronts and low-level winds at submesoscales from Radar Scatterometer data; as it has only been done at mesoscales so far[75–77]. Finally, following the recent work of Roman-Stork *et al.*[78], it would be interesting to explore new blended metrics based on both SST fronts and FSLEs and their co-variability with key biogeochemical elements (such as Chl-a, pH, nitrates, phosphates) to inform on where oceanic fronts are more likely to impact nutrient cycling in the upper and interior ocean[79].

Code availability

The code we used to produce thermal gradients with the BOA is available in R on Github thanks to Benjamin Galuardi (available on https://github.com/galuardi/boaR, accessed on May 19th, 2022). The code to produce the backward FSLEs belongs to Ismael Hernández-Carrasco, was extensively described in Hernández-Carrasco *et al.*[34], and is available upon request. An example jupyter notebook for front visualization is available on a github repository (https://github.com/FlorianeSudre/NOMAD_notebooks).

References

1. Mahadevan, A. & Archer, D. Modeling the impact of fronts and mesoscale circulation on the nutrient supply and biogeochemistry of the upper ocean. *Journal of Geophysical Research* **105**, 1209–1225, https://doi.org/10.1029/1999jc900216 (2000).
2. Taylor, J. R. & Ferrari, R. Ocean fronts trigger high latitude phytoplankton blooms. *Geophysical Research Letters* **38**, https://doi.org/10.1029/2011gl049312 (2011).
3. Lévy, M., Ferrari, R., Franks, P. J. S., Martin, A. P. & Rivière, P. Bringing physics to life at the submesoscale. *Geophysical Research Letters* **39**, https://doi.org/10.1029/2012gl052756 (2012).
4. Bakun, A. Fronts and eddies as key structures in the habitat of marine fish larvae: opportunity, adaptive response and competitive advantage. *Scientia marina* **70**, 105–122, https://doi.org/10.3989/scimar.2006.70s2105 (2006).
5. Lévy, M., Franks, P. J. S. & Smith, K. S. The role of submesoscale currents in structuring marine ecosystems. *Nature Communications* **9**, 4758, https://doi.org/10.1038/s41467-018-07059-3 (2018).
6. Chapman, C. C., Lea, M.-A., Meyer, A., Sallée, J.-B. & Hindell, M. Defining southern ocean fronts and their influence on biological and physical processes in a changing climate. *Nature Climate Change* **10**, 209–219, https://doi.org/10.1038/s41558-020-0705-4 (2020).
7. Lehahn, Y., d'Ovidio, F., Lévy, M. & Heifetz, E. Stirring of the northeast atlantic spring bloom: A lagrangian analysis based on multisatellite data. *Journal of Geophysical Research* **112**, https://doi.org/10.1029/2006jc003927 (2007).
8. Tzortzis, R. *et al.* Impact of moderately energetic fine-scale dynamics on the phytoplankton community structure in the western mediterranean sea. *Biogeosciences* **18**, 6455–6477, https://doi.org/10.5194/bg-18-6455-2021 (2021).
9. Weimerskirch, H., Le Corre, M., Jaquemet, S., Potier, M. & Marsac, F. Foraging strategy of a top predator in tropical waters: great frigatebirds in the mozambique channel. *Marine Ecology Progress Series* **275**, 297–308, https://doi.org/10.3354/meps275297 (2004).
10. Tew Kai, E. *et al.* Top marine predators track lagrangian coherent structures. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 8245–8250, https://doi.org/10.1073/pnas.0811034106 (2009).
11. Nieblas, A.-E., Demarcq, H., Drushka, K., Sloyan, B. & Bonhommeau, S. Front variability and surface ocean features of the presumed southern bluefin tuna spawning grounds in the tropical southeast indian ocean. *Deep-sea Research. Part II, Topical Studies in Oceanography* **107**, 64–76, https://doi.org/10.1016/j.dsr2.2013.11.007 (2014).
12. Nieto, K., Xu, Y., Teo, S. L. H., McClatchie, S. & Holmes, J. How important are coastal fronts to albacore tuna (thunnus alalunga) habitat in the northeast pacific ocean. *Progress in Oceanography* **150**, 62–71, https://doi.org/10.1016/j.pocean.2015.05.004 (2017).
13. Scales, K. L. *et al.* Mesoscale fronts as foraging habitats: composite front mapping reveals oceanographic drivers of habitat use for a pelagic seabird. *Journal of the Royal Society, Interface* **11**, 20140679, https://doi.org/10.1098/rsif.2014.0679 (2014).
14. Siegelman, L. *et al.* Enhanced upward heat transport at deep submesoscale ocean fronts. *Nature Geoscience* **13**, 50–55, https://doi.org/10.1038/s41561-019-0489-1 (2020).
15. Canny, J. A computational approach to edge detection. *IEEE transactions on pattern analysis and machine intelligence* **8**, 679–698 (1986).
16. Belkin, I. M. & O'Reilly, J. E. An algorithm for oceanic front detection in chlorophyll and sst satellite imagery. *Journal of marine systems: journal of the European Association of Marine Sciences and Techniques* **78**, 319–326, https://doi.org/10.1016/j.jmarsys.2008.11.018 (2009).
17. Cayula, J.-F. & Cornillon, P. Multi-image edge detection for sst images. *Journal of Atmospheric and Oceanic Technology* **12**, 821–829, 10.1175/1520-0426(1995)012<0821:miedfs>2.0.co;2 (1995).
18. Nieto, K., Demarcq, H. & McClatchie, S. Mesoscale frontal structures in the canary upwelling system: New front and filament detection algorithms applied to spatial and temporal patterns. *Remote Sensing of Environment* **123**, 339–346, https://doi.org/10.1016/j.rse.2012.03.028 (2012).
19. Shimada, T., Sakaida, F., Kawamura, H. & Okumura, T. Application of an edge detection method to satellite images for distinguishing sea surface temperature fronts near the japanese coast. *Remote Sensing of Environment* **98**, 21–34, https://doi.org/10.1016/j.rse.2005.05.018 (2005).
20. Torres Arriaza, J. A., Guindos Rojas, F., Peralta Lopez, M. & Canton, M. Competitive neural-net-based system for the automatic detection of oceanic mesoscalar structures on avhrr scenes. *IEEE transactions on geoscience and remote sensing: a publication of the IEEE Geoscience and Remote Sensing Society* **41**, 845–852, https://doi.org/10.1109/tgrs.2003.809929 (2003).
21. Turiel, A., Solé, J., Nieves, V., Ballabrera-Poy, J. & García-Ladona, E. Tracking oceanic currents by singularity analysis of microwave sea surface temperature images. *Remote Sensing of Environment* **112**, 2246–2260, https://doi.org/10.1016/j.rse.2007.10.007 (2008).
22. Tamim, A. *et al.* Detection of moroccan coastal upwelling fronts in sst images using the microcanonical multiscale formalism. *Pattern Recognition Letters* **55**, 28–33, https://doi.org/10.1016/j.patrec.2014.12.006 (2015).
23. Belkin, I. M., Cornillon, P. C. & Sherman, K. Fronts in large marine ecosystems. *Progress in Oceanography* **81**, 223–236, https://doi.org/10.1016/j.pocean.2009.04.015 (2009).
24. Belkin, I. M. Remote sensing of ocean fronts in marine ecology and fisheries. *Remote sensing* **13**, 883, https://doi.org/10.3390/rs13050883 (2021).
25. Park, S. & Chu, P. C. Thermal and haline fronts in the yellow/east china seas: Surface and subsurface seasonality comparison. *Journal of Oceanography* **62**, 617–638, https://doi.org/10.1007/s10872-006-0081-3 (2006).
26. McWilliams, J. C. Oceanic frontogenesis. *Annual Review of Marine Science* **13**, 227–253, https://doi.org/10.1146/annurev-marine-032320-120725 (2021).
27. Sudre, F. *et al.* Spatial and seasonal variability of horizontal temperature fronts in the mozambique channel for both epipelagic and mesopelagic realms. *Frontiers in Marine Science* **9**, https://doi.org/10.3389/fmars.2022.1045136 (2023).
28. Wang, D., Liu, Y., Qi, Y. & Shi, P. Seasonal variability of thermal fronts in the northern south china sea from satellite data. *Geophysical Research Letters* **28**, 3963–3966, https://doi.org/10.1029/2001gl013306 (2001).
29. Lin, L., Liu, D., Luo, O. & Xie, L. Double fronts in the yellow sea in summertime identified using sea surface temperature data of multi-scale ultra-high resolution analysis. *Continental Shelf Research* **175**, 76–86, https://doi.org/10.1016/j.csr.2019.02.004 (2019).
30. Liu, D., Wang, Y., Wang, Y. & Keesing, J. K. Ocean fronts construct spatial zonation in microfossil assemblages. *Global Ecology and Biogeography: a Journal of Macroecology* **27**, 1225–1237, https://doi.org/10.1111/geb.12779 (2018).
31. Aurell, E., Boffetta, G., Crisanti, A., Paladin, G. & Vulpiani, A. Predictability in the large: an extension of the concept of lyapunov exponent. *Journal of physics A: Mathematical and general* **30**, 1–26, https://doi.org/10.1088/0305-4470/30/1/003 (1997).
32. d'Ovidio, F., Fernández, V., Hernández-García, E. & López, C. Mixing structures in the mediterranean sea from finite-size lyapunov exponents. *Geophysical Research Letters* **31**, https://doi.org/10.1029/2004gl020328 (2004).
33. Hernández-Carrasco, I., López, C., Hernández-García, E. & Turiel, A. Seasonal and regional characterization of horizontal stirring in the global ocean. *Journal of Geophysical Research* **117**, https://doi.org/10.1029/2012jc008222 (2012).
34. Hernández-Carrasco, I., López, C., Hernández-García, E. & Turiel, A. How reliable are finite-size lyapunov exponents for the assessment of ocean dynamics. *Ocean Modelling* **36**, 208–218, https://doi.org/10.1016/j.ocemod.2010.12.006 (2011).
35. Bettencourt, J. H., López, C. & Hernández-García, E. Oceanic three-dimensional lagrangian coherent structures: A study of a mesoscale eddy in the benguela upwelling region. *Ocean Modelling* **51**, 73–83, https://doi.org/10.1016/j.ocemod.2012.04.004 (2012).
36. d'Ovidio, F., Isern-Fontanet, J., López, C., Hernández-García, E. & García-Ladona, E. Comparison between eulerian diagnostics and finite-size lyapunov exponents computed from altimetry in the algerian basin. *Deep-sea Research. Part I, Oceanographic Research Papers* **56**, 15–31, https://doi.org/10.1016/j.dsr.2008.07.014 (2009).

37. Bettencourt, J. H., Rossi, V., Hernández-García, E., Marta-Almeida, M. & López, C. Characterization of the structure and cross-shore transport properties of a coastal upwelling filament using three-dimensional finite-size lyapunov exponents. *Journal of Geophysical Research. Oceans* **122**, 7433–7448, https://doi.org/10.1002/2017jc012700 (2017).

38. Hernández-Carrasco, I., Orfila, A. & Rossi, V. & Garçon, V. Effect of small scale transport processes on phytoplankton distribution in coastal seas. *Scientific Reports* **8**, 8613, https://doi.org/10.1038/s41598-018-26857-9 (2018).

39. Mathur, M., David, M. J., Sharma, R. & Agarwal, N. Thermal fronts and attracting lagrangian coherent structures in the north bay of bengal during december 2015–march 2016. *Deep-sea Research. Part II, Topical Studies in Oceanography* **168**, 104636, https://doi.org/10.1016/j.dsr2.2019.104636 (2019).

40. Scales, K. L. *et al*. Fisheries bycatch risk to marine megafauna is intensified in lagrangian coherent structures. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 7362–7367, https://doi.org/10.1073/pnas.1801270115 (2018).

41. Abrahms, B. *et al*. Mesoscale activity facilitates energy gain in a top predator. *Proceedings. Biological sciences* **285**, 20181101, https://doi.org/10.1098/rspb.2018.1101 (2018).

42. Hernández-Carrasco, I. *et al*. Lagrangian flow effects on phytoplankton abundance and composition along filament-like structures. *Progress in Oceanography* **189**, 102469, https://doi.org/10.1016/j.pocean.2020.102469 (2020).

43. Sherman, K. *The large marine ecosystem approach for assessment and management of ocean coastal waters*, 3–16 (Elsevier, 2005).

44. LOCEAN/CLS/CTOH/CNES. Fsle - finite-size lyapunov exponents and orientations of the associated eigenvectors, https://doi.org/10.24400/527896/A01-2022.002 (2021).

45. Bates, A. E. *et al*. Biologists ignore ocean weather at their peril. *Nature* **560**, 299–301, https://doi.org/10.1038/d41586-018-05869-5 (2018).

46. Chin, T. M., Vazquez-Cuervo, J. & Armstrong, E. M. A multi-scale high-resolution analysis of global sea surface temperature. *Remote Sensing of Environment* **200**, 154–169, https://doi.org/10.1016/j.rse.2017.07.029 (2017).

47. Rio, M.-H., Mulet, S. & Picot, N. Beyond goce for the ocean circulation estimate: Synergetic use of altimetry, gravimetry, and *in situ* data provides new insight into geostrophic and ekman currents: Ocean circulation beyond goce. *Geophysical Research Letters* **41**, 8918–8925, https://doi.org/10.1002/2014gl061773 (2014).

48. Longhurst, A. Seasonal cycles of pelagic production and consumption. *Progress in oceanography* **36**, 77–167, https://doi.org/10.1016/0079-6611(95)00015-1 (1995).

49. Longhurst, A. *Ecological geography of the sea*, 2 edn (Academic Press, 2007).

50. Spalding, M. D. *et al*. Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience* **57**, 573–583, https://doi.org/10.1641/b570707 (2007).

51. Sudre, F. *et al*. Nomad: the ocean front dataset for the mediterranean sea and southwest indian ocean. *Sextant* https://doi.org/10.12770/3ea321a1-d9d4-49e5-a592-605b80dec240 (2023).

52. Dodge, K. L., Galuardi, B., Miller, T. J. & Lutcavage, M. E. Leatherback turtle movements, dive behavior, and habitat characteristics in ecoregions of the northwest atlantic ocean. *PloS one* **9**, e91726, https://doi.org/10.1371/journal.pone.0091726 (2014).

53. Merchant, C. J. *et al*. Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Nature Scientific Data* **6**, 223, https://doi.org/10.1038/s41597-019-0236-x (2019).

54. Rohner, C. A. *et al*. Satellite tagging highlights the importance of productive mozambican coastal waters to the ecology and conservation of whale sharks. *PeerJ* **6**, e4161, https://doi.org/10.7717/peerj.4161 (2018).

55. Luschi, P. *et al*. Marine turtles use geomagnetic cues during open-sea homing. *Current Biology* **17**, 126–133, https://doi.org/10.1016/j.cub.2006.11.062 (2007).

56. Weimerskirch, H., Corre, M. L., Kai, E. T. & Marsac, F. Foraging movements of great frigatebirds from aldabra island: Relationship with environmental variables and interactions with fisheries. *Progress in Oceanography* **86**, 204–213, https://doi.org/10.1016/j.pocean.2010.04.003 (2010).

57. Artale, V., Boffetta, G., Celani, A., Cencini, M. & Vulpiani, A. Dispersion of passive tracers in closed basins: Beyond the diffusion coefficient. *Physics of fluids (Woodbury, N.Y.: 1994)* **9**, 3162–3171, https://doi.org/10.1063/1.869433 (1997).

58. Nieblas, A.-E. *et al*. Defining mediterranean and black sea biogeochemical subprovinces and synthetic ocean indicators using mesoscale oceanographic features. *PloS one* **9**, e111251, https://doi.org/10.1371/journal.pone.0111251 (2014).

59. Haller, G. Lagrangian coherent structures. *Annual Review of Fluid Mechanics* **47**, 137–162, https://doi.org/10.1146/annurev-fluid-010313-141322 (2015).

60. Boffetta, G., Lacorata, G., Redaelli, G. & Vulpiani, A. Detecting barriers to transport: a review of different techniques. *Physica D. Nonlinear Phenomena* **159**, 58–70, https://doi.org/10.1016/s0167-2789(01)00330-x (2001).

61. Rousselet, L. *et al*. Large- to submesoscale surface circulation and its implications on biogeochemical/biological horizontal distributions during the outpace cruise (southwest pacific. *Biogeosciences* **15**, 2411–2431, https://doi.org/10.5194/bg-15-2411-2018 (2018).

62. Morales-Márquez, V., Hernández-Carrasco, I., Simarro, G., Rossi, V. & Orfila, A. Regionalizing the impacts of wind- and wave-induced currents on surface ocean dynamics: A long-term variability analysis in the mediterranean sea. *Journal of Geophysical Research. Oceans* **126**, https://doi.org/10.1029/2020jc017104 (2021).

63. Global ocean gridded l4 sea surface heights and derived variables reprocessed. copernicus climate change service (c3s). [dataset], https://doi.org/10.48670/MOI-00145 (2023).

64. Hernández-Carrasco, I. & Orfila, A. The role of an intense front on the connectivity of the western mediterranean sea: The cartagena-tenes front. *Journal of Geophysical Research. Oceans* **123**, 4398–4422, https://doi.org/10.1029/2017jc013613 (2018).

65. Tintore, J., La Violette, P. E., Blade, I. & Cruzado, A. A study of an intense density front in the eastern alboran sea: The almeria–oran front. *Journal of Physical Oceanography* **18**, 1384–1397, 10.1175/1520-0485(1988)018<1384:asoaid>2.0.co;2 (1988).

66. de Ruijter, W. P. M., Ridderinkhof, H., Lutjeharms, J. R. E., Schouten, M. W. & Veth, C. Observations of the flow in the mozambique channel. *Geophysical Research Letters* **29**, 140–1–140–3, https://doi.org/10.1029/2001gl013714 (2002).

67. Halo, I. *et al*. Eddy properties in the mozambique channel: A comparison between observations and two numerical ocean circulation models. *Deep-sea Research. Part II, Topical Studies in Oceanography* **100**, 38–53, https://doi.org/10.1016/j.dsr2.2013.10.015 (2014).

68. Lutjeharms, J. R. E. Three decades of research on the greater agulhas current. *Ocean Sci. Discuss.* **3**, 939–995, https://doi.org/10.5194/osd-3-939-2006 (2006).

69. García-Olivares, A., Isern-Fontanet, J. & García-Ladona, E. Dispersion of passive tracers and finite-scale lyapunov exponents in the western mediterranean sea. *Deep-sea Research. Part I, Oceanographic Research Papers* **54**, 253–268, https://doi.org/10.1016/j.dsr.2006.10.009 (2007).

70. Mangolte, I., Lévy, M., Haëck, C. & Ohman, M. D. Sub-frontal niches of plankton communities driven by transport and trophic interactions at ocean fronts. *Biogeosciences* **20**, 3273–3299, https://doi.org/10.5194/bg-20-3273-2023 (2023).

71. Harrison, C. S., Siegel, D. A. & Mitarai, S. Filamentation and eddy eddy interactions in marine larval accumulation and transport. *Marine ecology progress series* **472**, 27–44, https://doi.org/10.3354/meps10061 (2013).

72. Suaria, G. *et al*. *Dynamics of transport, accumulation, and export of plastics at oceanic fronts*, 1–51 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2021).

73. Pattiaratchi, C. *et al*. Plastics in the indian ocean – sources, transport, distribution, and impacts. *Ocean science* **18**, 1–28, https://doi.org/10.5194/os-18-1-2022 (2022).

74. Fifani, G. *et al*. Drifting speed of lagrangian fronts and oil spill dispersal at the ocean surface. *Remote sensing* **13**, 4499, https://doi.org/10.3390/rs13224499 (2021).

75. O'Neill, L. W., Chelton, D. B. & Esbensen, S. K. Covariability of surface wind and stress responses to sea surface temperature fronts. *Journal of climate* **25**, 5916–5942, https://doi.org/10.1175/jcli-d-11-00230.1 (2012).
76. Renault, L., McWilliams, J. C. & Masson, S. Satellite observations of imprint of oceanic current on wind stress by air-sea coupling. *Scientific reports* **7**, https://doi.org/10.1038/s41598-017-17939-1 (2017).
77. Seo, H. *et al.* Ocean mesoscale and frontal-scale ocean–atmosphere interactions and influence on large-scale climate: A review. *Journal of climate* **36**, 1981–2013, https://doi.org/10.1175/jcli-d-21-0982.1 (2023).
78. Roman-Stork, H. L., Byrne, D. A. & Leuliette, E. W. Mesi: A multiparameter eddy significance index. *Earth and space science (Hoboken, N.J.)* **10**, https://doi.org/10.1029/2022ea002583 (2023).
79. McKee, D. C. *et al.* Biophysical dynamics at ocean fronts revealed by bio-argo floats. *Journal of geophysical research. Oceans* **128**, https://doi.org/10.1029/2022jc019226 (2023).
80. Thyng, K., Greene, C., Hetland, R., Zimmerle, H. & DiMarco, S. True colors of oceanography: Guidelines for effective and accurate colormap selection. *Oceanography* **29**, 9–13, https://doi.org/10.5670/oceanog.2016.66 (2016).

Acknowledgements

Author contributions

F.S. with I.H.C. and V.R. designed the dataset. I.H.C. provided the FSLE code, and guidance for its execution and further analyses of FSLEs. J.S. retrieved the remotely sensed MUR SST data. F.S. produced the front dataset and subsequent analysis of the results. C.M. assisted with the High Performance Computing. V.R., B.D. and V.G. supervised the execution of this work and provided critical feedbacks. F.S. produced the figures and wrote the manuscript with input from all authors. All authors contributed to the article and approved the submitted version.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Additional information

Correspondence and requests for materials should be addressed to F.S. or V.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.