

---

## A fuzzy classification of the hydrodynamic forcings of the Rhone River plume: an application in case of accidental release of radionuclides

Delaval A. <sup>1,\*</sup>, Duffa Celine <sup>1</sup>, Pairaud Ivane <sup>2,3</sup>, Radakovitch O. <sup>1,4</sup>

<sup>1</sup> IRSN (Institut de Radioprotection et de Sûreté Nucléaire), PSE-ENV/SRTE/LRTA, 13115, Saint-Paul-Les-Durance, France

<sup>2</sup> UMR 6523 CNRS, IFREMER, IRD, UBO, Laboratoire D'Océanographie Physique et Spatiale, 29280, Plouzané, France

<sup>3</sup> IFREMER, Centre Méditerranée, Laboratoire Environnement Ressources Provence Azur Corse, F, 83507, La Seyne sur Mer, France

<sup>4</sup> Aix Marseille Univ, CNRS, IRD, INRAE, Coll France, CEREGE, Aix-en-Provence, France

\* Corresponding author : A. Delaval

---

### Abstract :

Assessing and modelling the coastal plume dispersion of nuclearized rivers is strategic in case of accidental releases, but taking into account the variation of main hydrodynamic forcings is challenging. This study uses fuzzy c-mean clustering of a 10 years series of discharge and wind speed at the Rhone River estuary (France) in order to explain the variability of its plume. The method allows to classify the data into 6 scenarios of hydrodynamic forcings that were related to different spatial extensions of the plume, as well as to surface currents measured in-situ. These scenarios were used to simulate the extension and dilution of a radioactive release issued from the river. Based on threshold values of the forcings, a decisional tree is proposed to provide a quick decision tool identifying, in real time, which climatological scenario occurs at the river mouth and the potential plume pattern.

### Highlights

► Analyze of 10 years records of discharge and wind conditions at Rhone River mouth. ► Fuzzy clustering performed to identify six different hydroclimatological trends. ► These trends induce different spreading of the Rhone river plume and contaminants. ► Plume shapes identification allows faster decision in case of accidental releases.

**Keywords :** Coastal plume, Fuzzy c-mean clustering, Accidental Release Scenarios, Rhone River, Radioprotection, Coastal management

## 1.Introduction :

The Rhone River catchment extends over 98000 km<sup>2</sup> and covers one fifth of the French metropolitan territory. It is the main source of particles and freshwater for the Gulf of Lion in the North Western Mediterranean sea (Durrieu De Madron et al., 2000), and all together one of the most important input to the Mediterranean sea (Ludwig et al., 2009).

The Rhone valley also hosts the largest concentration of nuclear power plants in Europe with 4 nuclear power plants in process and a spent fuel reprocessing center, under dismantlement since 1997. Eyrolle et al. (2020) recently synthesized the studies showing that this river carries artificial radionuclides from decades, resulting from authorized releases of low level radioactive liquid wastes and from the export of atmospheric deposits on watersheds consequently to nuclear weapons testing and Chernobyl accident.

France is presently ranked second in the world for the production of nuclear energy, and the total electricity production in the combined regions of Northern, Western and Southern Europe is projected to increase by 2050 ((IAEA, 2019)). Also, the risk of

incident on any kind of nuclear installations is still of concern in France and must be taken into account. As for any river, the transport of artificial radionuclides in case of accidental release occurs both in dissolved and particulate form, depending on the amount of suspended particulate matter and on the chemical properties of the radionuclides, and particularly their distribution coefficient (Tomczak et al., 2019). For the Rhone River, the prediction of dissolved vs particulate fluxes and the associated time scale for transit can be evaluated through numerical modeling (Launay et al., 2019), but the behavior of radionuclides once at sea is clearly less constrained, because it will primarily depend on the forcings governing the shape of the Rhone River plume.

The area of the Rhone river mouth is characterized by a very small tidal amplitude about 30 cm inducing the formation of a sedimentary delta. As usual in this case, the freshwater input forms a thin stratified plume of low salinity water (and higher turbidity) overlying the seawater and extending between 4 and 1000 km<sup>2</sup> (Estournel et al., 1997; Gangloff et al., 2017) with a thickness decreasing seaward (Pairaud et al., 2011; Gangloff et al., 2017). It is preferentially deflected westward in a clockwise orientation running East to West (Reffray et al., 2004) due to the general circulation induced by the Northern Current along the continental slope. Under north-northwest winds, the plume extends offshore towards the southwest, whereas it is pushed to the coast west of the river inlet in case of southeastern winds. Satellite and modelling results have also shown that the plume size increases with river discharge (Frayse et al., 2014; Gangloff et al., 2017). More episodic processes impact the plume pattern such as dense water formation and cascading (Ulses et al., 2008), upwelling cells and marine storms (Millot, 1990; Millot, 1999). As a result, this plume extends far beyond the coastal areas and may covers a large area in the GoL extending from the vicinity of the mouth up to the Cap de Creus at the french-spanish border (Sanchez-Cabeza et al., 1992). It can also reach the Gulf of Fos (Gontier et al., 1992; Charmasson et al., 1999) or the Bay of

Marseille (Pairaud et al., 2011 ; Fraysse et al., 2014) on the eastern side of river mouth. The Gulf of Fos is an important economic area with one of the biggest commercial port in Europe and a large shellfish area, and Marseille is one of the biggest Mediterranean coastal city with one million inhabitants. Due to the oligotrophic nature of the Mediterranean Sea, the region of freshwater influence (ROFI) of the Rhone river has a major influence on the distribution of plankton groups (Diaz et al., 2019) and thus pelagic catches on the GoL. Obviously, the inputs of chemical contaminants from the Rhone River can greatly affect the fishery activity.

The combination of meteorological and hydrodynamic forcings with the dynamics of the Rhone River discharge results in a large spatio-temporal variability in freshwater and associated pollutants delivery to the GoL (Martin et al., 2019). If an accidental release occurs in the Rhone River, the dissolved radionuclides may reach the estuary within 48h hours to few days depending on the source location and water discharge [unpublished results]. Once at sea, the different shapes that the plume may present will depend on hydrodynamic and weather conditions and will lead to contaminate different areas. Since one goal of radioprotection is to predict the transfer of radionuclides in the environment, there is a need to anticipate their dispersion at any time and in any kind of meteorological and hydrodynamical conditions.

Different numerical hydrodynamic models have been set up in the GoL, including the river mouth (Pairaud et al., 2011 ;Duffa et al., 2016), and they could be used actually in case of accidental release in order to predict the behavior of the freshwater input. However, the delay necessary for their implementation will range from few hours to few days, whereas very quick and concise information should be provided to experts and decision-makers as a first picture of the local issues. Also the potentially impacted zones will be better defined by performing a fine spatial scale simulation adequately centered, compared to a large-scale simulation

As a result, a preliminary study embracing all possible plume patterns is necessary, and a first step is to target the general behavior of the estuarine-plume system. Bárcena et al., (2015) explained that two approaches may be conducted for that: simulating several scenarios using constant conditions of hydrodynamic forcings or simulating few scenarios using the most frequent or extreme real hydrodynamic forcings during short-medium term periods (month to year).

These authors demonstrated that the first approach is not complete because real forcings cannot be deduced from the combination of simple idealized scenarios. The second approach relies on a subjective selection of scenarios by an expert and it will have an expensive computational cost for simulations if the need is to get an overview of the different kind of realistic responses of the estuarine-plume mean behavior. In this case, and to minimize subjectivity, a methodology based on data mining should be able to select the most relevant condensed hydrodynamic scenarios, taking into account the time evolution and the occurrence probability of the forcings.

Plume classifications based on satellite observations or hydrodynamic model output have been defined in several river-sea systems using Empirical Orthogonal Function or Self-Organizing-Map (Falcieri et al., 2014; Xu et al., 2019). Such classification method deals with large spatial scale but implies a heavy data pretreatment like « masking » to treat the satellite data or for the computation of the model. In addition, the need of long-term environmental databases (e.g., 10–20 years) to assess probabilities implies significant computational costs as well as long and multiple series of data to be used as boundary conditions and climatic forcings. Another approach is to classify the main hydrodynamics drivers by looking for example at the catchment discharge and the winds intensities and directions (Kaufmann & Whiteman 1999; Zhang et al., 2011). Since the plume response to these forcings can be longer than 24h (Demarcq & Wald, 1984 ; Estournel et al., 1997), the classification should work observation by observation but must also keep consistency

over longer temporal scales of few days in order to be accurate. Clustering performed on temporal series helps to assess the consistency of a trend over time, and a fuzzy clustering algorithm provides a continuous cluster membership function allowing to spot significant trend changes.

In this context, this paper presents a methodology based on statistical analysis and numerical modelling that was developed to address the limitations of the previously mentioned approaches. Firstly, we used a fuzzy c-mean algorithm to identify and classify combinations of winds and discharge at the mouth of the Rhone river in order to define “model scenarios” of realistic forcings. Secondly, the consequences for sea surface currents will be assessed and the resulting plume pattern will be modeled for each scenario, as well as the distribution of dissolved radionuclides due to a hypothetical and episodic release on the Rhone River. These *plumes scenario* can be used as a support for operational tools improvement and decision.

## 2. Material and Methods

### 2.1 Field study and data

The Rhone River hourly discharges have been provided by the C.N.R (Compagnie Nationale du Rhône) thanks to the Rhone Sediment Observatory (OSR program). They were measured at the SORA station, in the city of Arles located 47 km upstream of Rhone River mouth (Fig 1). It must be noted that the Rhone River splits in two branches upstream of this station: the Grand Rhone and Petit Rhone. The station reports the discharge for the Grand Rhone River only, which represents about 90 % of the total Rhone River discharge (Boudet et al., 2017). In our case we focus only on the river plume at the Grand Rhone outlet.

Weather data and subsurface marine currents data are issued from the MesuRho station (Pairaud et al., 2016), operational since June 2009 and located at the Buoy Float Immersed (BFI) maritime buoyage Roustan East (43 ° 19.2 N, 4 ° 52 E) on the Rhone prodelta (20 m water depth). It is about 1 mile southeast of the mouth and was configured to collect physico-chemical data in near real time and at high frequency (about 30 min) in the fresh/marine waters transition zone. It is equipped with a weather station at 10 m height and an Acoustic Current Doppler Profiler (ADCP). The instrumentation is connected by a cable to a controller located above the sea surface and powered by solar panels. The measurements are transmitted to the Coriolis data center via GPRS (about 1 transmission every 12h since 2015, 4h before).

Weather variables used are the average wind speed over 30 min and the gust wind speed. Gust wind speed is the maximal mean wind speed over 0.5 second observed during a period of 30 min.

The observations used were registered between 2009 and 2019 and result in a total of 128262 data. The subsurface currents (maximum depth of 1.5 m) from 2010 to 2019 were also used, when available and after quality control validation, leading to a total of 31826 observations.

In order to perform multivariate analysis and regression, wind and currents variables (expressed in terms of velocities  $u$  and directions  $\theta$ ) are described by an Eastward and a Northward component  $X$  and  $Y$  and calculated as follow :

$$\begin{cases} X = u \cdot \cos(\theta) \\ Y = u \cdot \sin(\theta) \end{cases} \quad (1)$$

The corresponding hourly discharges in Arles (64131 obs.) were shifted with a 24 h delay, which corresponds to the transit time between Arles and the river mouth for a mean liquid discharge.

## 2.2 Principal Component Analysis

Principal Component Analysis (PCA) has been widely used in environmental sciences including hydrologic and hydrodynamics (e.g. Hannah et al., 2000; Paireaud et al., 2008). The common goal to all principal component methods is to describe a data set ( $X$  with  $i$  individuals or observations and  $w$  variables) using a small number ( $p < w$ ) of uncorrelated variables, while retaining as much information (variance) as possible. The reduction is achieved by transforming the data into a new set of continuous variables named the principal components.

The reduction of dimensionality provides a framework to visualize data which is especially important for large datasets (Husson et al., 2010). This facilitates the analyses based on geometrical criteria such as separate observations into  $k$  distinct sub-groups (clustering) or determination of extreme points (Renner, 1993; Napoleon & Pavalakodi, 2011)).

Using PCA as a pre-processing tool in order to cluster presents two additional advantages. The reduction of dimensionality speeds up the convergence of classification algorithms, which usually depends on the square of  $p$  and  $i$  (Ben-Dor et al., 2004), and it reduces the noise, the essential of the information being on the first components whereas the noise is on the lasting ones (Husson et al., 2010). PCA has been performed using the R package “FactoMineR” (Lê et al., 2008).

## 2.3 Fuzzy c-mean algorithm

Clustering is a usual method for data mining when it comes to identify groups and classify individuals, but many algorithms exist and present different results and convergence speeds (Jain et al., 1999). The first goal is to find an algorithm based on geometrical criteria as simple as possible for a more realistic interpretability, and the

second one is to find a fast convergence algorithm in order to treat the important dataset.

The most usual method is the c-mean or k-mean (MacQueen, 1965 ;Yadav & Sharma, 2013) and its fuzzy alternative (Bezdek, 1981 ; Fu Lai & Tong, 1994). C-means are iterative algorithms that classify individuals of a dataset into C groups. The algorithm allows to randomly define C centroids in the same coordinate systems as the individuals. Each individual  $x$  (total of K) is then assigned to the closest centroid center  $c_i$ . The barycenter of each subgroup is then calculated and becomes the new centroid. Again, individuals are reassigned to the closest centroid. This iterative procedure minimizes the objective function (J) and the procedure ends when J reaches an inferior threshold in Eq (2).

$$J = \sum_{i=1}^C \sum_{X_k \in C_i} \|x_k - c_i\|^2 \quad (2)$$

This method is defined as “crisp”, which means that each observation is set to belong to its closest centroid cluster. Consequences are that observations with different distances from the nearest cluster are classified into this cluster without degree of uncertainty and the ambiguity of the data is eliminated.

Cluster boundaries are usually not sharp in environmental sciences (Zadeh et al., 1965), especially when ambiguous data exist, and membership degrees are more realistic than crisp assignments (Klawonn & Höppner, 2003). A priori we do not expect a crisp classification and it is important to have feedback on the confidence of classification for each individual. As a result, the ambiguity of the data can be preserved and his probability can be used later for post treatments (Kim et al., 2011).

The fuzzy alternative introduces two new parameters. The first one is the membership coefficient  $\mu_{ik}$ , the coefficient of the  $k$ th observation to the  $i$ th cluster. This membership

represents how closely the  $k$ th data object ( $x_k$ ) is located from the  $i$ th cluster center. It varies from 0 to 1 depending on the distance ( $\|x_k - c_i\|^2$ ), and a higher membership coefficient indicates stronger association between the  $k$ th data object to the  $i$ th cluster.

$$\mu_{ik} = \left[ \sum_{j=1}^c \left( \frac{\|x_k - c_i\|^2}{\|x_k - c_j\|^2} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (3)$$

The second parameter  $m$  is the fuzziness coefficient. It is greater than 1 and usually dependent on the dataset structure because it represents the degree of overlap of the clusters (Klawonn & Höppner, 2003). If we set  $m$  to a smaller value, more (less) weight is given to the objects that are located closer to (farther from) a cluster center. As  $m$  is close to 1,  $\mu_{ik}$  converges to 0 for the objects that are far from a cluster center, or 1 for those close to a cluster center, which implies less fuzziness (i.e. clearer cut).

The symbol  $\| \cdot \|$  denotes any vector norm that represents the distance between the data object and the cluster center. Here we use the 2-norm (Euclidean norm) which is widely used in the FCM.

The new  $c$ -mean function to minimize becomes:

$$J = \sum_{i=1}^c \sum_{k=1}^K (\mu_{ik})^m \|x_k - c_i\|^2 \quad (4)$$

The robustness brought by the fuzzy approach over the crisp classification is a significant improvement in term of efficiency and convergence. Because each individual (observations) has a probability to belong to each center, centers are adjusted faster and the algorithm converges faster (Fu Lai & Tong, 1994; Ferraro & Giordani, 2015).

Also, without any prior information on the cluster structure (sphericity of clusters, possible overlap) the fuzzy  $c$ -mean provides better results than its crisp counterpart (Selim & Kamel, 1992). As a result, hydrologic and climatologic combinations can be identified by fuzzy  $c$ -mean (Kim et al., 2011; Zhang et al., 2011; Bárcena et al., 2015).

In this study the fuzzy-cmean algorithm is performed using the “e1071” package from R software (Hornik et al., 2019).

#### 2.4 Choice of the number of clusters $C$ and the coefficient of fuzziness $m$

Fuzzy  $c$ -mean algorithm needs to be initialized with the number of clusters  $C$  and the coefficient of fuzziness  $m$ . The best combination of these parameters is not determined by the algorithm. One approach is to run different simulations with different  $\{C, m\}$  pairs and to check the efficiency of clustering with a quality criteria (Ramze Rezaee et al., 1998; Setnes & Babuška, 1999)

Many criteria and their efficiency are available in Wang & Zhang (2007) and Liu et al. (2010). Some have fast calculation like partition coefficient (PC) or partition entropy (PE), but they monotonously decrease with the number of clusters and the lack of direct connection to the geometry of the dataset. Others are more complete but computationally expensive, such as the Dunn index (Dunn, 1974) or the fuzzy silhouette (Campello & Hruschka, 2006), and they could not be calculated with this dataset. The Xie and Beni index (Xie & Beni, 1991) could be calculated based on Eq (5). XB has a direct connection to the geometrical property of dataset because it takes into account both compacity and separation of the clusters. It deals correctly with noisy datasets, size or density variations (Liu et al., 2010).

$$XB = \frac{\sum_{i=1}^C \sum_{k=1}^K \mu_{ik}^m \|x_k - c_i\|^2}{K \min_{i,j} \|c_j - c_i\|^2} \quad (5)$$

Calculation of XB is also fast for our dataset: from 4 to 10 seconds depending on  $\{C, m\}$  pair.

$C$  could be any integer number between 2 and 358, the last one being theoretically the square root of the dataset length (Chaimontree et al., 2010). The fuzzifier  $m$  can be in

theory any real number between 1 and  $\infty$ . In our case, the interval of  $\{C, m\}$  simulations has been restrained based on the following:

- Depending on the river-sea system involved, a different number of plume patterns exist. In literature, we found that a river plume can present up to 8 patterns (Xu et al. 2019). As a result, we do not expect our number of cluster to exceed 8 and the number of clusters  $C$  was set between 2 and 8.

- Previous studies report that values of  $m$  can range from 1 to 4. Most of them use  $m \in [1.5, 2.5]$  and as result  $m$  is usually set to 2 by default (Klawonn & Höppner, 2003). Overall  $m$  is lower for large datasets (Klawonn & Höppner, 2003), and the lower limit will be fix in our case to 1. Using the empirical threshold equation based on the length and dimensions of the dataset proposed by Schwämmle & Jensen (2010), we found that the superior threshold value of  $m$  for our dataset is around 2.5. By safety, this threshold value is increased by +0.25.

As a result, parameter  $m$  will be tested in the interval  $[1; 2.75]$  and  $C$  in the interval  $[2; 8]$ .

### 3. Results and discussion

#### 3.1 Principal Component Analysis

PCA successfully reduced the five original variables (Wind speed toward North and East, Gust wind speed toward North and East and Rhone discharge) into three components and gave a summed variance of 96.4% (Fig1 supplementary material). This is not surprising since the gust wind speed and mean wind speed are correlated due to same direction (Fig 2 supplementary material). The first axis contains 51% of variability with

the information on wind direction. The second axis with 25 % of variability contains the information on wind speed, and the last one (20% variability) corresponds to the Rhone discharge. The “elbow criteria”, the “Kaiser rule” and the interpretation of the components confirm without ambiguity these three components (Fig 1 supplementary material). The lasting 3.6% carried by the two remaining components concern really specific and scarce interactions like the anticorrelation between mean wind speed and gust wind speed.

As a result, 80 and 20% of the variability are due to the variations of winds and liquid discharge respectively.

### 3.2 Clustering results and performances

Fuzzy-c-mean clustering was performed on the 128262 observations and the three main dimensions resulting from PCA. A summary on classification performances based on XB index is shown Fig 2. All configurations performs reasonably well except the one with 3 clusters. An interesting result is that the 2 clusters configuration performed reasonably well, which confirms that the plume dynamics can be described as a first approach by considering only the wind direction that is South East against North West winds. This is in agreement with the 50% of variability held by the wind direction discussed hereunder. However, the configuration selected is the one giving the better result for XB, with 6 clusters optimized at  $m=2.45$  ( $XB=0.19$ ).

### 3.3 Characterization of the scenarios

A cluster gathers observations having close values for one or more variables. These properties on variables are specific to each cluster and are then interpreted hereunder as a scenario.

In order to interpret the clustering and to characterize the resulting scenarios we present the distribution of winds and discharges in Fig 3 and 4, whereas Fig 5 shows the percentage of occurrence of these scenarios for each month. The discharge distribution in each cluster was significantly different from the global distribution of Rhone discharge in Arles based on the Kolmogorov-Smirnov test (see Fig 4).

In the description below, a flood event for the Rhone River refers to a discharge above a threshold set at  $3900\text{m}^3/\text{s}$  in Arles (Boudet et al., 2017). A storm criteria is usually the significant wave height, but this parameter showed too many breaks in the time series transmitted in near real time by the buoy over the 2009-2019 period. Also, we defined sea storm here by using as a threshold the quantile 98 of our offshore gust wind speed dataset (50.3% of total dataset) which is  $27.8\text{ m/s}$  ( $100\text{ km/h}$ ) (Klawe & Ulbrich, 2003).

Cluster 1 gathers South-East winds (paragon  $126^\circ$ ) with  $9.2\text{ m/s}$  mean wind-speed and  $16\text{ m/s}$  mean gust. It contains 86% of all observed sea storm events (with the highest intensity) and 19% of the flood events. The distribution of the hourly water discharge does not characterize this cluster. Observations belonging to this cluster have less than 6% occurrence in July-September, rising up to 23% in October and November. This cluster can be interpreted as moderate to high waves scenario resulting from fresh breeze to violent storm South-East marine winds.

Observations in cluster 2 are winds with velocities around  $4.5\text{m/s}$ , and  $9.2\text{ m/s}$  gust fully coming from the South ( $171^\circ$ ). It contains 12% of all observed sea storm events. It gathers discharges values under  $2500\text{ m}^3/\text{s}$ , with a median at  $960\text{ m}^3/\text{s}$ . The observations mainly occur in August-September with 23% occurrence. This cluster can be interpreted as a Rhone River low flow scenario mainly associated with South-East marine breeze or sometimes a Sirocco wind coming from the South (Reiter, 1975).

Wind observations in cluster 3 show an important variability and are superimposed with clusters 4 and 5. Most representative winds present a mean speed of 8.5 m/s, and gust speed of 13.1 m/s. The Rhone discharge distribution for cluster 3 is very different from the reference distribution (highest Kolmogorov's D statistic). It gathers discharges higher than 2000 m<sup>3</sup>/s and contains most of the flood events (79% of them). This is also the cluster showing the highest contrast in seasonality, with an occurrence up to 33% from November to February, decreasing to 0-2% during the July-October period. Cluster 3 can be interpreted as the "high river flow" scenario with a combination of different winds coming from the North West.

In cluster 4, observations are usually winds with 7.2m/s mean wind-speed and 9.1 m/s mean gust coming from the West (272°). It contains 2% of all observed sea storm events. Discharges are below 2500 m<sup>3</sup>/s with a median around 1070 m<sup>3</sup>/s. These observations mainly occur in July-August-September with 33% of occurrence. A specific point is that their occurrence increases during the afternoon with a peak around 1h AM (Fig 3 supplementary material). Interpretation of this cluster is a Rhone River low flow scenario gathering moderate sea-breeze coming from the South-West (Cros et al., 2004), with sometimes a strong onshore gale from West.

Cluster 5 corresponds to winds with 12.3 m/s average speed and 18.6 m/s gust coming from a restricted area in the North (325°). The corresponding water discharge distribution is on the lower part of the global distribution (median of 1140 m<sup>3</sup>/s) and discharges are always below 3000 m<sup>3</sup>/s. The monthly occurrence is stable (15%) with a peak in February at 25%. The strong average wind intensities and gust speeds (highest at 340 °) combined with the restricted wind direction parallel to the Rhone valley stand

for the characteristics of the Mistral wind (Reiter, 1975). As a result, cluster 5 can be interpreted as a Mistral wind scenario (dry and strong breeze to strong gale) associated with low to moderate discharges.

Cluster 6 usually gathers winds with 5.7 m/s average speed, and gusts of 12.5 m/s coming from the North East (11°). However, winds coming from 340-360° (North/North-West) are also observed. The related hourly water discharges distribution is in the lower part of the global one (median of 1230 m<sup>3</sup>/s) and few discharges higher than 3000 m<sup>3</sup>/s are observed. Occurrence of cluster 6 observations is very stable all along the year, ranging between 14 and 18 %. This cluster presents the largest gap between the wind speed average and the gust wind speed, and shows an increasing occurrence in the early morning. The highest gusts reach 50 m/s and occur episodically in winter with an origin from 10 to 60° (North-East). These are the strongest gusts observed among all scenarios. We interpret cluster 6 as a scenario gathering land breeze or valley flow during summer and winds channeled by Pre Alps mountains (Cros et al., 2004; Duine et al., 2017) which become stronger in winter (“orsure” according Reiter, 1975)).

### 3.4 Consequences for surface currents

Six clear wind/discharge patterns have been identified, but did they correspond or induce different hydrodynamics responses of the surface currents in the vicinity of the Rhone River mouth? Consequences for subsurface currents observations issued from the ADCP on the MesuRho station are investigated through a least squares multiple regression. Observation membership to clusters  $C_i$  are the explanatory variables and currents in Eastward and Northward directions are the response variables.

Estimators  $X_i$  and  $Y_i$  are then used to calculate current orientation  $\theta$  (rad) and speed  $u$  (m/s) observed on each cluster with equation (1).

Confidence intervals are calculated with the robust White standard errors with “lmtest” package (White, 1980; Hothorn et al., 2019) to avoid heteroskedasticity and underestimation of confidence intervals. The main current direction for each cluster obtained by least square regression on memberships is presented on figure 6 (right) along with the global current rose (left).

Currents oriented at  $280^\circ$  correspond to scenario 1 and are in agreement with the more general modelling and satellite observations during similar south easterlies wind conditions showing the plume tackled to the Camargue coast (Marsaleix et al., 1998; Gangloff et al., 2017). Scenario 2 presents small currents not related to the wind direction. In this case, winds are probably too low and currents are driven by the general circulation, which has a current speed similar to those of this scenario: 10 cm/s.

In scenario 3 the current direction correspond to those at the Rhone River mouth, meaning that during high water events (discharges superior to  $2500 \text{ m}^3/\text{s}$ ) the river influence becomes significant.

Scenario 4, 5 and 6 seem to follow the surface Ekman transport, with a deflection to the left relative to the wind direction. Scenario 5 is the one presenting the largest interval of confidence despite having the straightest wind distribution. A closer look at the data shows that, in this scenario, the currents deeper than 1.2 m present an important heterogeneity in their direction. However, for wind average speeds superior to 15 m/s and gust wind speeds over 25 m/s this heterogeneity does no longer exist and all currents are oriented in a  $150^\circ$  direction. For comparison, scenario 5 paragon is an average wind speed of 12.3 m/s and gust wind speed of 18.6 m/s, values which are inferior to the two thresholds and may explain the currents discrepancies associated with this scenario.

To conclude, each scenario has its own current direction and intensity, statistically different and significant.

### 3.5 Application

The main objective of this work is to define the general trends of dispersal in the GoL that can be expected in the case of artificial radionuclides release within the river. Since releases may occur at any time in a year, we modeled the dispersion of a radioactive plume in the GoL for each of the previous hydrodynamic scenario, in order to get an overview of the potential impacts, whatever the hydrodynamic and climatological conditions.

The simulation code used at IRSN for the marine area is STERNE (“Simulation du Transport et du transfert d'Eléments Radioactifs dans l'environnement marin”, or “Simulation of radionuclide transport and transfer in marine environments”). It was designed to assess the radiological impact of accidental releases affecting the marine environment. Eulerian radionuclide dispersion is calculated using a tracer advection diffusion equation. More details on the code can be found in Duffa et al. (2016).

We use the 2010 hydrodynamic outputs provided by IFREMER with its MARS3D model implemented on the North-western Mediterranean Sea (Nicolle et al., 2009). The simulations assumed a release of 1 TBq of  $^{137}\text{Cs}$  dissolved activity in the river, over a temporal window of 48 h.  $^{137}\text{Cs}$  was chosen because this radionuclide is released at each nuclear accident and is also found in authorized releases from nuclear powerplants. This radionuclide presents a high radiotoxicity (Garnier-Laplace et al., 2011) and is relatively soluble in seawater with a  $K_d$  ranging from 450 to 2000 L/kg (Delaval et al., 2020). For comparison, the estimated average direct discharge of  $^{137}\text{Cs}$  to the ocean during the Fukushima Daiichi nuclear power plants was around 5000 TBq (Buesseler et al., 2017).

The simulations were done for each hydro-meteorological scenario defined on the basis of the 2009-2019 dataset and the plume extension in the GoL was modelled for each

scenario. Since the hydrodynamic inputs are only available for the year 2010, we selected in this input the most representative temporal window for each scenario simulation by integrating observations memberships over a sliding window of 48h. The temporal window presenting the highest summed membership values were selected for each cluster.

The results of the 6 simulations are very different in terms of plume shape and thus affected areas (Fig 7). Mean winds and discharges conditions over 48 h hours are indicated in the figure. It must be noted that these values are specific to the chosen temporal windows, and thus can be different from the parangons presented in the previous chapter.

In scenario 1 (strong marine wind with moderate discharge conditions), the plume is constrained to the coast and extends west in agreement with the currents at the buoy. This scenario has already been highlighted by Demarcq & Wald, (1984) and Many et al., (2018) or modelled by Estournel et al., (1997). A part of the activity remains blocked in the estuary due to winds in opposition with its flowing path and an increase in sea level at the mouth limiting the power of the jet.

The lowest expansion of the plume is observed with scenario 2 (weak wind with discharge slightly under the annual mean). The plume has the lowest surface spatial expansion among the 6 plumes and is nearly stagnant and remains with a high activity. This is in agreement with the currents observed at the buoy showing really low speeds. This scenario appears mostly during summer (Fig 4), and satellite images confirmed that the turbid plume present effectively its smallest area at this period (Gangloff et al. (2017).

Scenario 3 corresponds to a high Rhone River discharge and northwest winds conditions. The plume presents a large area but the northwesterly wind is powerful enough to carry

the plume offshore. For different wind stress simulations, Marsaleix et al. (1998) showed that winds around 30 km/h were sufficient to detach the plume from the coast. According to this author this threshold is independent of Rhone River discharge.

In scenario 4 (strong westerly wind and low discharge conditions), the plume extends over a large area favored by the presence of a summer stratification at low discharge. A part of the plume at the latitude of the buoy can be deflected eastwards in the Gulf of Fos, in agreement with current data at the buoy. According to Fraysse et al., (2014) this plume shape is the first step toward an intrusion in the Bay of Marseille if this scenario is followed by south east winds conditions. To note, such intrusions occurred in summer (Fraysse et al, 2014) when the probability of occurrence of this scenario are the highest.

In scenario 5, with a strong northwest wind (Mistral) and low discharge conditions the plume stands out from the coast as shown by Demarcq & Wald, (1984) and Gangloff et al. (2017) and modelled by Estournel et al. (1997). This case is favorable to an export of the plume far away from the coast, even at low discharge conditions and moderate Mistral (but the wind speed are however above the 30 km/h threshold proposed by Marsaleix et al. (1998)).

Finally, the northeast wind and moderate discharge conditions of scenario 6 maintain the plume towards the coast, similarly to scenario 1. It is here again in agreement with the currents measured at the buoy. This is confirmed by Gangloff et al. (2017) who showed that for the most northern winds (higher than  $340^\circ$ ) the plume tends to be tacked to the coast. The part of the plume going East towards the gulf of Fos is quite unexpected. A closer look at the dataset shows that this temporal window of 48 h, despite being classified as scenario 6, presented 16 consecutives observations (8 hours period) corresponding to scenario 2. This explains why the plume extends toward the gulf of Fos in a similar fashion as in scenario 2.

All these simulations shown that the activity plume may be maintained along the western coast of the Rhone river outlet (Scenarios 1 and 6), its eastern side (Scenario 4) or can extend far away from the coast (Scenario 5). Plume of different extents can also remain nearly stable (Scenarios 2 and 3) until a change of hydrodynamic, conditions and thus scenario (most likely shifts from 2 to 4, 5, 6 and from 3 to 5, 6, 1).

Each cluster of hydrodynamic forcings has thus its own patterns for Rhone River plume spreading in the GoL. Their hydrodynamic variables and resulting trends were presented in the previous chapters, and these results now allow to better evaluate the risk of propagation of  $^{137}\text{Cs}$  activity. However, it is also interesting to define thresholds values for these variables, in order to be able to select the most appropriate scenario to apply in case of alert on accidental release of radionuclides (or any kind of chemical contaminants). To summarize the differences obtained between the scenarios, a simplified decision tree has been constructed (Fig 8). From top to bottom it allows to outline practical separation criteria (Fig 8).

The classification of these 6 scenarios is based on the wind (and gust) direction, wind speed and water discharge. These *in situ* conditions can thus be associated with a scenario in near real-time, as they are available at this time scale from the websites of Coriolis Cotier (wind) and Vigicrues (river flow). The tree reproduces the classification using 80% of the hydrodynamic raw data (without PCA treatment) as training, and 20% as a validation set. It allows a fast crisp classification into one of the 6 established scenario with 83% of accuracy on both trained and tested data. As an example, in case of wind direction of  $270^\circ$  and a river discharge of 2500 m<sup>3</sup>/s, the shape of the plume will

correspond to scenario 3. Currents directions at the buoy (Coriolis Cotier) can also be used as an additional verification.

## Conclusion

In this paper, a 10-year period was considered in order to identify the main combinations of hydrodynamic forcings (wind and Rhone River discharge) using a fuzzy c-mean clustering. These combinations, called scenario summarized mean shelf behavior providing a very important information to estimate and understand the Rhone plume patterns in case of accidental release. In addition, existence of observations memberships allowed to spot the best temporal windows to run simulations covering all possible patterns.

6 scenarios have been identified and simulations showed that the plume behavior was different for each of them. These plume patterns are more or less critical in terms of radiologic risks regarding the areas affected and the dilution of the activity. If necessary, wind speed measured on one point are sufficient to extrapolate plume shape for 48 hours on this zone, and the surface currents measured at the Roustan buoy will give a first idea of plume orientation. This study provided a first global picture of main Rhone River plume patterns and consequences for radionuclides accidental releases, but the methodology may be applied to other estuaries.

## Content of the Paper

Fig 1 : Map of the GoL (Mediterranean sea) and Rhone River estuary indicating the locations of discharge and weather stations used in the fuzzy-clustering algorithm. Scheme of MesuRho buoy devices Weather station + PAR (1), solar panels and control and transmission automat ABIN (2), Multiparameters sonds (3 and 6) nitrates captor ISUS (4), benthic station (5), ADCP (7).

Fig 2 : Xie and Beni index values for different  $\{C,m\}$  simulations with fuzzy c-mean algorithm. A lower index indicates a better classification.

Fig 3: Grey dots: mean wind origin and speed (m/s) over 30 min from 2009 to 2019 measured at the Mesurho buoy, in front of the Rhone river. Color dots: cluster mean wind and speed sub-distributions.

Fig 4 : Rhône river hourly discharge (SORA, Arles) from 2009 to 2019 distribution (grey) and cluster Rhône discharge sub-distributions. Results from two-sample Kolmogorov-Smirnov test are shown. D statistic indicates how the distribution in each cluster is different from the reference distribution (all Rhône discharge values) Higher values of D indicates larger differences.

Fig 5: Monthly occurrence of the 6 different scenarios over the 2009-2019 period.

Fig 6: Surface currents (depth<1.5m) direction and module measured at the Mesurho (a) buoy and main direction (thick black dashes) and current speed for each scenario (b) obtained by least square regression. Confidence intervals are for  $\alpha=0.05$

Fig 7 : Radioactive surface plume shape in the GoL under the most representatives temporal windows of the 6 scenarios for a released activity of 1TBq of  $^{137}\text{Cs}$  in 48 hours. Wind orientation is shown by the yellow arrows.

Fig 8 : Simplified decision tree for scenario identification

## Acknowledgements

The authors are indebted to the Institute for Radiological Protection and Nuclear Safety (IRSN) and to Region Sud (Provence-Alpes-Côte d'Azur) authorities for the PhD funding. This study was conducted within the Rhône Sediment Observatory (OSR) program, a multi-partner research program funded through Plan Rhône of the European Regional Development Fund (ERDF), Agence de l'Eau Rhône Méditerranée Corse, CNR, EDF and three regional councils (Région Auvergne-Rhône-Alpes, PACA and Occitanie). The Mesurho station is part of the COAST-HF network and associates IFREMER, IRSN, CNRS, CETMEF, and CEREMA..

## References

- Bárcena, J. F., Camus, P., García, A., & Álvarez, C. (2015). Selecting model scenarios of real hydrodynamic forcings on mesotidal and macrotidal estuaries influenced by river discharges using K-means clustering. *Environmental Modelling and Software*, *68*, 70–82. <https://doi.org/10.1016/j.envsoft.2015.02.007>
- Ben-Dor, A., Shamir, R., & Yakhini, Z. (2004). Clustering Gene Expression Patterns. *Journal of Computational Biology*, *6*(3–4). <https://doi.org/https://doi.org/10.1089/106652799318274>
- Bezdek, J. C. (1981). *Pattern recognition with Fuzzy Objective Function Algorithms* (Springer (ed.); Plenum Pre). <https://doi.org/10.1007/978-1-4757-0450-1>
- Boudet, L., Sabatier, F., & Radakovitch, O. (2017). Modelling of sediment transport pattern in the mouth of the Rhone delta: Role of storm and flood events. *Estuarine, Coastal and Shelf Science*, *198*, 568–582. <https://doi.org/10.1016/j.ecss.2016.10.004>
- Buesseler, K., Dai, M., Aoyama, M., Benitez-Nelson, C., Charmasson, S., Higley, K., Maderich, V., Masqué, P., Morris, P. J., Oughton, D., & Smith, J. N. (2017). Fukushima Daiichi–Derived Radionuclides in the Ocean: Transport, Fate, and Impacts. *Annual Review of Marine Science*, *9* (1), 173–203. <https://doi.org/10.1146/annurev-marine-010816-060733>
- Campello, R. J. G. B., & Hruschka, E. R. (2006). A fuzzy extension of the silhouette width criterion for cluster analysis. *Fuzzy Sets and Systems*, *157*(21), 2858–2875. <https://doi.org/10.1016/j.fss.2006.07.006>
- Chaimontree, S., Atkinson, K., & Coenen, F. (2010). Advanced Data Mining and Applications Part 1. In Springer (Ed.), *Advanced Data Mining and Applications Part 1* (p. 625).
- Charmasson, S., Barker, E., Calmet, D., Pruchon, A. S., & Thébault, H. (1999). Long-

- term variations of man-made radionuclide concentrations in a bio-indicator *Mytilus galloprovincialis* from the French Mediterranean coast. *Science of the Total Environment*, 237–238, 93–103. [https://doi.org/10.1016/S0048-9697\(99\)00127-8](https://doi.org/10.1016/S0048-9697(99)00127-8)
- Cros, B., Durand, P., Cachier, H., Drobinski, P., Fréjafon, E., Kottmeier, C., Perros, P. E., Peuch, V. H., Ponche, J. L., Robin, D., Saïd, F., Toupance, G., & Wortham, H. (2004). The ESCOMPTE program: An overview. *Atmospheric Research*, 69 (3–4), 241–279. <https://doi.org/10.1016/j.atmosres.2003.05.001>
- Delaval, A., Duffa, C., & Radakovitch, O. (2020). A review on cesium desorption at the freshwater-seawater interface. *Journal of Environmental Radioactivity*, 218, 106255. <https://doi.org/10.1016/j.jenvrad.2020.106255>
- Demarcq, H., & Wald, L. (1984). La dynamique superficielle du panache du Rhône d'après l'imagerie infrarouge satellitaire. *Oceanologica Acta*, 7(2), 159–162. <http://hal.archives-ouvertes.fr/hal-00464192/>
- Diaz, F., Bănar, D., Verley, P., & Shin, Y. J. (2019). Implementation of an end-to-end model of the Gulf of Lions ecosystem (NW Mediterranean Sea). II. Investigating the effects of high trophic levels on nutrients and plankton dynamics and associated feedbacks. *Ecological Modelling*, 405, 51–68. <https://doi.org/10.1016/j.ecolmodel.2019.05.004>
- Duffa, C., Bailly, P., Caillaud, M., Charmasson, S., & Renaud, P. (2016). Development of emergency response tools for accidental radiological contamination of French coastal areas. *Journal of Environmental Radioactivity* 151, 487–494. <https://doi.org/10.1016/j.jenvrad.2015.04.019>
- Duine, G. J., Hedde, T., Roubin, P., Durand, P., Lothon, M., Lohou, F., Augustin, P., & Fourmentin, M. (2017). Characterization of valley flows within two confluent valleys under stable conditions: observations from the KASCADE field experiment.

- Quarterly Journal of the Royal Meteorological Society*, 143(705), 1886–1902.  
<https://doi.org/10.1002/qj.3049>
- Dunn, J. C. (1974). Well-Separated Clusters and Optimal Fuzzy Partitions. *Journal of Cybernetics*, 4(1), 95–104. <https://doi.org/10.1080/01969727408546059>
- Durrieu De Madron, X., Abassi, A., Heussner, S., Monaco, A., Aloisi, J. C., Radakovitch, O., Giresse, P., Buscail, R., & Kerherve, P. (2000). Particulate matter and organic carbon budgets for the Gulf of Lions (NW Mediterranean). *Oceanologica Acta*, 23(6), 717–730. [https://doi.org/10.1016/S0399-1784\(00\)00119-5](https://doi.org/10.1016/S0399-1784(00)00119-5)
- Estournel, C., Kondrachoff, V., Marsaleix, P., & Vehil, R. (1997). The plume of the Rhone: Numerical simulation and remote sensing. *Continental Shelf Research*, 17(8), 899–924. [https://doi.org/10.1016/S0278-4343\(96\)00064-7](https://doi.org/10.1016/S0278-4343(96)00064-7)
- Eyrolle, F., Lepage, H., Antonelli, C., Morereau, A., Cossonnet, C., Boyer, P., & Gurriaran, R. (2020). Radionuclides in waters and suspended sediments in the Rhone River (France) - Current contents, anthropic pressures and trajectories. *Science of The Total Environment*, 137873. <https://doi.org/10.1016/j.scitotenv.2020.137873>
- Falcieri, F. M., Benetazzo, A., Sclavo, M., Russo, A., & Carniel, S. (2014). Po River plume pattern variability investigated from model data. *Continental Shelf Research*, 87, 84–95. <https://doi.org/10.1016/j.csr.2013.11.001>
- Ferraro, M. B., & Giordani, P. (2015). A toolbox for fuzzy clustering using the R programming language. *Fuzzy Sets and Systems*, 279, 1–16. <https://doi.org/10.1016/j.fss.2015.05.001>
- Frayse, M., Pairaud, I., Ross, O. N., Faure, V. M., & Pinazo, C. (2014). Journal of Geophysical Research: Oceans Generation processes and impacts on ecosystem

- functioning. *Journal of Geophysical Research: Oceans*, 6535–6556.  
<https://doi.org/10.1002/2014JC010022>
- Fu Lai, C., & Tong, L. (1994). *Fuzzy Competitive Learning*. 7(3), 539–551.
- Gangloff, A., Verney, R., Doxaran, D., Ody, A., & Estournel, C. (2017). Investigating Rhône River plume (Gulf of Lions, France) dynamics using metrics analysis from the MERIS 300m Ocean Color archive (2002–2012). *Continental Shelf Research*, 144, 98–111. <https://doi.org/10.1016/j.csr.2017.06.024>
- Garnier-Laplace, J., Beaugelin-Seiller, K., & Hinton, T. G. (2011). Fukushima wildlife dose reconstruction signals ecological consequences. *Environmental Science and Technology*, 45(12), 5077–5078. <https://doi.org/10.1021/es201637c>
- Gontier, G., Grenz, C., Calmet, D., & Sacher, M. (1992). The contribution of *Mytilus* sp. in radionuclide transfer between water column and sediments in the estuarine and delta systems of the Rhône river. *Estuarine, Coastal and Shelf Science*, 34(6), 593–601. [https://doi.org/10.1016/S0272-7714\(05\)80064-2](https://doi.org/10.1016/S0272-7714(05)80064-2)
- Hannah, D. M., Smith, B. P. G., Gurnell, A. M., & McGregor, G. R. (2000). An approach to hydrograph classification. *Hydrological Processes*, 14(2), 317–338.  
[https://doi.org/10.1002/\(SICI\)1099-1085\(20000215\)14:2<317::AID-HYP929>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1099-1085(20000215)14:2<317::AID-HYP929>3.0.CO;2-T)
- Meyer D., Dimitriadou E., Hornik, K., Weingessel, A., Leisch, F., (2019). *Package ‘e1071.’*
- Hothorn, T., Zeileis, A., Farebrother, R. W., Cummins, C., Millo, G., & Mitchell, D. (2019). *Testing Linear Regression Models (Package ‘lmtest’)*.
- Husson, F., Josse, J., & Pagès, J. (2010). *Technical Report . Principal component methods - hierarchical clustering-partitional clustering: why would we need to choose for visualizing data ? Agrocampus.*

- International Atomic Energy Agency. (2019). Energy, Electricity and Nuclear Power Estimates for the Period up to 2050. *Reference Data Series No. 1*
- Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys*, *31*(3), 264–323. <https://doi.org/10.1145/331499.331504>
- Kaufmann, P., & Whiteman, C. D. (1999). Cluster-analysis classification of wintertime wind patterns in the Grand Canyon region. *Journal of Applied Meteorology*, *38*(8), 1131–1144. [https://doi.org/10.1175/1520-0450\(1999\)038<1131:cacoww>2.0.co;2](https://doi.org/10.1175/1520-0450(1999)038<1131:cacoww>2.0.co;2)
- Kim, H. S., Kim, J. H., Ho, C. H., & Chu, P. S. (2011). Pattern classification of typhoon tracks using the fuzzy c-means clustering method. *Journal of Climate*, *24*(2), 488–508. <https://doi.org/10.1175/2010JCLI3751.1>
- Klawa, M., & Ulbrich, U. (2003). A model for the estimation of storm losses and the identification of severe winter storms in Germany. *Natural Hazards and Earth System Science*, *3*(6), 725–732. <https://doi.org/10.5194/nhess-3-725-2003>
- Klawonn, F., & Höppner, F. (2003). What is fuzzy about fuzzy clustering? Understanding and improving the concept of the fuzzifier. In Springer (Ed.), *Advances in Intelligent Data Analysis* (Vol. 1716, pp. 254–263). <https://doi.org/10.1108/ir.1999.04926fae.001>
- Kovács, J., Bodnár, N., & Török, Á. (2016). The application of multivariate data analysis in the interpretation of engineering geological parameters. *Open Geosciences*, *8*(5), 52–61. <https://doi.org/10.1515/geo-2016-0005>
- Launay, M., Dugué, V., Faure, J. B., Coquery, M., Camenen, B., & Le Coz, J. (2019). Numerical modelling of the suspended particulate matter dynamics in a regulated river network. *Science of the Total Environment*, *665*, 591–605. <https://doi.org/10.1016/j.scitotenv.2019.02.015>

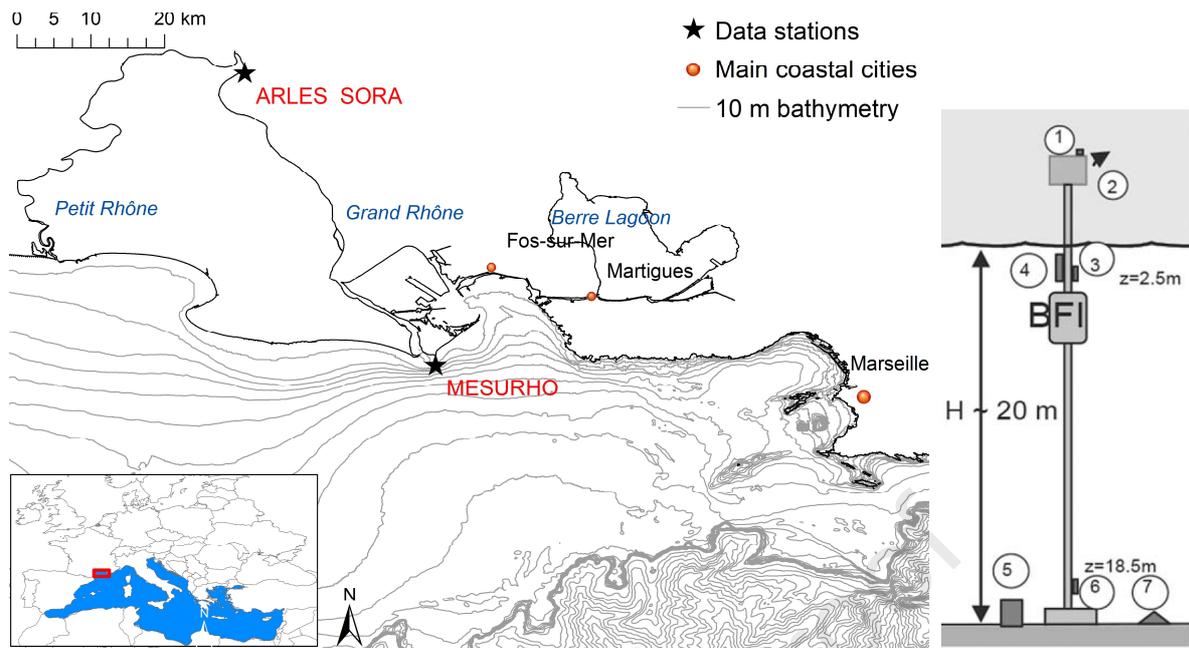
- Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, *25*(1), 1–18.  
<https://doi.org/10.18637/jss.v025.i01>
- Liu, Y., Li, Z., Xiong, H., Gao, X., & Wu, J. (2010). Understanding of internal clustering validation measures. *Proceedings - IEEE International Conference on Data Mining, ICDM*, 911–916. <https://doi.org/10.1109/ICDM.2010.35>
- Ludwig, W., Dumont, E., Meybeck, M., & Heussner, S. (2009). River discharges of water and nutrients to the Mediterranean and Black Sea: Major drivers for ecosystem changes during past and future decades? *Progress in Oceanography*, *80*(3–4), 199–217. <https://doi.org/10.1016/j.pocean.2009.02.001>
- MacQueen, J. (1965). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 281–297.
- Many, G., Bourrin, F., Durrieu de Madron, X., Ody, A., Doxaran, D., & Cauchy, P. (2018). Glider and satellite monitoring of the variability of the suspended particle distribution and size in the Rhône ROFI. *Progress in Oceanography*, *163*, 123–135.  
<https://doi.org/https://doi.org/10.1016/j.pocean.2017.05.006>
- Marsaleix, P., Estournel, C., Kondrachoff, V., & Vehil, R. (1998). A numerical study of the formation of the Rhone River plume. *Journal of Marine Systems*, *14*(1–2), 99–115. [https://doi.org/10.1016/S0924-7963\(97\)00011-0](https://doi.org/10.1016/S0924-7963(97)00011-0)
- Martin, D., Pititto, F., Gil, J., Mura, M. P., Bahamon, N., Romano, C., Thorin, S., Schvartz, T., Dutrieux, É., & Bocquet, Y. (2019). Long-distance influence of the Rhône River plume on the marine benthic ecosystem: Integrating descriptive ecology and predictive modelling. *Science of the Total Environment*, *673*, 790–809.  
<https://doi.org/10.1016/j.scitotenv.2019.04.010>

- Millot, C. (1990). The Gulf of Lions' hydrodynamics. *Continental Shelf Research*, 10(9–11), 885–894. [https://doi.org/10.1016/0278-4343\(90\)90065-T](https://doi.org/10.1016/0278-4343(90)90065-T)
- Millot, C. (1999). Circulation in the Western Mediterranean Sea. *Journal of Marine Systems*, 20(1–4), 423–442. [https://doi.org/10.1016/S0924-7963\(98\)00078-5](https://doi.org/10.1016/S0924-7963(98)00078-5)
- Napoleon, D., & Pavalakodi, S. (2011). A New Method for Dimensionality Reduction Using KMeans Clustering Algorithm for High Dimensional Data Set. *International Journal of Computer Applications*, 13(7), 41–46. <https://doi.org/10.5120/1789-2471>
- Nicolle, A., Garreau, P., & Liorzou, B. (2009). Modelling for anchovy recruitment studies in the Gulf of Lions (Western Mediterranean Sea). *Ocean Dynamics*, 59(6), 953–968. <https://doi.org/10.1007/s10236-009-0221-6>
- Pairaud, I. L., Gatti, J., Bensoussan, N., Verney, R., & Garreau, P. (2011). Hydrology and circulation in a coastal area off Marseille: Validation of a nested 3D model with observations. *Journal of Marine Systems*, 88(1), 20–33. <https://doi.org/10.1016/j.jmarsys.2011.02.010>
- Pairaud, I. L., Lyard, F., Auclair, F., Letellier, T., & Marsaleix, P. (2008). Dynamics of the semi-diurnal and quarter-diurnal internal tides in the Bay of Biscay. Part 1: Barotropic tides. *Continental Shelf Research*, 28(10–11), 1294–1315. <https://doi.org/10.1016/j.csr.2008.03.004>
- Pairaud, I., Repecaud, M., Ravel, M., Fuchs, R., Arnaud, M., Champelovier, A., Rabouille, C., Bombled, B., Toussaint, F., Garcia, F., Raimbault, P., Verney, R., Meule, S., Gaufres, P., Bonnat, A., & Cadiou, J. (2016). Plateforme instrumentée de suivi des paramètres environnementaux à l'embouchure du Rhône. In *Mesures à haute résolution dans l'environnement marin côtier* (CNRS Alpha, pp. 73–87).
- Ramze Rezaee, M., Lelieveldt, B. P. F., & Reiber, J. H. C. (1998). A new cluster validity

- index for the fuzzy c-mean. *Pattern Recognition Letters*, 19(3–4), 237–246.  
[https://doi.org/10.1016/S0167-8655\(97\)00168-2](https://doi.org/10.1016/S0167-8655(97)00168-2)
- Reffray, G., Fraunié, P., & Marsaleix, P. (2004). Secondary flows induced by wind forcing in the Rhône region of freshwater influence. *Ocean Dynamics*, 54(2), 179–196.  
<https://doi.org/10.1007/s10236-003-0079-y>
- Reiter, E. R. (1975). Technical report. *Weather phenomena of the mediterranean basin Part I: General description of the meteorological processes*. Environmental Prediction Research Facility, Naval Postgraduate School, Monterey
- Renner, R. M. (1993). The Resolution of a Compositional Data Set into Mixtures of a Fixed Source Compositions. *Applied Statistics*, 42(4), 615–631.  
<https://doi.org/https://doi.org/10.2307/2986179>
- Sanchez-Cabeza, J.-A., Molero, J., Merino, J., Pujol, L., & Mitchell, P. I. (1992). 137 Cs as a tracer of the Catalan current. *Oceanologica Acta*, 18(2), 221–226.
- Schwämmle, V., & Jensen, O. N. (2010). A simple and fast method to determine the parameters for fuzzy c-means cluster analysis. *Bioinformatics*, 26(22), 2841–2848.  
<https://doi.org/10.1093/bioinformatics/btq534>
- Selim, S. Z., & Kamel, M. . (1992). *On the mathematical and numerical properties of the fuzzy c-means algorithm*. 49, 181–191.
- Setnes, M., & Babuška, R. (1999). Fuzzy relational classifier trained by fuzzy clustering. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 29(5), 619–625. <https://doi.org/10.1109/3477.790444>
- Tomczak, W., Boyer, P., Krimissa, M., & Radakovitch, O. (2019). *Kd distributions in freshwater systems as a function of material type , mass- volume ratio , dissolved organic carbon and pH*. *Applied Geochemistry* 105, 68–77.

- Ulses, C., Estournel, C., Puig, P., Durrieu de Madron, X., & Marsaleix, P. (2008). Dense shelf water cascading in the northwestern Mediterranean during the cold winter 2005: Quantification of the export through the Gulf of Lion and the Catalan margin. *Geophysical Research Letters*, *35*(7), 2–7. <https://doi.org/10.1029/2008GL033257>
- Wang, W., & Zhang, Y. (2007). On fuzzy cluster validity indices. *Fuzzy Sets and Systems*, *158*(19), 2095–2117. <https://doi.org/10.1016/j.fss.2007.03.004>
- White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, *48*(4), 817. <https://doi.org/10.2307/1912934>
- Xie, X. L., & Beni, G. (1991). A validity measure for fuzzy Clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *13*(8), 841–847.
- Xu, C., Xu, Y., Hu, J., Li, S., & Wang, B. (2019). A numerical analysis of the summertime Pearl River plume from 1999 to 2010: Dispersal patterns and intraseasonal variability. *Journal of Marine Systems*, *192*, 15–27. <https://doi.org/10.1016/j.jmarsys.2018.12.010>
- Yadav, J., & Sharma, M. (2013). A Review of K-mean Algorithm. *International Journal of Engineering Trends and Technology*, *4*(7), 2972–2976.
- Zadeh, L. A. (1965). *Fuzzy Sets \* Information and Control*. *8* 353, 338–353.
- Zhang, G., Wang, Y., Zhang, J., & Niu, Y. (2011). Study on flood clustering and recognition methods based on fuzzy set theory. *2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 887–891. <https://doi.org/10.1109/FSKD.2011.6019681>

Journal Pre-proof



*Fig 1 : Map of the GoL (Mediterranean sea) and Rhone River estuary indicating the locations of discharge and weather stations used in the fuzzy-clustering algorithm. Scheme of MesuRho buoy devices Weather station + PAR (1), solar panels and control and transmission automat ABIN (2), Multiparameters sonds (3 and 6) nitrates captor ISUS (4), benthic station (5), ADCP (7).*

**2 columns fitting**

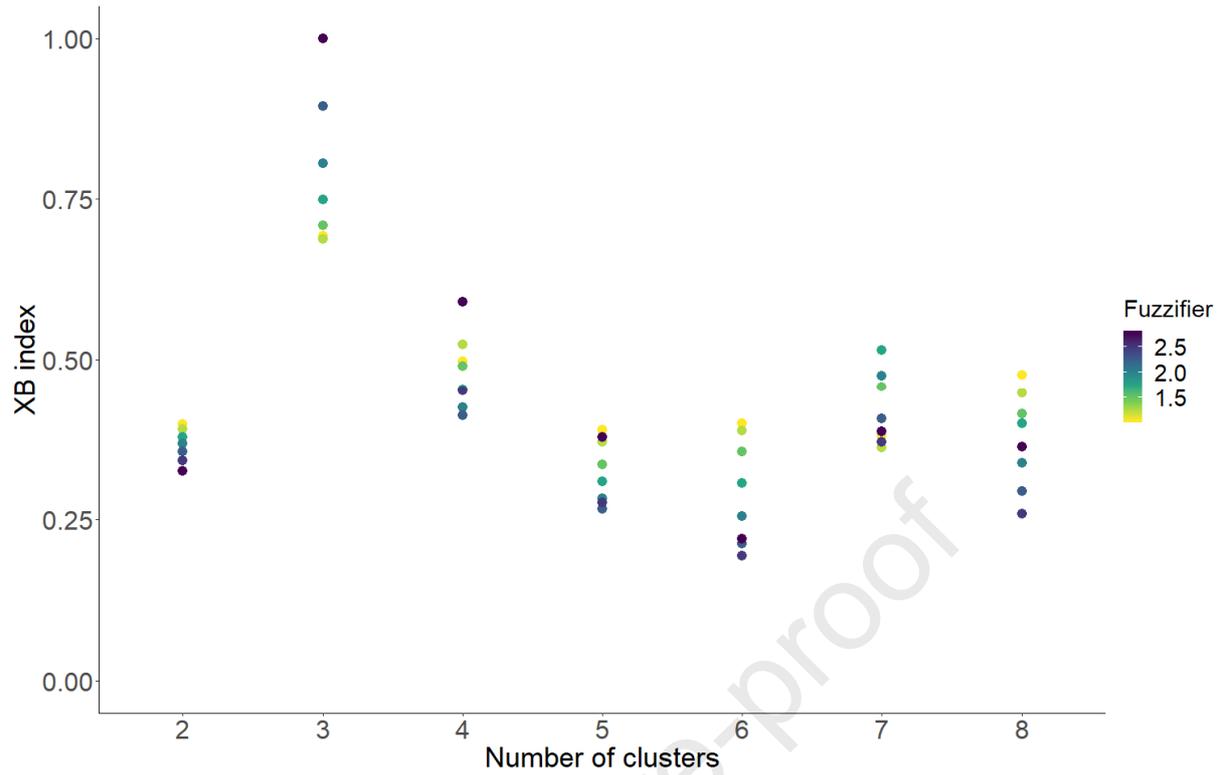


Fig 2 : Xie and Beni index values for different  $\{C,m\}$  simulations with fuzzy c-mean algorithm. A lower index indicates a better classification.

1 column fitting

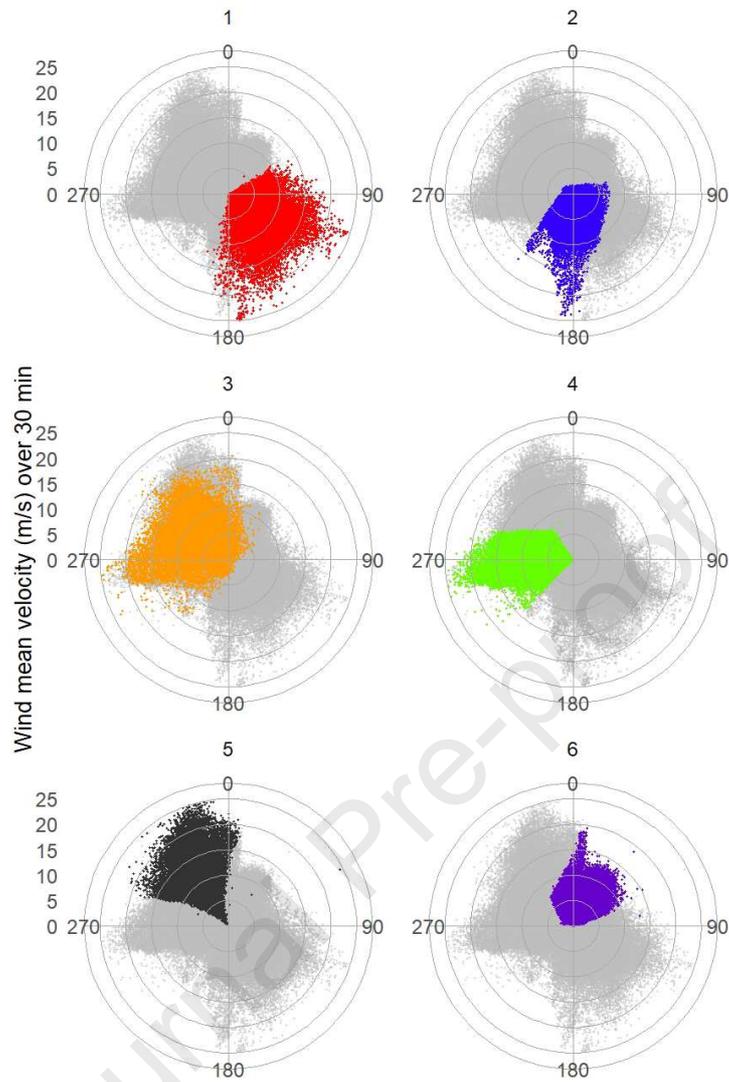


Fig 3: Grey dots: mean wind origin and speed (m/s) over 30 min from 2009 to 2019 measured at the Mesurho buoy, in front of the Rhone river. Color dots: cluster mean wind and speed sub-distributions.

**1 column fitting**

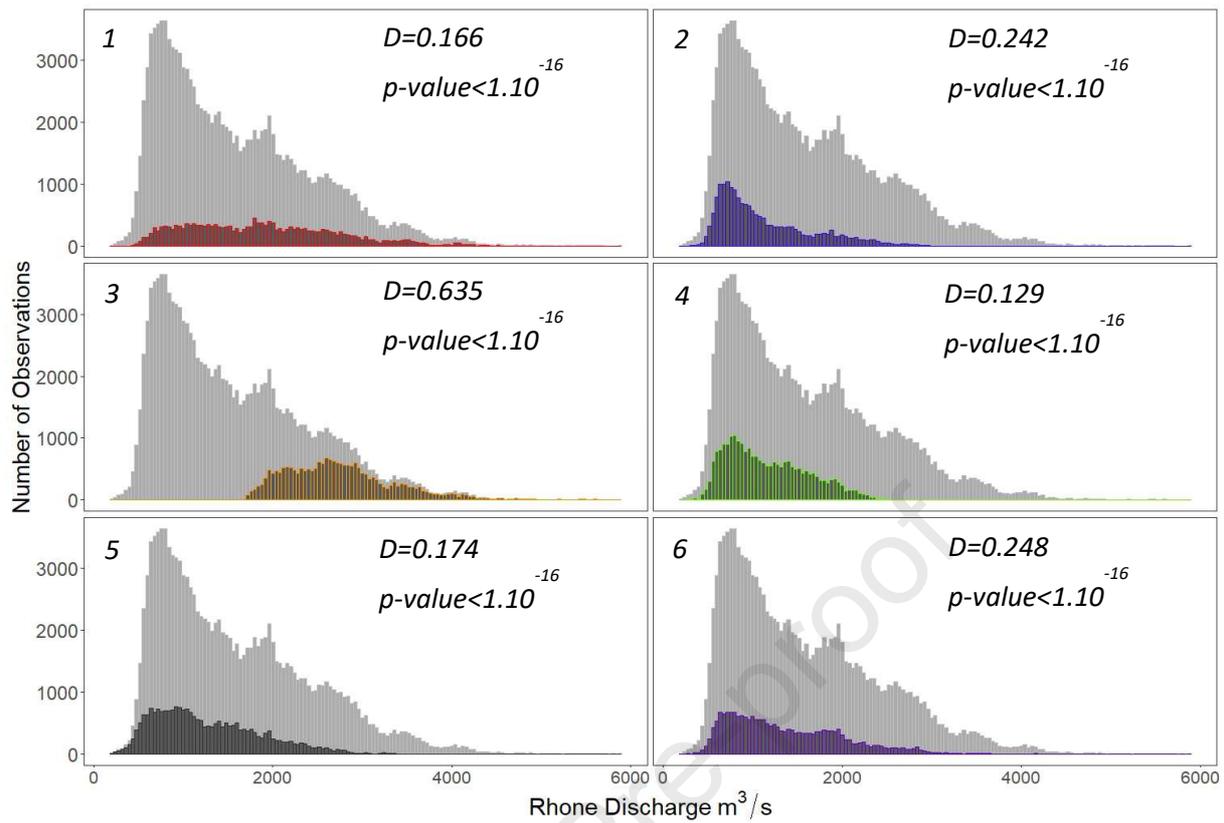


Fig 4 : Rhône river hourly discharge (SORA, Arles) from 2009 to 2019 distribution (grey) and cluster Rhône discharge sub-distributions. Results from two-sample Kolmogorov-Smirnov test are shown.  $D$  statistic indicates how the distribution in each cluster is different from the reference distribution (all Rhône discharge values) Higher values of  $D$  indicates larger differences.

1 column fitting

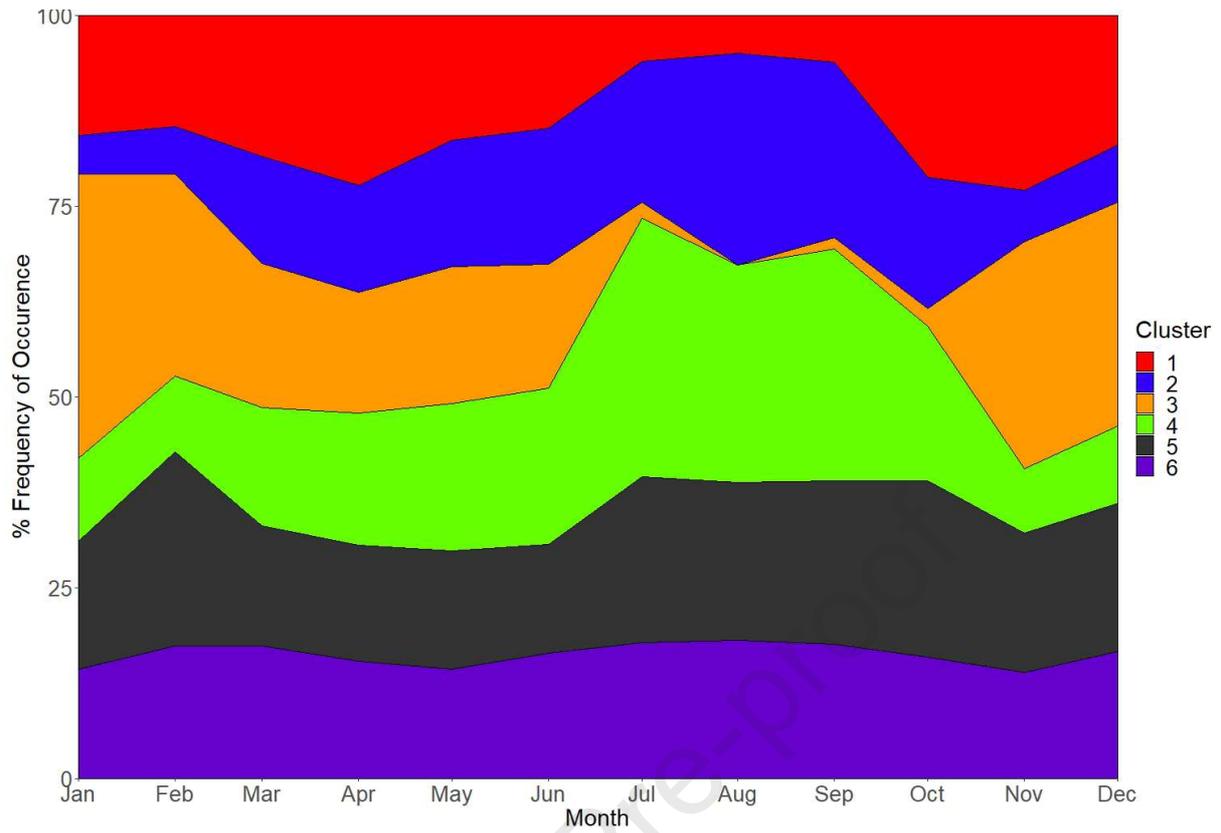


Fig 5: Monthly occurrence of the 6 different scenarios over the 2009-2019 period.

1 column fitting

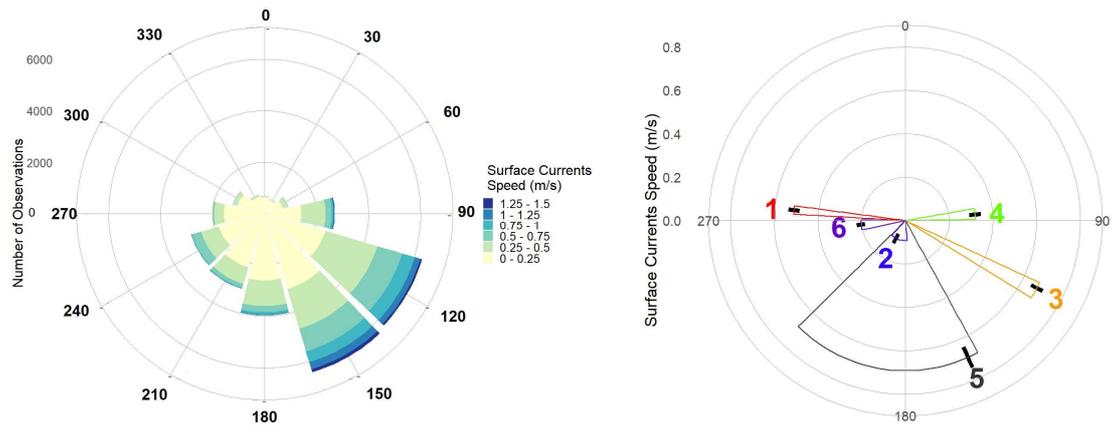
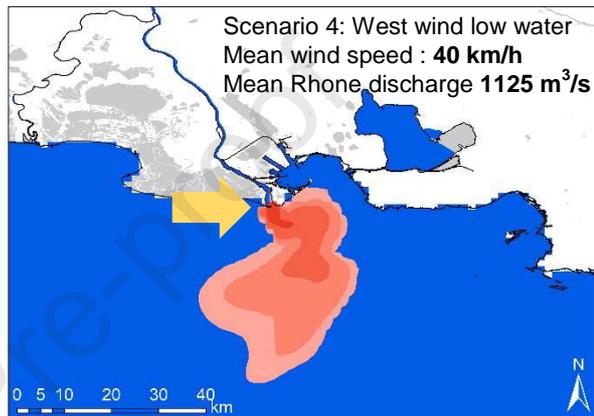
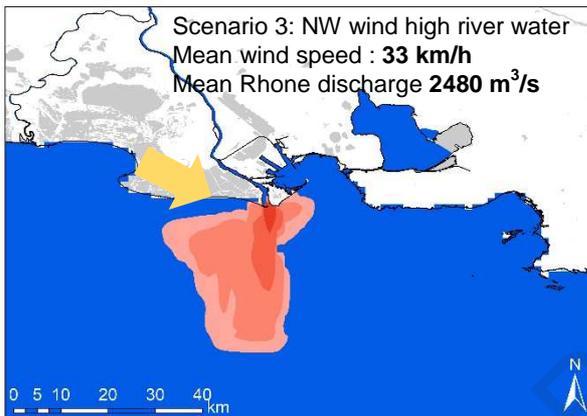
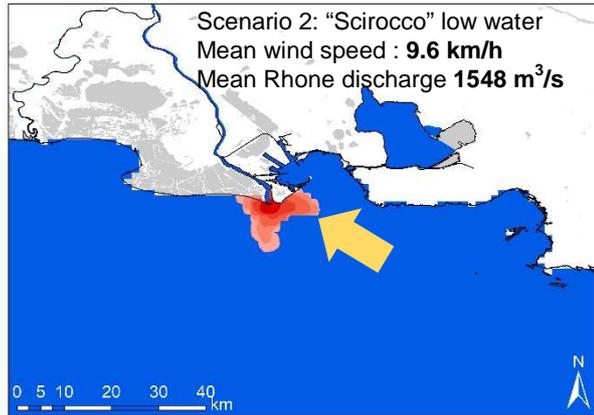
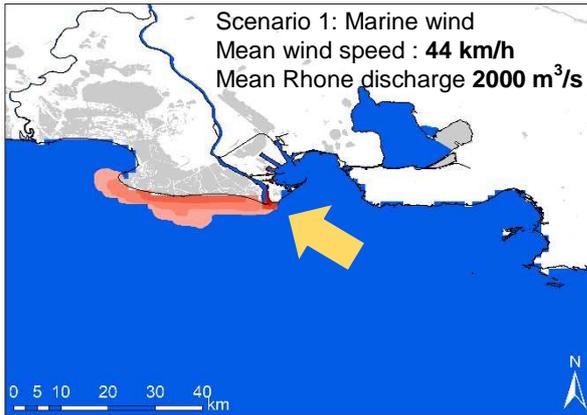
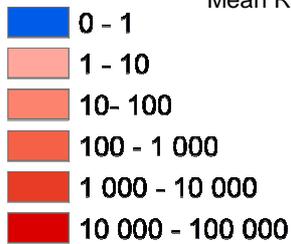


Fig 6: Surface currents (depth<1.5m) direction and module measured at the Mesurho (a) buoy and main direction (thick black dashes) and current speed for each scenario (b) obtained by least square regression. Confidence intervals are for  $\alpha=0.05$



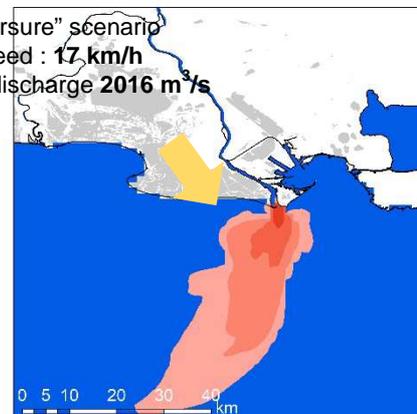
Activity in seawater (Bq/m<sup>3</sup>)

Scenario 5: "Mistral" scenario  
 Mean wind speed : **46 km/h**  
 Mean Rhone discharge **1170 m<sup>3</sup>/s**



Disconnected wet zones

Scenario 6: "Orsure" scenario  
 Mean wind speed : **17 km/h**  
 Mean Rhone discharge **2016 m<sup>3</sup>/s**



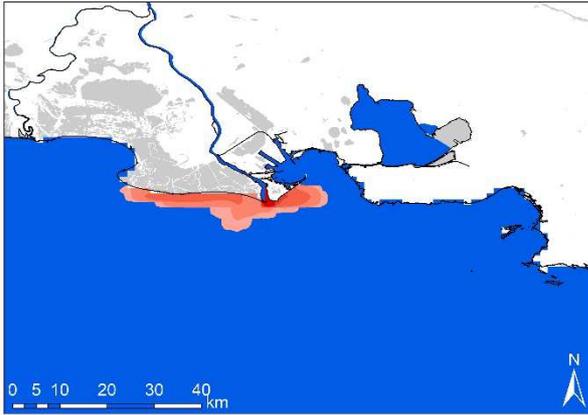
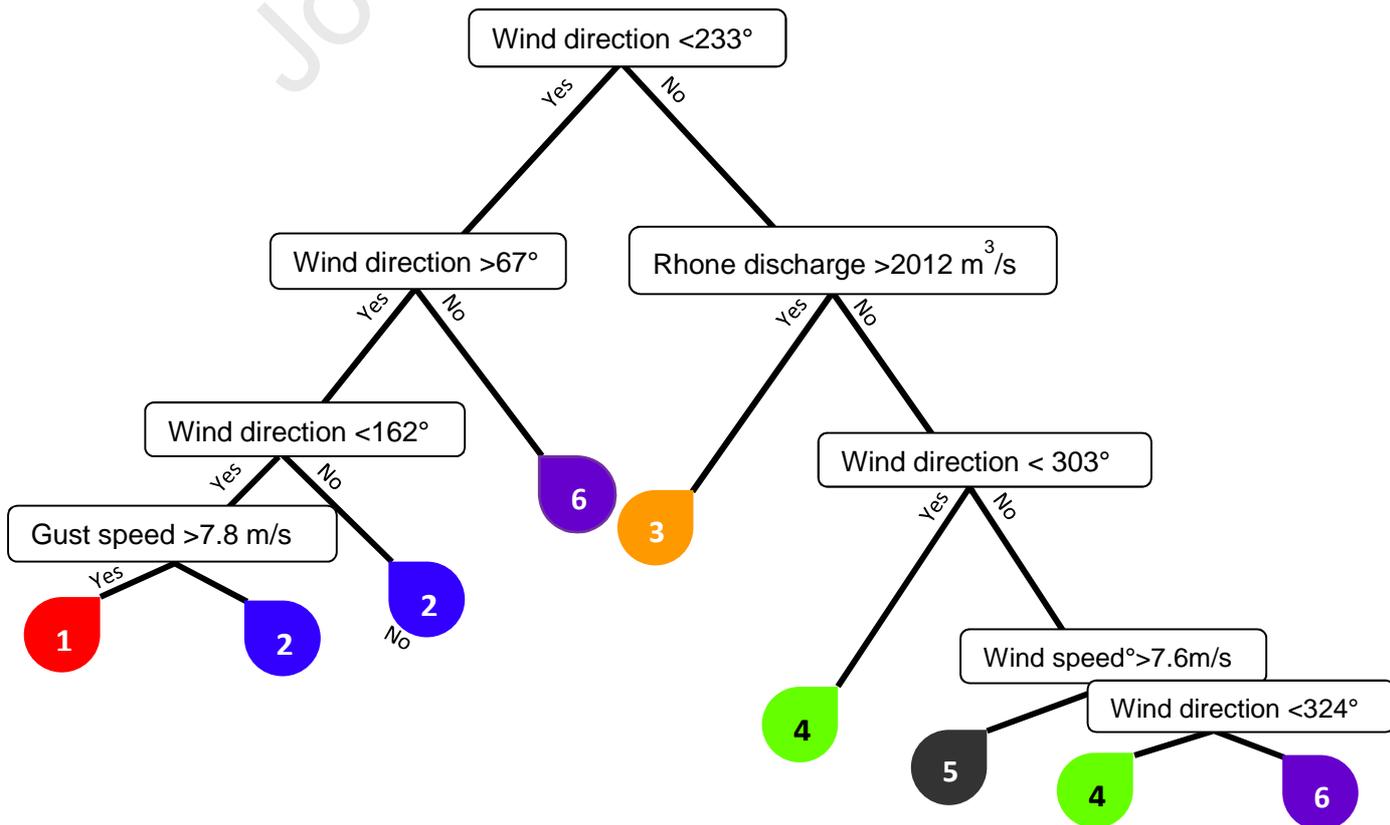


Fig 7 : Radioactive surface plume shape in the GoL under the most representatives temporal windows of the 6 scenarios for a released activity of 1TBq of  $^{137}\text{Cs}$  in 48 hours. Wind orientation is shown by the yellow arrows.

**2 columns fitting**



yes  
yes No

*Fig 8 : Simplified decision tree for scenario identification*

**2 columns fitting**

Journal Pre-proof

**Declaration of interests**

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

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof