



HELLENIC REPUBLIC
**National and Kapodistrian
University of Athens**
— EST. 1837 —

Inter-Institutional MSc in
"Oceanography and Management of the Marine Environment"

MASTER THESIS

Ensemble-based oil spill model
prediction using stochastic wind forcing

Garouniatis Paschalis
(22005)

Supervisor: Sarantis Sofianos, Assoc. Prof.

Athens, April 2023

Abstract

The Aegean Sea is one of the world's busiest trade routes throughout history in terms of maritime transport, which inevitably leads to the occurrence of mostly unintentional accidents causing oil pollution. Given its complex and intense weather and sea current patterns with strong seasonality, the uncertainty assessment of the oil spill forecasting systems in this region is of great interest. The objective of the present study is to assess the impact of wind forcing model uncertainties on the oil spill model prediction using the numerical model MEDSLIK-II. We will use stochastic modeling of the wind forcing based on Empirical Orthogonal Functions (EOF) modes. Ensemble members will be generated using EOF modes including the integration of a perturbation factor which will represent the uncertainty needed for the stochastic wind forcing. The results will focus on the oil spill model uncertainty, as approximated by an ensemble, and compared with a deterministic simulation.

Keywords: Aegean Sea, MEDSLIK II, oil spill, ensemble, uncertainty

Περίληψη

Το Αιγαίο είναι ένας από τους πιο πολυσύχναστους εμπορικούς διαύλους του κόσμου ιστορικά όσον αφορά τις θαλάσσιες μεταφορές, κάτι που αναπόφευκτα οδηγεί στο να συμβούν ατυχήματα τα οποία προκαλούν ρύπανση της θάλασσας με πετρέλαιο. Δεδομένου επίσης των πολύπλοκων και υψηλής έντασης καιρικών μοτίβων τα οποία παρουσιάζουν έντονη εποχικότητα, η εκτίμηση της αβεβαιότητας των συστημάτων πρόγνωσης στην περιοχή καθίσταται εξαιρετικά σημαντική. Σκοπός της παρούσας μελέτης είναι να εκτιμηθεί το αντίκτυπο της αβεβαιότητας των ατμοσφαιρικών δράσεων στην απόδοση του μοντέλου διασποράς πετρελαιοκηλίδας χρησιμοποιώντας το αριθμητικό μοντέλο MEDSLIK-II. Θα μοντελοποιήσουμε το ανεμολογικό πεδίο στοχαστικά βασιζόμενοι στην θεωρία EOF. Θα δημιουργηθούν μέλη ανσάμπλ χρησιμοποιώντας τα μοτίβα EOF στα οποία θα εμπεριέχεται ένας παράγοντας αβεβαιότητας, ο οποίος θα αποδίδει την αβεβαιότητα του ανεμολογικού πεδίου. Τα αποτελέσματα επικεντρώνονται στην αβεβαιότητα του μοντέλου πρόγνωσης διασποράς της πετρελαιοκηλίδας, που προσεγγίζεται με ανσάμπλ και γίνεται σύγκριση με τα αντίστοιχα που προκύπτουν από τα ντετερμινιστικά δεδομένα.

Λέξεις-κλειδιά: Αιγαίο Πέλαγος, MEDSLIK II, πετρελαιοκηλίδα, «ανσάμπλ», αβεβαιότητα

Acknowledgments

First and foremost, I would like to thank my professor and supervisor, Sarantis Sofianos, for the opportunity of working on such an interesting subject and for providing guidance and feedback throughout this thesis. I would also like to thank Vasileios Vervatis, Ioannis Karagiorgos and Konstantinos Kampouris for their assistance and dedicated involvement in every step throughout the process of researching and writing this thesis. Finally, I must express my gratitude to my family and friends, whose continuous support and encouragement have been invaluable.

Contents

Chapter 1 – Introduction	1
1.1 Study area.....	2
1.2 Objectives	3
Chapter 2 – Methodology.....	4
2.1 Empirical Orthogonal Functions (EOF)	4
2.1.1 Overview	4
2.1.2 Fundamentals.....	5
2.1.3 Empirical Orthogonal Functions (EOFs)	7
2.1.4 Singular Value Decomposition (SVD)	9
2.1.5 Generation of stochastic wind forcing.....	9
2.2 Model Description	11
2.3 Model inputs – Data.....	14
2.3.1 Oceanic forcing.....	14
2.3.2 Wind forcing.....	16
2.3.3 Bathymetry and coastlines	19
2.4 Metrics	20
2.4.1 Convex hull area	20
2.4.2 RMSE	21
2.4.3 Uncertainty index s	22
2.4.4 Oiling probability.....	23
2.5 Experiment setup.....	23
Chapter 3 – Results Discussion	25
3.1 Surface oil concentration maps	25
3.2 Uncertainty assessment for oil spill spread.....	34
3.3 Beached oil	39

3.3 Uncertainty assessment for beached oil	45
Chapter 4 – Conclusions and recommendations for future research	47
4.1 Summary and conclusions	47
4.2 Suggestions for future work.....	48
References.....	49

Chapter 1 – Introduction

Since shipping is amongst the factors with the greatest impact in worldwide economy and thus our societies, it is important to identify and address possible threats and downsides such as unintentional oil pollution. In the event of an oil spill, modelling predictions serve as the initial/forefront tools to assist regional and national contingency plans (Zodiatis et al., 2017).

The accuracy of oil spill predictions can be influenced by uncertainties related to meteorological or oceanographical conditions. Wind is a major source of uncertainty for oil spill modeling due to our incomplete knowledge of initial conditions in the atmosphere as well as simplifications of the equations used to describe the weather.

A solution to this problem is the usage of ensemble-based oil spill simulations. Wind forcing conditions can be represented by a number of different, but equally possible model forecasts, created by perturbed initial conditions and/or state variables. The ensemble spread can be regarded as a proxy of the model uncertainties in the forecast. A large spread indicates large model errors in the prediction and yields several forecast predictions increasing the possibility some of them to be closer to the true (unknown) state.

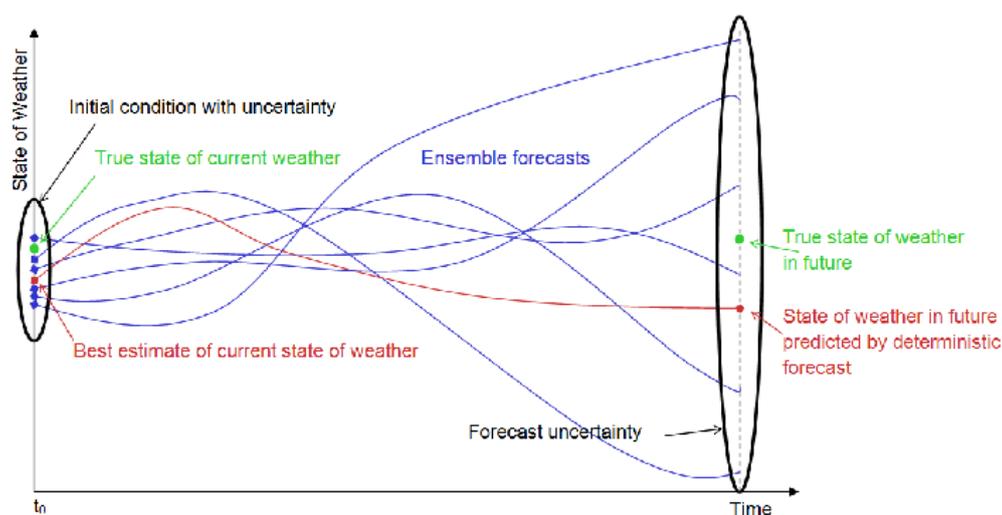


Figure 1.1: Schematic showing uncertainty captured in an Ensemble Weather Forecast

Source: https://www.researchgate.net/figure/Schematic-showing-uncertainty-captured-in-an-Ensemble-Weather-Forecast_fig1_291115693

1.1 Study area

In this study we will focus on the Aegean Sea. It is a basin with complex and intense weather and sea current patterns with strong seasonality, complicated coastline and bathymetry. Therefore, the uncertainty assessment of the oil spill forecasting systems in this region is of great interest.

The surface water circulation of the Aegean Sea (fig. 1.2) is characterized by a generally cyclonic circulation on the scale of the basin, cyclonic and anticyclonic mesoscale gyres and transient eddies. Seasonal changes in meteorological conditions, the complex topography and bathymetry of the area, the inflow of lower temperature and salinity water from the Black Sea and the outflow of rivers in Greece and Turkey create a complex circulation that changes temporally and seasonally. Perhaps the most important feature of the circulation is the low salinity waters of the Black Sea (BSW) that enter the Aegean through the straits of the Dardanelles. After its entry into the Aegean, this water mass generally follows a cyclonic course. From the Dardanelles it continues in a westerly direction, passing North of Lemnos, south of Halkidiki and then southwestward moving along the island of Euboea, as it finally enters the south Aegean through the strait of Kafireas and the strait of Mykonos-Ikaria.

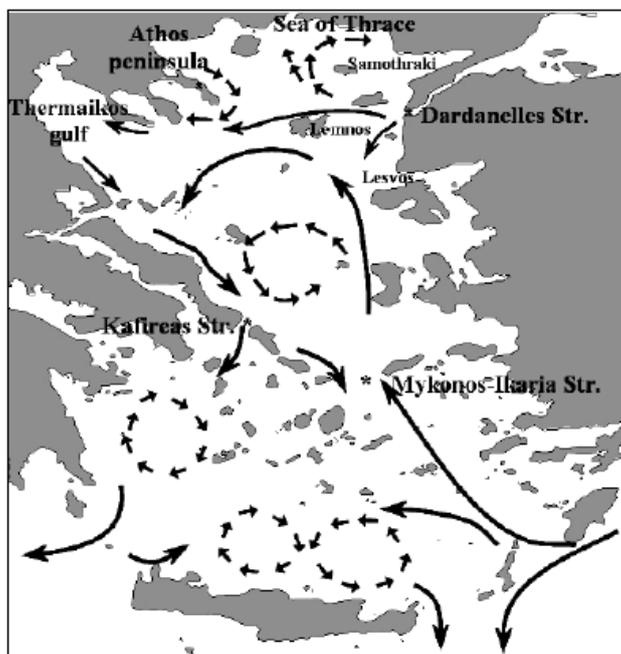


Figure 1.2:

Schematic representation of the Aegean Sea upper circulation. Reprinted from "Circulation and hydrological characteristics of the North Aegean Sea: a contribution from real-time buoy measurements" by NITTIS, K., & PERIVOLIOTIS, L. (2002), Mediterranean Marine Science.

The wind field of the Aegean Sea is dominated mainly by north winds with average monthly values ranging from 3 m/s to > 7.5 m/s. These north winds show an annual fluctuation with two maximums: one in winter during the period from December to February and one in summer from July to August. During the winter, strong, cold and dry north winds blow over the Aegean Sea, while occasionally south winds may occur. During the warm season, the wind field is dominated by the Etesian Winds, a system of strong and dry north winds (Poulos et al., 1997).

1.2 Objectives

The purpose of this study is to evaluate the impact of the atmospheric forcing uncertainty on the performance of the oil spill modeling and the transport of the pollutants in the marine environment. To that end, we use a probabilistic approach for the simulation of the oil spill by generating an ensemble of oil spill forecasts which incorporate a perturbation factor. Finally, after evaluating this impact, we try to answer if the ensemble approach for the oil spill prediction can provide additional information with respect to a deterministic approach, providing the decision-makers with a picture of several equally possible outcomes, to better plan the mitigation procedures in the event of an oil spill.

Chapter 2 – Methodology

Since the objective of this study is to assess the impact of wind forcing model uncertainties on the oil spill model prediction, the experiment setup will consist of two parts. The first part is a preprocess stage where we will generate stochastic wind forcing data based on Empirical Orthogonal Functions (EOF) modes and the second part is the model itself.

2.1 Empirical Orthogonal Functions (EOF)

2.1.1 Overview

Empirical Orthogonal Function (EOF) analysis is designed to find covariability within a data set and create new composite variables that capture that internal dependence, allowing a few composite uncorrelated variables to describe most of the variability (variance) in the data described in the much larger, dependent dataset. This method is typically applied to large space-time data (e.g., time series collected at numerous spatial locations) and its objective is to reveal relationships in space that share the same time variability and combines them into a single spatial pattern sharing a common time variability. This single pattern then captures all the replicate information otherwise stored in the many individual time series. This makes it easier to view how the variable varies in space and time and collapses a potentially huge data set into a minimum number of patterns that capture most of the variance. Each of the patterns identified has an associated time series describing how the amplitude of the pattern changes in time, known as principal components (or PC). Each pattern (EOF) and its time series (PC) together are called a mode.

2.1.2 Fundamentals

The heart of EOF analysis is the concept of eigenvectors, also referred to as characteristic vectors. These vectors are described by the following equations:

$$\mathbf{Ax} = \lambda\mathbf{x}$$

or rearranged,

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$$

where \mathbf{A} is a square matrix, λ is a scalar such that the new vector \mathbf{b} ($= \lambda\mathbf{x}$) is parallel to \mathbf{x} , since parallel vectors can differ only in length (or magnitude), and thus are equal to scalar multiples of one another.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$[\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & & a_{nn} - \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{0}$$

In this homogeneous form (i.e., the equation equals zero), we know that the system can only have a nontrivial solution when the determinant is equal to zero:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & & a_{nn} - \lambda \end{vmatrix} = 0$$

The determinant can be expanded to form a polynomial in λ , normalized by multiplying through by $(-1)^n$ as:

$$\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda^1 + a_0 = 0$$

The polynomial in λ is called the characteristic equation or characteristic polynomial. It is of order n which means that there are n roots (though all n roots needn't be unique). These n roots (n values of λ) are called the eigenvalues of the matrix A . From each eigenvalue we can obtain an eigenvector \mathbf{e} :

$$\begin{aligned} \lambda_1 \text{ yields } \mathbf{e}_1 &= [x_{11} \ x_{21} \ x_{31} \ \cdots \ x_{n1}]^T \\ \lambda_2 \text{ yields } \mathbf{e}_2 &= [x_{12} \ x_{22} \ x_{32} \ \cdots \ x_{n2}]^T \\ &\vdots \\ \lambda_n \text{ yields } \mathbf{e}_n &= [x_{1n} \ x_{2n} \ x_{3n} \ \cdots \ x_{nn}]^T \end{aligned}$$

All eigenvectors are orthogonal and orthonormal since they meet the following two conditions:

$$\mathbf{e}_i^T \mathbf{e}_j = 0 \quad \text{and} \quad \mathbf{e}_i^T \mathbf{e}_i = 1$$

Once we have all of the eigenvalues and corresponding eigenvectors, we can represent the complete eigenstructure – that is, the entire set of eigenvalues and eigenvectors – in a single equation of the form:

$$\mathbf{A}\mathbf{E} = \mathbf{E}\mathbf{\Lambda}$$

where \mathbf{A} is the original square matrix from which the eigenvalues and eigenvectors originate, \mathbf{E} is the square ($n \times n$) matrix for which each column contains one eigenvector, and $\mathbf{\Lambda}$ is a square diagonal matrix ($n \times n$) that contains an eigenvalue on each diagonal element, such that the eigenvalues are aligned with their corresponding eigenvector in the \mathbf{E} matrix.

2.1.3 Empirical Orthogonal Functions (EOFs)

These orthogonal eigenvectors (which result from all symmetric matrices A), are often termed empirical orthogonal functions (EOFs). Combining them into a single matrix \mathbf{E} gives an orthonormal matrix. All nondiagonal elements in $\mathbf{E}^T\mathbf{E}$ are the covariances between the vectors (zero for these orthogonal vectors) and the diagonals are the magnitude of the vectors, which were already normalized to equal one.

Therefore, any symmetric matrix A can be factored into the product of the diagonal matrix, Λ , containing the eigenvalues of the matrix which is pre- and post-multiplied by the orthogonal matrix, E , containing the eigenvectors (EOFs). Or, any symmetric matrix A can be reduced to a diagonal matrix, whose elements consist of its eigenvalues, by pre- and post-multiplying the A matrix with the orthogonal matrix containing its eigenvectors (this is called diagonalization).

Since the eigenvectors of any symmetrical matrix are orthogonal to one another, they can be combined with appropriate coefficients to produce any nonzero vector, z (where the order of the symmetric matrix is the same as that of the vector). That is, they can be used like any other orthogonal basis for interpolating, smoothing or any other purpose for which we have employed such functions previously. Specifically,

$$\mathbf{z} = c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_n\mathbf{e}_n = \mathbf{E}\mathbf{C}$$

where C is the matrix containing the c_i constants.

Pre-multiplying through by \mathbf{E}^T gives:

$$\mathbf{E}^T\mathbf{z} = \mathbf{E}^T\mathbf{E}\mathbf{C} = \mathbf{C}$$

So, the coefficients in vector \mathbf{C} are determined from $\mathbf{E}^T\mathbf{z}$, from which \mathbf{z} can be decomposed as $\mathbf{E}\mathbf{C}$. Thus, the eigenvectors form an orthogonal basis for \mathbf{z} , and the details (i.e., structure or shape) of the basis changes with the composition of the matrix \mathbf{A} from which it was derived.

Considering a zero-mean multidimensional dataset containing a variable sampled at n locations and m times (containing an m -dimensional time series for each row), and stored in data matrix \mathbf{X} whose sample covariance matrix is proportional to $\mathbf{X}\mathbf{X}^T$.

$$\mathbf{X} = \begin{matrix} & \begin{matrix} j = 1 & j = 2 & \dots & j = m \end{matrix} \\ & \underbrace{\hspace{10em}}_{\text{times}} \\ \begin{matrix} i = 1 \\ i = 2 \\ \vdots \\ i = n \end{matrix} & \left[\begin{array}{cccc} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & & x_{2m} \\ & \vdots & & \\ x_{n1} & x_{n2} & & x_{nm} \end{array} \right] \end{matrix} \quad \left. \vphantom{\begin{matrix} i = 1 \\ i = 2 \\ \vdots \\ i = n \end{matrix}} \right\} \text{locations}$$

$$\frac{\mathbf{X}\mathbf{X}^T}{n-1} = \hat{\Sigma}_X = \begin{bmatrix} s_1^2 & s_{12} & \dots & s_{1n} \\ s_{21} & s_2^2 & & s_{2n} \\ & & \vdots & \\ s_{n1} & s_{n2} & \dots & s_n^2 \end{bmatrix}$$

Where s_i^2 represents the sample temporal variance of the \mathbf{X} series over all sampled times at location 1, and s_{21} represents the sample covariance between the time series obtained at locations 2 and 1. As constructed here, each column of the data matrix \mathbf{X} contains the values of \mathbf{X} at a time $t = 1, 2, \dots$, or m across all n spatial locations. In other words, the time series occupy the rows (i^{th} row contains the time series for the i^{th} location), and each column contains a map of values for time j (a “map” because it shows the values of the variable x at multiple locations all at the same time).

From the above we can obtain the following:

$$\mathbf{C}_{nm} = \mathbf{E}_{nm}^T \mathbf{X}_{nm} \quad \text{and} \quad \mathbf{X}_{nm} = \mathbf{E}_{nm} \mathbf{C}_{nm}$$

where \mathbf{C} is now a fully populated ($n \times m$) matrix of coefficients. The second equation states that the spatial series in \mathbf{X} can be re-expressed as a linear sum of (spatial) eigenvectors as one map, j , per time: $i = 1, m$.

2.1.4 Singular Value Decomposition (SVD)

Considering the last two equations and the relationship defining SVD that any $n \times m$ matrix \mathbf{A} ($n \geq m$) can be decomposed as

$$\mathbf{A}_{nm} = \mathbf{U}_{nm} \mathbf{S}_{nm} \mathbf{V}_{nm}^T$$

where \mathbf{U} , \mathbf{V}^T are orthogonal and \mathbf{S} is a diagonal matrix, we can perform SVD on our data matrix \mathbf{X} and obtain $\mathbf{X} = \mathbf{U} \mathbf{S} \mathbf{V}^T$ and the sample covariance matrix $\mathbf{X} \mathbf{X}^T = \mathbf{U} \mathbf{S} \mathbf{V}^T \mathbf{V} \mathbf{S}^T \mathbf{U}^T = \mathbf{U} \mathbf{S} \mathbf{S}^T \mathbf{U}^T = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$.

The advantage with SVD analysis is that we can apply directly to the original (non-square) data matrix, \mathbf{X} , and obtain the eigenvectors and eigenvalues without ever having to actually construct the sample covariance matrix as seen in 2.1.3.

$$\mathbf{X}_{nm} = \mathbf{U}_{nm} \mathbf{S}_{nm} \mathbf{V}_{nm}^T$$

Furthermore, the SVD is a very robust decomposition method that typically provides a solution even if the covariance matrix is only marginally stable and it is recommended for multivariate datasets (in our case bivariate).

2.1.5 Generation of stochastic wind forcing

In this study we will perform a bivariate SVD analysis of the zonal and meridional wind components expressed by the matrix

$$\mathbf{U} \in \mathbb{R}^{k \times l} \text{ as } \mathbf{U} = \mathbf{X} \mathbf{\Lambda} \mathbf{Y}^T$$

where $\mathbf{X} \in \mathbb{R}^{k \times j}$ and $\mathbf{Y} \in \mathbb{R}^{l \times j}$ are each composed of orthonormal vectors, $\mathbf{\Lambda} \in \mathbb{R}^{j \times j}$ is diagonal, k is the temporal dimension, l is twice the spatial dimension considering zonal and meridional wind components, j is the low-rank dimension ($j \ll k \ll l$). The \mathbf{Y} matrix contains the spatial EOFs for the different modes and the $\mathbf{X} \mathbf{\Lambda}$ matrix denotes the Expansion Coefficients (ECs) of the wind's temporal variability. In our case we have chosen that the maximum number of modes j will be such to explain the percentage of 95% of the cumulative variance.

The next step is to reconstruct data of both wind components for the same period which will differ from the original ones due to the application of perturbation by applying a zero mean Gaussian pseudo-random perturbation factors on all EOFs

$$as \ U'_i = \sum_{j=1}^p \varepsilon_{i,j} (X_j \Lambda_j)(Y_j)^T$$

where $i = 1, 2, \dots, N$ denote the ensemble members and $j = 1, 2, \dots, p$ denotes the low-rank number of the EOFs. The standard deviation of $\varepsilon_{i,j}$ is set to a value of 0.3, assuming that the standard deviation of uncertainties on the wind velocity is 30% of the standard deviation of the wind, based on the ECMWF noise to signal ratio and studies for mid-latitude wind uncertainty assumptions which qualifies our choice as physically consistent.

It is important to note that the ensemble consists of pairs corresponding to odd/even members, derived from pairs of odd/even Gaussian factors so that $\varepsilon_{i,j} = -\varepsilon_{i+1,j}$

2.2 Model Description

The numerical model MEDSLIK II (De Dominicis et al., 2013a), (De Dominicis et al., 2013b), is a freely available community model, which is based on its precursor the oil spill model MEDSLIK (Lardner et al. 1998, Lardner et al. 2006, Zodiatis et al. 2005, Zodiatis et al. 2008). It is designed to predict the transport and weathering of an oil spill, caused by complex physical processes occurring at the sea surface, using a Lagrangian representation of the oil slick. This numerical representation requires the following different state variables: the oil slick, the particle and the structural state variables, which are all used for different calculations. The transformation and movement of an oil slick depend on many factors, the main ones being:

1. meteorological and oceanographic conditions at the air-sea interface
2. the marine currents in the oil spill area as well as the chemical characteristics
3. the initial volume
4. the rate of oil release

A brief description of the basic equations used by MEDSLIK II is given below based on De Dominicis et al. (2013a), De Dominicis et al. (2013b) and Zodiatis et al. (2017), and a schematic representation of the model's solution procedure methodology can be seen in fig. 2.1. Over time, as the oil moves, its concentration changes due to physical and chemical processes also known as "weathering", e.g. evaporation, emulsification, dispersion in the water column and viscosity changes.

The general equation for calculating the oil concentration $C(x, y, z, t)$ in the marine environment is:

$$\frac{\partial C}{\partial t} + \mathbf{U} \cdot \nabla C = \nabla \cdot (\mathbf{K} \nabla C) + \sum_{j=1}^M r_j(\mathbf{x}, C(\mathbf{x}, t), t) \quad (1)$$

where $\partial C/\partial t$ is the time rate of oil concentration change, \mathbf{U} is the sea current mean field with components (U, V, W), \mathbf{K} is the diffusivity tensor which parameterizes the turbulence and $r_j(C)$ are the number of M transformation rates that change the oil concentration due to physical and chemical transformation processes.

The solution of the above equation and the calculation of the evolution of the oil concentration based on a Lagrangian formalism, is based on the following fundamental assumptions:

- Water hydrodynamics and other processes generally are not influenced by the constituent particles.
- The constituent particles behave like water parcels, moving through infinitesimal displacements with the absence of inertia and with no interaction among themselves.
- Physical and chemical processes modify the volume associated with each particle, by acting on the entire slick rather than on the properties of every single particle.

Based on these assumptions, Eq. 1 is divided into two components:

1. The transformation equation due to "weathering":

$$\frac{\partial C_1}{\partial t} = \sum_{j=1}^M r_j(\mathbf{x}, C_1(\mathbf{x}, t), t) \quad (2)$$

where C_1 is the concentration of oil due to "weathering" processes. The transformation processes act on the total volume of the oil slick and the oil slick state variables are defined. The "weathering" processes are calculated through Mackay et al. (1980) fate algorithms. In order to be used, the surface volume of the oil slick is divided into a thin part, V_{TN} , at the edges of the oil slick, and a thick part, V_{TK} , near its center.

2. The advection-diffusion equation:

$$\frac{\partial C}{\partial t} = -\mathbf{U} \cdot \nabla C_1 + \nabla \cdot (\mathbf{K} \nabla C_1) \quad (3)$$

where the oil slick is discretized into a large number of particles, transported by sea currents, wind, waves and diffusion processes, with associated particle state variables, some of which are deduced from the oil slick state variables. Concentration C is subdivided into its following components, called structural state variables: The surface oil concentration, C_S , subsurface C_D , adsorbed on the shore, C_C and in bottom sediments C_B .

Finally, to fully solve the problem of advection-diffusion and transformation, of the general equation for calculating concentration (Eq. 1), a numerical grid must be determined where the particles can be measured and the concentration calculated.

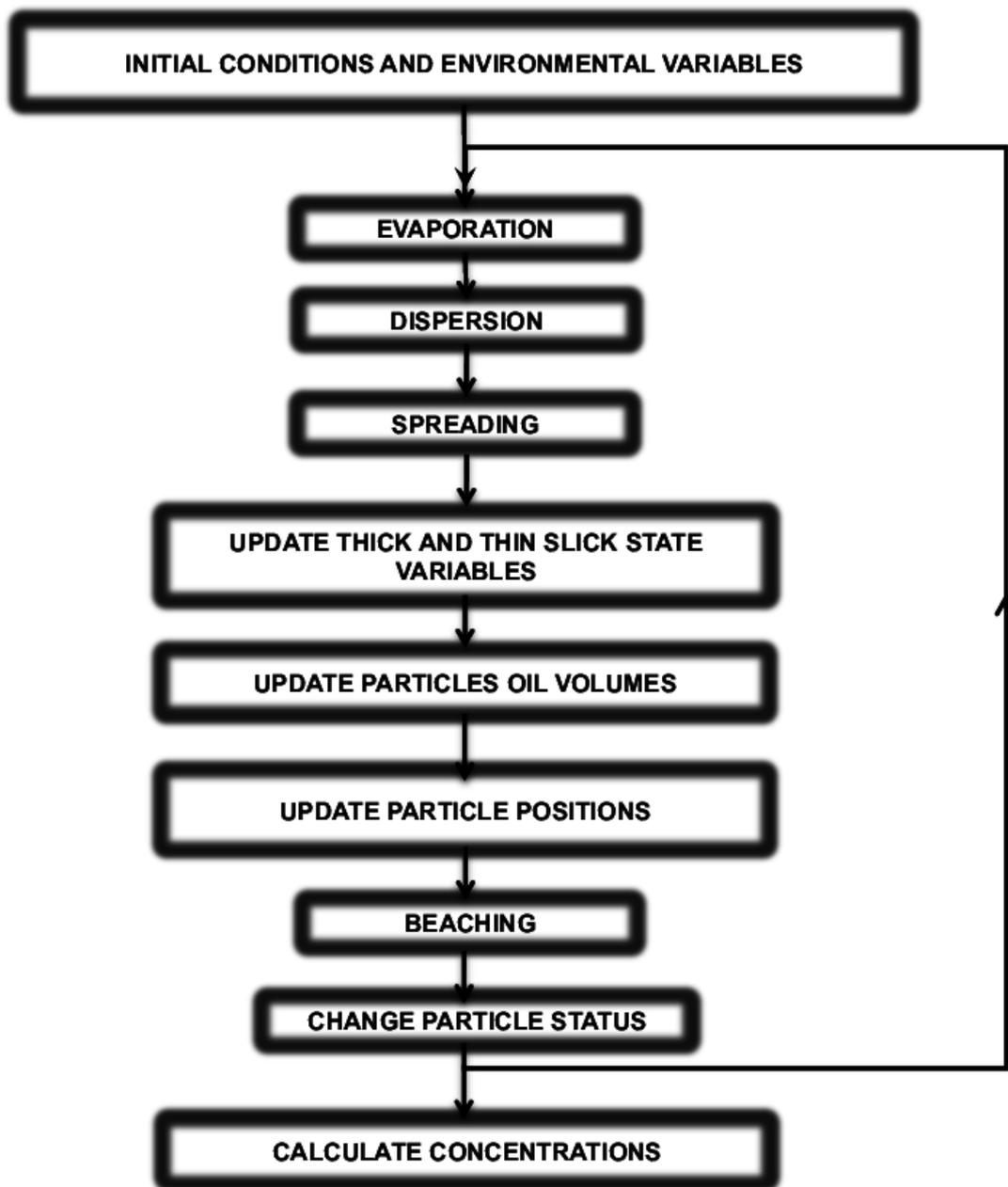


Figure 2.1: MEDSLIK-II model solution procedure methodology. Reprinted from “MEDSLIK-II, a Lagrangian marine surface oil spill model for short-term forecasting – Part 1: Theory” by M. De Dominicis et al. (2013), Geoscientific Model Development

2.3 Model inputs – Data

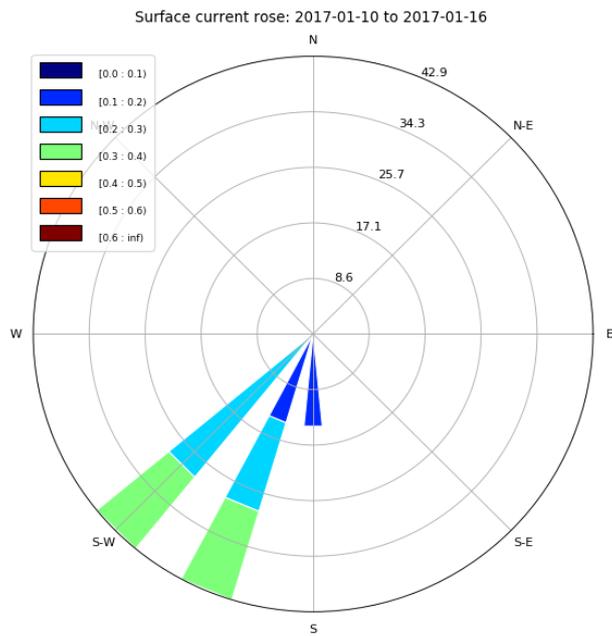
MEDSLIK-II requires the input of data on atmospheric winds, sea surface temperature, marine currents, and information about the oil spill, in order to calculate the oil transport and transformation processes.

2.3.1 Oceanic forcing

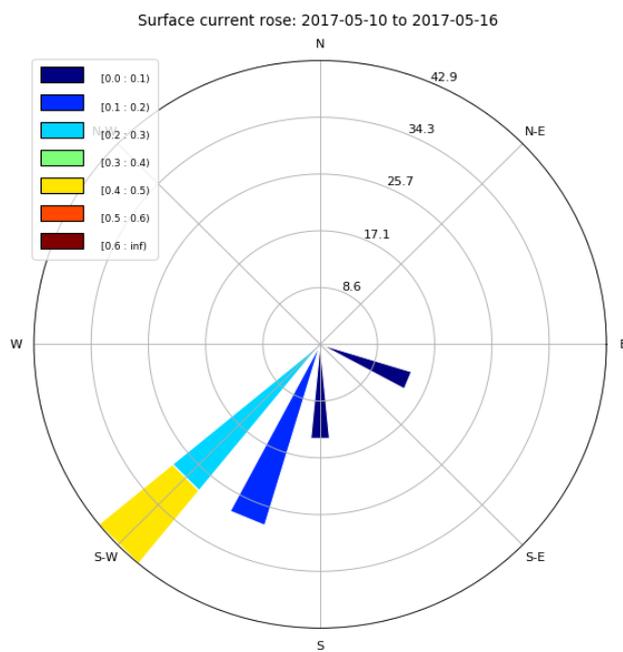
For the oceanic forcing, daily three-dimensional currents were used (velocity and temperature), covering the area 18 ° E to 32 ° E and 32 ° N to 43 ° N, with depths from 1 m up to 150 m. We selected two periods of seven-day simulations, the one in winter (mid-January 2017) and the other in spring (mid-May 2017). The input currents were interpolated to 1 h time step, which is used by the oil spill model MEDSLIK II.

The data were provided by the physical component of the Mediterranean Forecasting System (Med-Currents), a coupled hydrodynamic-wave model, whose outputs are freely available on the Copernicus Marine Environment Monitoring Service (CMEMS) portal (marine.copernicus.eu). The model covers the entire Mediterranean with a horizontal grid resolution of 1/24° (approximately 4 km) and 141 unevenly distributed vertical levels. Hydrodynamics is provided by NEMO v3.6 while the wave component is provided by Wave Watch-III.

Figure 2.2 presents the rose diagrams of the daily surface sea current velocity, for a point near the starting position of the oil slick. Figure 2.2a shows the speed and direction of the surface sea current for the period from 2017-01-10 to 2017-01-16 (winter), while fig. 2.2b for the period from 2017-05-10 to 2017-05-16 (spring). The prevailing direction of the surface sea currents is south-southwest for both time periods, which agrees with the bibliography about the surface water circulation of the Aegean Sea, with values ranging from 0.1 to 0.4 m/s in winter and 0 to 0.5 m/s in spring.



(a)



(b)

Figure 2.2: Surface current speed rose of the deterministic simulation for (a) winter and (b) spring period

2.3.2 Wind forcing

For the wind forcing, 50 ensemble members of 3-h wind velocities were used, covering the area of 18 ° E to 32 ° E and 32 ° N to 43 ° N and a horizontal grid resolution of 9 km for the deterministic and 18 km for the ensemble members. Two periods of seven-day simulations were selected (mid-January 2017 and mid-May 2017). The input winds, like the three-dimensional currents, were also interpolated to 1 h time step, which is used by the oil spill model MEDSLIK II.

The ensembles were generated with following procedure:

- Application of SVD on the deterministic dataset
- Usage of all modes describing the wind field information up to 95% of the deterministic scenario
- Generation of stochastic wind forcing with the incorporation of a perturbation factor

Figures 2.3 and 2.4 present the wind roses of the 3-h wind velocity at 10 m for a point near the starting position of the oil slick. Figure 2.3 shows the wind speed and direction of the deterministic and total ensemble simulation for the period from 2017-01-10 to 2017-01-16 (winter). The prevailing wind direction is north-northeast, nearly opposite to the currents in the area, with a maximum value above 12 m/s for both wind roses. The total ensemble wind displays a larger variability of wind directions. Likewise, figure 2.4 presents the wind speed and direction of the deterministic and total ensemble simulations for the period from 2017-05-10 to 2017-05-16 (spring). The prevailing wind direction is to the south, with a maximum value between 8 to 10 m/s. In spring, the intensity of the wind is lower than the winter and the prevailing wind direction is closer to that of the currents in the area. Also, the differences between the deterministic and the total ensemble are lower than the winter case.

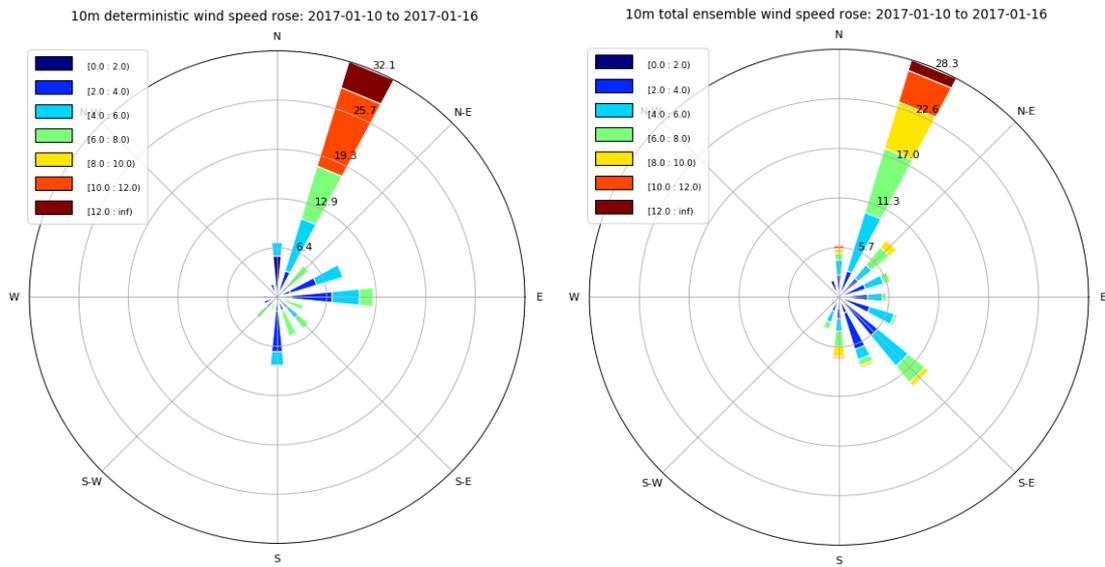


Figure 2.3: Wind speed rose of the deterministic (left) and total ensemble (right) simulation for winter period

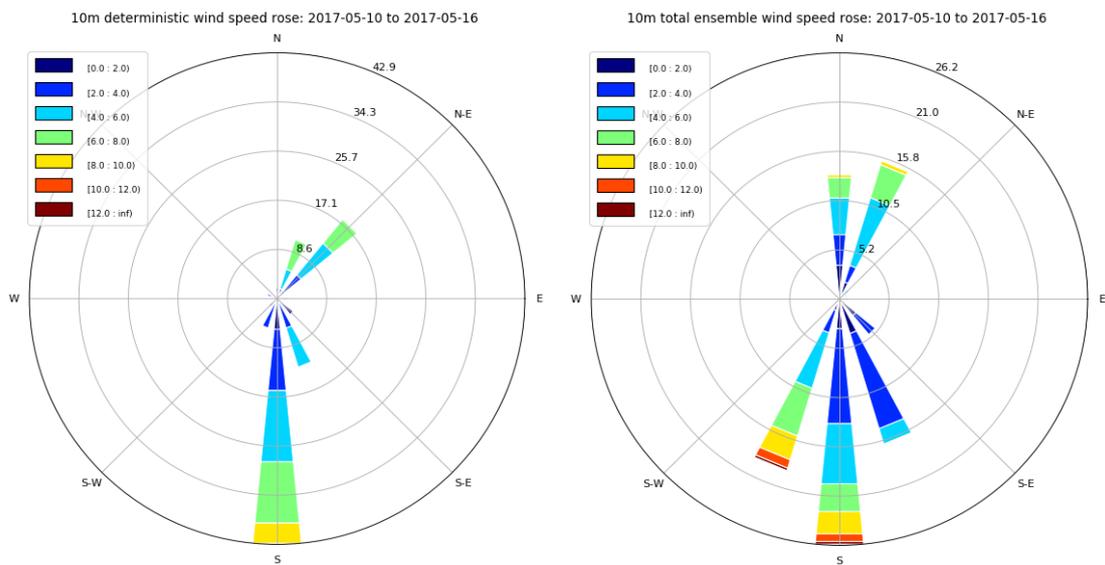


Figure 2.4: Wind speed rose of the deterministic (left) and total ensemble (right) simulation for spring period

Another way to visualize the variability of wind direction and speed is shown in figure 2.5. We can observe sudden changes of the wind in the winter as shown in figure 2.5a, especially between 40 to 80 hours and 120 to 168 hours, indicating high variability of the atmosphere and potentially a higher uncertainty in the forecasts by incorporating a perturbation factor in our data. In spring the lower wind intensity and the gradual change of the wind, as shown in figure 2.5b, indicate a less variable wind pattern and potentially lesser uncertainty.

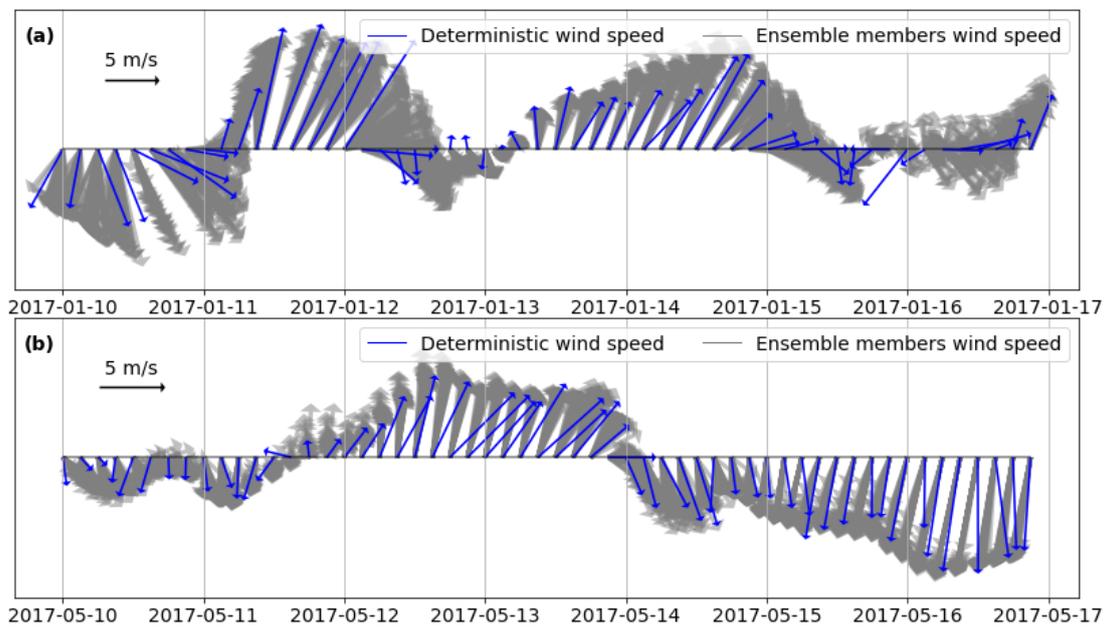


Figure 2.5: 10m wind speed of deterministic and ensemble mean for (a) winter and (b) spring

2.3.3 Bathymetry and coastlines

Bathymetry used in the MEDSLIK II simulations was obtained from the General Bathymetric Chart of the Oceans (GEBCO). The data set GEBCO_2014, a global grid at 30 arc-second intervals, was used for the defined study area (from 23 ° E to 26 ° E and 36 ° N to 39 ° N). For coastlines, version 2.3.7 of the high-resolution GSHHG geographic data set was used.

Figure 2.6 presents the bathymetry and the coastlines of the study area, along with the names of the locations and the starting point of the oil spill (shown as a red cross in the map).

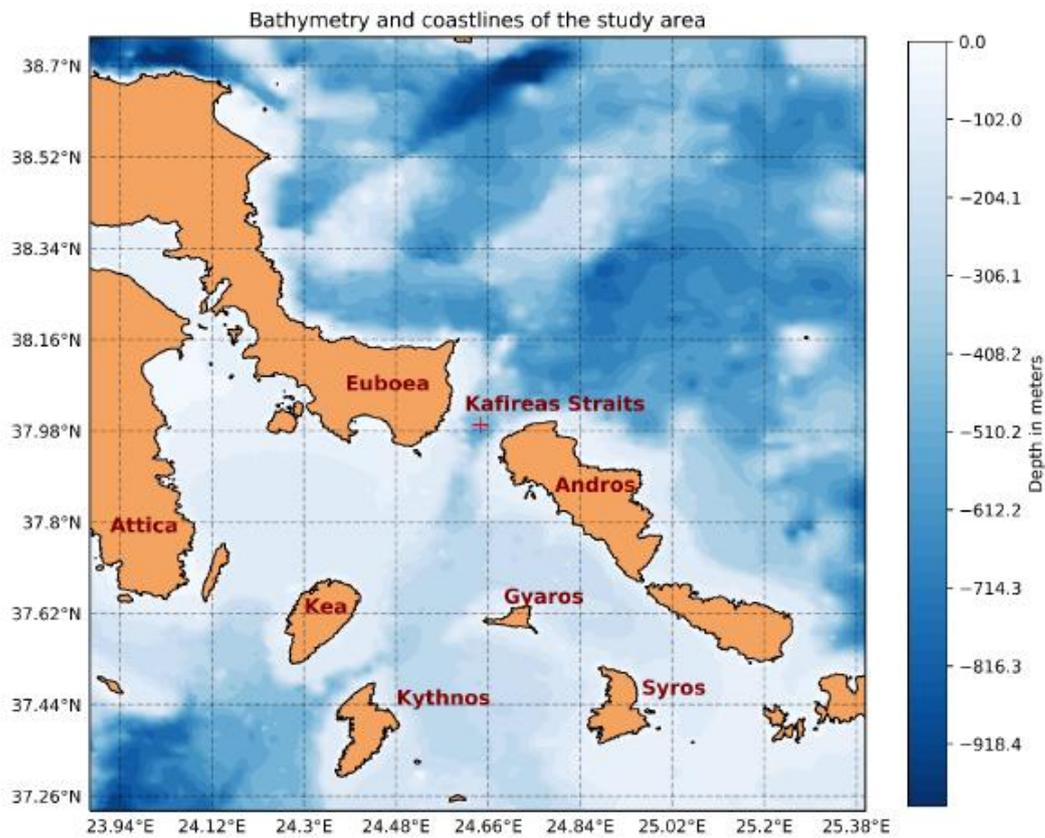


Figure 2.6: Bathymetry, coastlines and study area information

2.4 Metrics

2.4.1 Convex hull area

The convex hull of a given set of oil spill particle positions inside the area of interest is defined as the smallest convex polygon that contains all positions in the set of modelled particles. In this study convex hull is used to examine the spreading, transport and dispersion of the simulated oil spills and evaluate the uncertainty of the area affected by the oil particles simulated by the deterministic and the ensemble members.

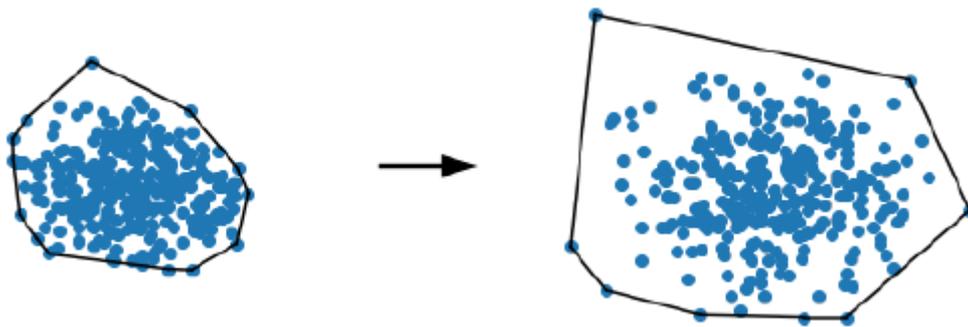


Figure: Example of convex hull of a set of points as they spread in time.

In addition, we use the convex hull to compute two more metrics, A_{exceed} and DA. A_{exceed} denotes the area of the deterministic convex hull that exceeds the area of one ensemble member's convex hull, while DA denotes the difference in the size between the deterministic convex hull and the convex hull of the combined ensemble oil spills.

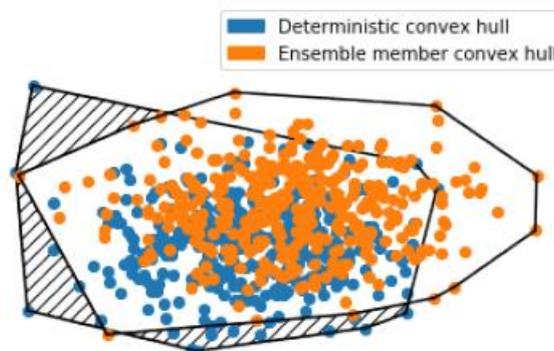


Figure: Area of deterministic that exceeds the area of the ensemble member convex hull (A_{exceed}).

Furthermore, apart from the DA metric, we use percentage change to calculate the percentage increase (or decrease) between the deterministic convex hull and the convex hull of the combined ensemble oil spills (total ensemble convex hull).

$$\text{Percentage change} = \frac{DA}{\text{Deterministic convex hull area}} * 100\%$$

2.4.2 RMSE

The root mean square error (RMSE) of the ensembles is estimated with respect to the deterministic simulation, calculating the separation distance between the deterministic and the ensemble means of oil spill Lagrangian trajectories, as a function of the forecast lead time (De Dominicis et al., 2013b). In this study, the Lagrangian trajectory definition refers to the mean trajectory geographically weighted by the number (and subsequently concentration) of the released oil spill particles.

$$RMSE(t_i) = \sqrt{\frac{\sum_{s=1}^S d_i(x_s(t_i), x_0(t_i))^2}{S}}$$

where d_i is the distance between the deterministic, x_0 , and the ensemble, x_s , mean positions respectively, at a given forecast lead time t_i following the mean trajectories after the initial release of particles, and S is the total number of the ensemble members.

2.4.3 Uncertainty index s

The non-dimensional index s according to De Dominicis et al. (2013b) and Liu & Weisberg (2011) is defined as:

$$s(t_i) = \frac{1}{S} \sum_{s=1}^S \frac{\sum_{t=t_0}^{t_i} d_i(x_s(t), x_0(t))}{\sum_{t=t_0}^{t_i} l_{0i}(x_0(t_0), x_0(t))}$$

where d_i and S have already been defined above and l_{0i} is the length of the simulated deterministic mean trajectory at a given forecast lead time t_i , following the mean trajectories after the initial release of particles at time t_0 .

The average of the separation distances between the simulated deterministic and ensemble members is weighted by the length of the deterministic trajectory, aiming at the reduction of possible evaluation errors that could arise by using only the Lagrangian separation distance (as in RMSE). For this reason, the uncertainty index s is used alongside the RMSE as it provides a more accurate quantification of the uncertainty in the oil spill trajectories.

In most studies, RMSE and index s are used in the comparison between observed and simulated trajectories, as negative oriented metrics, in order to evaluate the modelling system's capability in reproducing the observed trajectories (predictability of the trajectories). The lower the RMSE and s values, the better the performance and predictability of the model simulations, with 0 indicating a perfect fit between observation and simulation. In this study, we use RMSE and index s , as positive oriented metrics, to compare the deterministic with the ensemble trajectories. Higher RMSE and s values indicate more significant differences between the deterministic and ensemble trajectories and consequently the importance of the uncertainty generated by the ensemble simulations.

2.4.4 Oiling probability

The description of oiling probability is presented by Amir-Heidari et al. (2019) and Goldman et al. (2015). In the event of an oil spill, the oiling probability for a receptor (e.g. the coastline in our case) indicates the chance of the receptors' exposure to oil. The traditional approach for the calculation of oiling probability is based on a binary philosophy, i.e. oil spill events counted as "0" for nonexistent concentrations of oil in the beach or "1" for measured concentration regardless the amount of oil. The oiling probability for a total of n oil spill scenarios, with one source and one receptor only is according to Amir-Heidari et al. (2019):

$$P = \frac{\sum_{i=1}^N B_i}{N}$$

where B_i equals "1" or "0" if during the i^{th} simulation we measure oil concentrations or not, and N is the total number of simulations.

2.5 Experiment setup

A single oil release point was used performing simulations of 7 days forecast lead time, with continuous oil release and a rate of spillage of 5 tons per hour. The oil spill duration and the spill rate were chosen, taking into account significant accidents of the past, like for instance the Prestige case (Portman, 2016), (Sepp Neves et al., 2016). The number of parcels, used in the simulations to calculate diffusion and dispersion and to estimate the concentration of oil in the oil slick, was 105, while the values of horizontal and vertical diffusions remained constant during experiments and the default model values were used. Stokes drift was taken into account in the calculation of oil transport, and the depth of the mixing layer for the periods January 2017 and May 2017 was set at 50 m and 10 m respectively, according to the oceanographic data used.

The range of experiments performed included:

- An atmospheric forcing of 50 ensembles of hourly (1h) wind speed values at 10 m as well as the ECMWF-HRES deterministic simulations for the time period of January 2017 and May 2017 (a total of 50+1 simulations per experiment/period).
- The corresponding oceanographic data (marine current velocity and temperature) for the above two periods and for depths of 0 m, 10 m, 30 m, 120 m.
- The type of oil chosen for this study is categorized as API 31 representing medium oil spills

A total of 102 simulations were performed for a period of 168 hours.

Chapter 3 – Results Discussion

In this chapter we will present the results of MEDSLIK II simulations and quantify, using the metrics discussed in the previous chapter, the uncertainty generated in the oil spill forecasting by the stochastic wind forcing ensembles and their comparison with the deterministic simulation.

The oil type chosen for the simulations is API 31, which represents the most common type of oil (medium density oil). Furthermore, the oil slick in the simulations is represented by 10^5 independent Lagrangian particles.

3.1 Surface oil concentration maps

The direction of spreading and transportation of surface oil is greatly controlled by the direction of the wind in conjunction with the wind field changes. The concentrations of the ensemble oil spill, as simulated by the model, are constant, but there are variations in the transport and the evolution of the shape and size of the oil spills between members solely because of the wind spread.

Figures 3.1 and 3.2 illustrate the surface oil concentrations and spreading of the deterministic and ensemble simulation for winter respectively. The simulation run time for each occasion is 72 and 168 hours. Different colors represent the different concentrations of particles. 72 hours after the occurrence of the oil spill, the sudden changes in the wind field are responsible for the observed spread of the oil slick and its southwest transportation. Between 72 and 168 hours, the oil slick spreads further in the area, mainly in the downwind direction, and is transporting northeast while beginning to interact with the coast of the island of Andros.

Although the prevailing transport of the ensemble oil spill members is quite similar to the deterministic oil spill, there are significant differences concerning the spreading of the oil spill. These differences are caused by the sudden changes in the wind field associated with the incorporation of the perturbation factor, as described in 2.2, among the ensemble members.

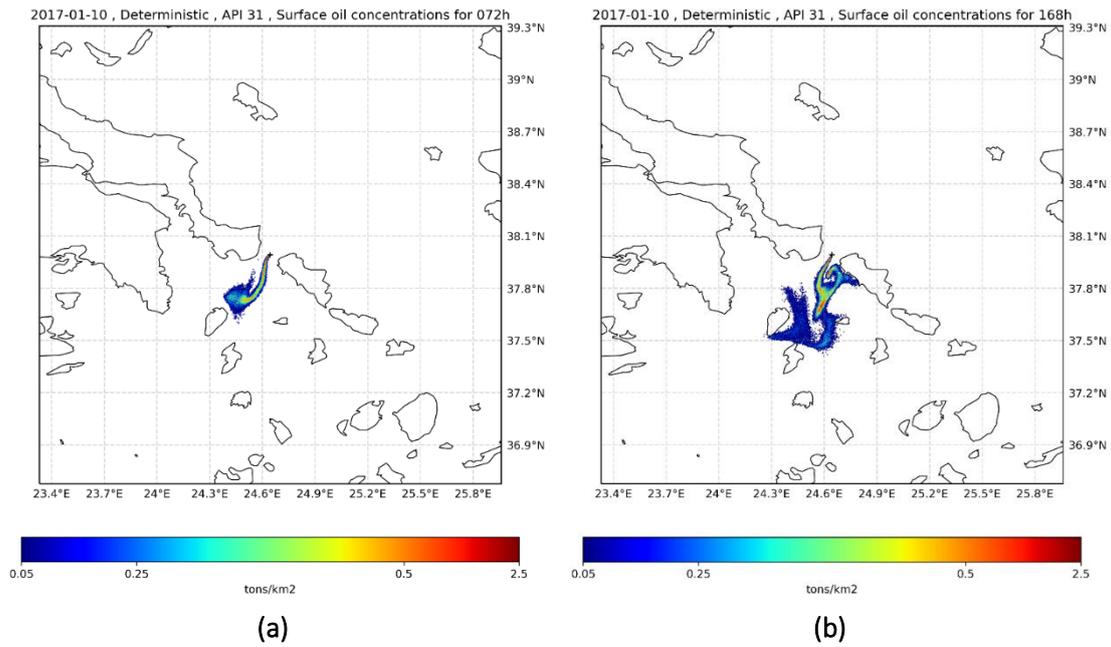


Figure 3.1: Surface oil concentrations of deterministic simulation for winter and run time of: (a) 72 hours, (b) 168 hours.

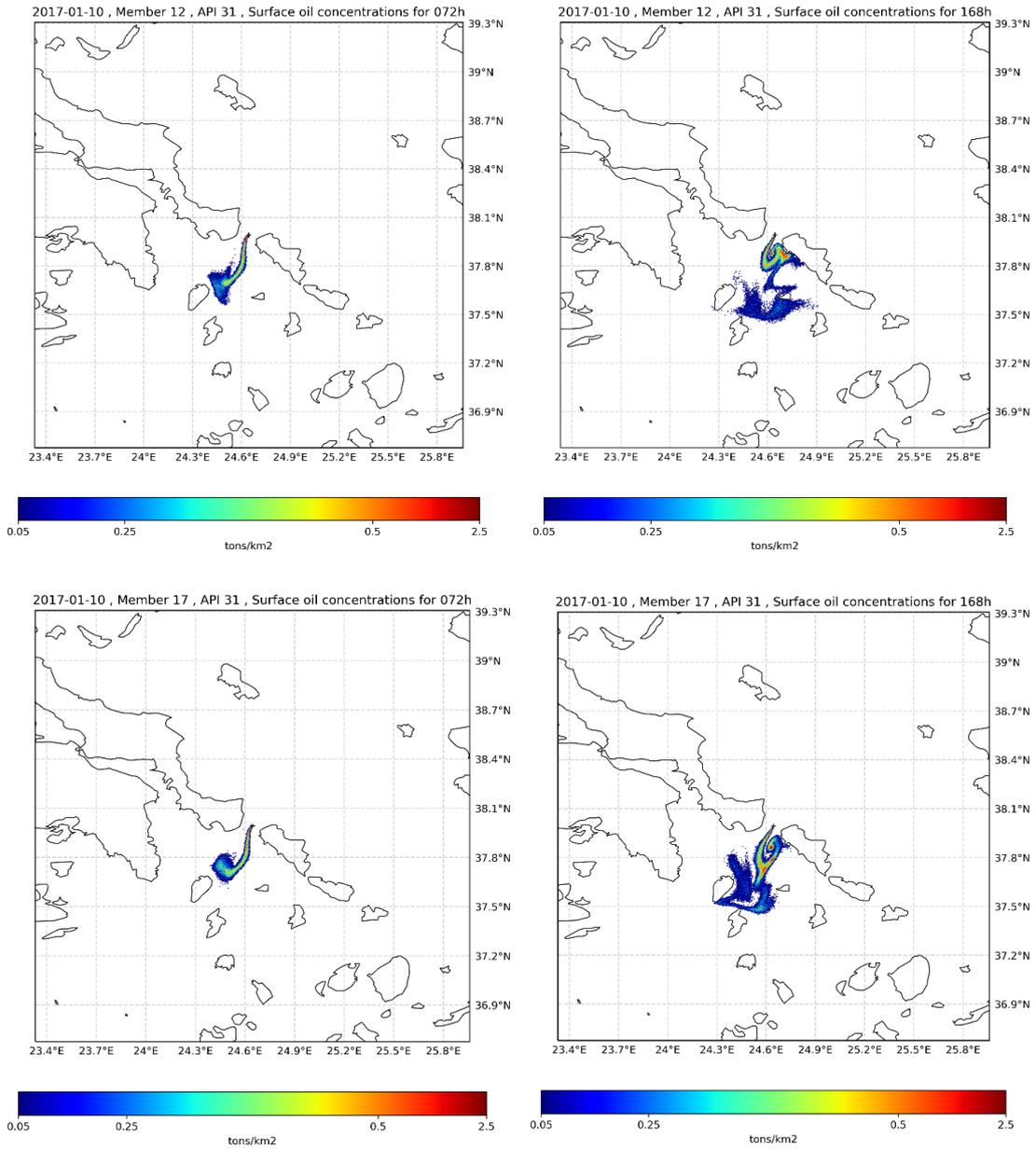


Figure 3.2: Surface oil concentrations of selected ensemble members for winter and run time of: 72 hours, 168 hours.

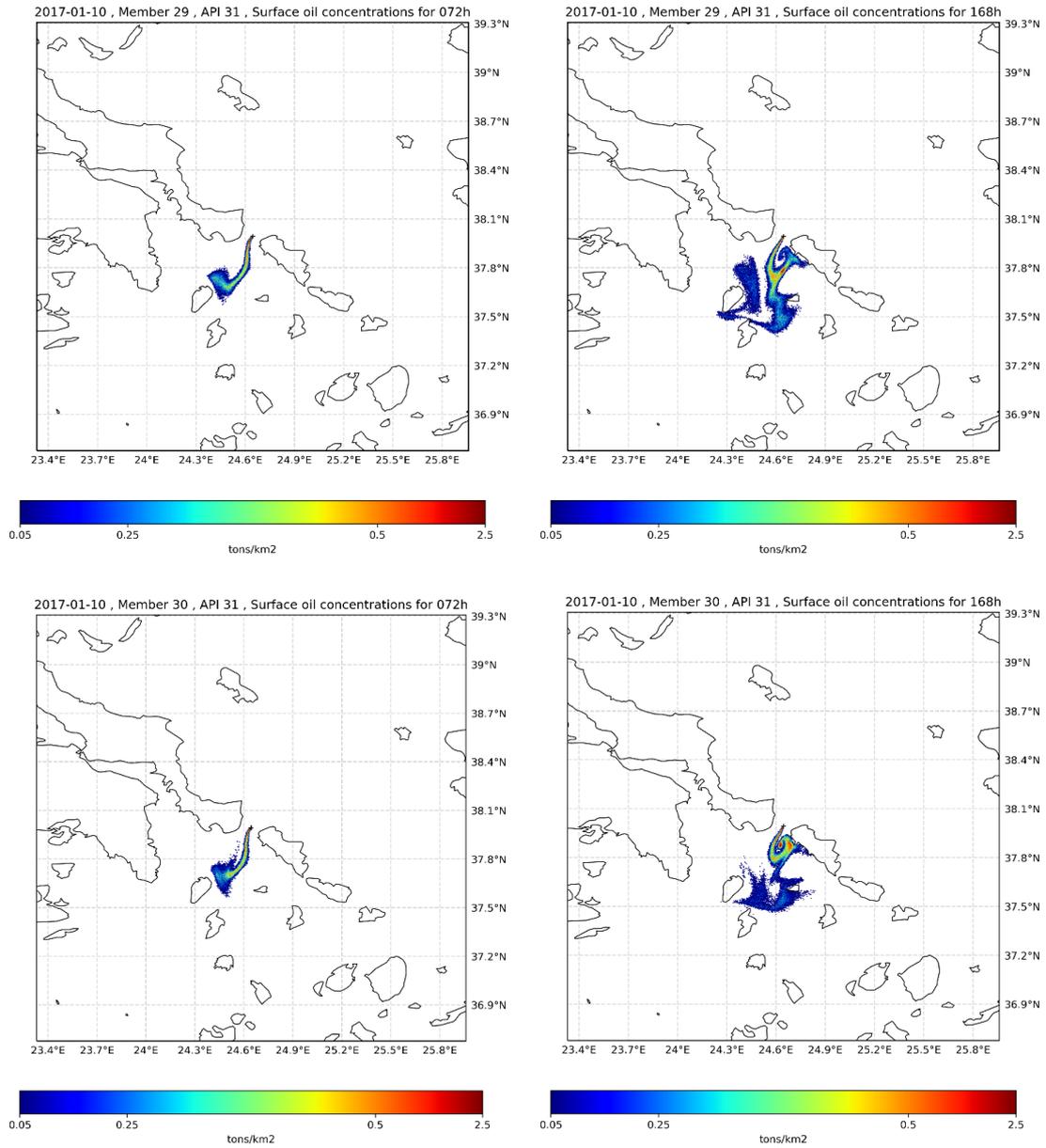


Figure 3.2: Surface oil concentrations of selected ensemble members for winter and run time of: 72 hours, 168 hours.

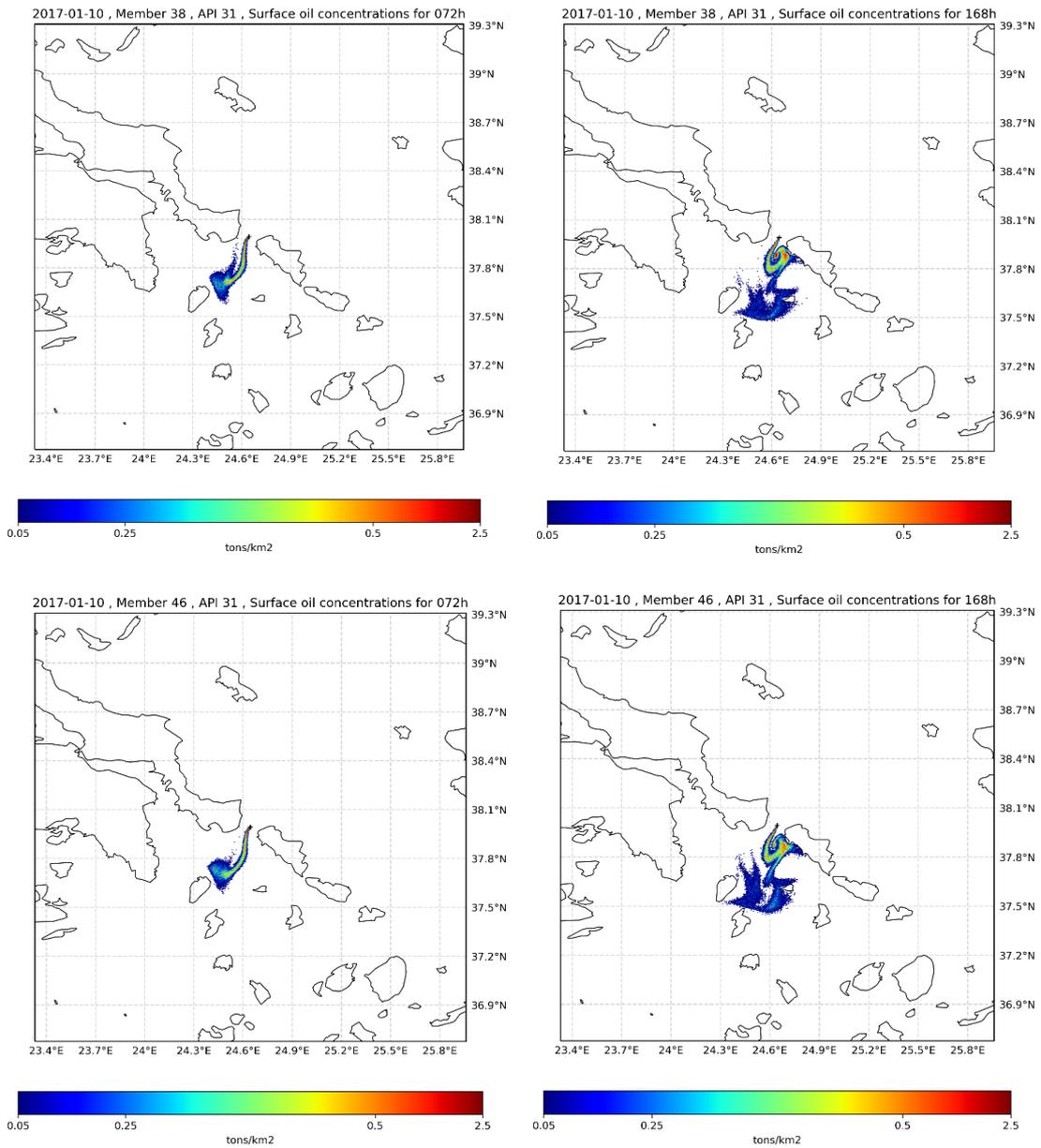


Figure 3.2: Surface oil concentrations of selected ensemble members for winter and run time of: 72 hours, 168 hours.

Similarly, figures 3.3 and 3.4 illustrate the surface oil concentrations and spreading of the deterministic and ensemble simulation for spring respectively. The simulation run time for each occasion is the same as before. After 72 hours of simulation, the oil slicks is transporting in a westward direction interacting with the southern coasts of the island of Euboea. Afterwards and between 72 and 168 hours, a change in the direction of the wind is observed causing the transportation of the oil slick southwest.

Differences in the oil spill distribution between the deterministic and the ensemble simulations are also present in spring, although on a lesser extent compared to winter, which is mainly attributed to the lower wind spread of the ensemble and the more gradual change of the wind.

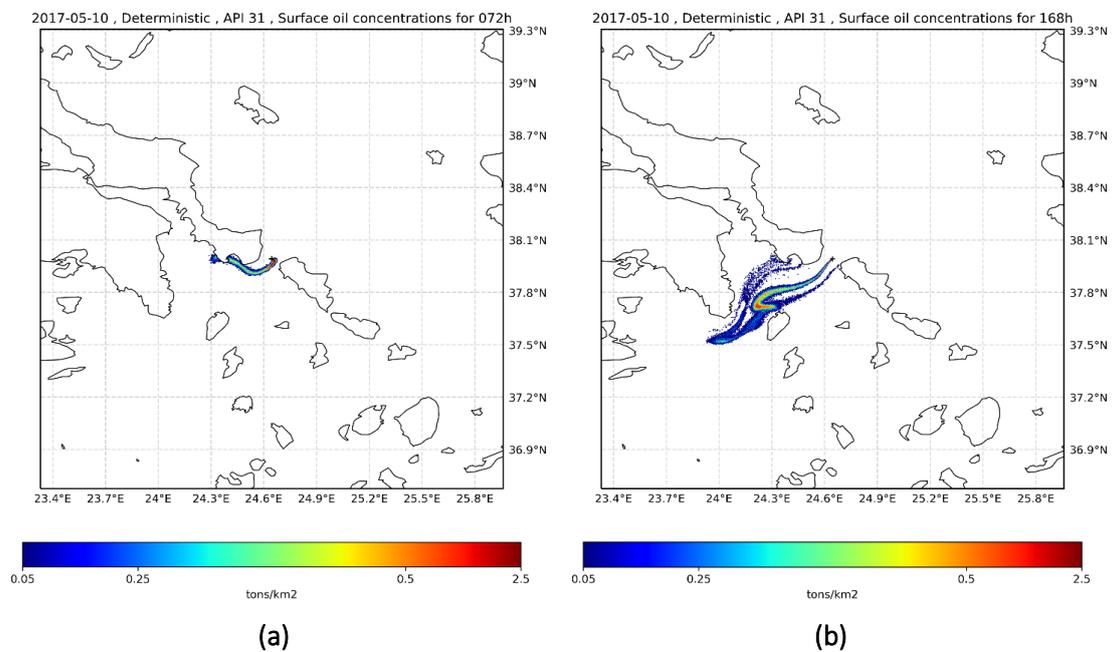


Figure 3.3: Surface oil concentrations of deterministic simulation for spring and run time of: (a) 72 hours, (b) 168 hours.

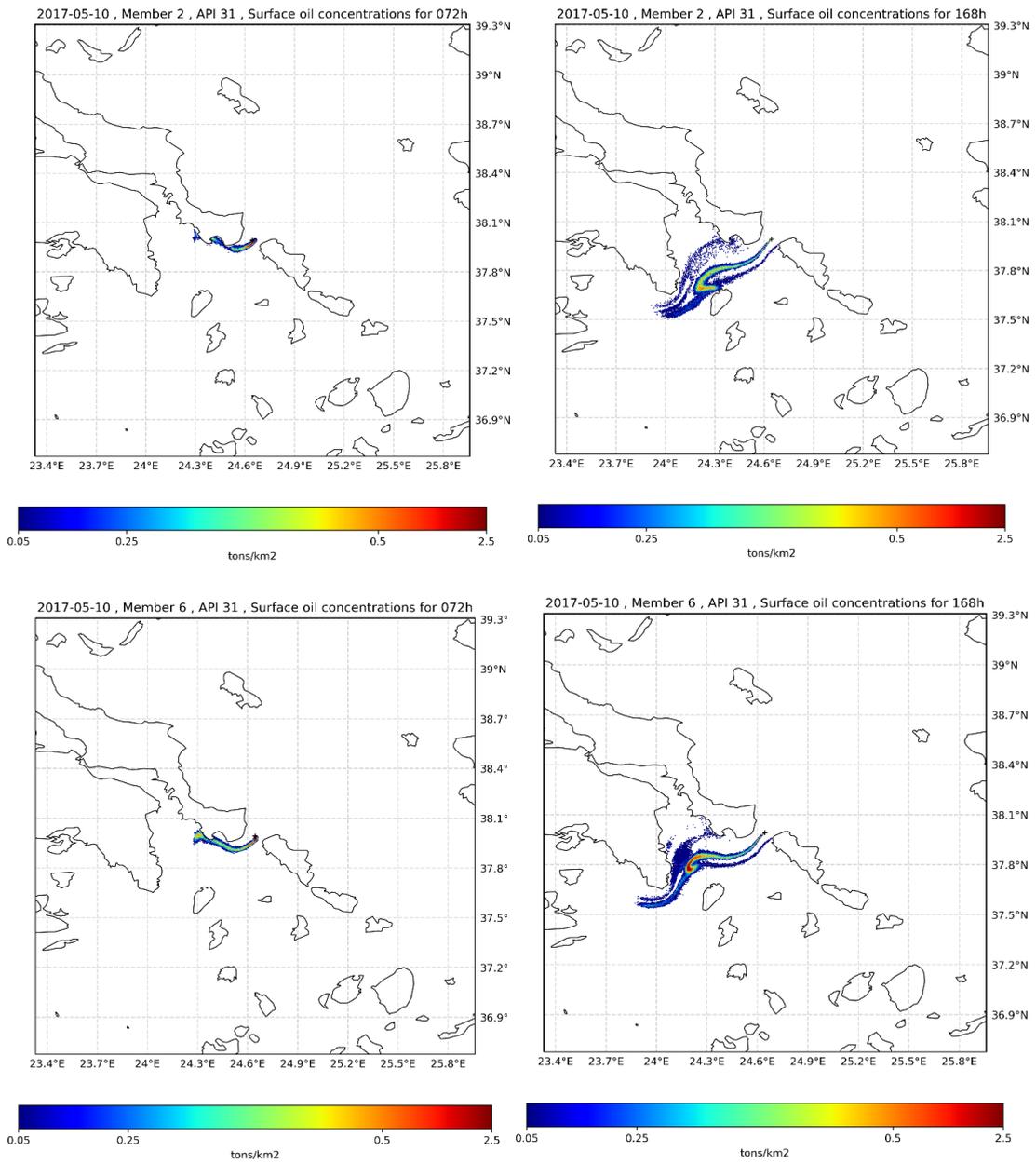


Figure 3.4: Surface oil concentrations of selected ensemble members for spring and run time of: 72 hours, 168 hours.

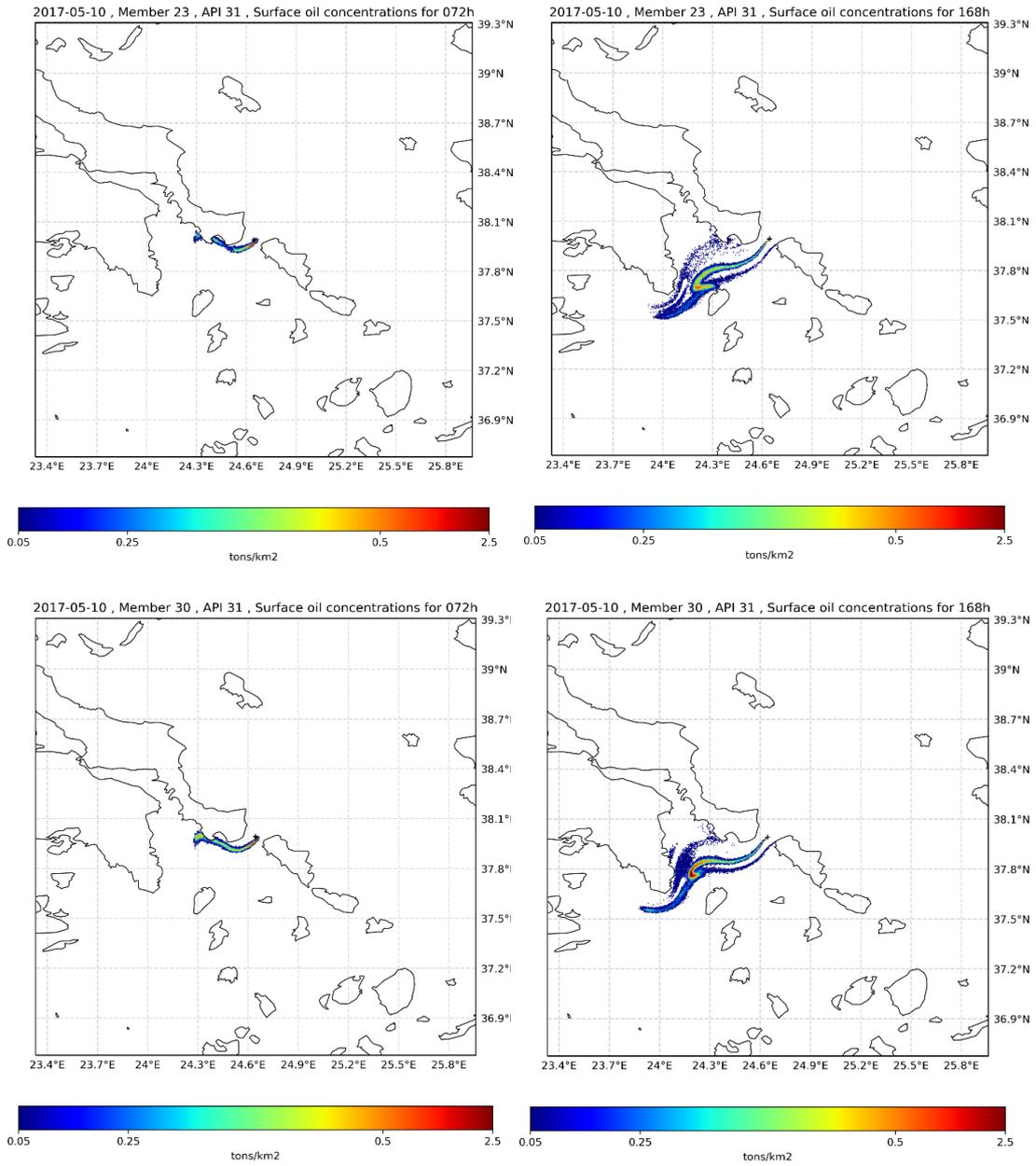


Figure 3.4: Surface oil concentrations of selected ensemble members for spring and run time of: 72 hours, 168 hours.

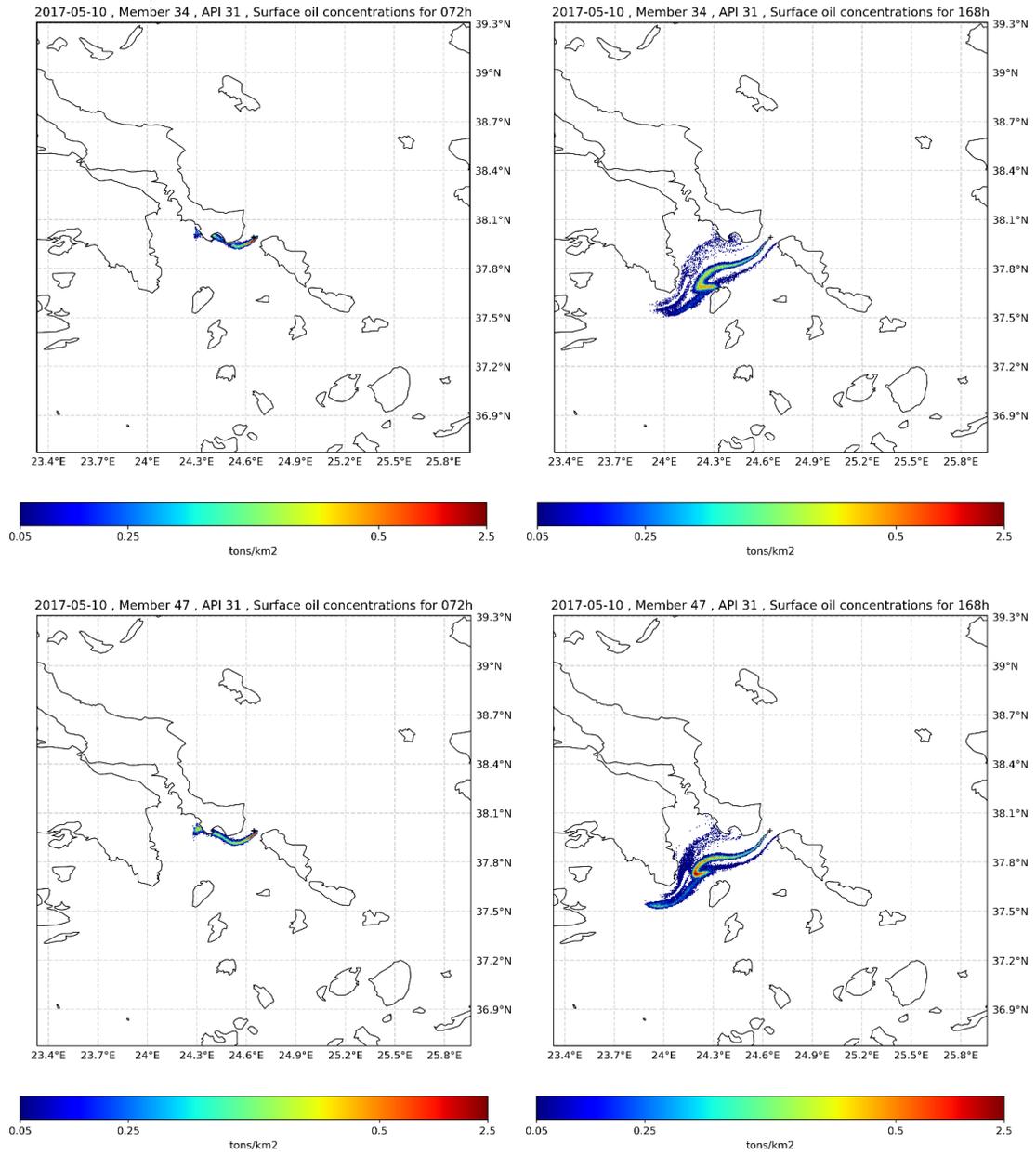


Figure 3.4: Surface oil concentrations of selected ensemble members for spring and run time of: 72 hours, 168 hours.

3.2 Uncertainty assessment for oil spill spread

Figure 3.5 presents the convex hull of the deterministic and ensemble members for winter and spring. The simulation run times are 72 and 168 hours. Both surface and subsurface parcels, as well as those deposited on the coast were used in the computation, to describe in the best way possible the extent of the oil spills simulated by MEDSLIK II and the extent of the area they affect.

Convex hull comparison of deterministic and ensemble simulations

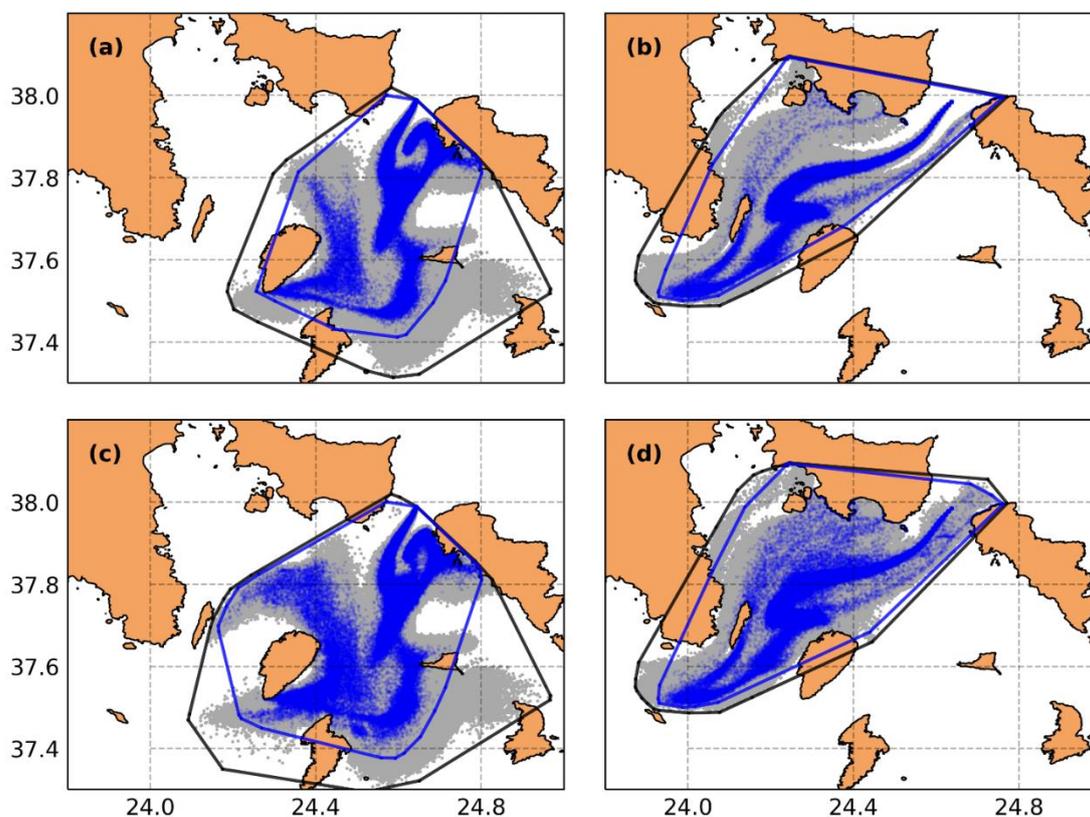


Figure 3.5: Convex hull comparison of deterministic and ensemble simulations after 168 hours:
(a) surface particles – winter 2017, (b) surface particles – spring 2017,
(c) all particles – winter 2017, (d) all particles – spring 2017
Blue color represents the deterministic case while grey color represents the ensemble members.

By examining figure 3.5 we can easily conclude that the most significant differences occur in winter season. This is due to the changes of the wind direction, which in the winter season are more frequent compared with the spring season. Another factor contributing to that conclusion is the climatology of the examined area. In particular, the prevailing wind direction, the maximum wind speed and the marine currents which have been discussed in chapter 2. For the winter season the prevailing wind direction is north – northeast rendering it opposite to the marine currents in the area, while the wind speed shows peak values above 12 m/sec. In contrast, for the spring season, the case is south and between 8-10 m/sec respectively.

Keeping this in mind, we proceed to quantify the difference between the convex hulls of figure 3.5 (cases a to d). The results are as follows:

➤ Winter season

The surface parcel difference between the total ensemble and deterministic convex hulls is 65.44% (figure 3.5a), while for all parcels it is 54.21% (figure 3.5b)

➤ Spring season

The surface parcel difference between the total ensemble and deterministic convex hulls is 24.92% (figure 3.5c), while for all parcels it is 21.83% (figure 3.5d)

The results are not surprising, since we expected a much greater uncertainty concerning the spreading of the oil spill in the winter season compared to the spring season case.

These variations in the extent of the area affected by the ensemble oil spills are also shown in figure 3.6, where we present the metrics A_{exceed} (discussed in chapter 2) and DA. The first quantifies the area of the deterministic convex hull that exceeds the area of each member's convex hull. The second quantifies the area difference and percentage change between the deterministic and the total ensemble convex hull. Both metrics present variations for the whole duration of the simulations.

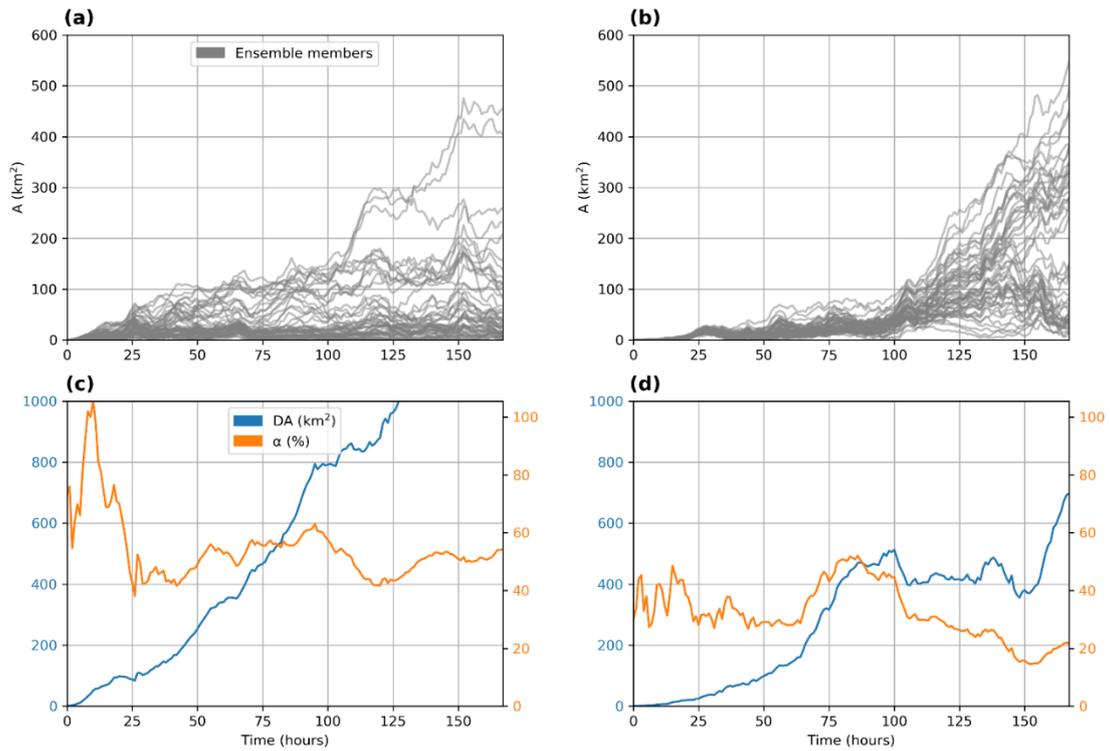


Figure 3.6

- (a) Area of each ensemble member exceeding the deterministic (A_{exceed}) for winter season
- (b) Area of each ensemble member exceeding the deterministic (A_{exceed}) for spring season
- (c) Area difference (DA) and percentage change between deterministic and total ensemble convex hull for winter season and through the whole runtime simulation
- (d) Area difference (DA) and percentage change between deterministic and total ensemble convex hull for spring season and through the whole runtime simulation

Observing figures 3.6a and 3.6b, which correspond to winter and spring season respectively, we can identify the differences in the wind field direction and its sudden changes as well as the existence of the perturbation factor that we have discussed before. In figure 3.6a the metric A_{exceed} increases rapidly for several members reaching 100 km² in less than 36 hours and reaches a maximum of around 480 km² in 152 hours. The moment the A_{exceed} spread increases rapidly and the difference in the extent of the oil spills and the area they affect, could prove important knowledge in the formation of the mitigation strategy and the expected length of the deployed booms. In contrast, in figure 36.b the metric A_{exceed} increases with a slower pace reaching 100 km² in about 100 hours and a maximum of 540 km² at the end of the simulation run time.

These differences are expected since in the winter period there are significant changes in the wind field direction intensified by the incorporation of the perturbation factor and the absence of early beaching, compared to the spring period where the case is quite the opposite.

Observing figures 3.6c and 3.6d we can conclude the following. Firstly, as expected, the difference between the size of the deterministic and the combined ensemble convex hulls (DA metric) for the winter period is much greater than the one for the spring period. In particular, DA in the winter period exceeds 1000 km² after barely 125 hours of run time, while in the spring period it reaches a maximum of almost 700 km² at the end of the simulation. Secondly, all major differences occur during the first day of the simulation run time, which is shown by the metric α and represents the percentage change between deterministic and total ensemble convex hull.

To further evaluate the uncertainty in the oil spill, we use the RMSE and the uncertainty index s , described in detail in chapter 2. As expected, RMSE is increasing with time (figure 3.7a), with higher values and a more gradual increase in the winter case. In spring (figure 3.7b), RMSE displays a discontinuous increase with a significant peak after about 70 hours of simulation run time, mostly due to the high interaction of the oil spill with the coastline around this time periods. Afterwards, the RMSE appears to be increasing again, although it doesn't reach the maximum value of the winter case.

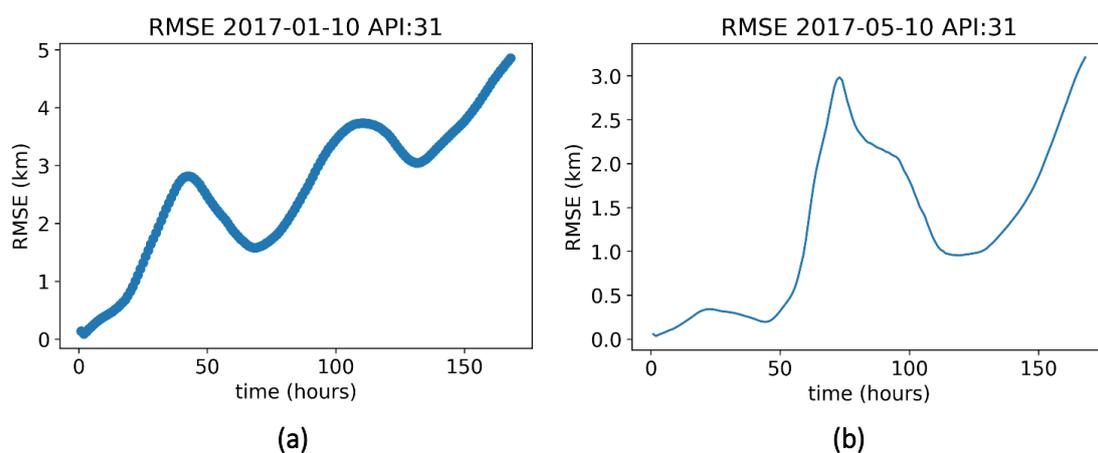


Figure 3.7: RMSE for winter (a) and spring (b) cases

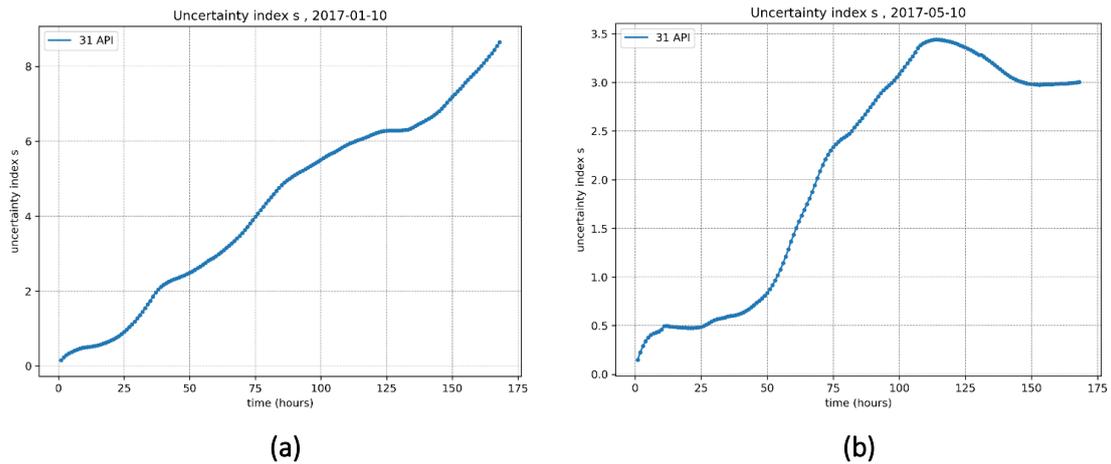


Figure 3.8: Uncertainty index s for winter (a) and spring (b) cases

These observations are supported by the uncertainty index s , presented in figure 3.8. In winter (figure 3.8a) the uncertainty index s increases over time presenting an almost linear growth. Unlike the winter case, in spring (figure 3.8b), uncertainty index s displays lower overall values, fluctuations in its growth, and a decrease in its value after the first 110 hours.

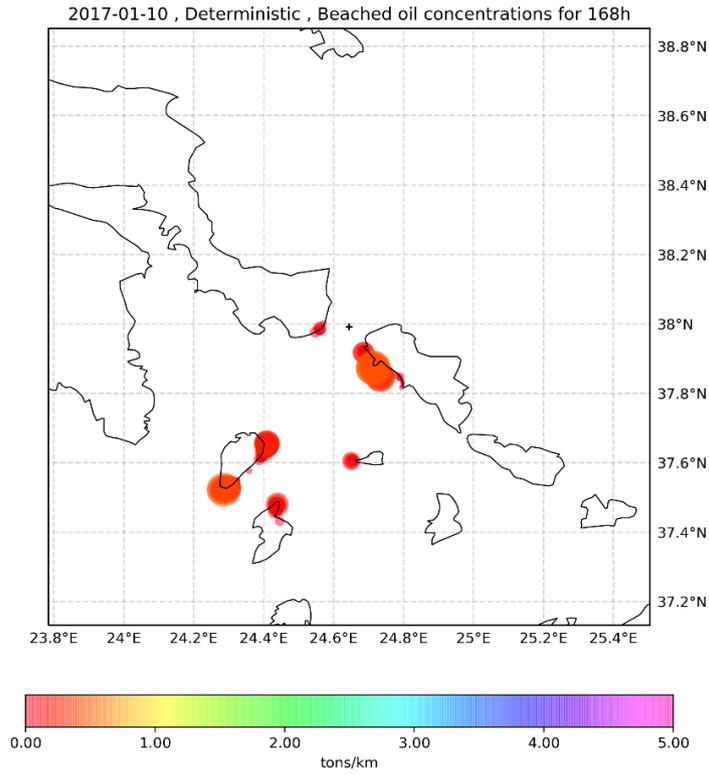
Summarizing our results we can conclude that the differences between spring and winter seasons can be explained by the variations of the direction of the wind field which are intensified by the incorporation of the perturbation factor. Of all the metrics used, the s index especially is used to show the contribution of the atmospheric ensemble in terms of increasing the oil spill uncertainty, so as the ensemble prediction be more useful operationally with respect to the deterministic prediction.

3.3 Beached oil

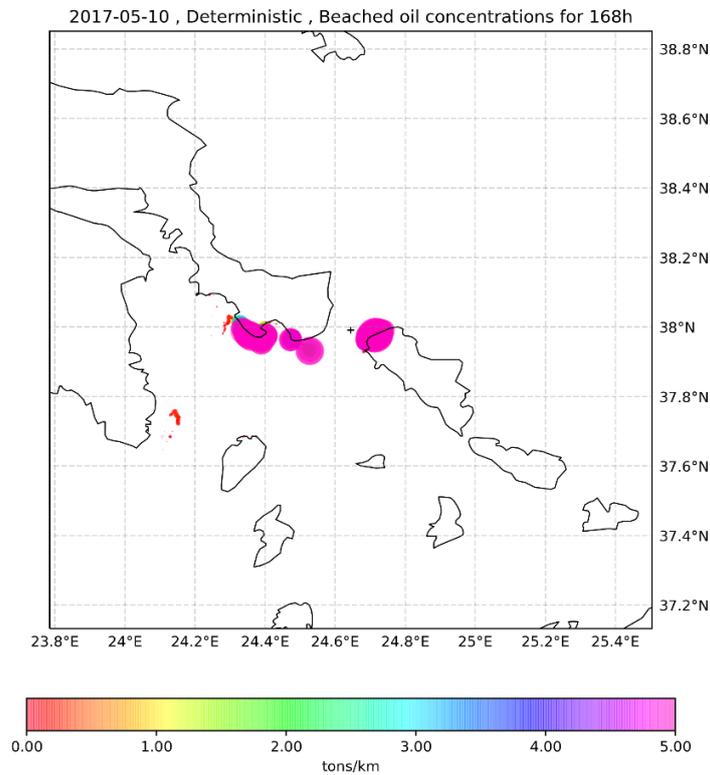
Figure 3.9 illustrates the beached oil concentrations of the deterministic simulation for winter and spring cases at the end of the simulation run time (168 hours). The different concentrations of beached oil particles are represented by different colors. By comparison, greater concentration of oil is expected in the spring case. As seen before, this happens due to the prevailing wind direction which at this case is west causing early interaction of oil with the shoreline of Euboea.

Figures 3.10 and 3.11 illustrate the beached oil concentrations of the ensemble simulation for both season cases. By comparison, in figure 3.10 we can observe significant differences amongst the ensemble members of the winter case as well as with the deterministic case (figure 3.9a) indicating a high degree of uncertainty in the amount of beached oil concentrations. These differences occur because of the differences and uncertainty of the ensemble spread which is discussed in section 3.2. In contrast, observing figure 3.11 all ensemble members display greater oil concentrations compared to figure 3.10 and nearly no differences in the areas of beaching compared to the deterministic case (figure 3.9b).

Nevertheless, in both cases, the uncertainties generated using the ensemble approach for the wind forcing and the information about the possible concentration values and possible locations of the beached oil, provide a useful tool to assess the impact on the coastal area and take the necessary precautions.



(a)



(b)

Figure 3.9: Beached oil concentrations of deterministic simulation for winter (a) and spring (b) cases

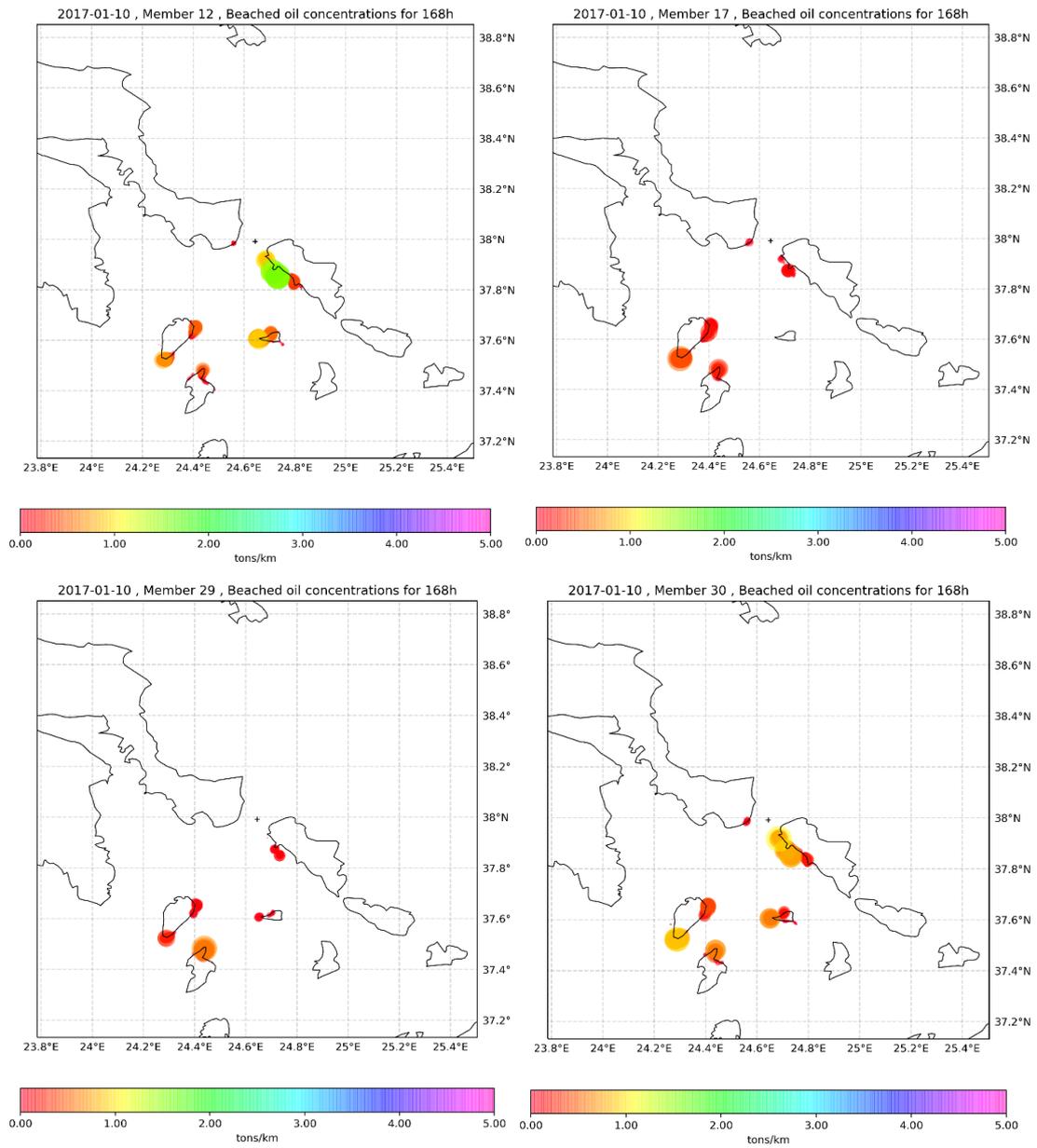


Figure 3.10: Beached oil concentrations of some ensemble members for winter case

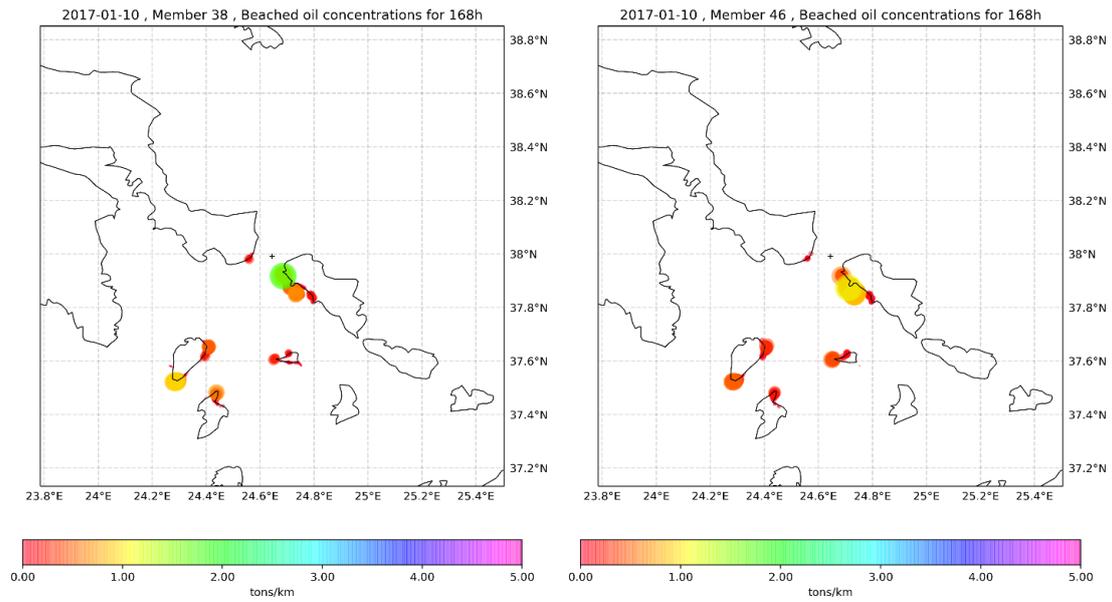


Figure 3.10: Beached oil concentrations of some ensemble members for winter case

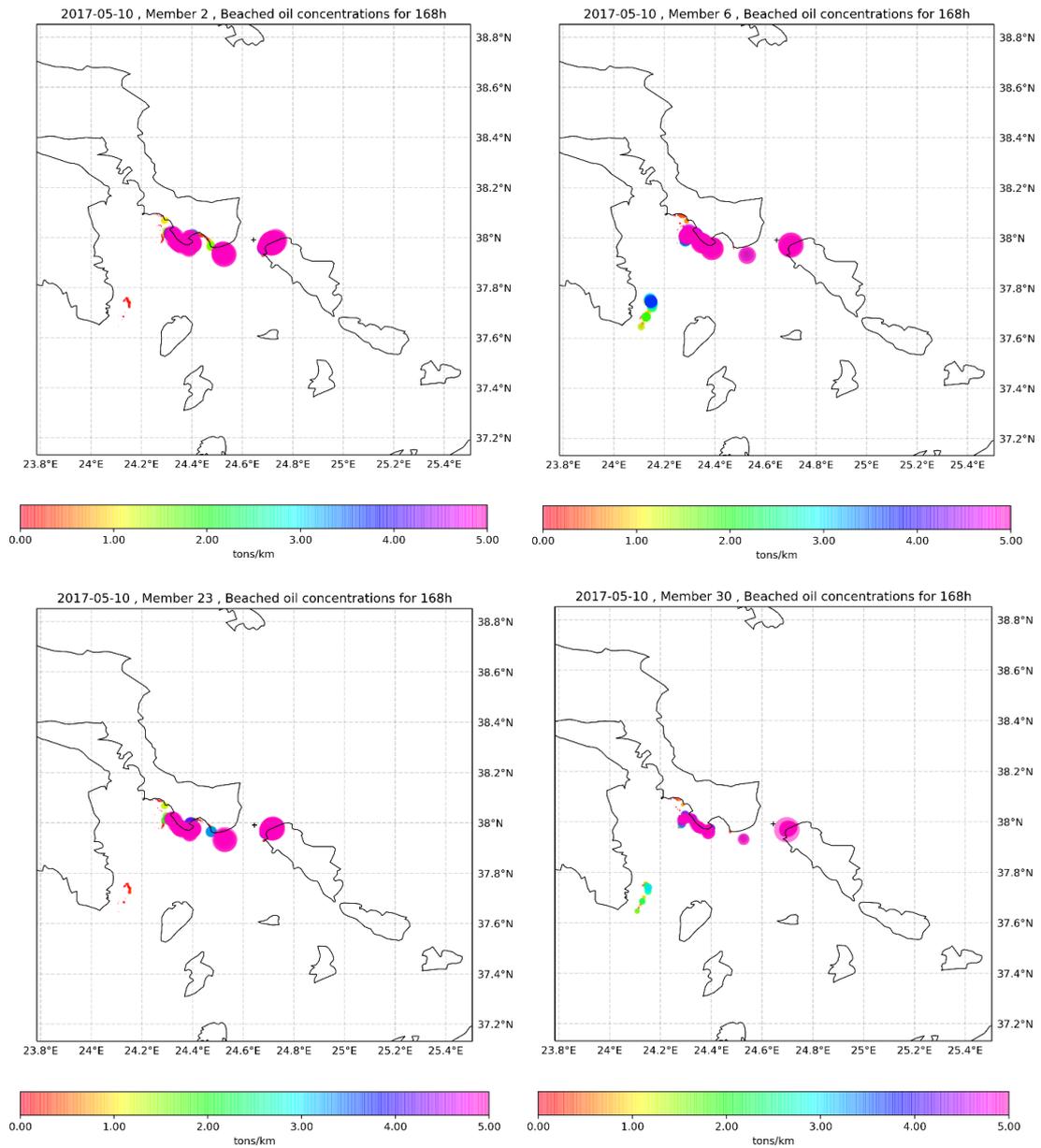


Figure 3.11: Beached oil concentrations of some ensemble members for spring case

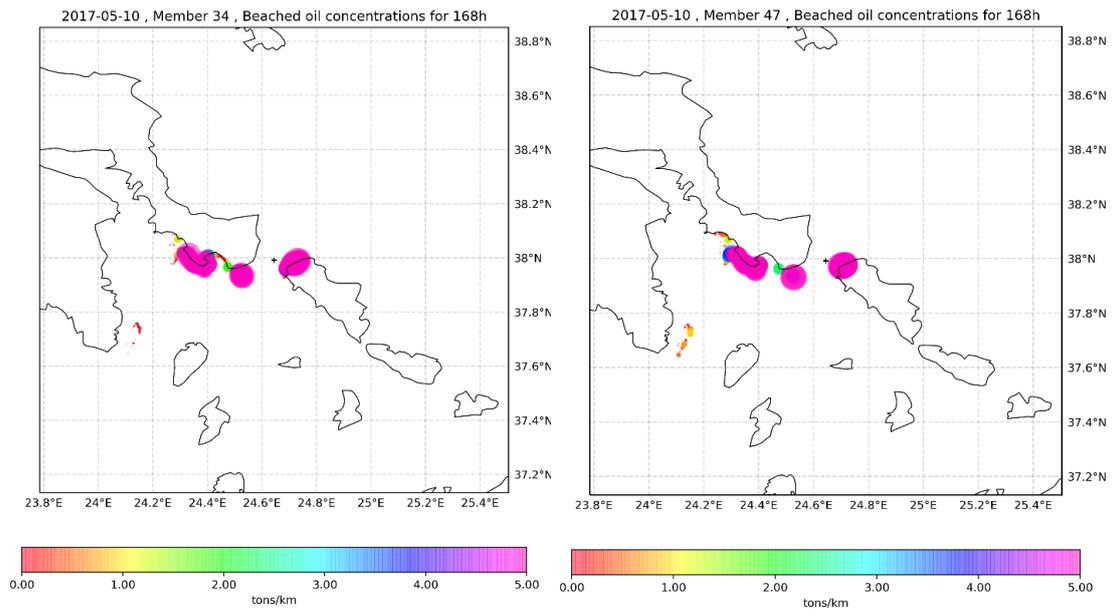
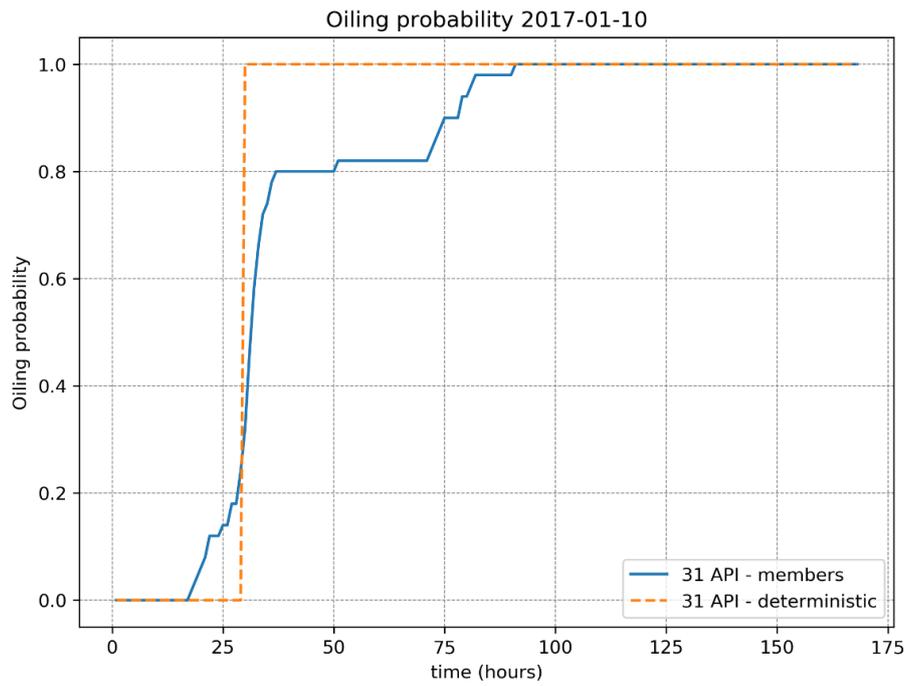


Figure 3.11: Beached oil concentrations of some ensemble members for spring case

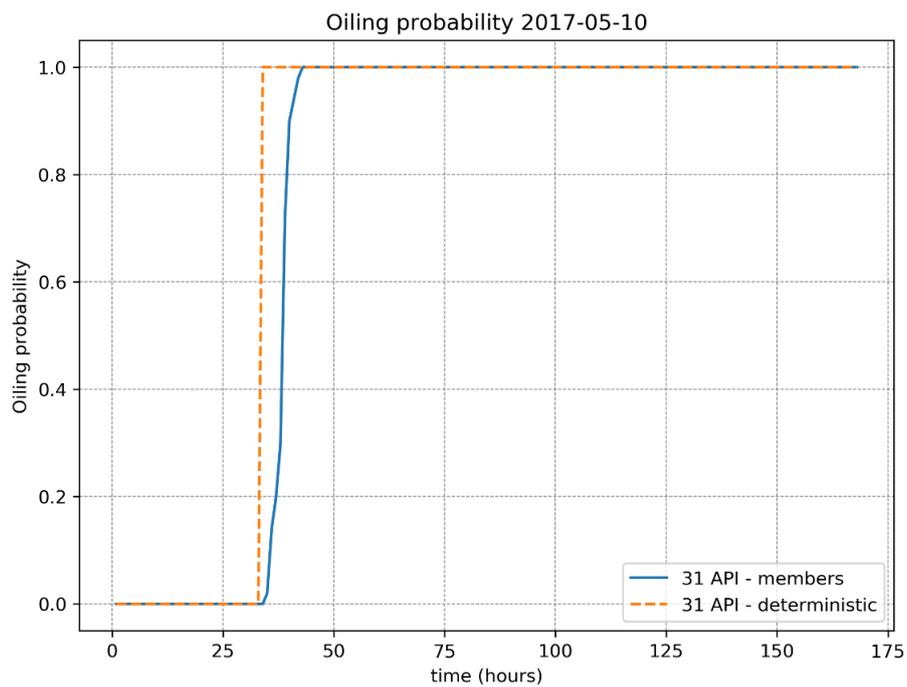
3.3 Uncertainty assessment for beached oil

In order to quantify the temporal uncertainty for the beached oil, we used oiling probability as defined in chapter 2, to examine the probability of the presence of fixed oil on the coast, across the 50 ensemble members in total, for every time step. Oiling probability shows an uncertainty in the hit time for beached oil for the ensemble runs, in contrast to the binary event of a single deterministic run.

As shown in figure 3.12a, in winter, the uncertainty in the temporal window for the beached oil for the ensemble members begins at 15 hours and ends at 90 hours, having a duration of 75 hours. This means that according to the performed simulations, before the 15-hour time mark and after the 90-hour time mark, the probability of the presence of total fixed oil on the coast (oiling probability) is 0% and 100% respectively, while for the duration between these two time marks the probability varies in time as more members predict oil beaching. In spring, as shown in figure 3.12b, the uncertainty window for the ensemble members is quite shorter than in winter, most likely due to the lower ensemble spread of the oil spill, as discussed previously. The duration of this uncertainty temporal window is approximately 8 hours, beginning at the 33-hour mark and finishing at the 41-hour mark.



(a)



(b)

Figure 3.12: Oiling probability of deterministic and ensemble members for (a) winter and (b) spring

Chapter 4 – Conclusions and recommendations for future research

4.1 Summary and conclusions

This study aims at evaluating the impact of the atmospheric forcing uncertainty on the performance of the oil spill modelling and the dispersion of the pollutants in the marine environment of the Aegean Sea, following a probabilistic approach for the simulation of the oil spill. Stochastic wind forcing based on Empirical Orthogonal Functions (EOF) modes was used and generated an ensemble of oil spill forecasts using oil spill model MEDSLIK II, in order to better represent the predictability of the atmospheric forcing. We investigated uncertainties like the spreading of the oil slick, the extent of the oil spill, including both surface and subsurface as well as the oil beached on the coast, the amount of total fixed oil on the coast and finally temporal uncertainties regarding the oil beaching time. Finally, we examined, whether the uncertainty information generated by the ensembles is important, and therefore, if an ensemble approach for the atmospheric forcing can improve the information provided by the oil spill forecasting.

Wind forcing influences heavily the oil transportation in the study area. The sensitivity analysis performed indicates the importance of the wind on the performance of the oil spill modelling especially in nearshore areas. All the ensemble oil slicks simulated by MEDSLIK II, spread mainly in the downwind direction, but there are variations in the transport and the evolution of the shape and size of the oil spills between the ensembles. Oil spill ensemble spread increases over time, representing the uncertainty generated from the incorporation of the perturbation factor in the wind forcing data. The extent of the area affected by the ensemble oil spills in total, is found to be greater than the area only affected by the deterministic simulation, providing additional information with respect to the deterministic approach.

Our results show clearly that the uncertainty increases over time causing a deterministic approach for the wind forcing to limit the accuracy and the information provided by the oil spill model forecasts, regardless of how good the forecast model is. Thus, the use and the added value of a probabilistic approach for the wind forcing that takes into account the predictability of the atmospheric forcing, in relation to the deterministic, and the simulation of ensemble oil spill forecasts that predict several equally possible oil spill states becomes more important over time.

In conclusion, the ensemble approach as described above has great potential benefits and seems to improve the possibilities of the oil spill modelling, by predicting the possible extent of the oil spill and the subsequent area that may be affected, the volume and location of the beached oil, and by providing a possible time window for the appearance of oil in the coastal area. This knowledge could prove an important tool to better plan and direct the available resources for the control and mitigation procedures, in the event of an oil spill.

4.2 Suggestions for future work

- MEDSLIK II calculates the wave-induced velocity (Stokes drift) using an empirical formulation that depends on wind amplitude ([De Dominicis et al., 2013a](#)). Future studies could fruitfully explore this issue further by using complex numerical wave models to calculate the Stokes drift and examine the potential effects it may have on the ensemble approach for the atmospheric forcing.
- Also, it will be important that future research verify our results against observations.

References

- Bjornsson H., Venegas S. A. (1997). A manual for EOF and SVD analysis of climatic data.
<http://muenchow.cms.udel.edu/classes/MAST811/eof.pdf>
- De Dominicis, M., Pinardi, N., Zodiatis, G., & Archetti, R. (2013b). MEDSLIK-II, a Lagrangian marine surface oil spill model for short-term forecasting – Part 2: Numerical simulations and validations. *Geoscientific Model Development*, 6(6), 1871–1888.
<https://doi.org/10.5194/gmd-6-1871-2013>
- De Dominicis, M., Pinardi, N., Zodiatis, G., & Lardner, R. (2013a). MEDSLIK-II, a Lagrangian marine surface oil spill model for short-term forecasting – Part 1: Theory. *Geoscientific Model Development*, 6(6), 1851–1869.
<https://doi.org/10.5194/gmd-6-1851-2013>
- Douglas M. G. (2018). *Quantitative Methods of Data Analysis for the Physical Sciences and Engineering*
<https://doi.org/10.1017/9781139342568>
- GEBCO Compilation Group. (2014). GEBCO 2014 Grid. Retrieved from
<https://www.gebco.net/>
- Kampouris, K., Vervatis, V., Sofianos, S. (2021). Oil spill model uncertainty quantification using an atmospheric ensemble. *Ocean Sci. Discuss.* 17, 919–934.
<https://doi.org/10.5194/os-17-919-2021>
- Lardner, R., Zodiatis, G., Hayes, D., Pinardi, N. (2006). Application of the MEDSLIK oil spill model to the Lebanese Spill of July 2006. European Group of Experts on Satellite Monitoring of Sea Based Oil Pollution. European Communities.
- Lardner, R., Zodiatis, G., Loizides, L., & Demetropoulos, A. (1998). An operational oil spill model for the Levantine Basin (Eastern Mediterranean Sea). *International Symposium on Marine Pollution*.

- Liu, Y., & Weisberg, R. H. (2011). Evaluation of trajectory modeling in different dynamic regions using normalized cumulative Lagrangian separation. *Journal of Geophysical Research*, 116(C9), C09013.
<https://doi.org/10.1029/2010JC006837>
- Liubartseva, S., De Dominicis, M., Oddo, P., Coppini, G., Pinardi, N., & Greggio, N. (2015). Oil spill hazard from dispersal of oil along shipping lanes in the Southern Adriatic and Northern Ionian Seas. *Marine Pollution Bulletin*, 90(1–2), 259–272.
<https://doi.org/10.1016/j.marpolbul.2014.10.039>
- MarineTraffic – Global Ship Tracking Intelligence
www.marinetraffic.com).
- NITTIS, K., & PERIVOLIOTIS, L. (2002). Circulation and hydrological characteristics of the North Aegean Sea: a contribution from real-time buoy measurements. *Mediterranean Marine Science*, 3(1), 21.
<https://doi.org/10.12681/mms.255>
- Preisendorfer R. W., Mobley C. D. (1988). *Principal Component Analysis in Meteorology and Oceanography*
https://www.researchgate.net/profile/Curtis-Mobley-2/publication/44529688_Principal_component_analysis_in_meteorology_and_oceanography_by_Rudolph_W_Preisendorfer_posthumously_compiled_and_edited_by_Curtis_D_Mobley/links/541875e90cf2218008bf3da1/Principal-component-analysis-in-meteorology-and-oceanography-by-Rudolph-W-Preisendorfer-posthumously-compiled-and-edited-by-Curtis-D-Mobley.pdf
- Vervatis V., Testut C. E., De Mey-Frémaux P., Ayoub N. (2016). Data assimilative twin-experiment in a high-resolution Bay of Biscay configuration: 4DEnOI based on stochastic modeling of the wind forcing
<https://doi.org/10.1016/j.ocemod.2016.01.003>
- Zodiatis G., Lardner R., Georgiou G., Kallos G., P. N. (2005). Operational oil spill modeling predictions in the Mediterranean. 4th EuroGOOS Conference: European Operational Oceanography Present and Future, Brest, 6-9 June, 131–132.

Zodiatis, G., Coppini, G., Perivoliotis, L., Lardner, R., Alves, T., Pinardi, N., ... Sepp Neves, A. A. (2017). Numerical Modeling of Oil Pollution in the Eastern Mediterranean Sea. In A. Carpenter & A. G. Kostianoy (Eds.), Handbook of Environmental Chemistry (Vol. 83, pp. 215–254). https://doi.org/10.1007/698_2017_131

Zodiatis, G., Lardner, R., Alves, T. M., Krestenitis, Y., Perivoliotis, L., Sofianos, S., & Spanoudaki, K. (2017). Oil spill forecasting (prediction). Journal of Marine Research, 75(6), 923–953. <https://doi.org/10.1357/002224017823523982>