How sensitive is a simulated river plume to uncertainties in wind forcing? A case study for the Red River plume (Vietnam)

Tung Nguyen-Duy, Nadia K. Ayoub, Pierre de Mey-Frémaux, Thanh Ngo-Duc

Ocean Modelling 186 (2023) 102256

Contents lists available at ScienceDirect

Ocean Modelling

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

How sensitive is a simulated river plume to uncertainties in wind forcing? A case study for the Red River plume (Vietnam)

Tung Nguyen-Duy [a,b,*], Nadia K. Ayoub [a], Pierre De-Mey-Frémaux [a], Thanh Ngo-Duc [b]

[a] Université de Toulouse, LEGOS (CNRS, IRD, CNES, UPS), Toulouse, France
[b] LOTUS International Joint Laboratory, University of Science and Technology of Hanoi, Vietnam Academy of Science and Technology, Hanoi, Viet Nam

ARTICLE INFO

Keywords:
Ensemble modeling
Atmospheric perturbation
Unsupervised classification
River plume
Gulf of Tonkin
Red River delta

ABSTRACT

This paper aims at characterizing the sensitivity of a simulated plume's properties to uncertainties in the wind fields which force the ocean model using an ensemble method. The case study is the Red River plume in the Gulf of Tonkin. The variability of the Red River plume in the mid and far field is described in a previous paper using a clustering analysis (Nguyen-Duy et al., 2021) and is shown to be mostly driven by monsoon winds and tides. In the present study, we also aim at assessing the robustness of the classification with respect to the wind forcing uncertainties. The variability of the wind uncertainty is estimated as 60% of the wind variability with a higher variability near the coast. Based on that estimation, two ensembles of 50 simulations each with perturbed wind forcing are run over the summer 2015 period. Then, we examine the ensemble spread (defined as the standard deviation across the members) of the wind stress and of the ocean variables. The coastal current shows similar spread for both meridional and zonal flows, with the highest spread related to the highest wind stress spread. The sensitivity is the largest at the surface for salinity and at the base of the mixed-layer for temperature. The properties of the river plume are analyzed. The spread of the plume area is maximum in August, which is the same time as when the plume is the most spread out. The clustering analysis applied to the ensemble members shows some cluster attribution shifts between different members, but the cluster that is most likely to occur is still the one from the reference simulation (with unperturbed wind). These limited changes suggest that the cluster analysis of the reference simulation in Nguyen-Duy et al. (2021) is indeed robust to the wind forcing errors. The uncertainty on the plume thickness is typically less than 2m, sometimes reaching 4m (for a total thickness of 10m). The freshwater transport mainly follows the variations of the current due to the changes of wind. Possible implications of this study for the assimilation of high-frequency radar data are discussed at the end of the paper. Firstly, the relevance of the ensemble in simulating the model errors is assessed: the comparison with the data suggests that the model suffers from systematic errors that are not represented by the ensemble (by construction). Secondly, the ensemble is used to provide examples of model correction in a hypothetical data assimilation, highlighting its potential to constrain the plume by correcting directly the surface salinity, but also correcting the surface coastal current and the wind stress.

1. Introduction

The Gulf of Tonkin (GOT) is a shelf sea located east of Vietnam and south of China (Fig. 1**A**). It is a shallow area, with depth less than 100m. It connects with the open sea through the Hainan Strait and southwest of the Hainan island. The Red River is the main source of freshwater in the gulf, providing a mean runoff around 3300 m^3/s (Luu et al., 2010), and is the second-largest river in Vietnam in terms of discharge.

Rivers, in general, are vectors of exchange between the continent and the ocean. They bring freshwater that impacts the buoyancy of the coastal waters and therefore the coastal dynamics. Rivers carry matter and sediments that influence the coastal morphology (bathymetry, coastline). They also bring along other kinds of contaminants such as plastic debris. It is therefore crucial to fully understand the fate of the river water from the estuaries to the ocean, especially in the mid-field and far field, where the plume is mostly driven by its buoyancy, Earth rotation, wind and ambient currents in addition to the initial momentum of the river flow. In GOT, the variability of the Red River plume has been examined by several authors. Using numerical results combined with observational data, Gao et al. (2013) and Rogowski et al. (2019) show

* Corresponding author at: LEGOS (CNRS, IRD, CNES, UPS), Université de Toulouse, Toulouse, France.
E-mail address: nguyenduytung@hotmail.com (T. Nguyen-Duy).

https://doi.org/10.1016/j.ocemod.2023.102256
Received 12 March 2023; Received in revised form 10 June 2023; Accepted 19 August 2023
Available online 21 August 2023

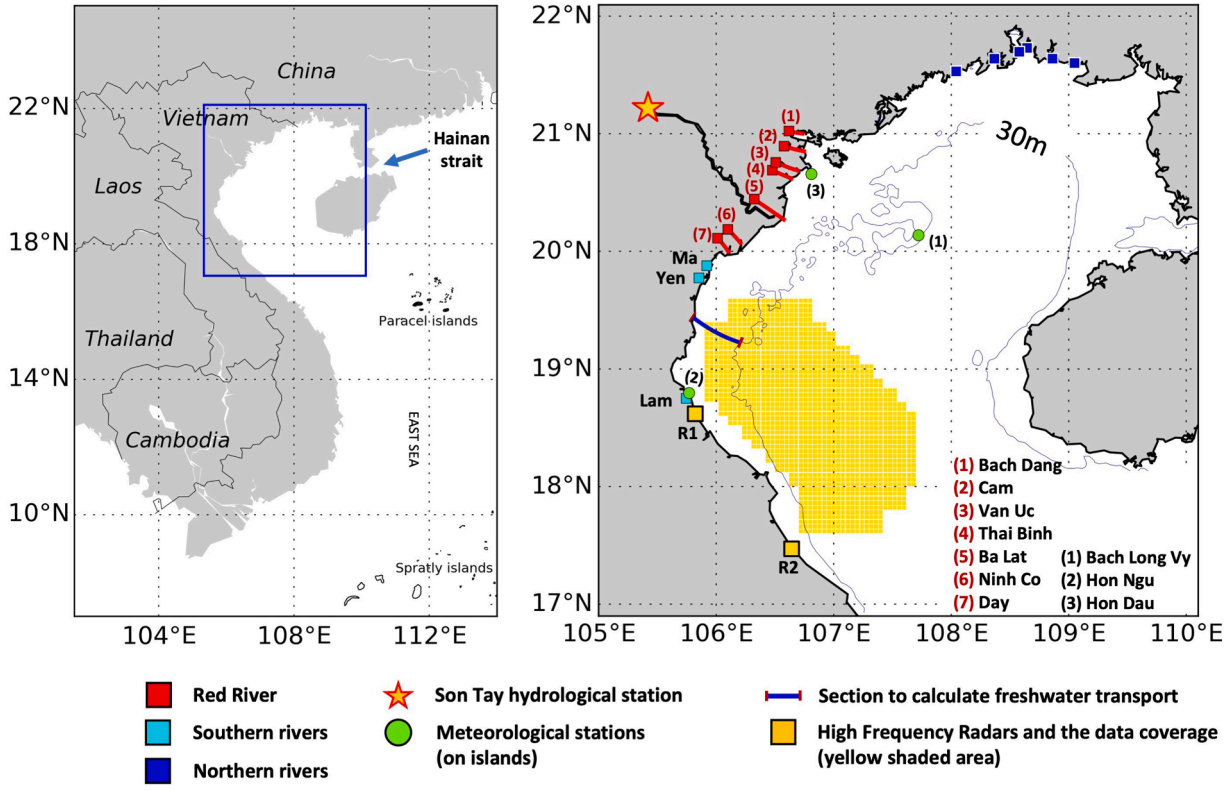

Fig. 1. **(A)** Location of the Gulf of Tonkin. **(B)** Location of the rivers, section and data used in the paper. The thin grey line indicates the bathymetry isoline of 30m. Red numbers indicate the name of the Red River mouths, while black numbers show the name of the meteorological stations (on islands).

that in winter, the plume forms a coastal buoyant current while spreading more offshore in summer. The difference between the two seasons is due to the monsoon wind, which is northeasterly in winter and southwesterly in summer, and to the runoff variability with the highest (lowest) Red River runoff in summer (winter).

The variability of the plumes from the Red River and from three rivers south of the Red River (namely the Ma, Yen and Lam rivers) was further examined by Nguyen-Duy et al. (2021, hereinafter referred to as ND21), with the help of a high resolution, multi-year 3D simulation. This simulation was later used by Tran et al. (2022) together with high-frequency radar observations to examine the impact of the freshwater on the dispersive properties of the coastal circulation. In ND21, the plumes of the Red River and of the southern rivers are identified using a passive tracer. Most of the time, they are mixed with each other and are therefore studied as a single plume that is called for simplicity 'the Red River plume' in ND21 and in this study. The results of ND21 confirm that the plume has a large variability, both in terms of surface coverage and thickness of the plume. A cluster analysis was performed in order to classify the main patterns of the Red River plume over the period 2011–2016. The classification is based on the K-means method and is described in detail in ND21 (see also their Appendix C in the Supplementary Material) and in Nguyen-Duy (2022). Four clusters were identified. The first one represents winter conditions, when the runoff is lowest, the current is southward, and the wind is northeasterly. In this cluster, the plume has the smallest area and flows mainly in a ~25 km-wide band downstream (southward) along the Vietnamese coast. In the second cluster, the plume extends a little bit more, and represents the transition period between the two monsoon seasons. From June to August, the plume patterns are mainly represented by clusters 3 and 4. Fig. 2 displays the tracer concentration at surface, sea surface salinity (SSS), and wind conditions, corresponding to cluster 3 and cluster 4, but averaged from June to August 2015 (the focus period in this article) instead of over the whole period (2011 – 2016) adopted in ND21. The third cluster features the summer monsoon, when the wind is southerly and spreads the plume more offshore and northward; the plumes from the different rivers are then disconnected. The fourth cluster is present when the runoff is the highest and the wind is still upwelling favorable; it has the largest surface plume coverage. The transition from a downstream coastal current to an upstream current under the effect of upwelling-favourable winds is observed in other off-equatorial plumes (e.g. Hickey et al., 2009; Yankovsky et al., 2022).

The plume thickness is also strongly affected by the wind, as illustrated in ND21. In winter, when the wind is downwelling favorable, the plume sticks to the coast and reaches the bottom. In summer, the runoff is three times higher than in winter, but since the plume spreads offshore it forms a thinner layer (between 5 and 10 m, see ND21).

The work of ND21, as many studies on river plumes in other areas (e.g. Vaz et al., 2018; Wang et al., 2014; Jia et al., 2019), relies on numerical modeling. However, the model results are known with some uncertainties that can come from the model itself (physical assumptions, parameterizations, etc.), from the simulations' configuration (prescribed coastline and bathymetry, parameters, etc.), or from the forcing (atmospheric forcing, lateral boundary forcing, initial conditions, etc.). Understanding the uncertainties and their sources can help us better assess the model results in different scenarios, and quantifying the uncertainties is a necessary first step toward data assimilation. One way to estimate model uncertainties is to use ensembles of simulations, to compute statistics from the ensembles, such as spread and quantiles, and to analyze the probabilistic distribution. Ensemble methods in oceanography were first introduced for data assimilation purposes (e.g. Evensen, 2003), and then have been used as well for uncertainties estimates or budgets in open ocean (e.g. Lucas et al., 2008) or coastal

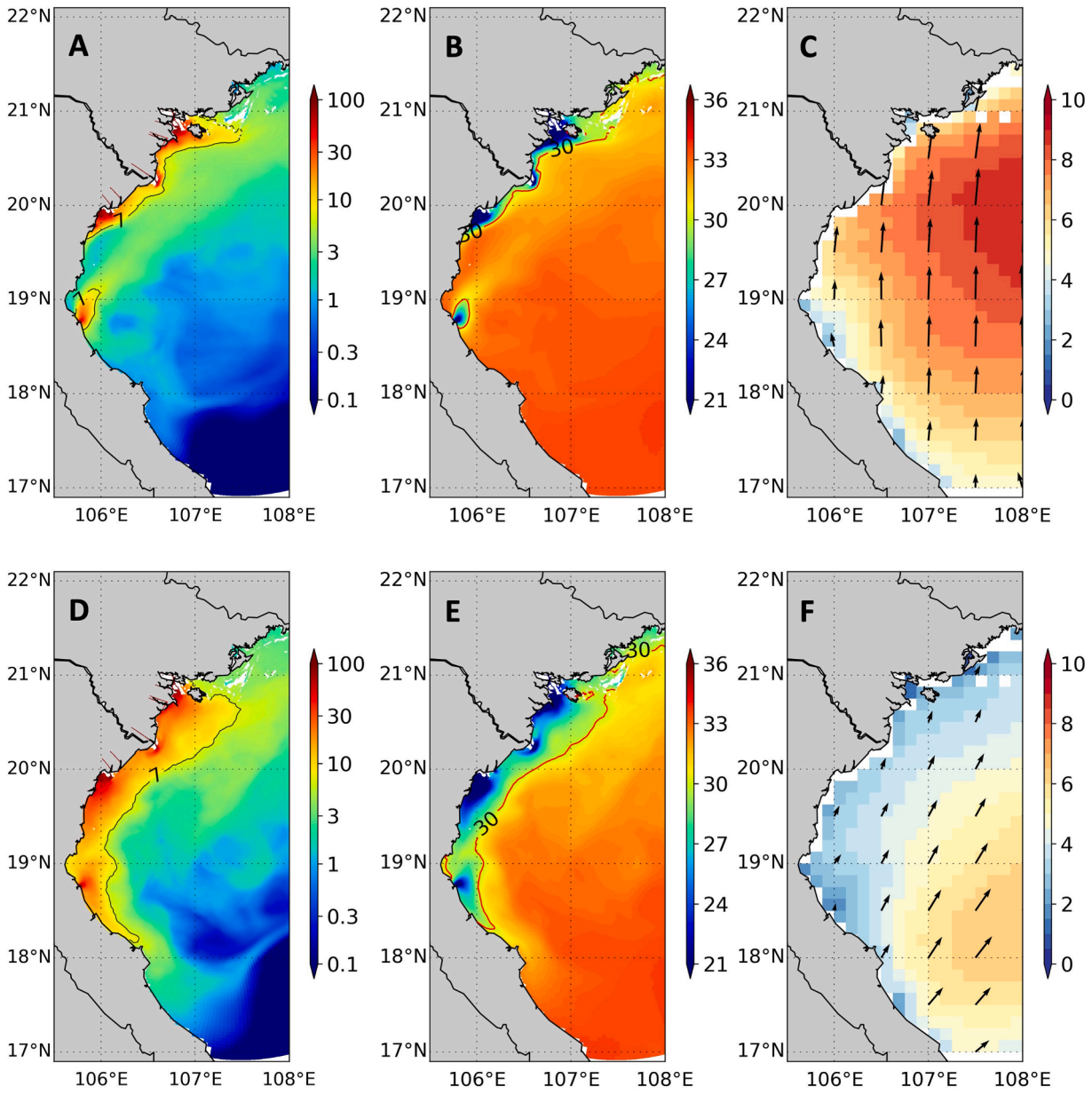

Fig. 2. **(A):** Surface tracer concentration (arbitrary unit/m^3), averaged from June to August 2015 over the days when the plume is classified into cluster 3. **(B):** Same as **(A)** but for SSS. **(C):** Same as **(A)** but for 10m wind from ECMWF (in m/s). **(D–F):** same as **(A–C)** but averaged over the days when the plume is classified into cluster 4. The contour lines in (A, D) show the area where the surface concentration is equal to or higher than 7 unit/m^3, which is the threshold to identify the river plume.

ocean studies (e.g. Auclair et al., 2003; Kim et al., 2011), and generalized to coupled models such as physical-biogeochemical ones (e.g. Vervatis et al., 2021a). Probabilistic modeling is also implemented in numerical weather prediction systems, such as ECMWF[1], while ensembles of models are used to assess seasonal to interannual predictions (e.g. Kirtman et al., 2014; Herrmann et al., 2021).

Previous studies have evidenced the large impact of errors in the surface forcing on coastal simulations (e.g; Auclair et al., 2003; Vervatis et al., 2021b). Given the role of wind forcing on the Red River plume properties evidenced in ND21, we aim in the present study at analyzing the sensitivity of the coastal dynamics to uncertainties on wind forcing and to evaluate the robustness of the plume characteristics described in ND21. More specifically, we use stochastic modelling to address the following questions: what are the spatial and time structures of the sensitivity of the coastal dynamics and tracer dilution pathways to wind perturbations? Can we identify the physical processes at work? Are the plume properties, and in particular the typical patterns identified by the cluster analysis of ND21, robust to wind errors? In the final section, we take a preliminary look at the problem of data assimilation: could the model errors due to wind uncertainties estimated from the ensembles be used in the context of HF radar data assimilation?

Our approach is based on the analysis of two ensembles of simulations generated by perturbing the wind forcing. A first step is to create the wind perturbations and to examine if the wind uncertainty is realistically reproduced by comparing the ensemble wind speed with observational wind data from meteorological stations and a satellite product. We then examine the response of the model to those uncertainties, at the surface and in the subsurface layer. We focus on the

[1] https://www.ecmwf.int/en/about/media-centre/focus/2022/30-years-ensemble-forecasting-ecmwf

plume properties by revisiting the results of ND21. We also discuss the impact of wind uncertainties on freshwater transport through a coastal section. Finally, we explore the possibility of using the ensembles as a proxy of prior errors for assimilating surface velocity observations from high-frequency radars. We start by assessing whether the estimated uncertainties from the ensemble are statistically consistent with the model error deduced from model-observation misfits. We then conduct a very preliminary analysis of the impact of such observation on model variables.

The period of study is from June to August 2015. It corresponds to the rainy season with high runoff and the period of strong wind variations.

This paper's outline is as follows. In Section 2, the method to create the wind perturbations is described; then, the model's ability to reproduce the wind uncertainty is assessed. The results from the ensemble simulation on the sensitivity of the ocean variables and of the plume variability are presented in Section 3. In Section 4, we discuss the use of the ensembles for data assimilation. The conclusion and perspectives are provided in Section 5.

2. Methods

2.1. SYMPHONIE model and reference configurations

The simulations are conducted with the SYMPHONIE model implemented over the Gulf of Tonkin. Both SYMPHONIE and the Gulf of Tonkin configuration (hereafter named GOT_REF) have been described in detail in ND21. We just recall essential information here.

SYMPHONIE is a numerical model that solves the primitive, Boussinesq, hydrostatic equations of ocean circulation (Marsaleix et al., 2006; 2008) and has been specifically designed to simulate the coastal ocean. For instance, SYMPHONIE can accurately represent tides, intertidal areas, and the oceanic response to high-frequency atmospheric forcing in shallow areas. It can run on highly variable polar or bipolar grids in order to refine the resolution locally, for instance at river mouths or inside estuaries.

The GOT_REF configuration has first been set-up by Piton et al. (2020) and has then been updated by ND21. It has a variable horizontal grid, with a fine horizontal resolution of 300m near the Red River mouths and a coarser resolution of 4.5km near the open boundary. The vertical discretization consists of 20 sigma levels. At the open boundaries, tidal surface elevation and current at K1, O1, P1, Q1, K2, M2, N2, S2, M4 frequencies from the tidal atlas FES2014 (Lyard et al., 2021) are taken into account. The model is also forced by daily averages of sea surface height, 3D zonal velocity, meridional velocity, temperature, and salinity fields, from the global analysis (hereafter 'OGCM') produced by Mercator-Océan International and provided by Copernicus Marine Systems (CMEMS) at a resolution of 1/12° (Mercator Océan, 2020). We adopted the implementation of open-boundary conditions as described in Toublanc et al. (2018). At the surface, operational European Centre for Medium-Range Weather Forecasts (ECMWF) analyses (with a spatial resolution of 1/8°) are used to provide 3–hourly wind, precipitation, solar energy, atmospheric temperature, dew-point temperature, and surface pressure. Fluxes of momentum, heat, and freshwater are then computed using the bulk formulae of Large and Yeager (2004). The K-ε turbulence closure scheme is used, with the implementation described in Michaud et al. (2012). The QUICKEST scheme is used for the advection and diffusion of tracers (Neumann et al., 2011). Horizontal advection and diffusion of momentum are respectively computed with a 4th-order centered scheme and a bi-harmonic scheme, while a 2nd-order centered scheme is used for vertical advection of momentum (Damien et al., 2017).

There are 3 main river systems in the gulf: (1) the Red River, (2) the Ma, Yen, and Lam rivers south of the Red River delta and hereafter referred to as the 'southern rivers', and (3) six rivers north of the delta and along the Chinese coast, hereafter referred to as the 'northern rivers' (Fig. 1**B**). In GOT_REF, the Red River is configured with daily discharge data obtained from the National Hydro-Meteorological Service of Vietnam (NHMS) at Son Tay hydrological station (close to the apex of the delta); the discharge is then split into 7 distributaries using the results of Vinh et al. (2014). In this configuration, the delta is represented by its main 7 distributaries. Each distributary is connected to the ocean through an estuary with a realistic bathymetry. The estuary is extended upstream by a channel with an idealized rectangular shape and sloping depth to damp the tidal propagation (the channels are illustrated by the red boxes numbered (1) – (7) in Fig. 1). The length of the channel is chosen to exceed the salt water intrusion as explained in ND21 and illustrated in Nguyen-Duy, (2022). For the smaller rivers at the north and south of the delta (blue and cyan points in Fig. 1), there are no estuaries and channels; the salinity and temperature are prescribed to seasonally varying values. The boundary condition at the most upstream entry point of the channel or directly at the river mouth into the ocean (southern and northern rivers) is described in detail in Appendix B of ND21 (provided in supplementary material) and is not repeated here. In particular, the current is adjusted to the dynamics of the downstream flow and is not necessarily uniform on the vertical. For the southern and northern rivers, monthly mean climatological runoffs from NHMS and from Gao et al. (2013) respectively are used. According to these data, the Red River, southern and northern rivers account for 60%, 30%, and 10% of the total runoff, respectively. To simulate the fate of the freshwater, we inject a passive tracer at the Red River and southern rivers' entry points, with a concentration equal to 100 arbitrary unit/m^3. The amount of tracer units entering the ocean is therefore varying according to the discharge. As in ND21, the river plume is identified as the area where the tracer concentration is equal to or higher than 7 unit/m^3. Using the tracer (and not surface salinity) to detect the plume is intended to avoid the freshening effect due to the high precipitation in the summer. The threshold of the tracer is chosen so that the plume area fits with the area identified by a criterion on sea surface salinity (sea surface salinity less than 30) in the low precipitation period, as suggested by ND21. We have checked that the impact of evaporation (which tends to increase the tracer surface concentration) is not significant: we made a rough estimate of the ratio of the change of surface concentration over the surface concentration, averaged over the area of interest, and found that it varies around 4‰ and does not exceed 15‰. This is much smaller than the impact of the perturbations on the wind presented in the next sections and is therefore neglected in the rest of the paper.

The long reference simulation (GOT_REF) has been assessed with several datasets (Piton et al., 2021; Nguyen-Duy et al., 2021; Nguyen--Duy, 2022) and shows adequate skills in reproducing the ocean state, both at the surface and sub-surface. Nothing indicates a major deficiency in the model or configuration. Therefore, in this paper, we do not duplicate the assessment published earlier but analyze the sensitivity of the simulations to wind forcing uncertainties.

2.2. Ensemble method

2.2.1. Generation of wind perturbation

The ensemble is generated by perturbing the ECMWF wind forcing. The perturbations are meant to represent uncertainties of the ECMWF wind analyses. Several methods have been developed in past studies, (e.g., Ayoub et al., 2015, Storto and Yang, 2023, Vervatis et al., 2021a). It is difficult to compare their performance as they are not applied in the same purpose nor region and period of study. Here, we assume that the space and time structure of the wind uncertainties are the same as those of the wind variability as in some previous studies (Le Hénaff et al., 2009; Barth et al., 2011; Quattrocchi et al., 2014; Ghantous et al., 2020). This allows us to build the perturbations from the bivariate (zonal, meridional) Empirical Orthogonal Functions (EOFs) of the wind variability as described below.

The perturbed wind condition:

$$\boldsymbol{W_p}(x,y,t) = \boldsymbol{W}(x,y,t) + \sum_{k=1}^{n} \phi_k(x,y)\theta_k(t)\sigma_k(t) \quad (1)$$

where:

$\boldsymbol{W_p}(x,y,t)$: perturbed wind vector
$\boldsymbol{W}(x,y,t)$: original wind vector (ECMWF field)
$\phi_k(x,y)$: spatial component of EOF mode k
$\theta_k(t)$: temporal component of EOF mode k
$\sigma_k(t)$: pseudo-random number
n: number of EOFs used to compute the perturbation (see Section 2.2.2)

Values of σ_k are pseudo-randomly drawn for each EOF mode k and each member; the distribution of σ_k is assumed to be Gaussian, with a zero mean and a given standard deviation ε. The choice of ε is discussed in Section 2.2.3. The σ_k are updated every 5 days independently from the previous drawn 5 days earlier, which reflects the synoptic weather time scale over the region (e.g. tropical cyclones); in between, they are linearly interpolated.

The perturbations are computed at the time step of the ECMWF fields, that is every 3 h. The resulting wind field is then linearly interpolated at every model time step and spatially interpolated onto the model grid.

The ensemble simulation is composed of 50 members, divided into 25 pairs. In each pair, the absolute value of the perturbations is the same, but the signs are opposite. Therefore, at any time, the average of the ensemble perturbations is 0; the average response in the model will be different from zero because of non-linearities.

The ensemble is run from 01 May 2015 to 31 August 2015, starting from initial conditions from the reference run. We then allow for a 1-month spin-up time of the ensemble. All the analyses are carried out from 01 June 2015 to 31 August 2015.

In summary, to create the wind perturbations, the wind EOFs and

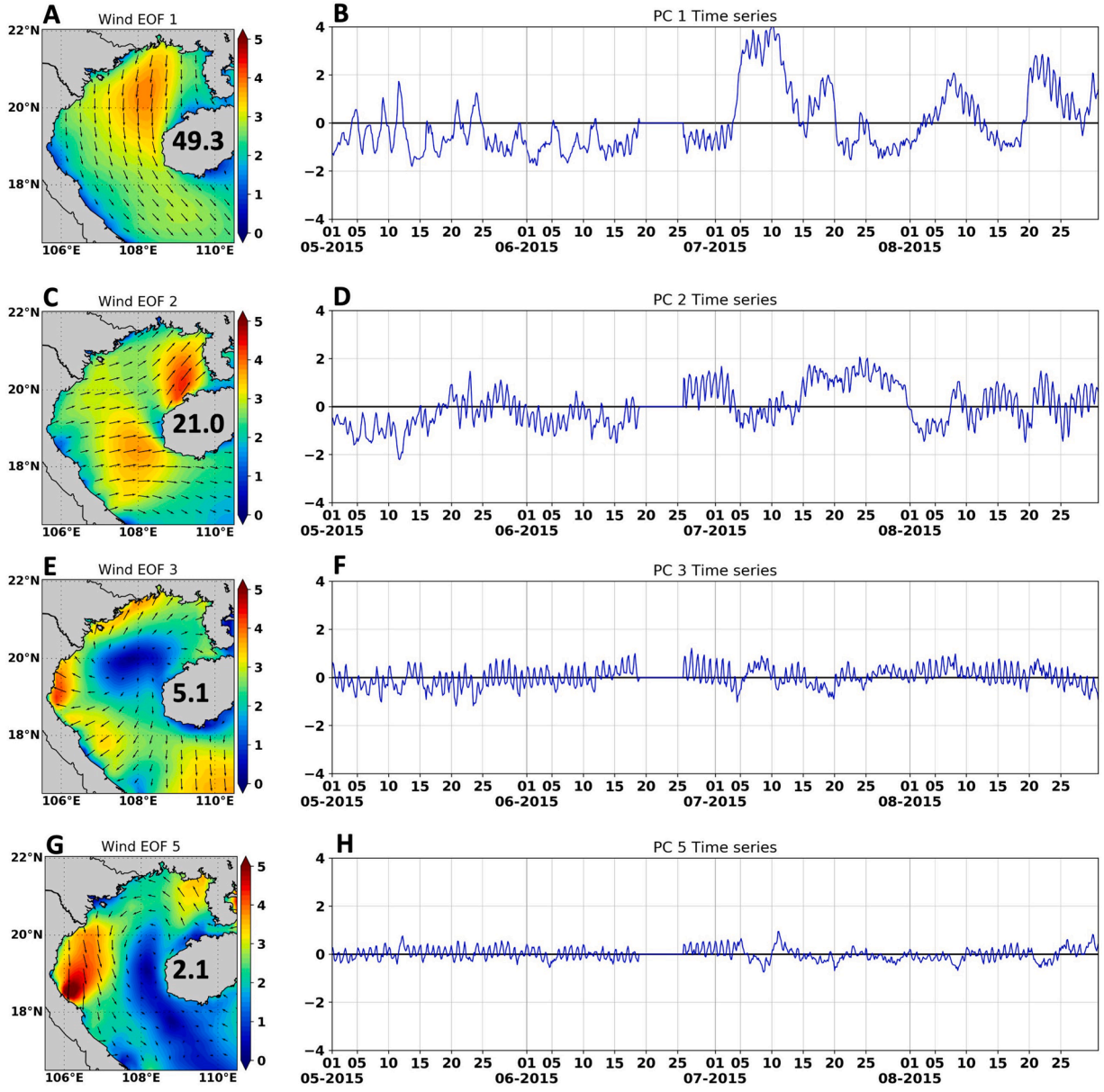

Fig. 3. (A): Spatial component of EOF1 (Unit: m/s). The arrows show the direction while the color map shows the amplitude. The grid over which the arrows are shown is subsampled for the figure, but EOFs are calculated over the full-resolution ECMWF grid. The number indicates the explained variance. **(B)**: Temporal component of EOF1. **(C, D), (E, F), (G, H)**: same as **(A, B)** but for EOF 2, 3, 5, respectively.

pseudo-random numbers are needed. In the next two sections, we examine the choices applied for each of them.

2.2.2. EOF analysis

Fig. 3 shows the spatial and temporal components of the dominant bivariate EOFs that are computed from 3-hourly, 0.25 degree, ECMWF winds from 01 May to 31 August 2015, over the GOT domain and over the ocean points only. From 19 to 26 June, there was a tropical cyclone in the gulf. This extreme event is not within the scope of this study and not well represented by second-order statistics such as EOFs[2], therefore we decided to exclude that period before calculating the EOFs. It means that we chose to set the perturbation to 0 during that period. Therefore, in the ensemble simulations, the model will run with unperturbed wind forcing.

EOF 1 is the main contribution to the wind variability, explaining 49.3% of the total variance. Its spatial component shows a wind mostly in the north-south (northwest-southeast) direction in the northern (southern) part of the GOT (Fig. 3**A**). The velocity is lower near the coast and reaches its highest values near the center - north of the gulf. The temporal components in May and June show negative values, which means that the wind direction is mostly northward, and corresponds to the summer monsoon (Fig. 3**B**). In early July, the wind direction reverses and a strong southward wind event is observed between 5–12 July. After this event, alternating southerlies and northerlies are observed.

EOF 2 accounts for 21% of the total variance, representing mostly a zonal wind pattern with the largest velocity located north of the Hainan island. EOF 3 explains 5.1% of the total variance, with the lowest values at the center of the gulf (near 20°N), and largest at the coast.

The wind perturbations are built from Eq. (1) with a limited number n of EOFs. Here, it would seem reasonable at first glance to keep the first 5 EOFs to reach at least 80% of the total variance. If we included higher modes, the percentage of accumulated variance would not change much (from EOF 5 to EOF 12 it would only increase from 80.3% to 88.6%, not shown). However, EOF 4 shows a very strong wind near the Hainan Strait (not shown). We suspect that its geometric pattern may not be realistic (due to an artifact of data assimilation). Therefore, in this study, EOF 4 is discarded and four EOFs are used (1, 2, 3, 5) with the total variance of ~ 77.5%.

2.2.3. Estimation of the amplitude of the random error

Another challenge of the perturbation generation is the choice of the amplitude of random errors. Quattrocchi et al. (2014); Vervatis et al. (2016), and Ghantous et al. (2020), in their stochastic studies in the Bay of Biscay, generated the ensemble with a Gaussian pseudo-random number generator with standard deviation $\varepsilon = 0.3$, which assumes that the standard deviation of uncertainty on the wind velocity is 30% of the standard deviation of the wind velocity. However, this value is not necessarily general as it is based on the assumptions of the authors for a specific situation. Pasmans and Kurapov, (2019) estimated ε from satellite (ASCAT) and *in situ* observations of wind and from a Bayesian Hierarchical Model; in their case, ε differs according to the rank of the EOFs. They found that in their area of study, the value 30% is usually an underestimation, especially for the high modes.

In this study, we use ASCAT wind products under the following assumption: the difference between ECMWF and ASCAT winds can be used as a proxy of the uncertainty on ECMWF fields. This allows us first to estimate ε, as described below, and then to verify the generated wind perturbations ensemble as presented in Section 2.3.

The ASCAT product distributed by EUMETSAT (EUMETSAT, OSI SAF, 2021) contains level 2 wind data, retrieved from the Advanced Scatterometer (ASCAT) on MetOp-B satellite, at 12.5 km sampling resolution (with an effective resolution of 25 km). It is available daily and can be downloaded from the JPL website: https://podaac.jpl.nasa.gov/dataset/ASCATB-L2-Coastal. The data are downloaded for the period from 01 May to 31 August 2015. The ASCAT User Manual indicates that the accuracy is better than 2 m/s in wind component standard deviation with a bias of less than 0.5 m/s in wind speed.

Since ASCAT data are provided on different tracks on different days, they are first linearly interpolated onto the ECMWF grid. Then, the daily fields where less than 50% of the data are valid over the study region are filtered out; over the 123 days of this period, only 72 days (~58.5%) are used for further calculation.

To estimate ε, our approach is much simpler than the one of Pasmans and Kurapov, (2019) and does not introduce modal dependency. Since ε is built as the ratio between the standard deviation of uncertainty on the wind velocity and the standard deviation of the wind, then here ε is computed as the standard deviation of the difference between ECMWF and ASCAT data (std(diff)) divided by the standard deviation of the ECMWF field (std(ECMWF)). For both the zonal (U_w) and meridional (V_w) components, this ratio is higher at the Red River coast and can be locally higher than 1 (not shown). Further offshore, this ratio decreases to less than 0.5. On average over the domain, this ratio is around 0.8 for U_w and 0.7 for V_w. In conclusion, these comparisons show that it seems reasonable to choose the value of ε within the range [0.5-1.0]. However, we did not take into account the ASCAT measurement error in the above calculations. This is equivalent to assuming that the differences between ECMWF and ASCAT winds are due to uncertainties on ECMWF winds only; this implies that the variability of ECMWF wind uncertainties are likely overestimated. We therefore choose the conservative value of $\varepsilon = 0.6$ for the reference ensemble called ENS_REF.

Higher values of ε near the coast are expected mainly because of the insufficient effective resolution of the ECMWF model for solving the atmospheric dynamics at the land-sea interface. Let's note however that, by constructing the perturbations from the wind fields EOFs, we consider only the uncertainties in the ECMWF model space; in other words, we cannot represent sub-grid scale errors or other omitted processes. Nevertheless, to better reproduce the higher wind uncertainty close to the coast and to assess the sensitivity of the river plume due to coastal wind errors, we configure another ensemble, called ENS_COAST, by doubling in (Eq. (1)) the contributions of EOF 3 and EOF 5, which are both locally intensified at the coast (Fig. 3).

Both ensembles have 50 members. The size of the ensembles has been determined (not shown) as the best compromise between cost (computing and storage) and keeping a large enough number of members to yield reasonably robust second-order statistics.

Table 1 summarizes the list of simulations used in this paper.

Table 1
List of simulations.

Name	Number of members	Characteristics
GOT_REF	1	reference simulation (with unperturbed winds) over 2010-2016 and described in ND21
ENS_REF	50	Reference ensemble with $\varepsilon = 0.6$, over May-August 2015
ENS_COAST	50	Same as ENS_REF but the amplitude of the contributions of EOF 3 and EOF 5 is multiplied by 2

2.3. Ensemble verification

In this section, we will analyze the simulated wind uncertainties from the ensemble using wind measurements first from satellite (ASCAT mission) and then from marine stations. The purpose is to assess whether the ensemble wind can represent the error of the wind forcing. In other words, we address the meaningfulness of the wind perturbations.

We use the ASCAT observations first, considering that the misfits between ASCAT and ECMWF wind velocity components are representative of ECMWF wind velocity errors, as stated in Section 2.2.3. Our

[2] Extreme events contribute asymmetrically to statistical distributions, and as such are better represented by odd-order moments such as skewness.

approach is explained in the Appendix. For each wind component, the variance over the ensemble and time samples (hereafter 'ensemble-time variance') (left-hand side term of Eq. (4) in the Appendix) is compared to the variance of the differences between ASCAT and ECMWF over the study period minus the variance of ASCAT errors (right-hand side terms of Eq. (4)) (Fig. 4). For the latter, we take the maximum value, i.e. 4 m^2/s^2 (see 2.2.3). For both U_w and V_w components, the variance of the misfits varies between 4 and 8 m^2/s^2 offshore and is the largest near the Red River delta, with values higher than 10 m^2/s^2 (Fig. 4**A, B**).

The variance of ENS_COAST shows different spatial patterns (Fig. 4**B, D**) which are consistent with those of the main EOFs. It is characterized by low values near the Red River delta (5 - 9 m^2/s^2 and 10 - 20 m^2/s^2 for U_w and V_w respectively), which is consistent with the variance of the misfits minus the variance of the error. The ensemble-time variance for U_w reaches its maximum near the southern boundary (around 17 m^2/s^2), which is much higher than the variance of the misfits minus the variance of the error. The ensemble-time variance for V_w is maximum in the north of the gulf (more than 30 m^2/s^2). In ENS_REF (not shown), the variance is slightly smaller along the Vietnamese coast than in ENS_COAST, which is consistent with the lower magnitude of EOF 3 and 4 added to the perturbation by structure. Overall, the ensemble-time variance is larger than the variance of the misfits (minus the variance of ASCAT errors). We interpret these inconsistencies as an indication of two possible limitations of our approach: (1) ECMWF errors are biased in the region (their mean over the period of study is non null) which means that we do not represent with our ensemble-time samples all the ECMWF errors (since our samples represent their variable part only, and not the systematic errors), (2) the error variance is not stationary in time and so we cannot adopt the formalism described in the Appendix to assess the ensemble consistency using ASCAT observations.

Locally, in the center of the gulf for U_w and south of 19°N for V_w, the variance of the misfits is lower than 4 m^2/s^2, meaning that the right-hand side term of Eq. (3) (see Appendix) is negative, and Eq. (3) cannot hold. This suggests that ASCAT errors are overestimated there or that ECMWF errors have a non-zero mean over the period of study (contrary to our hypothesis).

With ENS_COAST, the ECMWF errors are better represented than in ENS_REF close to the Vietnamese coast (not shown) but, as ENS_REF, the ensemble is underdispersive.

Another way to assess the ensemble is to check its ability to represent the wind conditions from different observational datasets. Here, wind data (hereafter WIND_ISLAND) is provided by Viet Nam Meteorological and Hydrological Administration. These data contain the wind speed measured at 10m every 6 h; the instrumental error is not reported. We use observations from 3 stations, located on islands (Fig. 1**B**): Bach Long Vy (~120 km from the coast), Hon Ngu (~4 km from the coast), Hon Dau (~1 km from the coast). WIND_ISLAND is used as a 'truth' value to validate the wind ensemble. Since WIND_ISLAND represents the wind velocity at a station, while the ensemble wind represents that in a grid point, WIND_ISLAND velocity is compared to the ensemble wind velocity inside a box of 0.25 degrees around the station. We also use ASCAT wind speed data at the closest point to the station within 0.25 degrees.

In Bach Long Vy, most of the time (71 over 85 days), ENS_REF is able

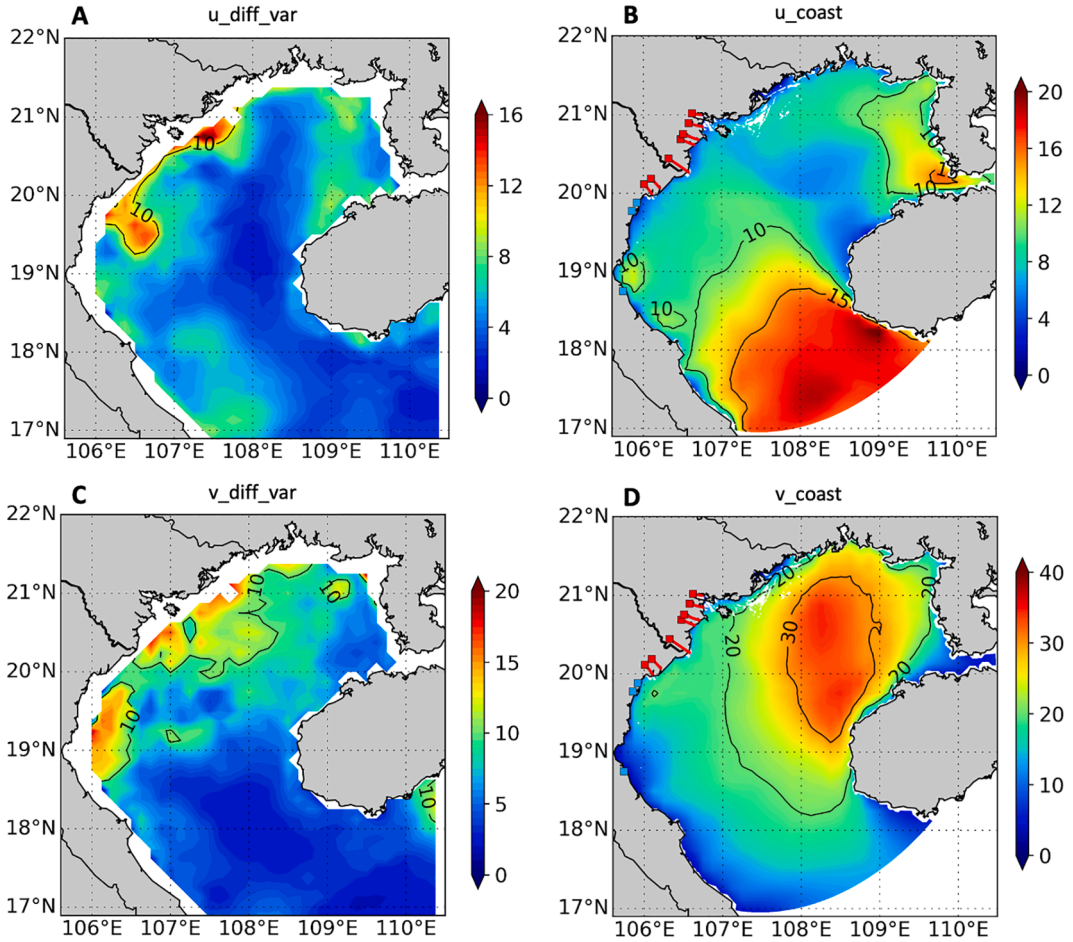

Fig. 4. **(A)** Variance of the difference between ASCAT and ECMWF wind, for U_w component. **(B)**: Ensemble variance of ENS_COAST, for Uw component. **(C–D)**: same as **(A–B)** but for V_w component. Units are m^2/s^2. Different color scales are used for U_w and V_w, and for the variance of the ASCAT-ECMWF difference and the ensemble variance.

to span the full range of the WIND_ISLAND (Fig. 5**A**). This is likely due to the fact that the wind data from Bach Long Vy is assimilated into the ECMWF forecasting system. Also, the variance budget analysis shows that this station is located in the area where the variance of the ensemble is consistent with the variance of the ECMWF-ASCAT misfit. ENS_COAST does not produce a much larger ensemble range, which is as expected since the amplitude of EOF 3 and 5 is small near Bach Long Vy. When considering the range between the first and the third quartile only, the numbers of days that the speed of WIND_ISLAND falls within this range are similar for ENS_REF (31 days) and ENS_COAST (33 days). Similarly, ASCAT follows the same trend as WIND_ISLAND. The mean of the absolute bias between the ASCAT and WIND_ISLAND is 1.87 m/s,

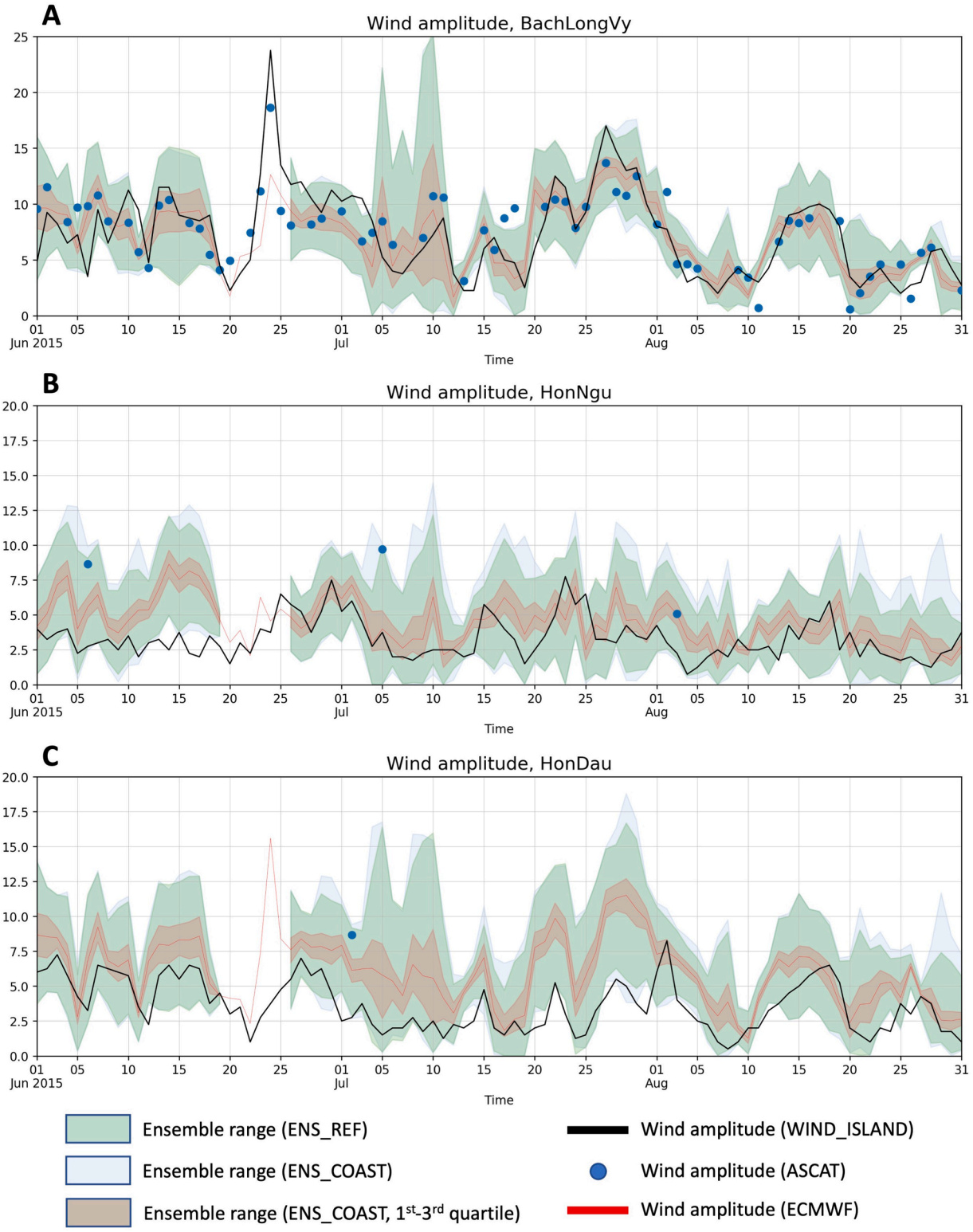

Fig. 5. The ensemble range of the daily mean wind amplitude from ENS_REF (shaded area), and the daily mean wind of WIND_ISLAND (black line) in Bach Long Vy **(A)**, Hon Ngu **(B)**, and Hon Dau **(C)**. Units are m/s.

equivalent to 26.6% of the ECMWF mean speed.

In Hon Ngu, the ECMWF wind seems to overestimate the wind speed of the WIND_ISLAND in June, but it is getting closer from July (Fig. 5**B**). The effect of the stronger perturbation near the coast is clearer, with the larger ensemble range shown in ENS_COAST. The number of days that the WIND_ISLAND stays within the ensemble range is 71 (for ENS_REF) and 77 (for ENS_COAST). Within the 1st–3rd quartile, the numbers of days are 18 and 20, for ENS_REF and ENS_COAST, respectively. The ASCAT data availability is limited here since the station location is adjacent to the coast. In all three data points, the wind speed from ASCAT overestimates WIND_ISLAND, however, it is still lower than the upper value of the ensemble range.

In Hon Dau, ECMWF and WIND_ISLAND wind velocities show similar variations, however, ECMWF data overestimate the observed one (Fig. 5**C**). Compared to other stations, the number of days when the WIND_ISLAND falls between the ensemble range is lower, reaching 63 (for ENS_REF) and 67 (for ENS_COAST), which correspond to 74% and 79% of the total days. The limited performance of the ensemble in this station can be explained by the underdispersion near the coast, as evidenced in the variance budget analysis. Similarly, due to the location of the station, most of the time, ASCAT data is not available in Hon Dau.

Overall, although both ensembles are underdispersive near the coast, they still show some skill at reproducing the wind uncertainty. Firstly, the ensemble can exhibit sharpness, meaning that the spread is small where the uncertainty is small and vice versa. Here, the range of the wind perturbations is smaller offshore and larger near the coast, which is consistent with the spatial error patterns estimated by ECMWF and ASCAT. Secondly, there is a bias between the two observational data sources (ASCAT and WIND_ISLAND), however, most of the time (at least 74%), the WIND_ISLAND amplitude stays within the ensemble range. Finally, ENS_COAST shows a better performance in the coastal area, which is consistent with the perturbations construction. Therefore, in the next sections, we will focus mostly on analyzing the results from ENS_COAST.

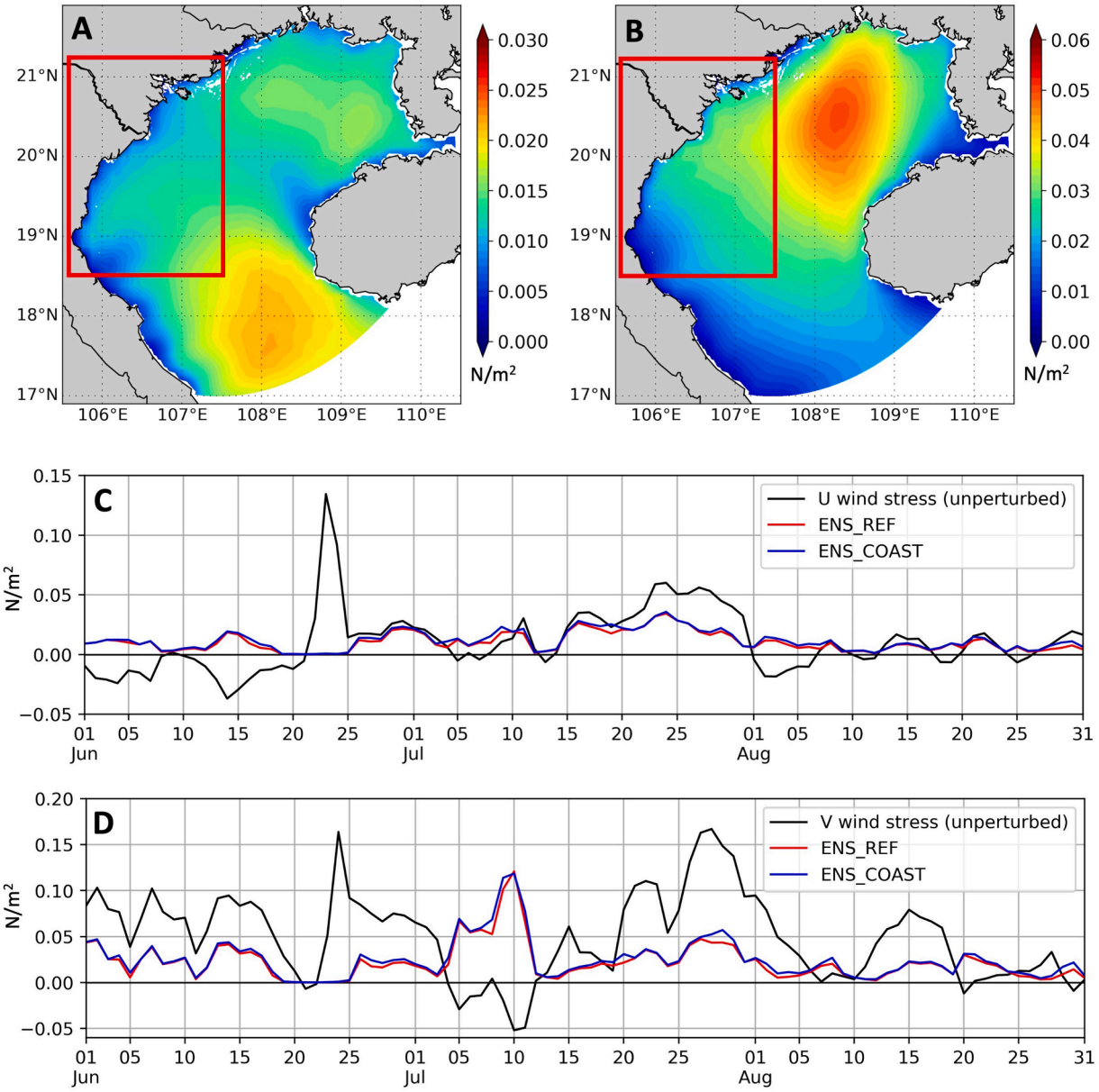

Fig. 6. Time-mean of ENS_COAST spread for wind stress, for the zonal **(A)** and for meridional (**B**) components. Spatial mean of the wind stress spread inside the red box, for the zonal **(C)** and meridional component **(D)**. Units are N/m^2.

3. Sensitivity analysis

In this section, we will present the main results of the paper. We will first examine the sensitivity of the wind stress, then of various ocean variables, and, finally, of different characteristics of the river plume.

3.1. Wind stress

Fig. 6**A** and **6B** show the mean ensemble spread of wind stress over the period, for the zonal (τ_x) and meridional (τ_y) components, respectively. The spread of τ_x is higher in the south of the gulf, while the spread of τ_y is higher in the north. This spatial structure is consistent with the variability of ECMWF wind velocity as depicted by EOFs 1 and 2, for τ_y and τ_x, respectively (Fig. 3). The maximum spread of τ_y is twice as large as the spread of τ_x, indicating that the spread affects not only the amplitude but the N-S direction as well.

Fig. 6**C** and 6**D** show the spatial mean over the red box of the unperturbed wind stress and of the ensemble spread as a function of time for the zonal and meridional components respectively. We first note that the time series for both τ_x and τ_y of the two ensembles only differ over short periods of a few days and, in these days, the spread of ENS_COAST is always larger than ENS_REF. Therefore, in the description below we will not differentiate the two ensembles.

Most of the time, the mean meridional wind component over the study domain is from the south. In the first three weeks, the zonal component is from the east; then it veers from the west. The wind stress is stronger in June-July than in August and corresponds to the summer monsoon. Disregarding the tropical cyclone occurring during June

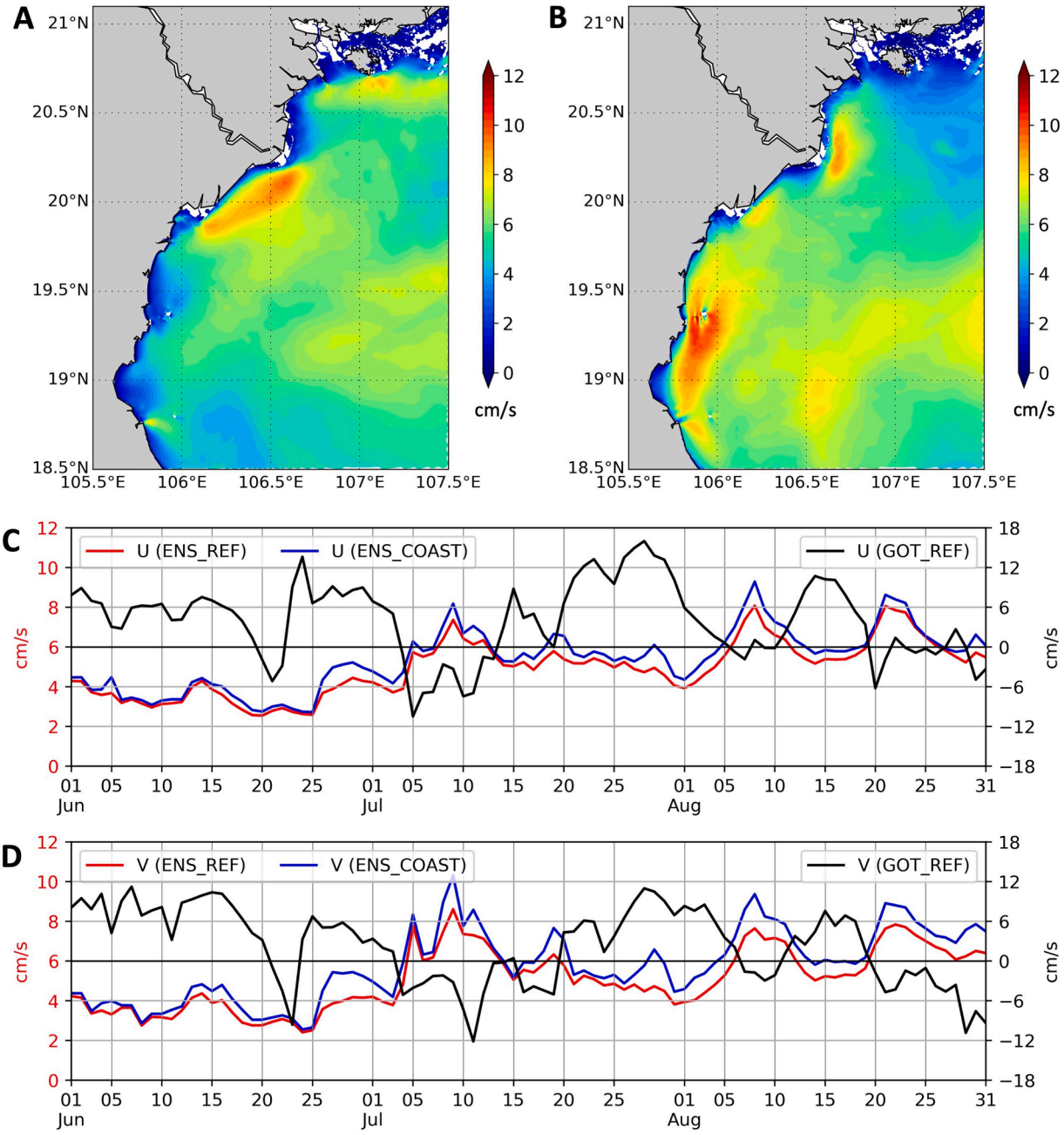

Fig. 7. (A): time-mean spread of the zonal surface current of ENS_COAST (cm/s). **(B)**: same as **(A)** but for the meridional component. **(C)**: Spatial mean of the ensemble spread of the zonal surface current, for ENS_REF and ENS_COAST (red and blue line, respectively, left axis). Spatial mean of the surface current in GOT_REF (unperturbed simulation, right axis). **(D)**: same as **(C)** but for the meridional component. Units are cm/s.

20–26, we can distinguish four main periods. In June, the southeasterly wind stress is moderate ($\tau_x < 0.04$ N/m^2 and $\tau_y < 0.1$ N/m^2); the mean spread does not exceed 0.03 N/m^2 for τ_x and 0.05 N/m^2 for τ_y. Then, from July 5–12, the wind stress weakens, but the meridional component reverses and the wind is mainly southward ($\tau_x < 0.03$ N/m^2 and $\tau_y < 0.06$ N/m^2). The ensemble spread of τ_y reaches about 0.12 N/m^2 which is by far the largest value of the ensemble spread for both τ_x and τ_y. In concrete terms, this means that some members have opposite meridional wind directions. From July 20 – Aug 1, we observe strong wind stress from the southwest, with the zonal and meridional components larger than 0.05 N/m^2 and 0.1 N/m^2 respectively with a peak of 0.17 N/m^2. The mean spread is large as well, reaching ~50% of the mean wind stress amplitude for τ_x, and ~30% for τ_y. Finally, from August 5–31, the wind stress is weak, mostly lower than 0.04 N/m^2 for both components, except around August 15, when the southwesterlies strengthen for a few days. The spread is relatively weak as well but can exceed the wind stress amplitude (for τ_y only).

In the rest of Section 3, we characterize the impact of the wind stress perturbations on the main ocean variables of interest for each of these four periods (Section 3.2), then the impact on the Red River plume properties (Section 3.3) and on the freshwater transport (Section 3.4).

3.2. Ocean variables

3.2.1. Surface current

Fig. 7**A** and 7**B** show the mean ensemble spread of the surface current over the period, for U and V components, respectively. The spread is about 6-8 cm/s on average over the area. The largest values (> 10 cm/s) are found for the coastal current, between 19.8–20.2°N for U and 18.8–20.5°N for V. As seen in Section 3.3, for a given day, the coastal current can have different directions for different members.

Fig. 7**C** and 7**D** show the temporal evolution of the spatial mean of the ensemble spread of the surface current over the area shown in Fig. 7**A, B** for ENS_REF and ENS_COAST. Both ensembles show a similar trend, as for the wind stress shown in Fig. 6. For ENS_COAST, in June, the spread of U and V components is similar, approximately 4 cm/s, while the mean current is around 6 cm/s. Then, from July 5, the spread increases to around 6 cm/s and can reach more than 8 cm/s in some events.

From July 05–10, the spread exceeds 8 cm/s for U and 10 cm/s for V. This peak occurs at the time when the wind stress amplitude is weak (Fig. 6**C, D**). The spread of the bottom current also shows a maximum (2–4 cm/s) showing that strong wind perturbations have a considerable impact on the wind-driven current both at the surface and subsurface.

Other high mean spread events occur from Aug 5–10 and Aug 20–25, but there is no clear signal of increased wind stress spread. The spread (higher than 8 cm/s for both components) is larger than the magnitude of the current (fluctuating around 0 for U and less than 5 cm/s for V) suggesting that the simulated result of the instantaneous current is not robust (at least in the low wind or changing wind conditions). Interestingly, in all 3 events, high spreads occur when the current direction changes. Indeed, during these transition periods, the current weakens and becomes more sensitive to wind perturbations.

Over the whole period, the spread of ENS_COAST is higher than ENS_REF, for both components. We, therefore, expect ENS_COAST to feature larger uncertainties on the plume.

3.2.2. Sea surface salinity

The time mean spread of the SSS is the largest along the vietnamese coast; further offshore it falls below 0.5 (Fig. 8**A**). At the river mouths, the spread is very low due to the homogeneity within the ensemble of the riverine water properties which are set by the runoff and the tidal mixing. Downstream, the spread increases and can reach 2 near the Yen river. The high spread in the plume can be due to two reasons. Firstly, wind perturbations affect the wind-driven current at the surface, and therefore the horizontal advection. Secondly, they have an impact on the vertical mixing. In that case, the uncertainties on the plume should also be significant at the subsurface. In the next section, the subsurface

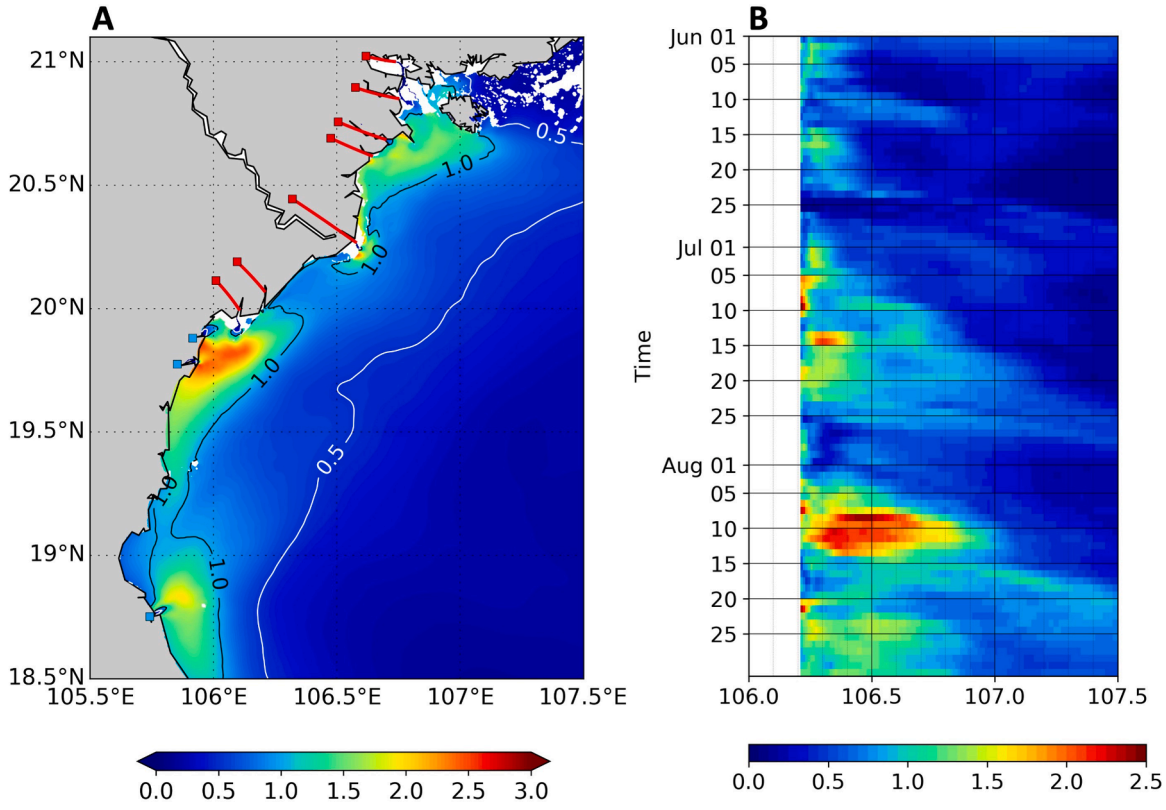

Fig. 8. (A) Temporal mean of the spread of the ENS_COAST SSS. **(B)**: Hovmöller diagram of the ENS_COAST spread of SSS at 20°N.

spread will be examined.

Fig. 8**B** shows that the spread is also highly variable in time. From June to August, it increases all along the coast with the largest values north and south of 20°N in August. At 20°N, interestingly the spread decreases close to the coast between July 25 and August 1, likely because the plumes from the different rivers are not connected anymore during this period. The maximum spread occurs around August 10, just after a peak in the spread of the coastal current.

3.2.3. *Subsurface temperature and salinity*

The vertical section at 19°N of the ensemble spread for temperature on August 20 shows a maximum of 1.9°C at around 25m depth (Fig. 9**A**). By comparing with the same section of temperature from GOT_REF, we find that the maximum spread occurs at the maximum gradient depth (base of the mixed layer) (Fig. 9**B**) consistently with the findings of past studies (e.g. Andreu-Burillo et al. (2002) and Ayoub et al. (2015)).

For salinity, the maximum spread occurs near the surface (Fig. 9**C**), where the salinity is low due to runoff. Similar to temperature, the maximum spread should occur where the gradient is the highest. For salinity, the horizontal gradient is large where the river plume is present (Fig. 9**D**), therefore the spread is highest at the surface. The vertical gradient of salinity is small except in the area strongly influenced by runoff. Uncertainties in temperature and in salinity have thus a distinct spatial distribution due to the different underlying driving processes.

3.3. *River plume*

In this section, we will examine the sensitivity of the river plume to the wind forcing uncertainty, including the surface plume area and the plume thickness. Here, the plume is identified as the area where the passive tracer concentration is equal or higher than 7 arbitrary units/m^3, as explained in Section 2.1.

3.3.1. *Plume surface area*

The runoff of the Red River and southern rivers, the plume area in the GOT_REF and in the ENS_COAST members are shown in Fig. 10**A**.

The river runoff is almost stable around 4000 m^3/s until July 25, apart from two small peaks in early July. It then increases, exceeds 8000 m^3/s between August 1-5 and then stabilizes around 6000 m^3/s from August 10 until the end of the study period.

In GOT_REF, the plume area fluctuates around 5000 km^2 until July 5, when it starts increasing, reaching ~12000 km^2 on July 15. It then decreases, down to a value below 5000 km^2 at the end of July, then increases again until August 22-25 when it reaches a maximum value of ~20000 km^2. These plume area variations are clearly linked to the runoff variations, with a lag in time of 10 to 15 days, but, as discussed in ND21, wind strongly impacts the plume area through dilution and horizontal transport. Interestingly, the members have a different behavior from August 13-15 (when the wind stress increases for a few days): in some members the plume area continues increasing, while in others, it decreases more or less and, in some cases, stays below its August 15 value (less than 10000 km^2) until the end of the run. In one member, the plume area reaches up to 25000 km^2 on August 25.

On average, the spread of the plume area of ENS_REF is around 1452 km^2 (not shown), which is equal to 18.6% of the plume area in GOT_REF. The corresponding values for ENS_COAST are 1607 km^2 and 20.4%, suggesting that the plume area is also sensitive to the coastal wind error. Most of the time, the spread is smaller than 2000 km^2 (Fig. 10**B**). The peak of the spread (~4000 km^2) occurs between 20 - 26 August; it then represents about 20% of the plume area. The spread reaches another peak (~2000 km^2) between July 20-28, corresponding to about 40% of the plume area. The spread of the plume area is therefore large in two situations: when the wind stress spread is large (July 20-28) as expected, and when the plume is the most extended (end of August).

3.3.2. *Sensitivity of the plume pattern with respect to the wind uncertainty*

In ND21, using daily fields of the tracer concentration from a 6-year simulation, the surface plume patterns have been classified into 4 clusters, as reminded in the introduction. In this section, we examine the

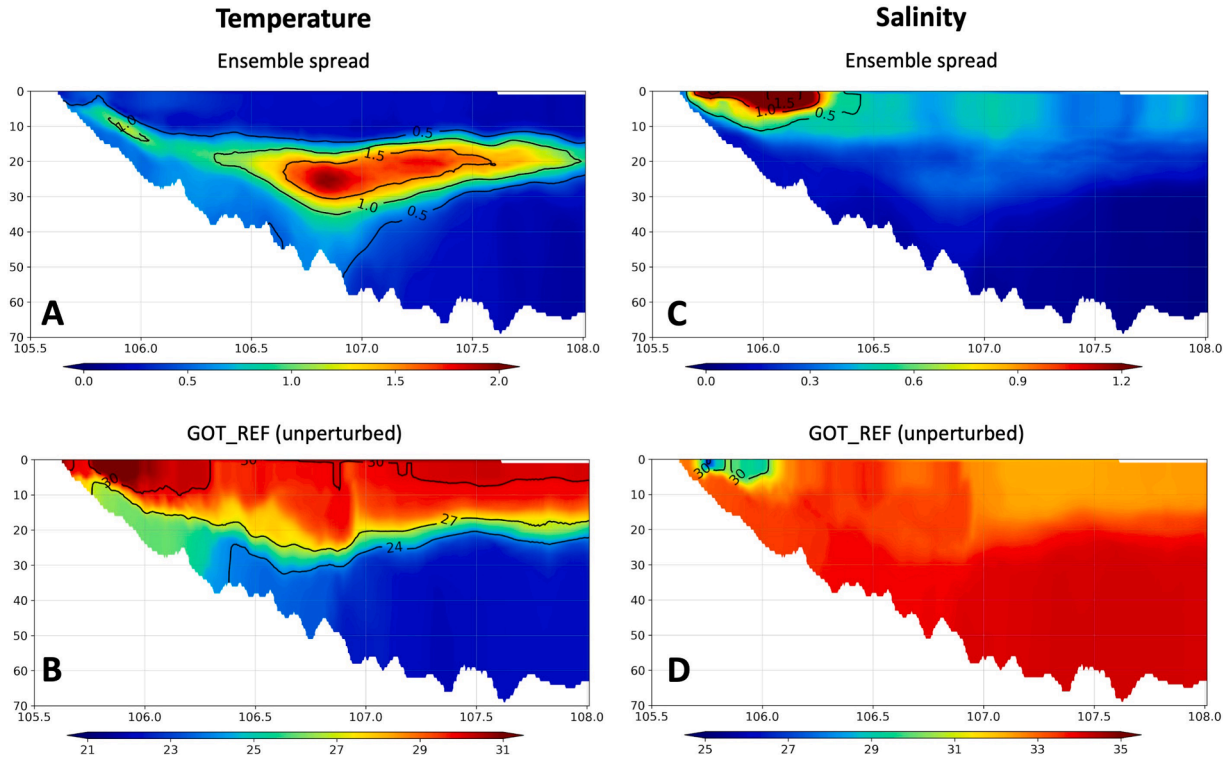

Fig. 9. (A): Ensemble spread of temperature at 19°N on August 20, 2015 (°C). **(B)**: Vertical section on the same day from GOT_REF (unperturbed simulation) (°C). **(C, D)**: same as **(A, B)** but for salinity.

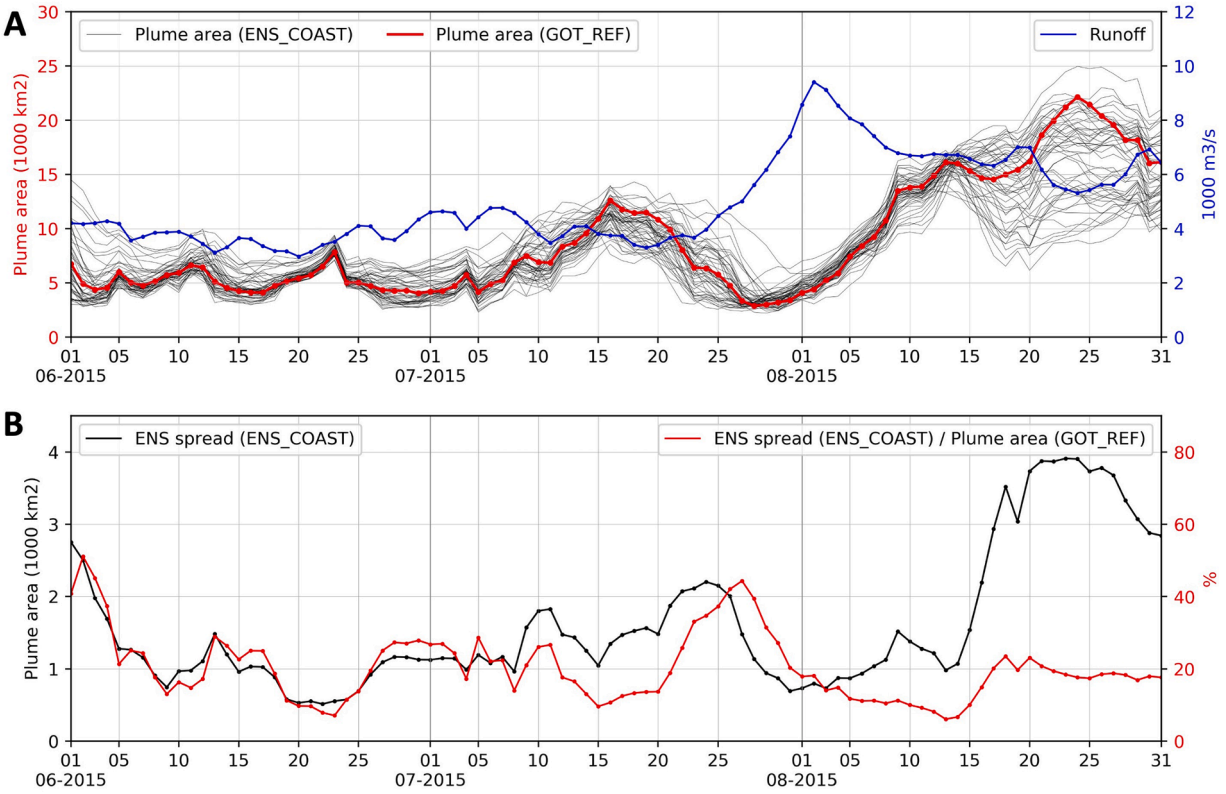

Fig. 10. **(A)**: Plume area (in km^2) from GOT_REF (red line, left axis). Plume area from all the members of ENS_COAST (black line, left axis). Daily runoff from the Red River and southern rivers (blue line, right axis, in m^3/s). **(B)**: Ensemble spread of the plume area of ENS_COAST in km^2 (black line, left axis). Ratio between the ensemble spread (of ENS_COAST) and the plume area from GOT_REF (red line, right axis).

robustness of this classification with respect to the wind perturbations. More precisely, we consider the clusters of ND21 trained over 2011-2016, hereafter called the reference clusters: by construction, each daily plume pattern is associated to a cluster. In this paper, we focus on the June-August 2015 period, and check whether the daily patterns in each ensemble member of ENS_COAST are associated to the same cluster as the reference run pattern.

To do that, for each member, the daily plume pattern is identified by the tracer concentration threshold (>=7 unit/m^3). Then, the distances between this daily pattern and the centroids of the 4 reference clusters are computed. Finally, the daily pattern is attributed to the cluster with the minimum distance.

We define a simple Maximum Likelihood estimate by choosing, for each date, the cluster that occurrs the most frequently across the ensemble (denoted as MAX estimate). Fig. 11**A** shows that the MAX clusters are the same as the reference clusters (REF on the figure), except on July 20 and 21 when the most likely cluster changes from 4 to 2 and 3, respectively. Therefore, we conclude that within our experimental conditions the findings of ND21 are yet unchallenged and robust with respect to wind perturbations.

In GOT_REF, the most frequent cluster over the study period is cluster 3 (54/92 days, Fig. 11**A**). It is also the most stable one, in the following sense: across the ensemble, we do not observe many changes from cluster 3 to other clusters; if it changes, it mostly changes to/from cluster 1 (5 July) or cluster 2 (5 August). The other two most frequent clusters in GOT_REF are cluster 2, which is mainly associated with low wind conditions, and cluster 4 which also corresponds to low wind but with a larger runoff.

Cluster 1 happens in just 2 days (06 and 07 July) in GOT_REF. In ND21, it is described as the pattern that occurs in the lowest runoff period and winter monsoon, so it is not expected to occur much in the rainy season in July. When submitted to the wind perturbations, the classification fluctuates between clusters 1 and 2 suggesting that the plume pattern on 06 and 07 July in GOT_REF is on the edge of cluster 1 and cluster 2. In that case, the attribution to cluster 1 in GOT_REF may not be as robust as the one to cluster 3.

The plume pattern appears to be the least robust to the wind perturbations over 3 periods: 5-22 July, 5-12 and 21-28 August, when clusters 2 and 4 occur. During these periods, at least 3 different clusters occur in different members on the same day (Fig. 11**B**). For example, on August 21, the reference cluster is 2; it only happens in 21 members, and 17 other times it is classified as cluster 4. This suggests that, these days, the plume patterns vary significantly, and using the MAX estimate may not be sufficient to conclude about the sensitivity of the plume pattern.

In order to have a more in-depth view about how much the cluster attribution changes across the ensemble and quantifying the uncertainty of the classification, we calculate the ratio D = D1/D2 * 100%, while D1 is the distance from the plume to the nearest cluster, and D2 is the distance from the plume to the second-nearest cluster (Fig. 11**C**). If D is small, the plume pattern robustly belongs to the specific cluster; in contrast, a larger value of D indicates that it stays near the edge of the two clusters. It is partly similar to the membership coefficient presented in Delaval et al. (2021), which is used to describe how closely the data is located to the cluster center, taking into account the pre-defined overlap of the cluster (fuzziness coefficient) in the Fuzzy C-Means (FCM) clustering algorithm. In contrast, our classification approach does not allow overlapping between different clusters. We first find that in GOT_REF, D can exceed 70% meaning that the classification is then less robust. For instance cluster 3 on June 21 or on July 27 is less robust than on August 1st. The periods of high value of D coincide with the periods of highly variable clusters across the ensemble (Fig. 11**A, B)** (5-22 July, 5-12 and 21-28 August), hinting at the fact that another possible measure of classification uncertainty (not tested) could be the spread of cluster attribution.

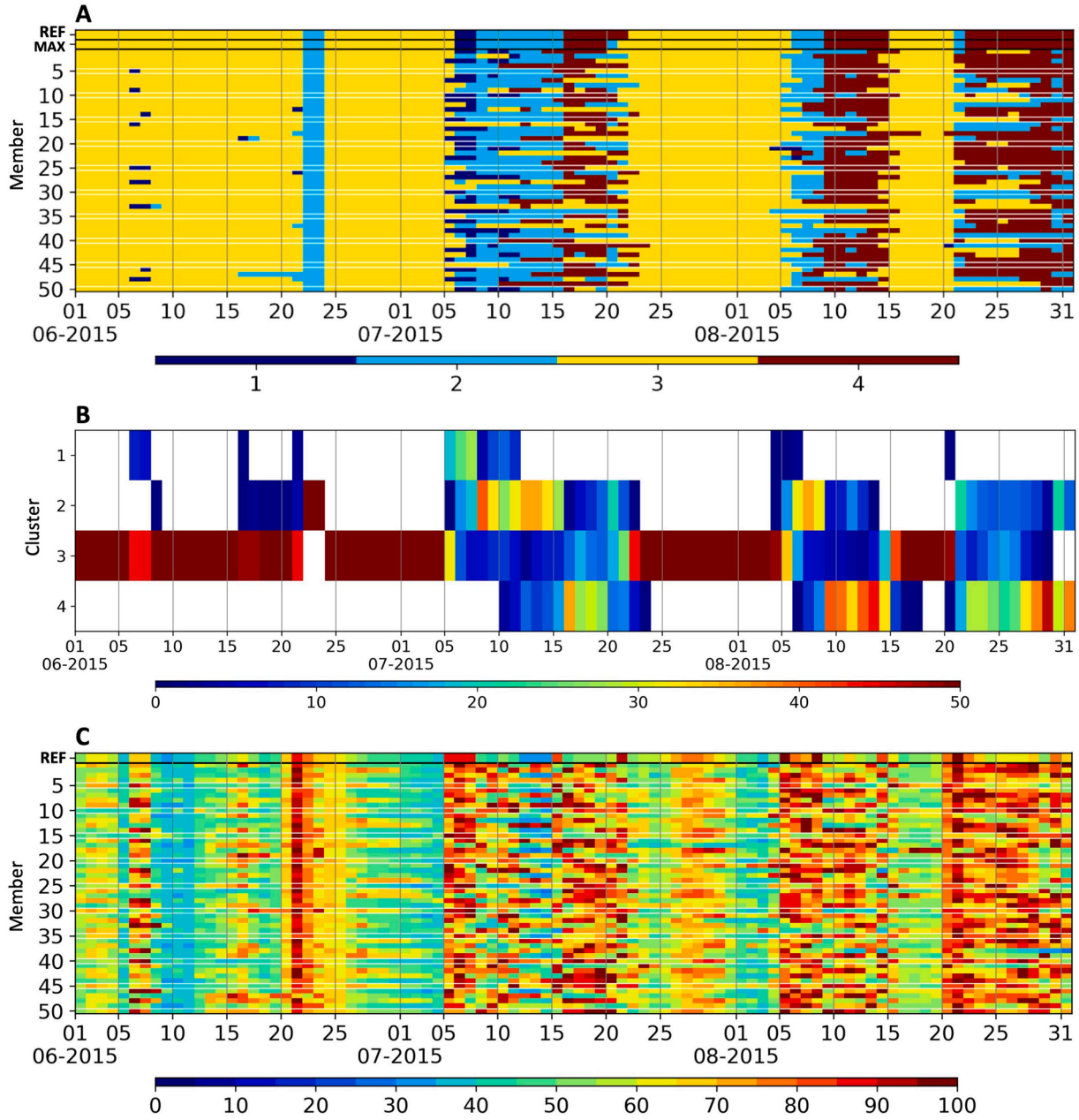

Fig. 11. **(A)**: Classification of the daily plume patterns in June-July 2015 across the ENS_COAST ensemble, using the reference clusters from Nguyen-Duy et al. (2021). **REF** shows the same classification for the (unperturbed) GOT_REF run. **MAX** is the cluster with the MAXimum likelihood, defined as the most frequent cluster drawn within the 50 samples shown underneath. **(B)**: The number of members in each cluster over the period. **(C)**: The ratio D = D1/D2 * 100%, where D1 is the distance from the plume pattern to the nearest cluster centroid, and D2 is the distance to the second-nearest cluster centroid.

We now illustrate through an example a change of attribution, examining two members of the same pair. In member 33 (of ENS_COAST), the plume is classified as cluster 3 most of the time unlike GOT_REF. Fig. 12 shows the mean surface tracer concentration, current and wind velocity over 5-10 July. The plume is advected northward; the plume from the Lam river is disjoint from the other rivers plumes, which is indeed a characteristic of reference cluster 3 (Fig. 12**A**). Close to the Red River delta, the wind is stronger and favors a northward advection, consistently with the surface tracer distribution. In member 34, the plume extends in a narrow band along the coast and is classified as cluster 1 (Fig. 12**D**). The strong northeasterly and westerly winds in the north and south of the GOT respectively trigger the southward coastal current, which explains why the tracer is trapped near the coast.

Fig. 11 evidences the impact of wind perturbations on the shape of the plume. It might be convenient for applications to further quantify the uncertainties on the plume location at given dates. This can be achieved by mapping the probability of occurrences of the plume, as done for three specific events on Fig. 13.

From 8-12 July, the spread of meridional wind stress is the highest (Fig. 6**D**) and there is also an increase of the plume area spread (Fig. 10**B**). On July 10, the plume is advected southward, with half of the members extending to around 25 km from the coast. This pattern is consistent with the fact that cluster 2 is the most likely in that period (Fig. 11**A**), but that all 4 different clusters are found in the ensemble on this day (Fig. 11**B**).

From 25 - 30 July, both the wind stress of GOT_REF and the spread of wind stress reach high values (Fig. 6**C, D**). It is also the time when the ratio between the spread of the plume area and the unperturbed plume area is the largest over July and August (Fig. 10**B**). On 25 July, along the coast from 20 – 21°N, the plume seems to be stable across the ensemble,

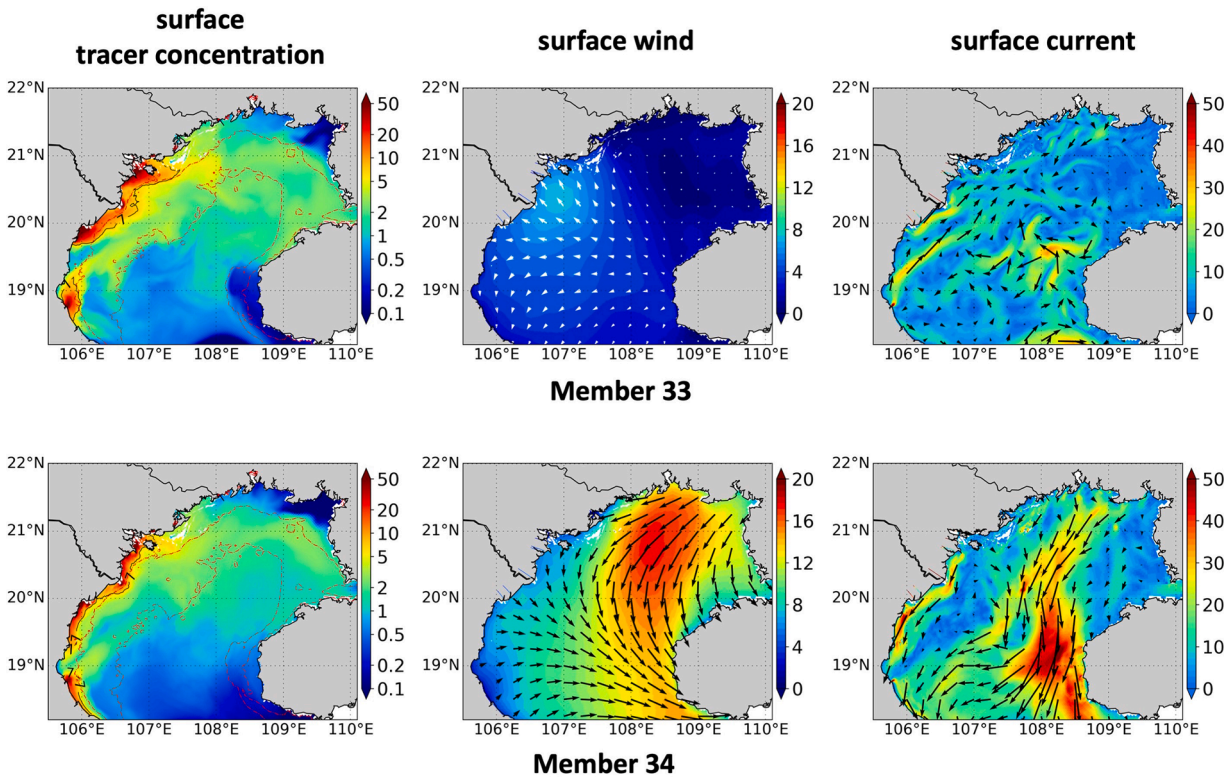

Fig. 12. **(A)**: mean surface concentration (unit.m^{-3}) over 5 - 10 July for member 33. The dashed lines show the isobaths of 20m and 40m. **(B)**: same as **(A)** but for surface wind velocity (m/s). **(C):** same as **(A)** but for surface current (cm/s). **(D–F)**: same as **(A–C)** but for member 34.

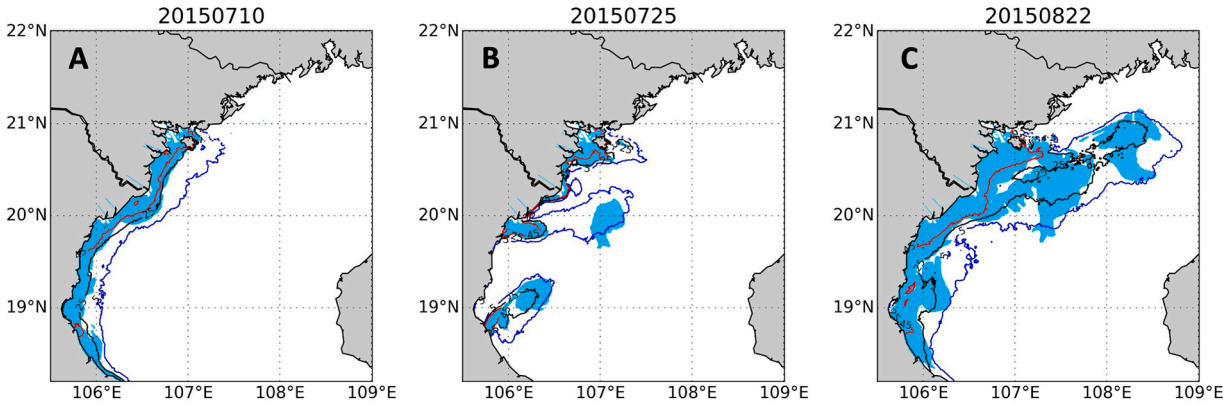

Fig. 13. **(A)**: Quantiles of plume occurrence on 10 July 2015. The contour lines show the area where the plume occurs in 5 (blue), 25 (black) and 45 (red) times over 50 members of ENS_REF. Cyan color shows the plume area from GOT_REF. **(B)**: Same as **(A)** but for 25 July. **(C)**: Same as **(A)** but for 22 August.

as shown by the small difference between the 5-member and 45-member plume occurrence (Fig. 13**B**). It is consistent with Fig. 11**A** where all members belong to cluster 3. In contrast, along 20°N, the spread is large: for half of the members, the plume extends to 106.3°E, while 5 members reach ~107.4°E, which is more than 100km from the coast. A large spread is also observed on the Lam river plume near 19°N. As discussed previously, cluster 3 is found to encompass plumes with similar shapes but with different areas. Note that the plume in the unperturbed run corresponds to the first decile of occurrence, with a large extension offshore.

On August 22, the plume in GOT_REF extends both offshore and along the coast and is associated to cluster 4. From the analysis of the surface current variability over the preceding days (not shown), we attribute this pattern to the succession of two types of wind conditions: (1) summer monsoon winds conditions and large runoff ($>$ 6000 m^3/s) over August 15-20 (Figs. 6D and 10A) leading to an offshore extension of the plume over the shelf; the plume's pattern is associated to cluster 3. (2) the wind stress intensity decreases after August 20 and the plume flows downstream along the coast. On August 22, the plume offshore represents a thick layer (thickness of 8-12 m) of freshwater which is older than the newer plume along the coast whose thickness does not exceed 5 m. The plume occurrence shows strong variations (Fig. 13**C**). Along around 20.6N, 5 members can extend to 108.8E, while 45 members extend to 107.3E only. A large spread is also observed near 19.5°N. The plume pattern classification shows that on this day, the plume mostly belongs to cluster 4, but it can also be attributed to cluster 2 and cluster 3. It is the time when the wind stress and its spread are relatively low (Fig. 6), however, the runoff is high and the plumes span a large area in all members (Fig. 10) and the mean spread of the surface current reaches a local maximum on August 20-23 (Fig. 7**C, D**). We suggest that

the sensitivity to the wind is itself sensitive to the runoff, and the variation of the plume occurrence is strengthened by the high spread of the current.

3.3.3. Plume thickness

The plume thickness at 19°N in GOT_REF is shown in Fig. 14**A** as a function of time. It is computed as the depth where the concentration is equal to 7 units/m^3. In June, the plume is about 2 m deep and detached from the coast. Then, the thickness increases and the plume extends further offshore, reaching a maximum thickness of around 10m on 20-25 July. In August, the plume thickness varies from 2 to 7m.

The number of members in ENS_COAST where the plume is present is shown in Fig. 14**B** and the spread of the plume thickness, is illustrated in Fig. 14**C**. To better estimate the spread of the plume thickness, we look at the points where the plume is present in at least 30 members (dotted red in Fig. 14**C**). In June, the spread is less than 1m. The peak of the spread occurs on 09-13 July (> 3m), which is also when the wind stress spread is the largest, suggesting that the change of the wind condition also has a considerable impact on mixing. Other high values of the spread occur also in the period of 5-12 August and 21-28 August at times of large variations of the plume pattern (Fig. 10), and of peaks in the surface current spread (Fig. 7).

In the period of 20-25 July, the plume in GOT_REF is the thickest, increasing from 3m near the coast to 10m offshore. The deepening of the plume is associated with an eddy (Fig. 15**A**), such as the one described in ND21. The map of relative surface vorticity on 24 July 2015 shows the signature of an anticyclonic eddy; part of the plume is drawn into the eddy where it undergoes some mixing and deepens. In this period, the spread of the plume thickness at 19°N is around 1-2m (Fig. 14**C**). Our first idea is that the processes causing or involved in the eddy have a critical impact on the mixing inside the eddy, so the direct effect of the wind perturbations on the mixing there is relatively weak. However, the surface vorticity computed from each member of the ENS_REF shows that the wind perturbation can alter the current, change the shape of the eddy and finally modify the plume thickness. Fig. 15**B–D** show the relative surface vorticity and plume thickness on the same day but from 3 members of ENS_COAST. Under different wind conditions, the plume can be either thicker (Fig. 15**B**), thinner (Fig. 15**C**), or its thickness is smaller than 8m (Fig. 15**D**). Another interesting finding is that the eddy in GOT_REF is located in the area where the plume is present in less than 20 members of ENS_REF, which means that the eddy's location and intensity cannot be considered as robust.

Now, we examine the impact of the wind perturbation on the mixing for a specific event. Fig. 16**A–C** shows the surface tracer concentration from GOT_REF, from the members with the largest plume area (member 41) and with the smallest plume area (member 42) on 25 July. The daily wind conditions corresponding to each simulation are shown in Fig. 16D–F. The wind patterns are similar in the three simulations: the winds blow from the southeast in the south and from the southwest in the north of the gulf, which favors northward and offshore transport of the tracer. However, the wind intensity differs between the simulations: it is the weakest in member 41 and the largest in member 42 on average over the whole domain.

In all 3 simulations, the passive tracer tends to extend offshore consistently with the wind direction, but with very different concentrations at surface (Fig. 16A–C) and subsurface (Fig. 16G–I). Between 18.5N and 19.5N, the stronger the wind, the more the influence of river waters extends offshore (with concentration below the threshold of 7 units/m^3 though). Between the isobaths 20m and 40m, the surface concentration is weak (< 2 units/m^3) under strong wind (member 42), while it locally exceeds 7 units/m^3 in GOT_REF and member 41. North of 21°N, the 3 simulations show little differences. The wind perturbations impact (1) the horizontal advection by wind-induced currents and (2) the vertical mixing by wind-induced vertical shear of the current; the distribution of the tracer concentration results from these two competing processes. In member 42, the surface currents south of 20°N

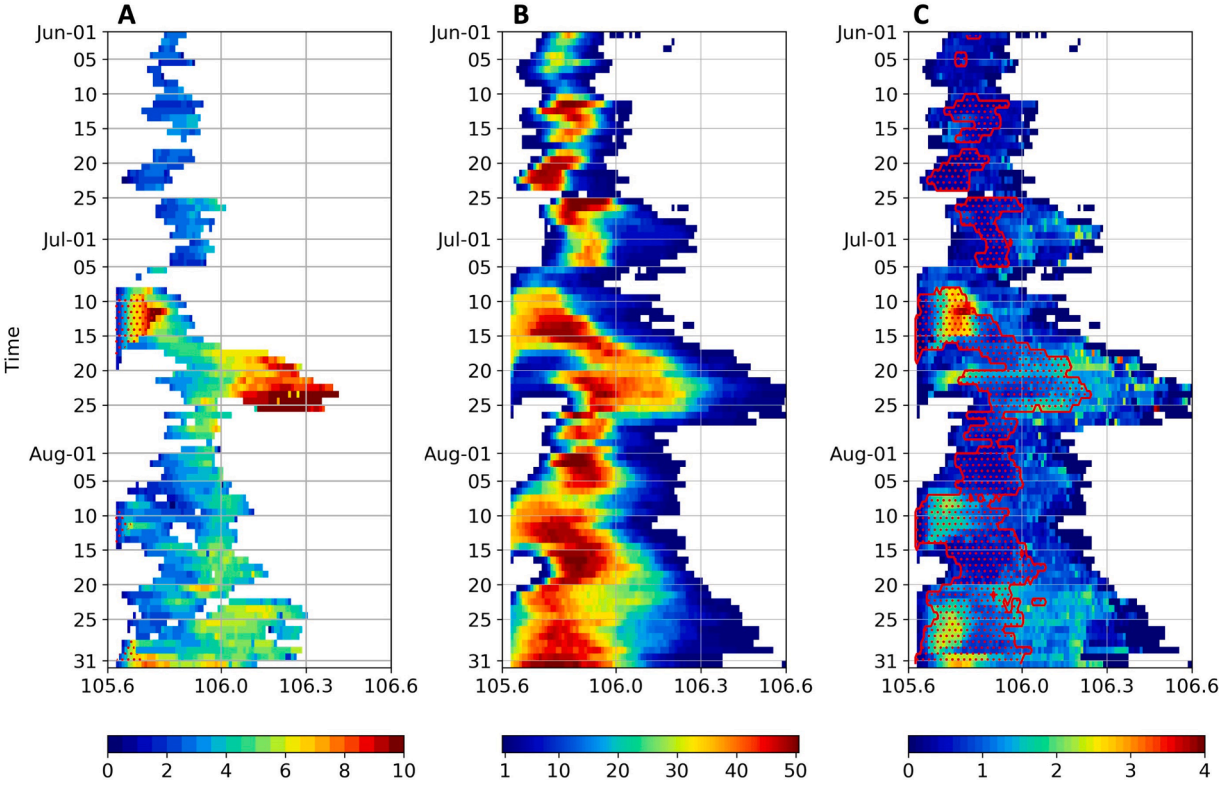

Fig. 14. (A): Plume thickness in GOT_REF (in m) at 19°N as a function of time and longitude. Red dots represent the area where the plume reaches the bottom. **(B)**: the number of members for which the plume is present and is used to compute the spread. **(C)**: spread of plume thickness (m). Red dots indicate the area where the plume is present in at least 30 members of the ensemble. **(B)** and **(C)** are computed from ENS_COAST.

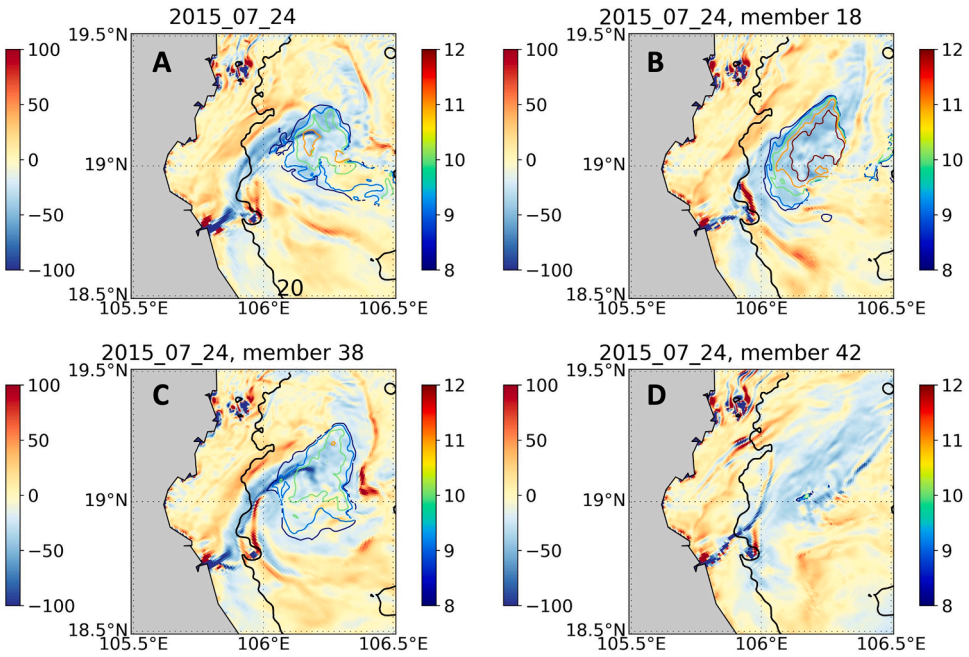

Fig. 15. (A): Relative surface vorticity (10^{-6}/s) (contour fill, left color bar) and plume thickness (m) (line, right color bar) on 24 July 2015 from GOT_REF. **(B–D)**: same as (A) but from members 18, 38, and 42 of ENS_COAST. The black solid line indicates the 20m isobath.

are indeed stronger than in member 41, not only on July 25 but also on average over the 4 previous days (not shown). They are probably responsible for the export further offshore between 18.5°N and 19.5°N (discussed previously). But they mostly lead to a more intense vertical mixing hence a dilution in surface and a deeper penetration of the tracer in the center of the basin, as depicted at 20°N (Fig. 16G–I).

To assess the mixing between different ensemble members, the histogram of the vertical diffusivity across the ensemble at a point (20°N, 107°E) at different depths is shown in Fig. 16**J–M**. At 4.4m, the diffusivity is distributed differently between members, mostly ranging between 0.01 and 0.025 m^2/s. Then, it decreases to less than 0.001 m^2/s in more than 10 members at 7.9 m, and more than 30 members at 13.0m. At 19.9m, all members show diffusivity smaller than 0.001 m^2/s, and 48 of them show diffusivity smaller than 0.0001 m^2/s, suggesting that the impact of the perturbation on mixing is limited at deeper depth.

However, it should be noted that the ocean state results from the integrated effect of wind forcing over time, which implies examining the wind conditions over at least several days preceding the event. Besides, the effect of the wind induced mixing on the ambient shelf water may also influence the path of the plume, as suggested by Toublanc et al. (2023) in their study of the Gironde River plume in the Bay of Biscay (France). From model sensitivity experiments, these authors suggest that the density gradient between the ambient ocean and estuarine waters conditions the dynamics of the plume, hence the development of a bulge in the case of the Gironde River. Understanding all the processes involved in the event of Fig. 16 would clearly require a very in-depth analysis, which is beyond the scope of this study.

3.4. Freshwater transport

The sensitivity of the plume (area, pattern, thickness) to the wind perturbations results from the combined effect of the sensitivity of the dilution (vertical mixing mainly) and of the horizontal transport (horizontal advection). As reported in the review papers of Horner-Devine et al. (2015) and Hetland and Hsu, (2013), both processes interplay. The horizontal transport is not only wind-driven but also buoyancy driven, therefore, it also depends on vertical mixing. Similarly, when the plume spreads, shear-induced vertical mixing is enhanced while the spatial extent on which mixing is active increases; the net mixing therefore increases. Wind induced mixing is produced by shear in the Ekman surface layer. As noted by Hetland and Hsu, (2013), upwelling and downwelling winds do not have equivalent impacts on mixing: upwelling winds which lead to more extended plumes are a priori more efficient in mixing the freshwater. This mechanism could explain why the spread of the plume properties (area, pattern, thickness) is large from August 15 whereas the spread of the wind stress is moderate: as the plume is well spread (due to a large runoff and low wind) and surface trapped, it is more sensitive to perturbations on wind-induced mixing.

The last property of the plume that we examine in this paper is the freshwater transport by the coastal current through a section south of the delta, where the water depth is less than 30 m (Fig. 1**B**).

The freshwater transport through a vertical section, for instance along the y-direction, is usually computed as follows (e.g. Schiller et al., 2011):

$$Q = \int \int_{-h}^{\eta} fw_f u dy dz \tag{3}$$

where fw_f is the freshwater fraction defined by: $fw_f = (S\text{-}S_b)/S_b$ with S the local salinity and S_b the background salinity (representative of the open ocean). u is the horizontal velocity in the direction perpendicular to the section. η and -h are the sea surface elevation and bottom floor respectively. We have found that this calculation is highly sensitive to the choice of S_b. Besides since S is also influenced by precipitation, we find more convenient and reliable to estimate the freshwater transport from the tracer concentration, using the definition below:

$$Q = \int \int_{-h}^{\eta} \frac{c}{100} u dy dz \tag{4}$$

where c is the local concentration of the tracer.

The freshwater transports from GOT_REF and ENS_COAST through the section are shown in Fig. 17**A**. From June to August, the freshwater transport is changing repeatedly between southward and northward transports. As expected, it is northward when the wind is mostly northward (end of June, end of July and around August 15, corresponding to cluster 3 periods), and reverses when the summer monsoon relaxes (clusters 2 and 4). The spread of the transport is 1502 m^3/s on

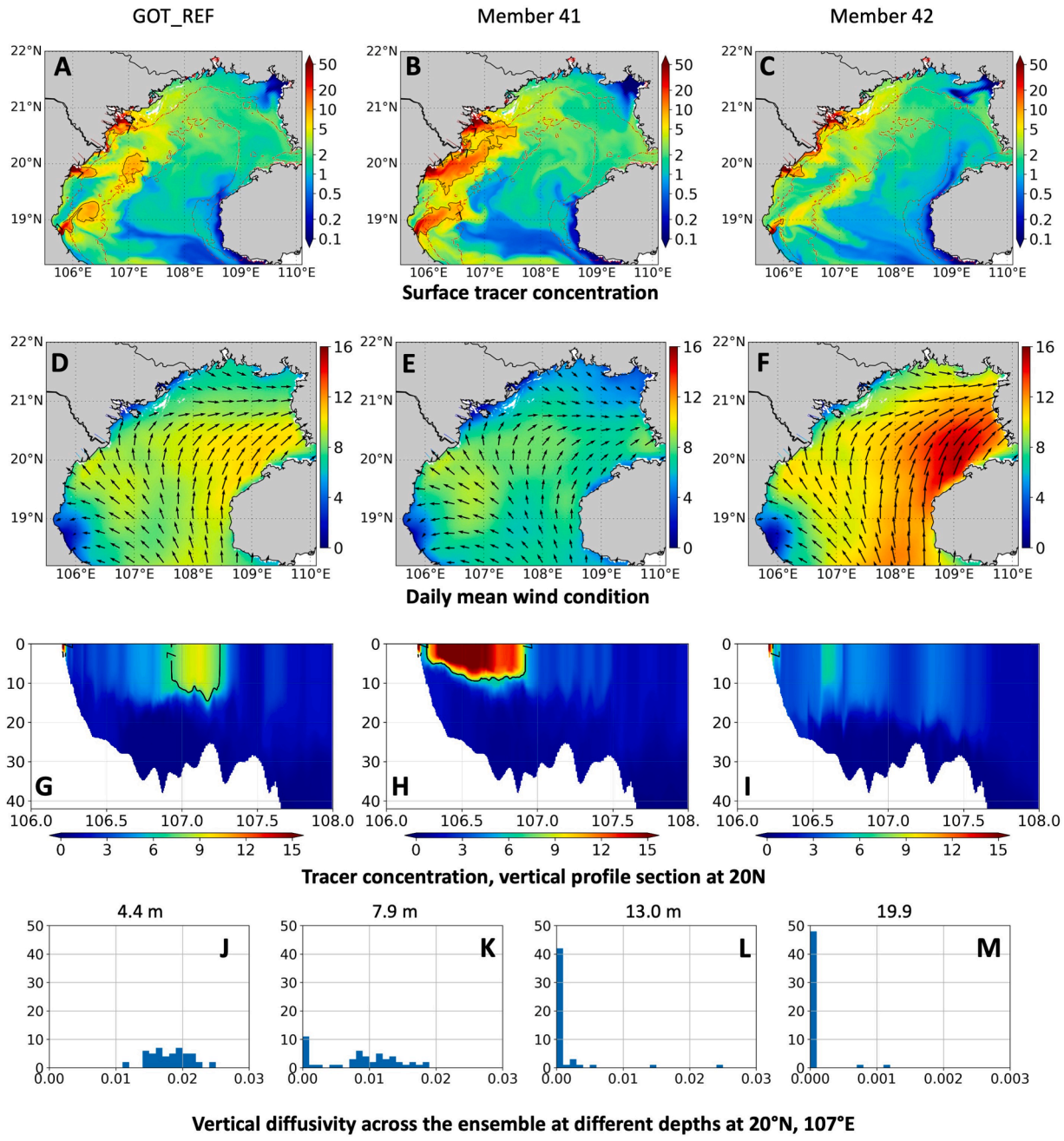

Fig. 16. Surface tracer concentration of GOT_REF **(A)**, the member with the largest plume area **(B)**, the member with the smallest surface plume area **(C)** on 25 July (arbitrary unit/m^3). The dashed lines show the isobaths of 20m and 40m. **(D–F)**: Daily mean wind conditions (m/s). **(G–I)**: Vertical profiles of tracer concentration at 20°N (arbitrary unit/m^3). **(J–M)**: Histogram of vertical diffusivity (m^2/s) across the ensemble at a point (20°N, 107°E) at 4.4 m, 7.9 m, 13.0m, 19.9m, respectively. The width of each bin is 0.001 m^2/s for **J, K, L** and 0.0001 for **M**.

average; it is minimal in June, and then increases from July. The maximum of the spread happens around 9 July, 18 July, 8 August and late August. It is highly correlated to the spread of the meridional current velocity (Fig. 7**D**).

To assess the contribution of the different factors to the spread of the transport, we decompose the freshwater transport into different terms (for each member):

$$Q = \int\int c * u \; dydz = \int\int (c_{mean} + c') * (u_{mean} + u') dy \; dz \tag{5}$$

$$Q = \int\int c_{mean} * u_{mean} dydz + \int\int c_{mean} * u' dydz + \int\int c' * u_{mean} dydz + \int\int c' * u' dydz = Q1 + Q2 + Q3 + Q4 \tag{6}$$

With

c_{mean}: ensemble mean of the tracer concentration (divided by 100)
c': difference between the concentration from the member and from the ensemble mean (divided by 100)
u_{mean}: ensemble mean of the velocity

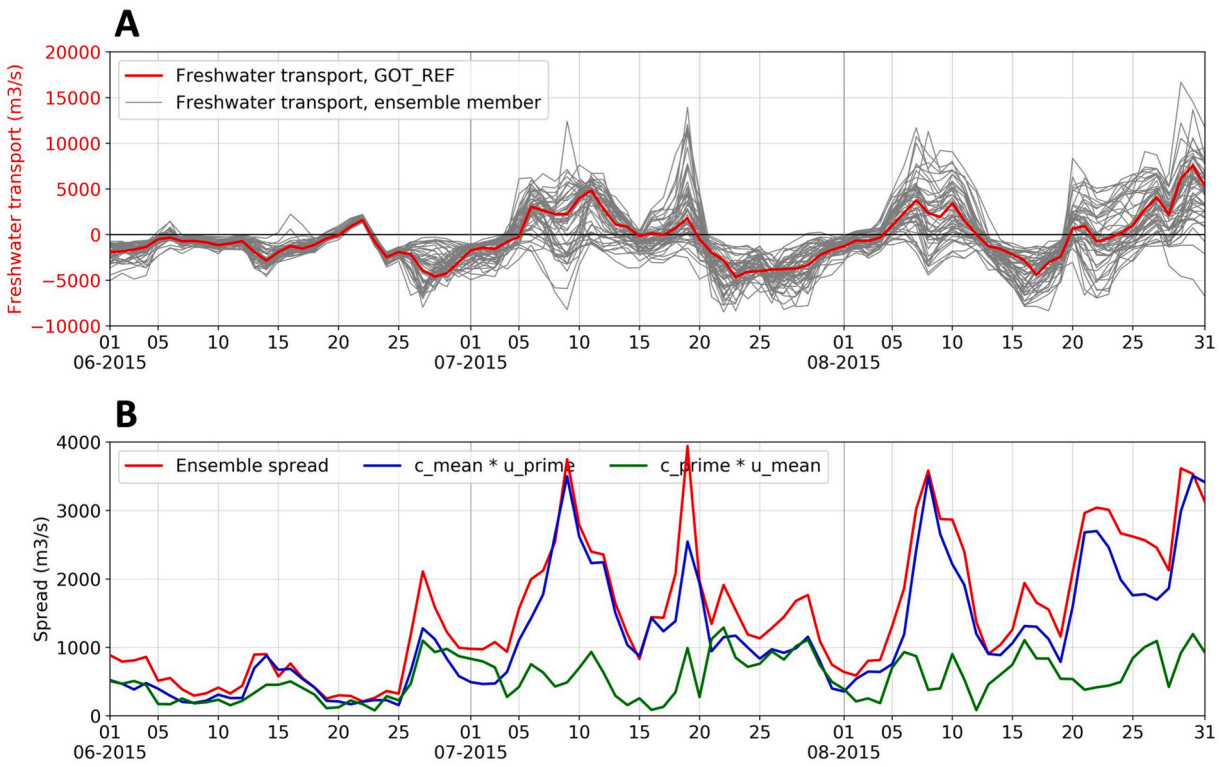

Fig. 17. (A): Freshwater transport of GOT_REF (red) and ENS_COAST (black) at a section (see Fig. 1**B**), positive southward (in m^3/s). **(B)**: Ensemble spread of the freshwater transport (Q, red), and of component 2 (Q2, blue) and component 3 (Q3, green) of the freshwater transport (in m^3/s), from ENS_COAST.

u': difference between the velocity from the member and from the ensemble mean

The freshwater transport is controlled by the tracer concentration and by the current. To assess the relative contributions of the tracer and of the current sensitivity to the total transport spread, we analyze Q2 and Q3 corresponding respectively to the flux of an ensemble mean tracer concentration by the anomalous current of each member with respect to the ensemble mean and to the flux of the anomalous concentration by the ensemble mean current. The time series of the ensemble spread of Q2 and Q3 at Section 1 are shown in Fig. 17**B**. First, the increase of the total freshwater transport spread at the end of June is due to the increase of spread of both current and concentration. Then, the spread of Q2 grows larger than the spread of Q3: at the peak periods, Q2 reaches three times the value of Q3. We can deduce that the freshwater transport across this particular section is more sensitive to uncertainties on the current than on the tracer concentration. On average, the mean spreads of Q2 and Q3 for ENS_REF are 996 m^3/s and 497 m^3/s, respectively. The corresponding values for ENS_COAST are 1187 m^3/s (+19%) and 557 m^3/s (+12%). The differences between ENS_REF and ENS_COAST indicate that the transport is also sensitive to the coastal wind errors.

4. Discussion on the use of the ensembles for data assimilation

Ensemble statistics are used as proxies of model errors in ensemble-based data assimilation methods, such as the Ensemble Kalman Filter (e.g. Evensen, 2003; Sakov and Oke, 2008). In this section we propose a preliminary study of the possible use of ENS_COAST in the case HF radar observations were assimilated into the model to constrain the plume. We first evaluate the consistency between the ensemble and the misfit between the observations to be assimilated and the model, within the observations uncertainties. Then we analyze the influence function of a single observations onto several model variables.

We use surface velocity data from the high–frequency radar (hereafter HFR) system based on two antennas located at 18.62°N and 17.47°N (R1 and R2 site, respectively, see Fig. 1B) along the coast and operated by the Center for Oceanography, Vietnam Administration of Sea and Islands (CFO, VASI). The data consists of daily maps of zonal and meridional components of the surface current, over the year 2015, built and provided by Tran et al. (2021). HFR data are representative of currents at 2.4 m below the surface (Tran et al., 2021). To compare observed and simulated currents, we first estimate the latter at 2.4 m depth. and then interpolate them onto the HFR grid (yellow area of Fig. 1B).

4.1. Empirical ensemble assessment from high-frequency radar data

We use rank histograms (Talagrand et al., 1997) to evaluate if the ensemble provides reliable probabilities of the observed quantity (Hamill, 2000); in other words we assess the compatibility between the distribution of the ensemble and the distribution of the observational data, as described in Vervatis et al. (2021b).

The rank histograms are computed for both components of the currents taking into account all the model and data values over the HFR grid. The computation is repeated every day, leading to the Hovmöller diagrams of Fig. 18.

In general, the ensemble is found to be underdispersive. In addition, the relative behavior of the observed and simulated currents show nonstationarity and biases. For the U component, the rank histogram is U-shaped in June, but it shows biases in mid July and in August, with the simulated current being too weak and too strong respectively with respect to the observations. For the V component, a bias is found in the first half of June, then the histograms have a U shape most of the time. This could mean that increasing the perturbation may improve the performance of the ensemble at describing the observed variance in statistical terms. However, as noted by Hamill, (2000), the U-shape of the histogram may be due to errors in the observations that are not taken

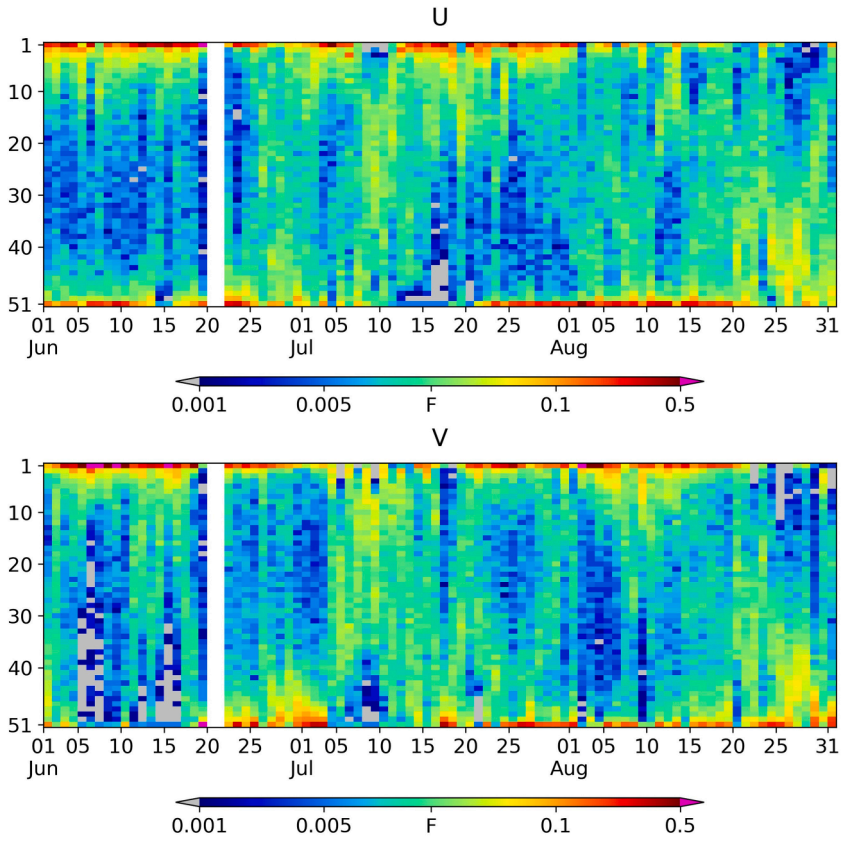

Fig. 18. Rank histograms of ENS_COAST near surface current versus HFR data over the HFR data grid, using 51 ranks (Rank 1 and 51 are open rank). F defines the value in which the histogram is flat (i.e. the rank is equi-distributed, F = 1/51 ~ 0.0196). The y-axis shows the rank. Gray (pink) color indicates pdf with lower (higher) values than the color scale. There is no data on 20 and 21 June.

into account. The histogram seems flatter and even overdispersive in July 5 - 15, which is the same time as the large spread of wind stress and surface current.

Another way to assess the relevance of the ensemble in simulating the model errors is to evaluate the skill of the ensemble variance at reproducing model errors. This method shares the same formalism as the method in Section 2.3 (see Appendix). We first compute the ensemble variance (ENSVAR) for the zonal and meridional components of the current over the HF radar domain and over time. We also compute the difference of the current between the ensemble mean and the observed one (for each component) at each radar grid point (hereafter innovation); we then compute the variance of these differences (VARDIFF) over the radar domain and over time. Assuming that the model error and observational error are uncorrelated, and non-biased (i.e. their expectancy over the period of study is null), VARDIFF should be close to the sum of ENSVAR and the observational error variance.

Table 2 shows that for both components, there is a large difference between VARDIFF and the sum of ENSVAR and the observational error variance. This comparison therefore reveals an inconsistency between the two estimates of the model errors variance. One hypothesis that is likely to be questioned is the absence of systematic model errors; if such errors were indeed present, the square of their expectancy would add to VARDIFF and may balance the budget discussed above. Besides, the rank histograms described above evidence the presence of biases over period of several days, indicating that VARDIFF may not have the required stationarity property. At last, it is also possible that the observations errors are imprecisely specified for our period of study.

Table 2
Empirical assessment of ENS_COAST. The variance of the observational errors is computed from the information given by Tran et al. (2021) (their Table 1).

	Variance of the difference (cm^2/s^2) (VARDIFF)	Ensemble variance (cm^2/s^2) (ENSVAR)	Variance of the observational error (cm^2/s^2)
Zonal current	91	83	135
Meridional current	154	141	91

4.2. *Impact of a surface current observation on the simulation of the plume*

In order to explore further the use of HFR observations to constrain the model, we compute the representers, that is the influence functions (covariances) in the model state of a single assimilated observation, and the associated corrections to the model state within some assumption on the innovation. We follow the method of Echevin et al. (2000).

We analyze the corrections onto different variables when an observation of the zonal or of the meridional current component is assimilated, with an innovation of +15 cm/s. This value is chosen arbitrarily but based on comparisons between the simulation and the HFR data. The observation error is assumed to be 10 cm/s based on estimates provided in Tran et al. (2021). The HFR observation is located at 19°N, 106°E within the coastal current (when present); for comparison we also compute the correction from an observation of the meridional current further offshore (outside the HFR grid). The data is assimilated on July 10 and we analyze the zero-lag correction on the same day. This date corresponds to a period of large spread of the current (6-8 cm/s) and a situation with a coastal current to the south along the Vietnamese coasts with an amplitude larger than 20 cm/s (not shown). A positive innovation at 19°N corresponds to a simulated coastal current which has a stronger southward component and a weaker component in the offshore direction than the observation.

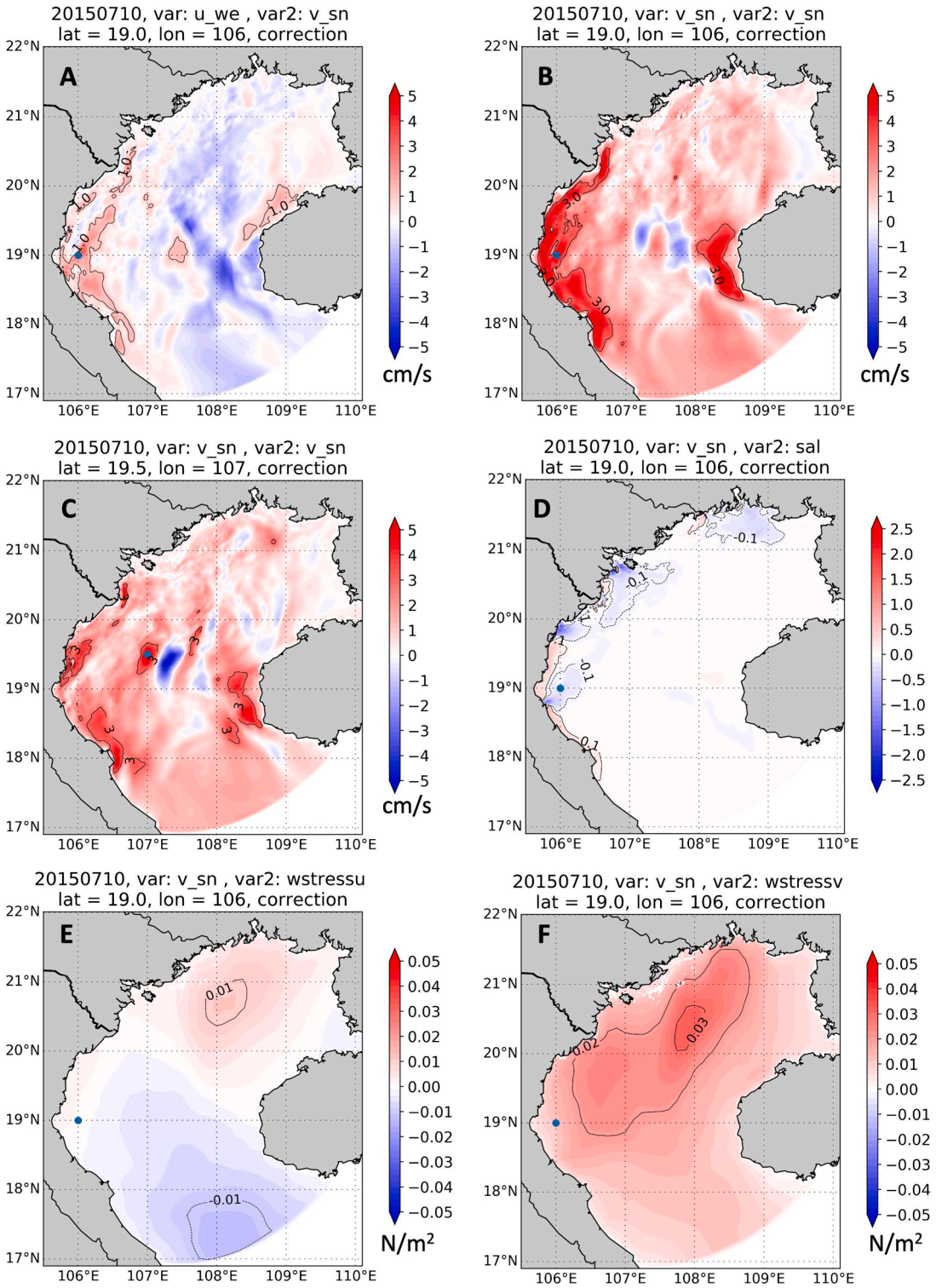

Fig. 19. **(A)** Correction (in cm/s) on the meridional velocity (V) in response to a 15 cm/s innovation in zonal velocity (U) at 19°N, 106°E (blue point) on July 10. **(B)** Corrections on V in response to an innovation in V (15 cm/s) at the same point. **(C)**: The correction on V in response to an innovation in V at the different location (19.5°N, 107°E). **(D–F)**: Corrections on surface salinity, τ_x, τ_y, respectively, in response to an innovation in V (15 cm/s) at the same point as **(A, B)**.

As expected, the corrections applied to the meridional velocity field if an observation of the meridional velocity at (19°N, 106°E) is assimilated are positive on the coastal current (reaching 5 cm/s, Fig. 19B). If the observation is on the zonal velocity (Fig. 19A), the correction on the coastal current is smaller by ~2-4 cm/s and is positive: the current is deflected offshore and consequently the coastal component decreases. Fig. 19C shows that an observation of the meridional component out of the coastal current leads to corrections with similar pattern and amplitude at the basin scale than an observation close to the coast; its impact on the coastal current is however smaller by several cm/s.

Fig. 19**D** shows the correction on sea surface salinity resulting from the assimilation of the meridional component of the current at (19°N, 106°E). As expected, a decrease of the southward coastal current leads to an increase of SSS along the coast south of 19.5°N. On the contrary the

correction is negative (corresponding to an increase of freshwater) in the northeastward direction off the mouths of the Red River and of the Lam and Ma rivers. The correction does not exceed 0.5. We also note an impact on the SSS in the northernmost part of the basin, probably due to the general increase of the northward velocity component (Fig. 19**B)** that prevents a southward spread of the freshwater from the northern rivers.

Previous results (e.g. Lamouroux, 2006; Barth et al., 2011) suggested that to efficiently constrain the high frequency dynamics over the shelf, correcting the ocean state sequentially may not be sufficient: a proposed complement or alternative is to correct the wind forcing (and possibly the atmospheric pressure forcing as well). Fig. 19**E, F** show the corrections on the zonal and meridional components of the wind stress that would result if the same observation of the meridional surface current at (19°N, 106°E) were assimilated. We observe an increase of the northward wind stress component all over the basin with a maximum correction north of 19N of ~0.03 N/m^2 (Fig. 19F), consistently with the pattern of the mean wind stress spread (Fig. 6B). The correction on the zonal wind stress leads to an intensification of the easterlies in the south and westerlies in the north but does not exceed 0.01 N/m^2 (Fig. 19E).

In conclusion to this section, our first attempt to estimate the impact of assimilating HFR observations using the ENS_COAST ensemble and an EnKF-like assimilation method shows that indeed the assimilation would be useful to constrain the river plume by correcting directly the surface salinity, but also the surface coastal current and the wind stress. Of course this experiment remains 'theoretical' since single observations only are considered. However, they pave the way for configuring more realistic DA experiments in the future, by evidencing the possible scales and amplitude of the correction.

5. Conclusions and perspectives

This study was meant as a complement to previous work (ND21) where we analyzed the variability at different time scales of the Red River plume in the far field, using numerical simulations and an unsupervised classification method. In particular, we had evidenced the impact of the monsoon reversal on the plume pattern, thickness and on the associated coastal circulation. In this study, ensemble simulations are used to assess the impact of an important source of uncertainties, namely the uncertainties on the forcing wind field (ECMWF analyses), on the Red River plume and transport of freshwater. In particular, we aim at assessing the robustness of the clustering results of ND21 and identify the physical processes at work to explain the sensitivity. We further explore the potential impacts of the assimilation of single observations of surface current stemming from high-frequency radars. This paper is dedicated to the Red River plume but we believe that the results, which are summarized below, contain valuable information for the modeling of other river plume systems in shelf seas.

One preliminary methodological question that we address is: how to generate wind perturbations that are indeed representative of ECMWF wind errors? By comparing the daily ECMWF analyses to satellite (ASCAT) wind data, the uncertainty of the ECMWF wind in this area is estimated as 60% of the wind variability. Two ensembles of 50 members each are generated: ENS-REF and ENS_COAST which represents enhanced uncertainties of the wind in a coastal strip along Vietnam. We then check if both ensembles are indeed significantly representative of ECMWF wind errors by comparing the statistical distributions of the perturbed winds with satellite and *in situ* observations.

We start the analysis of the ensemble by examining the spatial averages and time variability of the ensemble spread (defined as the standard deviation across the members) of the wind stress and of the ocean variables. We identify three typical situations. The first one (5-12 July) corresponds to a temporary weakening and reversal of the summer monsoon wind and to the largest wind stress spread. The plume is advected southward along the coast, but the spread of the current is large, both at surface and at the bottom: 8-10 cm/s and 2-4 cm/s respectively on average in the area. The second period (July 20 - August 5) corresponds to strong summer monsoon winds from the southwest; the mean spread of the wind stress is also high. Consistently, the surface current has large components to the north and east with a mean spread of 4-6 cm/s. The third period (5-30 August) is characterized by weaker winds (except around Aug 15), weaker wind spread and weak surface currents. The spread of the surface current reaches its maximum when the current is the lowest, suggesting that weak currents are more sensitive to the wind perturbations during this period. This period also corresponds to the largest runoff and to the largest spread of sea surface salinity.

The sensitivity of the plume is then examined relying on the metrics used by ND21 to characterize the plume main properties. We first find that wind uncertainties induce large uncertainties on the plume area that are highly variable in time. They reach 40% of the plume area in late July, but are the largest in absolute at the end of August when the plume area itself is the largest. During this period, the plume area evolves very differently depending on the members: the difference between the largest and the smallest areas reaches about 15000 km^2 when the plume area is about 20000 km^2. It also shows that the plume area is sensitive to the coastal wind error in ENS_COAST.

A difficulty in analyzing the ensemble behavior in terms of plume properties is that the occurrence of the plume is expressed as a categorical variable (with two values 1 and 0), making usual metrics such as ensemble mean or spread less suitable. The uncertainties on the plume location are thus presented as maps of probability of occurrence. The clustering method is another option. The patterns of the plume are classified using as reference (i.e the centroids) the clusters from the unperturbed run (i.e. the clusters of ND21). The classification is robust with respect to the wind uncertainties since the most likely cluster to occur is consistently found to be the one of the unperturbed simulation. However, depending on the dates, different members can be attributed to different clusters at the same date. The representativity of each cluster is also questioned in the sense that the membership to a cluster is more or less strong and thus more or less significant. This leads us to consider for future work a classification approach taking into account at least two criteria to characterize the plume, for instance both the occupied surface and the thickness. Other classification methods, such as the 'fuzzy classification' of Delaval et al. (2021), could lead to a more useful notion of membership; the counterpart is the need to choose additional parameters (e.g. the coefficient of fuzziness), often using arbitrary choices. One idea to be tested would be to use the cluster attribution spread in this study to build the fuzziness coefficient.

The plume thickness is also shown to be highly sensitive to the wind perturbations with typical variations of 2-3 m across the ensemble. We evidenced several processes. First, when the wind is weak, the perturbations can lead to wind reversal which, in turn, changes drastically the coastal circulation and therefore the advection of the freshwater. This results in different shapes and sizes of the plume at the surface. Changes in wind amplitudes also directly impact the local vertical mixing. Interactions between these processes explains the large time variability of the sensitivity. For instance, when the plume is surface advected offshore it is more sensitive to wind mixing than when it is confined to the coast with a deeper extent. The impact of the wind forcing on the mesoscale activity is also found to impact the plume dilution and transport. Another expected driver of the plume mixing is the frontal activity (as discussed for instance by Horner-Devine et al., 2015); we have chosen not to examine this process, which should be the subject of a full separate study.

Similarly, in this study we focus on the far field of the Red River plume but it would be interesting to study in detail the impact of wind on the near and mid field regions of the plume, as previous studies evidence different behaviors. For instance, Kakoulaki et al. (2014) suggested that the wind speed of more than 4m/s can affect the Merrimack River plume trajectory in all regions including the near field where the flow is critical. In contrast, Toublanc et al. (2023) find that the wind variability

does not impact the near field of the Gironde River plume.

At last we estimate the impact of the wind uncertainties on the freshwater transport by the coastal current. We found that the sensitivity of the transport is large, sometimes almost as large as the transport itself; this suggests that the transport estimate from a single simulation is not a robust quantity with respect to uncertainties in the wind forcing. We also show that the uncertainties on the transport are mainly due to uncertainties on the surface current, at least in the section studied.

In the prospect of assimilating HFR observations, we assessed the ensemble consistency with respect to the HFR data and found that the range of values reached by the model in both ensembles (ENS_REF and ENS_COAST) and by the HFR data (given their uncertainties) are not disjoint (at least not always and not everywhere); however, it is likely that the model suffers from systematic errors that are not represented by the ensemble (by construction). We then examine the impact of assimilating single observations of surface current at a given point and dates on ocean variables and wind stress. We found that indeed HFR data could constrain the simulation of the plume both directly (through correction on the surface salinity and current) and indirectly (through corrections on the wind stress).

In this study, due to constraint in computing resources, the ensemble size does not exceed 50 members; clearly, it would be relevant in the future to investigate the robustness of the ensemble statistics for much larger ensembles (e.g. 100 members). Besides, we limited the source of stochastic perturbations to the surface wind. Although it is likely that the wind is an essential, if not dominant, source of ocean uncertainties (as shown in past studies on different regions) in a shelf sea such as the Gulf of Tonkin, we could have considered other sources (e.g. Vervatis et al., 2021a; Matsuzaki and Inoue, 2022), including forcings such as boundary conditions or parameterization (e.g on vertical mixing). In a future study, we intend to focus on the plume uncertainties due to uncertainties on the Red River distributaries discharge at daily to seasonal scales.

CRediT authorship contribution statement

Tung Nguyen-Duy: Formal analysis, Methodology, Visualization, Data curation, Writing – original draft. **Nadia K. Ayoub:** Conceptualization, Formal analysis, Methodology, Resources, Writing – original draft, Writing – review & editing. **Pierre De-Mey-Frémaux:** Conceptualization, Methodology, Writing – review & editing. **Thanh Ngo-Duc:** Methodology, Resources, Writing – review & editing.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgements

The Ph.D. thesis of Tung Nguyen-Duy has been funded by the “Programme d’Excellence” of the French Embassy of Vietnam and by IRD ARTS program. This work is a contribution to the LMI LOTUS project (http://lotus.usth.edu.vn) and to the SWOT COCTO-FO project funded by CNES (French Spatial Agency). First we gratefully thank P. Marsaleix (LEGOS) for his help with the SYMPHONIE model and the modelling strategy. Many thanks as well to M. Ghantous (CLS) for discussions at the beginning of this work on the ensemble generation. We thank Manh-Cuong Tran and Alexei Sentchev (LOG, France) for providing the mapped HFR data. This study has been conducted using ECMWF products (www.ecmwf.int) and E.U. Copernicus Marine Service product (GLOBAL_ANALYSIS_FORECAST_PHY_001_024, available at https://resources.marine.copernicus.eu/?option=com_csw&view=details&product_id=GLOBAL_ANALYSIS_FORECAST_PHY_001_024). The SYMPHONIE code is distributed by the SIROCCO French national service website (https://sirocco.obs-mip.fr/), hosted by the Observatoire Midi-Pyrénées, Toulouse, France. This work has used HPC resources from CALMIP/France (grant 2022-p1119), from the HILO cluster at the University of Science and Technology of Hanoi (USTH) and from the computing and storage facilities at the Observatoire Midi-Pyrénées. Part of the work has been done while N. Ayoub was a visiting scientist at USTH in the framework of the LMI LOTUS. Support from IRD, CNES, CNRS, and USTH is gratefully acknowledged.

Appendix A. Ensemble assessment using observations

In Section 2.3, we aim at verifying if the variance of the generated wind perturbations is consistent with ECMWF wind error variance as estimated from the comparison with satellite data. In Section 4.1, we use HF radar observations to verify if the model error variance that is computed from the ensemble simulations is consistent with errors estimated from the model-data misfits. In both sections, the consistency check is based on the same approach which is described in this appendix. Such approach is used in data assimilation studies, as for example described in Appendix A of Evensen, (2003).

Let ψ^m and ψ^o be the estimates of the true variable ψ^t from the model and from observations respectively; the variables are expressed in the observation space. We can write:

$$\psi^m = \psi^t + \varepsilon^m \tag{1a}$$

$$\psi^o = \psi^t + \varepsilon^o \tag{1b}$$

where ε^m,$\varepsilon^o$are the errors from the model and observations respectively.

In the case of Section 2.3, ψ^t is a zonal or meridional wind component at 10 m, ψ^o is the corresponding component from ASCAT data interpolated onto the ECMWF grid, ψ^m is the corresponding component from ECMWF. The wind perturbations that we generate aim at representing the ECMWF model uncertainty ε^m.

In the case of Section 4.1, ψ^t is a zonal or meridional component of the surface current, ψ^o is the corresponding component from HF radar data, ψ^m is the corresponding simulated component with the SYMPHONIE model interpolated onto the HF radar grid. The ensemble of simulations allows us to estimate ε^m.

The ensemble method assessment relies on the statement that two estimates of the model errors are available: one direct estimate from the ensemble, one indirect estimate from the model-data misfits, $\psi^o - \psi^m$, (the so-called innovation).

Assuming that the errors from the model and the data are uncorrelated and non-biased (i.e. their expectancy is null), the variance of the innovation writes:

$$\overline{\left[(\psi^o - \psi^m) - \overline{(\psi^o - \psi^m)}\right]^2} = \overline{\varepsilon^{o^2}} + \overline{\varepsilon^{m^2}} \tag{2}$$

where the overline $\overline{(.)}$ denotes an expectancy value. (2) can be rewritten:

$$\overline{\varepsilon^{m^2}} = \overline{\left[(\psi^o - \psi^m) - \overline{(\psi^o - \psi^m)}\right]^2} = \overline{\varepsilon^{o^2}} \tag{3}$$

The left-hand side term of (3) is the variance of the model errors; we choose here to estimate it from the ensemble-time variance (the variance over ensemble and time samples) computed in the observation space (that is, in our case, at the observations grid points). Hereafter, we note $\overline{(.)}^E$ the expectancy when estimated from the ensemble and time (including both member samples and time samples), i.e. as a discrete sum over the members and over time.

The first term on the right-hand side of (3) is the variance of the innovation which we compute as the variance over time (i.e. the expectancy is estimated as a discrete sum over the time samples). Hereafter, we note $\overline{(.)}^T$ the expectancy when estimated over time.

$\overline{\varepsilon^{o^2}}$ is the variance of the observation errors, as provided with the observations.

Finally (3) can be re-written, using our choice of statistical estimators:

$$\overline{\varepsilon^{m^2}}^E = \overline{\left[(\psi^o - \psi^m) - \overline{(\psi^o - \psi^m)}\right]^2}^T - \overline{\varepsilon^{o^2}} \tag{4}$$

Any misfits between the left-hand and right-hand side of (4) indicate inconsistencies that may have several origins: (a) the observations errors are incorrectly specified, (b) there is a systematic error in the model, (c) the generated perturbations are incorrectly calibrated, (d) the model errors are due to different sources than the one(s) represented by the ensemble (in our case uncertainties on the forcing wind field), (e) the ensemble statistics have not converged yet (and a larger size ensemble should be considered), (f) the time statistics have not converged, e.g. because of the inherent non-stationarity of 1st- or 2nd-order statistics.

References

Andreu-Burillo, I., Caniaux, G., Gavart, M., De Mey, P., Baraille, R., 2002. Assessing ocean-model sensitivity to wind forcing uncertainties. Geophys. Res. Lett. 29 (18), 5.1–5.4. https://doi.org/10.1029/2001GL014473.

Auclair, F., Marsaleix, P., De Mey, P., 2003. Space-time structure and dynamics of the forecast error in a coastal circulation model of the Gulf of Lions. Dyn. Atmos. Oceans 36 (4), 309–346. https://doi.org/10.1016/S0377-0265(02)00068-4.

Ayoub, N.K., Lucas, M., De Mey, P., 2015. Estimating uncertainties on a Gulf Stream mixed-layer heat budget from stochastic modeling. J. Mar. Syst. 150, 66–79. https://doi.org/10.1016/j.jmarsys.2015.04.008.

Barth, A., Alvera-Azcárate, A., Beckers, J.M., et al., 2011. Correcting surface winds by assimilating high-frequency radar surface currents in the German Bight. Ocean Dyn. 61, 599–610. https://doi.org/10.1007/s10236-010-0369-0.

Damien, P., Bosse, A., Testor, P., Marsaleix, P., Estournel, C., 2017. Modeling postconvective submesoscale coherent vortices in the Northwestern Mediterranean Sea. J. Geophys. Res. Oceans 122, 9937–9961. https://doi.org/10.1002/2016JC012114.

Delaval, A., Duffa, C., Pairaud, I., Radakovitch, O., 2021. A fuzzy classification of the hydrodynamic forcings of the Rhone River plume: an application in case of accidental release of radionuclides. Environ. Model. Softw. 140, 105005 https://doi.org/10.1016/j.envsoft.2021.105005.

Echevin, V., Mey, P.D., Evensen, G., 2000. Horizontal and vertical structure of the representer functions for sea surface measurements in a coastal circulation model. J. Phys. Oceanogr. 30 (10), 2627–2635. https://doi.org/10.1175/1520-0485(2000)030<2627:HAVSOT>2.0.CO;2.

EUMETSAT, OSI SAF 2021. ASCAT wind product user manual, SAF/OSI/CDOP/KNMI/TEC/MA/126, Version 1.17.1.

Evensen, G., 2003. The Ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn. 53, 343–367. https://doi.org/10.1007/s10236-003-0036-9.

Gao, J., Xue, H., Chai, F., Shi, M., 2013. Modeling the circulation in the Gulf of Tonkin, South China Sea. Topical Collection on21–24 May 2012Proceedings of the4th International Workshop on Modelling the Ocean in Yokohama, Japan Ocean Dyn. 63 (8), 979–993. https://doi.org/10.1007/s10236-013-0636-y, 21–24 May 2012.

Ghantous, M., Ayoub, N., De Mey-Frémaux, P., Vervatis, V., Marsaleix, P., 2020. Ensemble downscaling of a regional ocean model. Ocean Model. 145, 101511 https://doi.org/10.1016/j.ocemod.2019.101511.

Hamill, T.M., 2000. Interpretation of rank histograms dor verifying ensemble forecasts. Mon. Weather Rev. 129 https://doi.org/10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2.

Herrmann, M., Nguyen-Duy, T., Ngo-Duc, T., Tangang, F., 2021. Climate change impact on sea surface winds in Southeast Asia. Int. J. Climatol. https://doi.org/10.1002/joc.7433 joc.7433.

Hetland, R.D., Hsu, T.J., 2013. Freshwater and sediment dispersal in large river plumes. Biogeochemical Dynamics at Large River–Coastal Interfaces: Linkages with Global Climate Change. Springer, New York, pp. 55–85 ed. TS Bianchi, MA Allison, W-J Cai.

Hickey, B., McCabe, R., Geier, S., Dever, E., Kachel, N., 2009. Three interacting freshwater plumes in the northern California current system. J. Geophys. Res. 114, C00B03. https://doi.org/10.1029/2008JC004907.

Horner-Devine, A.R., Hetland, R.D., MacDonald, D.G., 2015. Mixing and transport in coastal river plumes. Annu. Rev. Fluid Mech. 47, 569–594. https://doi.org/10.1146/annurev-fluid-010313-14140.

Jia, Y., Whitney, M.M., 2019. Summertime Connecticut river water pathways and wind impacts. J. Geophys. Res. Oceans 124 (3), 1897–1914. https://doi.org/10.1029/2018JC014486.

Kim, S., Samelson, R.M., Snyder, C., 2011. Toward an uncertainty budget for a coastal ocean model. Mon. Weather Rev. 139 https://doi.org/10.1175/2010MWR3352.1.

Kakoulaki, G., MacDonald, D., Horner-Devine, A.R., 2014. The role of wind in the near field and mid field of a river plume. Geophys. Res. Lett. 41, 5132–5138. https://doi.org/10.1002/2014GL060606.

Kirtman, B., Min, D., Infanti, J., Kinter, J., Paolino, D., Zhang, Q., van den Dool, H., Saha, S., Mendez, M., Becker, E., Peng, P., Tripp, P., Huang, J., DeWitt, D., Tippett, M., Barnston, A., Li, S., Rosati, A., Schubert, S., Rienecker, M., Suarez, M., Li, Z., Marshak, J., Lim, Y., Tribbia, J., Pegion, K., Merryfield, W., Denis, B., Wood, E., 2014. The North American multimodel ensemble: phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction. Bull. Am. Meteorol. Soc. 95, 585–601. https://doi.org/10.1175/BAMS-D-12-00050.1.

Large, W. G., & Yeager, S. (2004). Diurnal to decadal global forcing for ocean and sea–ice models: the data sets and flux climatologies. doi:10.5065/D6KK98Q6.

Lamouroux J., 2006. Erreurs de prévision d'un modèle océanique barotrope du Golfe de Gascogne en réponse aux incertitudes sur les forçages atmosphériques : caractérisation et utilisation dans un schéma d'assimilation de données à ordre réduit. Océan, Atmosphère. Université Paul Sabatier - Toulouse III. https://theses.hal.science/tel-00107174.

Le Hénaff, M., De Mey, P., Marsaleix, P., 2009. Assessment of observational networks with the Representer Matrix Spectra method-application to a 3D coastal model of the Bay of Biscay. Ocean Dyn. 59, 3–20. https://doi.org/10.1007/s10236-008-0144-7.

Lucas, M., Ayoub, N., Barnier, B., Penduff, T., De Mey, P., 2008. Stochastic study of the temperature response of the upper ocean to uncertainties in the atmospheric forcing in an Atlantic OGCM. Ocean Model. 20 (1), 90–113. https://doi.org/10.1016/j.ocemod.2007.07.006.

Luu, T.N.M., Garnier, J., Billen, G., Orange, D., Némery, J., Le, T.P.Q., Tran, H.T., Le, L.A., 2010. Hydrological regime and water budget of the Red River Delta (Northern vietnam). J. Asian Earth Sci. 37 (3), 219–228. https://doi.org/10.1016/j.jseaes.2009.08.004.

Lyard, F.H., Allain, D.J., Cancet, M., Carrère, L., Picot, N., 2021. FES2014 global ocean tide atlas: design and performance. Ocean Sci. 17, 615–649. https://doi.org/10.5194/os-17-615-2021.

Marsaleix, P., Auclair, F., Estournel, C., 2006. Considerations on open boundary conditions for regional and coastal ocean models. J. Atmos. Ocean. Technol. 23, 1604–1613. https://doi.org/10.1175/JTECH1930.1.

Marsaleix, P., Auclair, F., Floor, J.W., Herrmann, M.J., Estournel, C., Pairaud, I., Ulses, C., 2008. Energy conservation issues in sigma-coordinate free-surface ocean models. Ocean Model. 20 (1), 61–89. https://doi.org/10.1016/j.ocemod.2007.07.005.

Matsuzaki, Y., Inoue, T., 2022. Perturbation of boundary conditions to create appropriate ensembles for regional data assimilation in coastal estuary modeling. J. Geophys. Res. Oceans 127 (4). https://doi.org/10.1029/2021JC017911.

Michaud, H., Marsaleix, P., Leredde, Y., Estournel, C., Bourrin, F., Lyard, F., Mayet, C., Ardhuin, F., 2012. Three–dimensional modelling of wave–induced current from the surf zone to the inner shelf. Ocean Sci. 8 (4), 657–681. https://doi.org/10.5194/os-8-657-2012.

Neumann, L.E., Šimůnek, J., Cook, F.J., 2011. Implementation of quadratic upstream interpolation schemes for solute transport into HYDRUS–1D. Environ. Model. Softw. 26, 11. https://doi.org/10.1016/j.envsoft.2011.05.010.

Nguyen-Duy, T., Ayoub, N.K., Marsaleix, P., Toublanc, F., De Mey-Frémaux, P., Piton, V., Herrmann, M., Duhaut, T., Tran, M.C., Ngo-Duc, T., 2021. Variability of the Red

River Plume in the Gulf of Tonkin as revealed by numerical modeling and clustering analysis. Front. Mar. Sci. 8, 772139 https://doi.org/10.3389/fmars.2021.772139.
Nguyen-Duy D.T. (2022). Etude du panache du fleuve Rouge dans le golfe du Tonkin à partir d'une analyse en clusters et de simulations d'ensemble [These de doctorat, Toulouse 3]. https://www.theses.fr/2022TOU30089.
Mercator Océan (2020), Global Ocean 1/12° Physics Analysis and Forecast updated Daily, doi:10.48670/moi-00016.
Pasmans, I., Kurapov, A.L., 2019. Ensemble of 4DVARs (En4DVar) data assimilation in a coastal ocean circulation model, Part I: Methodology and ensemble statistics. Ocean Model. 144 https://doi.org/10.1016/j.ocemod.2019.101493.
Piton, V., Herrmann, M., Lyard, F., Marsaleix, P., Duhaut, T., Allain, D., Ouillon, S., 2020. Sensitivity study on the main tidal constituents of the Gulf of Tonkin by using the frequency–domain tidal solver in T–UGOm. Geosci. Model. Dev. 13 (3), 1583–1607. https://doi.org/10.5194/gmd-13-1583-2020.
Piton, V., Herrmann, M., Marsaleix, P., Duhaut, T., Ngoc, T.B., Tran, M.C., Shearman, K., Ouillon, S., 2021. Influence of winds, geostrophy and typhoons on the seasonal variability of the circulation in the Gulf of Tonkin: a high–resolution 3D regional modeling study. Reg. Stud. Mar. Sci. 45, 101849 https://doi.org/10.1016/j.rsma.2021.101849.
Quattrocchi, G., De Mey, P., Testut, C.E., Ayoub, N., Reffray, G., Chanut, J., Drillet, Y., Vervatis, V., 2014. Characterization of errors of a regional model of the Bay of Biscay in response to wind uncertainties: a first step toward a data assimilation system suitable for coastal sea domains. J. Oper. Oceanogr. 7, 25–34. https://doi.org/10.1080/1755876X.2014.11020156.
Rogowski, P., Zavala-Garay, J., Shearman, K., Terrill, E., Wilkin, J., Tran, H.L., 2019. Air–Sea–Land forcing in the Gulf of Tonkin: assessing seasonal variability using modern tools. Oceanography 32, 150–161. https://doi.org/10.5670/oceanog.2019.223.
Sakov, P., Oke, P.R., 2008. A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. Tellus A Dyn. Meteorol. Oceanogr. 60 (2), 361. https://doi.org/10.1111/j.1600-0870.2007.00299.x.
Schiller, R.V., Kourafalou, V.H., Hogan, P., Walker, N.D., 2011. The dynamics of the Mississipi River plume: impact of topography, wind and offshore forcing on the fate of plume waters. J. Geophys. Res. 116 https://doi.org/10.1029/2010JC006883.
Storto, A., Yang, C., 2023. Stochastic schemes for the perturbation of the atmospheric boundary conditions in ocean general circulation models. Front. Mar. Sci. https://doi.org/10.3389/fmars.2023.1155803.
Talagrand, O., Vautard, R., Strauss, B., 1997. Evaluation of probabilistic prediction systems. In: Proceedings of the ECMWF Workshop on Predictability. ECMWF, pp. 1–25.
Toublanc, F., Ayoub, N.K., Lyard, F., Marsaleix, P., Allain, D.J., 2018. Tidal downscaling from the open ocean to the coast: a new approach applied to the Bay of Biscay. Ocean Model. 124, 16–32. https://doi.org/10.1016/j.ocemod.2018.02.001.
Toublanc, F., Ayoub, N.K., Marsaleix, P., 2023. On the role of wind and tides in shaping the Gironde River plume (Bay of Biscay). Cont. Shelf Res. 253 https://doi.org/10.1016/j.csr.2022.104891.
Tran, M.C., Sentchev, A., Nguyen, K.C., 2021. Multi–scale variability of circulation in the Gulf of Tonkin from remote sensing of surface currents by high–frequency radars. Ocean Dyn. 71 (2), 175–194. https://doi.org/10.1007/s10236-020-01440-x.
Tran, M.C., Sentchev, A., Berti, S., Ayoub, N.K., Nguyen-Duy, T., Kim Cuong, Nguyen, 2022. Assessment of relative dispersion in the Gulf of Tonkin using numerical modeling and HF radar observations of surface currents. Cont. Shelf Res. 245. https://doi.org/10.1016/j.csr.2022.104784. Volume.
Vaz, N., Rodrigues, J.G., Mateus, M., Franz, G., Campuzano, F., Neves, R., Dias, J.M., 2018. Subtidal variability of the Tagus River Plume in Winter 2013. Sci. Total Environ. 627, 1353–1362. https://doi.org/10.1016/j.scitotenv.2018.01.325.
Vervatis, V., Testut, C.E., De Mey, P., Ayoub, N., Chanut, J., Quattrocchi, G., 2016. Data assimilative twin-experiment in a high-resolution Bay of Biscay configuration: 4DEnOI based on stochastic modeling of the wind forcing. Ocean Model. 100, 1–19. https://doi.org/10.1016/j.ocemod.2016.01.003.
Vervatis, V.D., De Mey-Frémaux, P., Ayoub, N., Karagiorgos, J., Ghantous, M., Kailas, M., Testut, C.E., Sofianos, S., 2021a. Assessment of a regional physical–biogeochemical stochastic ocean model. Part 1: ensemble generation. Ocean Model. 160, 101781 https://doi.org/10.1016/j.ocemod.2021.101781.
Vervatis, V.D., De Mey-Frémaux, P., Ayoub, N., Karagiorgos, J., Ciavatta, S., Brewin, R.J.W., Sofianos, S., 2021b. Assessment of a regional physical–biogeochemical stochastic ocean model. Part 2: empirical consistency. Ocean Model. 160, 101770 https://doi.org/10.1016/j.ocemod.2021.101770.
Vinh, V.D., Ouillon, S., Thanh, T.D., Chu, L.V., 2014. Impact of the Hoa Binh dam (Vietnam) on water and sediment budgets in the Red River basin and delta. Hydrol. Earth Syst. Sci. 18 (10), 3987–4005. https://doi.org/10.5194/hess-18-3987-2014.
Yankovsky, A.E., Fribance, D.B., Cahl, D., Voulgaris, G., 2022. Offshore spreading of a supercritical plume under upwelling wind forcing: a case study of the Winyah Bay outflow. Front. Mar. Sci. 8, 785967 https://doi.org/10.3389/fmars.2021.785967.
Wang, Y., Xue, H., Chai, F., Chao, Y., Farrara, J., 2014. A model study of the Copper River plume and its effects on the northern Gulf of Alaska. Ocean Dyn. 64 (2), 241–258. https://doi.org/10.1007/s10236-013-0684-3.