

# Comparison of different incremental analysis update schemes in a realistic assimilation system with Ensemble Kalman Filter



Y. Yan<sup>a,\*</sup>, A. Barth<sup>b</sup>, J.M. Beckers<sup>b</sup>, J.M. Brankart<sup>c</sup>, P. Brasseur<sup>c</sup>, G. Candille<sup>c</sup>

<sup>a</sup> LISTIC, Polytech Annecy-Chambéry, Université Savoie Mont-Blanc, Annecy-le-Vieux, France

<sup>b</sup> GHER, University of Liège, Liège, Belgium

<sup>c</sup> CNRS LGGE, Université de Grenoble 1, Grenoble, France

## ARTICLE INFO

### Article history:

Received 24 May 2016

Revised 16 April 2017

Accepted 13 May 2017

Available online 15 May 2017

### Keywords:

Incremental analysis update

Ensemble Kalman Filter

Probabilistic metrics

Comparison

## ABSTRACT

In this paper, three incremental analysis update schemes (IAU 0, IAU 50 and IAU 100) are compared in the same assimilation experiments with a realistic eddy permitting primitive equation model of the North Atlantic Ocean using the Ensemble Kalman Filter. The difference between the three IAU schemes lies on the position of the increment update window. The relevance of each IAU scheme is evaluated through analyses on both thermohaline and dynamical variables. The validation of the assimilation results is performed according to both deterministic and probabilistic metrics against different sources of observations. For deterministic validation, the ensemble mean and the ensemble spread are compared to the observations. For probabilistic validation, the continuous ranked probability score (CRPS) is used to evaluate the ensemble forecast system according to reliability and resolution. The reliability is further decomposed into bias and dispersion by the reduced centred random variable (RCRV) score. The obtained results show that 1) the IAU 50 scheme has the same performance as the IAU 100 scheme 2) the IAU 50/100 schemes outperform the IAU 0 scheme in error covariance propagation for thermohaline variables in relatively stable region, while the IAU 0 scheme outperforms the IAU 50/100 schemes in dynamical variables estimation in dynamically active region 3) in case with sufficient number of observations and good error specification, the impact of IAU schemes is negligible. The differences between the IAU 0 scheme and the IAU 50/100 schemes are mainly due to different model integration time and different instability (density inversion, large vertical velocity, etc.) induced by the increment update. The longer model integration time with the IAU 50/100 schemes, especially the free model integration, on one hand, allows for better re-establishment of the equilibrium model state, on the other hand, smooths the strong gradients in dynamically active region.

© 2017 Elsevier Ltd. All rights reserved.

## 1. Introduction

The incremental analysis update (IAU) scheme, consists of incorporating analysis increments in a gradual way, is originally proposed by Bloom et al. (1996). It acts like a low-pass filter, with which the strong gradients are smoothed and the spin up effects of the model forecast such as the spurious waves are filtered out. Thus, a more appropriate balance is achieved with the IAU scheme. Moreover, the IAU scheme does not dissipate the model's mesoscale signal, because the IAU forcing is significant only in areas where an analysis increment is available. Since 2000, it has been frequently used in data assimilation with ocean general circulation model (OGCM). Later, diverse varieties have been developed and implemented (Carton et al., 2000; Huang et al., 2002;

Alves et al., 2004; Ourmières et al., 2006). The main difference between these IAU schemes lies on the position of the increment update window. Carton et al. (2000) used an adaptive scheme very similar to the one in Bloom et al. (1996) for a free surface OGCM (IAU 50). In their scheme, an increment is calculated at the end of each assimilation window and applied at half of the assimilation window length before and after the analysis step. In this way, 50% of the model integration is repeated inside each assimilation window. In Huang et al. (2002) and Alves et al. (2004), the IAU schemes used are very similar (IAU 0), but different from the one proposed by Bloom et al. (1996). The increment is applied at every time step inside the next assimilation window without any model integration repeat. In Ourmières et al. (2006), another different scheme is proposed and implemented (IAU 100). The increment is applied at every time step inside the current assimilation window by re-running the model. In all these works, satisfactory results have been obtained and the capacity of the IAU techniques to act

\* Corresponding author.

E-mail address: [yajing.yan@univ-smb.fr](mailto:yajing.yan@univ-smb.fr) (Y. Yan).

like a continuous assimilation and to reduce the high frequency analysis-induced oscillations has been proven.

Yan et al. (2014) compared three IAU schemes (IAU 0, IAU 50 and IAU 100) to the conventional intermittent scheme with an idealised configuration (the so called square-box configuration (Cosme et al., 2010)) of the NEMO (Nucleus for European Modelling of the Ocean) primitive equation ocean model in a twin experiment using the Ensemble Kalman Filter (EnKF) (Evensen, 2004). They concluded that 1) in case with the same ensemble size, the three IAU schemes outperform the intermittent scheme. 2) in case with the same computational cost, thus smaller ensemble size for the IAU 50/100 schemes, the reduction of ensemble size degrades the performance of the IAU 50/100 schemes. Taken into account the reduction in analysis-induced oscillation and instability, as well as the computational cost, the IAU 0 scheme is recommended by authors. However, in these experiments, the model used is an idealised, thus non-realistic ocean circulation model and synthetic observations are generated from the model simulation by adding random noise. It is thus unlikely that the systematic bias exists between the observations and the model forecasts. The analysis state is the forecast state added to corrections which are the residuals projected onto the model state. Therefore, these corrections all live within the "ensemble universe". Although the ensemble mean is not necessarily dynamically balanced, the corrections are expected to be generally dynamically balanced and become smaller especially for well-observed systems like twin experiments. This results of small differences between the three IAU schemes. In realistic forecast systems, the corrections might be large and dependent on the localisation. Moreover, the corrections could be dynamically unbalanced to various degrees. What's the relative performance of these IAU schemes under these conditions? Are there situations where the IAU 50 or IAU 100 scheme is significantly better than the IAU 0 scheme and thus might be necessary? To answer these questions, the comparison of these IAU schemes in a realistic assimilation system seems necessary, which constitutes the first motivation of this paper.

Furthermore, in Yan et al. (2014), the assimilation results are only assessed according to the RMS error which provides a score to evaluate the deterministic system. The richness of the probabilistic ensemble distribution is thus not investigated exhaustively. Indeed, sophisticated probabilistic metrics that better make use of the probabilistic characteristics of the ensemble have been extensively promoted in recent publications (Hamill, 2000; Hersbach, 2000; Candille et al., 2006; Casati et al., 2008; Candille et al., 2014; Yan et al., 2015; Garnier et al., 2015). Probabilistic verification relies on two criteria: reliability and resolution, which correspond to the main attributes of an ensemble forecast system. The reliability corresponds to the statistical consistency between a priori predicted probabilities and a posteriori observed frequencies of the occurrence of the event under consideration. The resolution indicates the ability of a forecast system to separate a priori cases when the event under consideration occurs more or less frequently than the climatological frequency (Toth et al., 2003). The continuous ranked probability score (CRPS) (Stanski et al., 1989) allows for an evaluation of an ensemble forecast system according to these two criteria. The reliability can also be evaluated by the reduced centred random variable (RCRV) score (Talagrand et al., 1999). In Yan et al. (2015), a joint analysis of both deterministic and probabilistic verification demonstrates the relevance of the probabilistic metrics in case where the deterministic metrics alone seem insufficient to judge the assimilation results. Therefore, it will be interesting to compare the results obtained with different IAU schemes according to the probabilistic metrics. This constitutes the second motivation of this paper.

In this paper, the IAU 0, IAU 50 and IAU 100 schemes are implemented with an operational ocean circulation model over

the North Atlantic Ocean described in Section 2. Observations, the ensemble generation and the assimilation experiment setups with the EnKF are introduced in Section 3. Both deterministic and probabilistic metrics introduced in Section 4 are used to diagnose the assimilation results. Through comparisons between different IAU schemes, the relevance and the weakness of each IAU scheme are highlighted. High reliable situations where the three IAU schemes give similar results, are identified. Results discussions are given in Section 5 and Section 6. The conclusions are derived in Section 7.

## 2. Model

The circulation of the North Atlantic is simulated by the OPA (Océan PArallellisé<sup>1</sup>) code (NATL025 configuration of the NEMO model (Barnier et al., 2006)) using free surface formulation. Prognostic variables are the three-dimensional velocity fields and the thermohaline variables. The model domain covers the North Atlantic basin from 20°S to 80°N and from 98°W to 23°E. The primitive equations are discretized on an Arakawa C grid, with a horizontal resolution of  $1/4^\circ \times 1/4^\circ \cos(\phi)$  (where  $\phi$  is the latitude), which is considered as eddy-permitting in the mid-latitudes where the Rossby radius of deformation is about 100 km. The effective resolution, which becomes finer with increasing latitude, is  $\sim 27.75$  km at the equator and  $\sim 13.8$  km at 60°S or 60°N. Vertical discretization takes place on 46 geopotential levels, with a grid spacing increasing from 6 m at the surface to 250 m at the bottom. The maximum depth in the model is 5844 m.

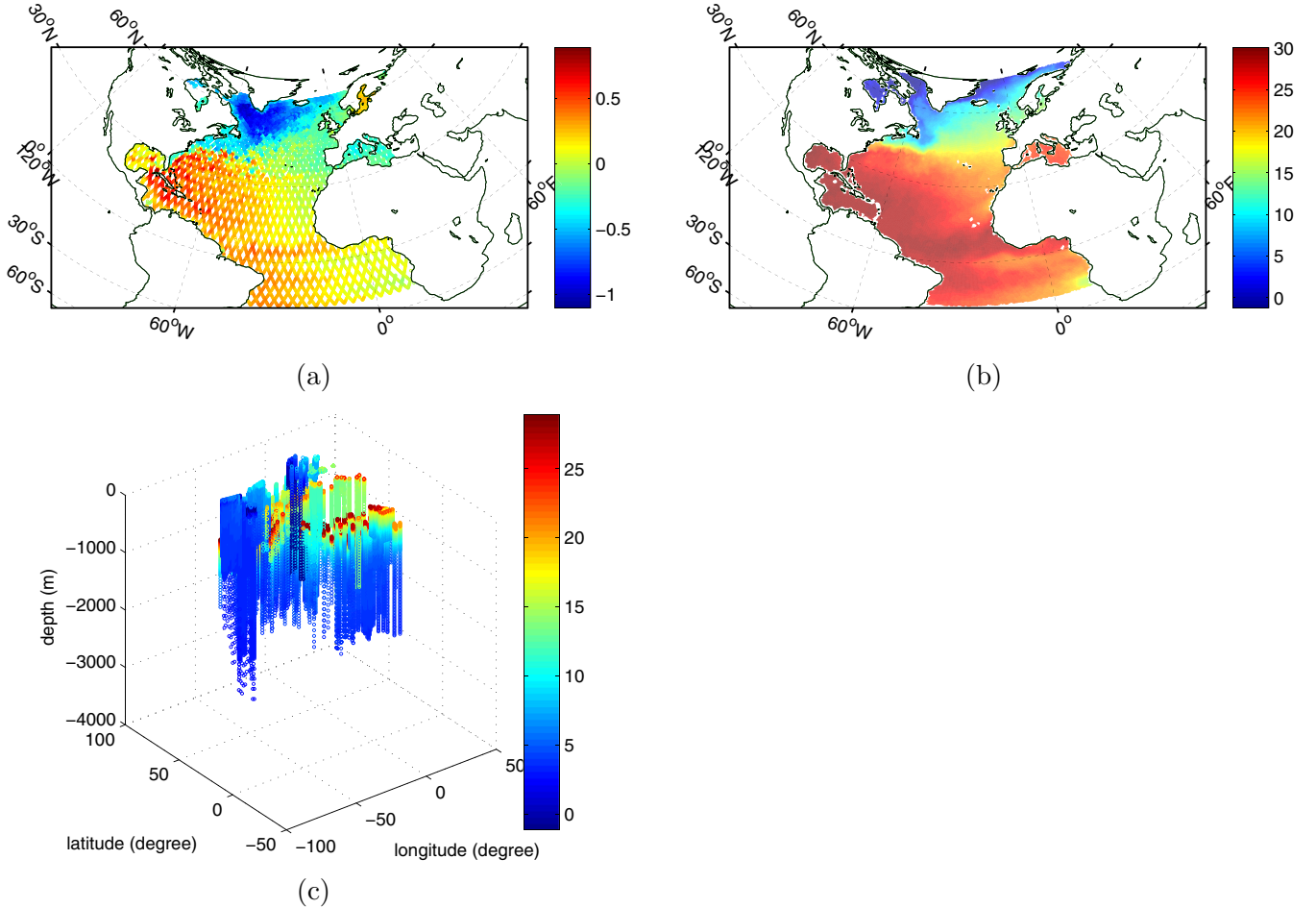
Partial step (PS) topography (Adcroft et al., 1997), that makes the depth of the bottom cell variable and adjustable to the real depth of the ocean, is used to represent flow-topography interactions. Momentum advection is realised with an energy and enstrophy conserving (EEN) numerical scheme (Arakawa and Lamb, 1981; Barnier et al., 2006) in vector form, with an additional term in the momentum equation to damp the faster external gravity waves. Tracer advection is carried out with a total variance diminishing advection scheme to avoid the overshoots generation in case of sharp gradients. Lateral mixing of tracers is modelled with a Laplacian lateral isopycnal diffusion operator,  $300 \text{ m}^2 \text{ s}^{-1}$  at the equator and decreasing poleward, proportional to the grid size. Lateral mixing of momentum is modelled with a horizontal bi-harmonic viscosity operator,  $-1.5 \times 10^{11} \text{ m}^4 \text{ s}^{-1}$  at the equator and decreasing poleward by the cube of the grid size. Surface boundary layer mixing and interior vertical mixing are parametrized according to a turbulence kinetic energy (TKE) turbulence closure model (Blanke and Delecluse, 1993). In case of static instability, a viscosity/diffusivity enhancement of up to  $10 \text{ m}^2 \text{ s}^{-1}$  is used. The forcing fluxes are calculated via bulk formulations, using the ERAinterim atmospheric forcing fields (Dee et al., 2011). The temperature and salinity fields are initialised by the Levitus climatology (Levitus et al., 1998). The horizontal and vertical velocity fields and SSH are initially set to zero. From the initial values of the prognostic variables and the atmospheric forcing field of the year 1989, the model spin up time is started. According to Testut et al. (2003), the numerical integration should neither be too long nor too short. Therefore, a model spin-up of 16 years is chosen (Kantha and Clayson, 2000). Time stepping is performed with a leap frog scheme, with  $\Delta t = 2400 \text{ s}$  as time step.

## 3. Assimilation experiments setup

### 3.1. Observation

Three types of observations, Jason-1 SSH data (Ménarda et al., 2003), AVHRR SST data (Casey et al., 2010) and ARGO temperature

<sup>1</sup> Paralleled Ocean.



**Fig. 1.** Example of (a) SSH (b) SST (c) temperature profile observation distribution.

profiles (Davis, 1991), are available at all assimilation steps. The SSH observation grids correspond to a typical along/across track resolution ( $11.2 \text{ km} \times 5.1 \text{ km}$ ), covering the North Atlantic basin except the subpolar area (Fig. 1(a)). The SST observation grids are very dense with a resolution of  $1/4^\circ \times 1/4^\circ$ , covering the whole North Atlantic basin (Fig. 1(b)). The observation grids of temperature profile are sparse, with observation points located between the surface and 2000 m depth only in part of the North Atlantic basin (Fig. 1(c)). Based on the nominal error and the representative error found in the literature (Testut et al., 2003), 5 cm and  $0.3^\circ$  are set as the observational error for SSH and temperature profile respectively. The observational error for SST is the standard deviation map associated with the temperature value, with a mean value of  $0.2^\circ$ . Moreover, ENVISAT SSH data (Resti et al., 1999), Mercator reanalysis SST data (Ferry et al., 2012) and ARGO salinity profiles are used as validation observations to evaluate the assimilation results. Note that, the Mercator reanalysis SST data are not rigorously independent of the AVHRR SST data assimilated in the experiments, because AVHRR data are used in the production of the Mercator reanalysis. An analysis localisation method is used to rule out corrections due to distant observations. We apply an approach similar to Testut et al. (2003) adapted to the square root analysis scheme of EnKF (Evensen, 2004). To compute the correction at each water column, the observations are weighted by a factor of  $\exp(-r^2/d^2)$  with  $d$  the localisation length scale. The localisation length scale is determined according to the auto-correlation length of SST and SSH, here 300 km.

### 3.2. Ensemble generation

The ensemble, with 60 members, is generated by adding realistic noise in the atmospheric forcing variables related to the temperature. Surface ocean variables are thus more involved. Perturbation for deep layers are mainly propagated from the surface perturbation through the model dynamics during the ensemble spin up time. The ensemble size is determined based on the trade off between the appropriate model error representation and the available computational capacity. The air temperature at 2 m ( $t_2$ ), wind velocities at 10 m ( $u_{10}$ ,  $v_{10}$ ), the long wave radiation ( $radlw$ ) and short wave radiation ( $radsw$ ) are considered. The temporal variability of the forcing variable is obtained by the Fourier decomposition of the forcing variable vector. The principle is as follows:

Let  $\mathbf{p}$  be the vector of forcing variables of the year 2005,

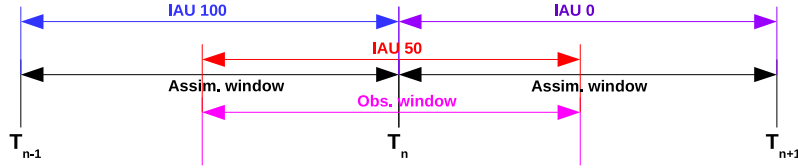
$$\mathbf{p}(x, y, t) = \sum_k a_k(x, y) \exp(i\omega_k t) \quad (1)$$

where  $\omega_k$  is the  $k$ th angular frequency and  $a_k(x, y)$  is the complex field corresponding to the Fourier coefficient of the angular frequency  $\omega_k$ .

With

$$\omega_k = \frac{2\pi k}{\Delta t} k = -\frac{k_{max}}{2}, \dots, \frac{k_{max}}{2} - 1 \quad (2)$$

$\Delta t$  is the time interval of the forcing variables, for example, every 3 h, and  $k_{max}$  is the total number of 3-hourly fields during the time span of the forcing variables under consideration, here one year.



**Fig. 2.** Illustration of different IAU schemes. The intervals in black correspond to assimilation windows.  $T_n$  represents the analysis time under consideration. The interval in pink corresponds to the observation window for the assimilation step  $T_n$ . Suppose that an analysis is performed and an increment is calculated at  $T_n$  using the observations inside the observation window in pink. This increment is then inserted gradually to the model integration inside the increment update window in purple with the IAU 0 scheme, inside the increment update window in red with the IAU 50 scheme, inside the increment update window in blue with the IAU 100 scheme. The intervals in black correspond to assimilation windows. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The perturbation is generated by combining the angular frequencies that we are interested in and a random time series with a temporal decorrelation scale determined by the corresponding angular frequency (Eq. (3)). Different frequency corresponds to different variability. For a realistic ocean circulation model, the ensemble should be representative of the impact of forcing errors on monthly time scale ocean dynamics. The monthly variability is thus taken into account during the perturbation computation.

$$\mathbf{p}_p = \alpha \text{Re} \left( \sum_k a_k(x, y) z_k(t) \right) \quad (3)$$

where  $z_k$  is a complex random time series with a temporal correlation scale of  $T_k = 2\pi/|\omega_k|$ , zero mean and unit variance. The factor  $\alpha$  takes into account that the expected perturbation is in general smaller than the temporal variability. It is determined in order that the perturbation is realistic.

Further details of the principle are explained in Barth et al. (2011) and Marmain et al. (2014). In order to validate the ensemble, the ensemble spread at the end of the ensemble spin up, is compared to the difference between the model prediction without perturbation and the observations.

### 3.3. Assimilation setup

The main difference between the IAU 0, IAU 50 and IAU 100 schemes lies on the increment update window position (Fig. 2). For a given analysis time step, the increment corresponds to the difference between the analysis and the forecast. The increment update window has the same length as the assimilation window (analysis interval or assimilation cycle.), which is the same as in Bloom et al. (1996), Alves et al. (2004) and Ourmières et al. (2006), in order that the correction interval corresponds to the observation interval in the case of the IAU 50 scheme. For the other two schemes, the same increment update window length is kept. Indeed, a too small increment update window will increase the correction quantity at each model integration time, thus induce instability of the model dynamics, while a too large increment update window seems inappropriate, because the model state may change a lot. In the IAU 0 scheme, the increment update window is located on the assimilation window after the analysis time step. Therefore, there is no model integration repeat for each assimilation window. For the IAU 50 scheme, the increment update window is located at half of the assimilation window length before and after the analysis time step. Therefore, for each assimilation window, 50% of the model integration is repeated. Regarding the IAU 100 scheme, the increment update window is located on the assimilation window before the analysis time step. For each assimilation window, 100% of the model integration is repeated. More detailed illustrations of these three IAU schemes can be found in Yan et al. (2014).

The assimilation cycle is 10 days. It is determined based on the fact that it must be long enough to accumulate a sufficient amount of observations to correct the model state accordingly. The 10-day interval corresponds to the characteristic time scale of the ARGO

data collection (Roemmich and the Argo Steering Team, 2009) and the Jason-1 SSH data cycle. Observations 5 days before and after each analysis step are used. The first 180 days (from 1st January to 29th June) in 2005 are ensemble spin up time. It is important to integrate the ensemble over a time interval covering a few characteristic time scales of the dynamical system to ensure dynamic stability and to correct multivariate correlations before beginning the assimilation. Afterwards, on one hand, the free ensemble run is performed over the last 180 days in order to compare to the assimilation. On the other hand, the EnKF is activated during the last 180 days in order to assimilate the observations into the model.

The assimilation tool used is the Ocean Assimilation Kit (OAK) (a suite of data assimilation software developed by the GHER group at University of Liège and distributed under GPL license) (Vandenbulcke et al., 2006; Barth et al., 2007; 2008; Vandenbulcke et al., 2010; Yan et al., 2015). The assimilation method provided in OAK is the square root analysis scheme of EnKF (Evensen, 2004). The model state vector for assimilation consists of 3 variables: SSH, temperature and salinity. The model state here is referred to the assimilation tool, and different from the prognostic variables of the ocean circulation model mentioned in Section 2. Only the temperature and salinity increments are incorporated in the model integration, which is different from Yan et al. (2014) where all the prognostic variables are observed. The main idea underlying this strategy is to correct the slow processes (baroclinic), not the fast processes (barotropic), because only the former can be effectively controlled given the time coverage of the observations. The zonal and meridional velocities include both slow and fast processes, thus they cannot be used directly. An alternative would be to recompute the velocity increments from thermohaline increments, but this would not help much to improve the method.

## 4. Metrics

For an objective evaluation of the assimilation results, we rely on both deterministic and probabilistic metrics. For deterministic validation, firstly, the root-mean-square (RMS) errors of the ensemble mean of SSH, SST, temperature and salinity profiles are calculated against different sources of observations. The ensemble mean/spread versus observations plots are also performed. Secondly, the horizontal (zonal and meridional) velocity is assessed by comparison to the OSCAR (Ocean Surface Current Analysis Real-time) third degree sea surface velocity (Bonjean and Lagerloef, 2002).

For probabilistic validation, the performance of the ensemble forecast system is diagnosed according to reliability and resolution. The CRPS measures the distance between the ensemble distribution and the verifying observations. According to Hersbach (2000) and Candille et al. (2006), the CRPS can be decomposed into CRPS\_Reli and CRPS\_pot: where  $\text{CRPS} = \text{CRPS\_Reli} + \text{CRPS\_pot}$ . CRPS\_Reli measures the reliability of an ensemble system, while CRPS\_pot provides information on the resolution. A detailed illustration of the computation of CRPS and its decomposition is given

in Appendix A. CRPS and its decomposition are negatively oriented. The smaller they are, the better an ensemble is. An ensemble system with CRPS value of 0 always exactly reproduces the verifying observation without any ensemble spread. CRPS\_Reli is equal to 0 if the system is perfectly reliable. A significant positive value of CRPS\_Reli indicates the lack of reliability. CRPS\_pot reaches its minimum for a perfect deterministic system and positive values quantify a lack of resolution (Candille et al., 2006).

The RCRV score provides further insight into the reliability of an ensemble system. Its definition is given in Eq. (4). It allows for the decomposition of the reliability into bias and dispersion. The average of RCRV, referred to as RCRV\_bias, is computed over all realisations of the system and represents the weighted bias between the ensemble and the observation. The standard deviation of RCRV, referred to as RCRV\_dispersion, constitutes an indicator of systematic over and under dispersion of the ensemble. It measures the agreement of the ensemble spread and the specified observational error with the observed amplitude of the forecast error. A perfectly reliable system has no bias and a dispersion equal to 1. A significant negative (positive) value of bias indicates a positive (negative) bias. A value of dispersion significantly larger (smaller) than 1 characterises the under-dispersion (over-dispersion) of the system.

$$\text{RCRV} = \frac{y_o - \bar{x}}{\sqrt{\sigma_o^2 + \sigma^2}} \quad (4)$$

where  $y_o$  is the observation,  $\sigma_o$  represents the observation error,  $\bar{x}$  corresponds to the ensemble mean and  $\sigma$  denotes the ensemble spread.

## 5. Comparison of results issued from different IAU schemes

In this section, the assimilation results are analysed and discussed. Comparison is performed between the three IAU schemes according to both deterministic and probabilistic metrics mentioned previously in Section 4. Thermohaline variables (temperature and salinity) and dynamical variables (SSH, horizontal velocity) are considered. Among these variables, SSH, SST and temperature profile are assimilated variables, while salinity profile and horizontal velocity are unassimilated variables. First, the ensemble mean and the ensemble spread of SSH, SST, temperature and salinity profiles are compared to the observations of validation. For horizontal velocity, the analyses are focused on the Gulf Stream region. The surface horizontal velocity is assessed by comparison to the OSCAR third degree sea surface velocity (Bonjean and Lagerloef, 2002). Second, probabilistic scores such as CRPS and RCRV are computed for SSH, SST, temperature and salinity profiles, in complementary to the deterministic metrics.

Note that, first, for SSH and salinity profile, ENVISAT altimetric data and ARGO profiles are used for validation. For SST, Mercator reanalysis data are used for validation. While for temperature profile, assimilated observations are used since no other sources of observations are available for validation. The comparison thus measures how efficiently these increment update strategies force the model towards the observations. Second, the scores for temperature and salinity are calculated without taking into account the volume (surface  $\times$  depth) represented by the model grid point, because the degradation of salinity by assimilation is mainly located at the surface and its contribution will be very small if the volume represented by the point is taken into account. Third, both deterministic and probabilistic scores shown in this section are calculated from the analyses averaged over the whole North Atlantic basin unless otherwise indicated. For temperature and salinity profiles, the depth in consideration is mainly above 2030 m (limited by the depth of the available observations), with exceptions for some temperature points located at more than 4500 m depth.

### 5.1. Deterministic validation

The spatially averaged RMS errors of the ensemble means of SSH, SST, temperature profile and salinity profile in the free run and with the three IAU schemes are shown in Fig. 3. The RMS errors are reduced by the three IAU schemes for the four variables, compared to those of the free run. Almost no difference is observed between the IAU 50 scheme and the IAU 100 scheme for the four variables. For SST and temperature profile, very small difference is observed between the IAU 0 scheme and the IAU 50/100 schemes. Larger difference between these schemes exists for SSH and salinity profile. For SSH, the RMS errors of the IAU 0 scheme are slightly larger than those of the IAU 50/100 schemes before the 9th step. Detailed inspection shows that larger RMS error with the IAU 0 scheme is mainly located in the subpolar area where no Jason-1 observations are available for assimilation. Due to errors present in the error covariance matrix and the lack of SSH observations, the SSH correction in the subpolar area is inefficient. With the IAU 0 scheme, sharper discontinuities exist between the SSH values in this area and those in the neighbouring areas. While with the IAU 50/100 schemes, a more smoothed SSH field is obtained. As a result, the stronger smoothing effect of the IAU 50/100 schemes can also help mitigating spurious corrections. For salinity profile, the improvement in RMS reduction of the IAU 50/100 schemes with respect to the IAU 0 scheme is observed from the 9th step and it becomes more and more pronounced towards the end of the assimilation experiments. According to detailed investigation, the improvement is mainly located in the equator and subpolar area in the upper layers. In the Gulf Stream region, the improvement is not so significant and it is position dependent.

The ensemble mean and the ensemble spread are compared to the observations and the associated observation errors for salinity profile, since larger difference is observed for this variable. According to Fig. 4, negative bias exists in the free run and this bias is reduced with the three IAU schemes, which corresponds to the RMS reduction with the three IAU schemes. The distances between the ensemble means and the observations with the IAU 50/100 schemes are smaller than those with the IAU 0 scheme, especially after the 5th step, which explains the more efficient RMS reduction of the IAU 50/100 schemes observed in Fig. 3(d). Furthermore, a significant difference between the IAU 0 scheme and the IAU 50/100 schemes lies on the ensemble spread: the ensemble spread is much smaller with the IAU 0 scheme. In order to further understand the degradation of the efficiency of salinity correction with the IAU 0 scheme, the covariance between SST and a surface salinity point (located at 25.97W, 43.80N) at three different analysis steps with the IAU 0 and IAU 50 schemes is analysed (Fig. 5). The covariance is only shown between the IAU 0 scheme and the IAU 50 scheme, because almost no difference is observed between the IAU 50 scheme and the IAU 100 scheme. The SST and sea surface salinity (SSS) covariance chosen should have positive value as shown in Fig. 5(a), (b), (d) and (f). With the IAU 0 scheme, the covariance becomes close to zero from the third analysis step, which is not appropriate. Thus, the efficiency of the salinity correction by SST observations is significantly degraded. Moreover, Fig. 6 shows the ensemble mean and the ensemble spread of salinity averaged over the whole domain before analysis at each step with the IAU 0 and IAU 50 schemes. The ensemble spread is always much smaller with the IAU 0 scheme compared to that with the IAU 50 scheme, which implies that with the IAU 0 scheme, the salinity is not appropriately adjusted before each analysis due to lack of free model integration.

Fig. 7 shows the OSCAR third degree sea surface velocity and the surface horizontal velocities in the Gulf Stream region on 6th November, 2005, obtained in the free run and with three IAU schemes. Compared to the free run, the main jet is intensified

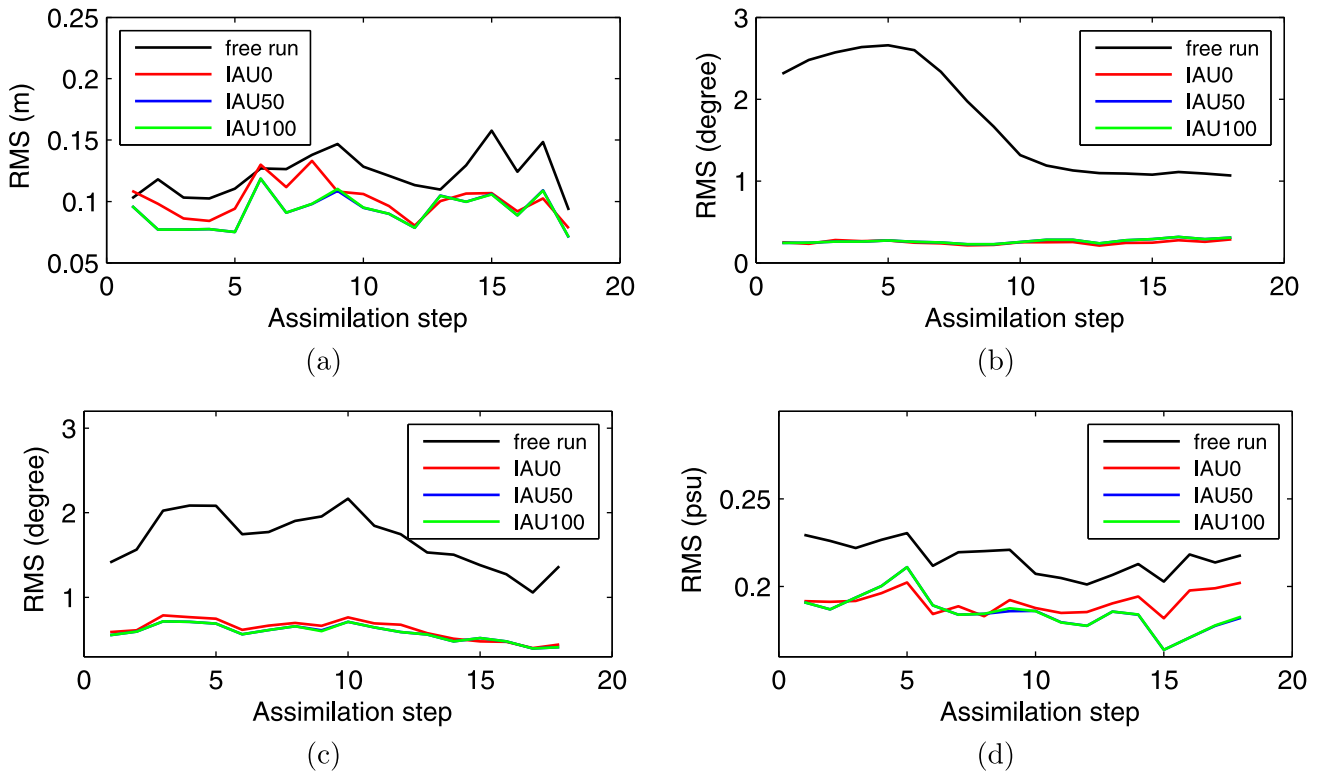


Fig. 3. Spatially averaged RMS error of (a) SSH (b) SST (c) temperature profile (d) salinity profile with different IAU schemes.

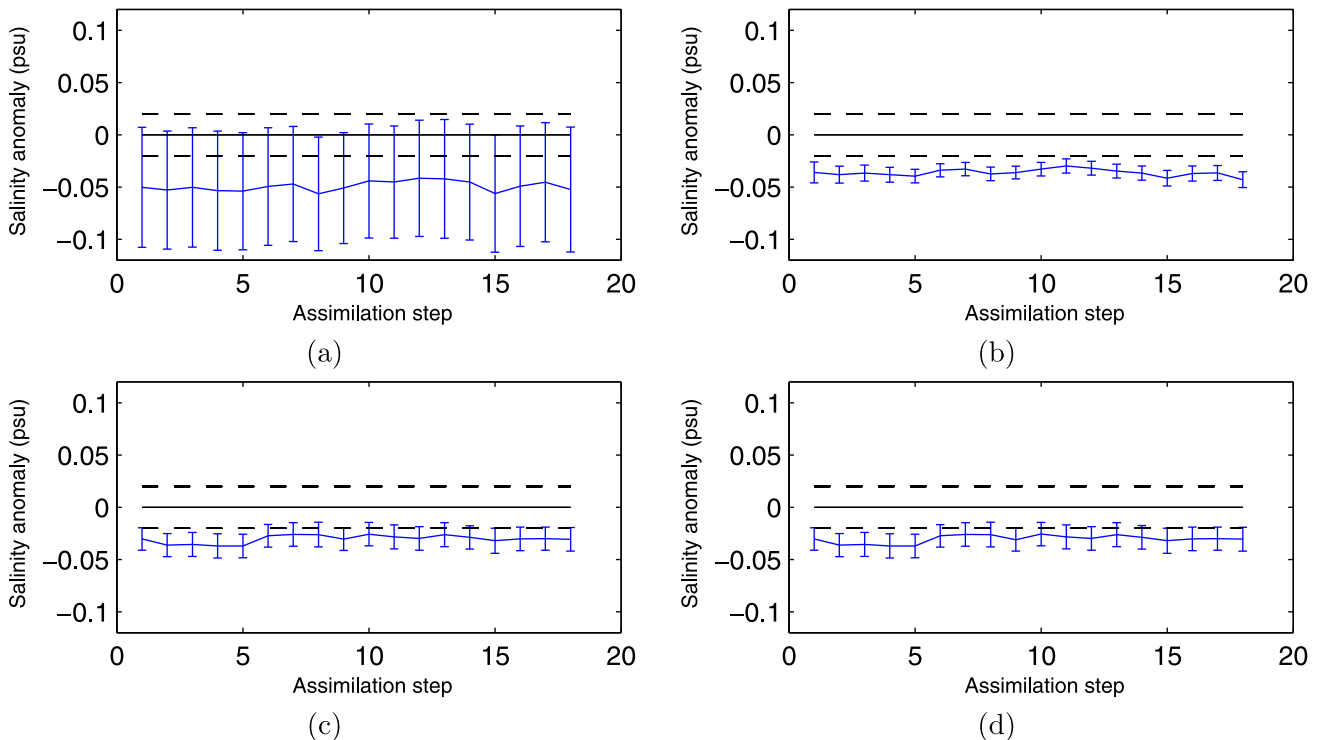
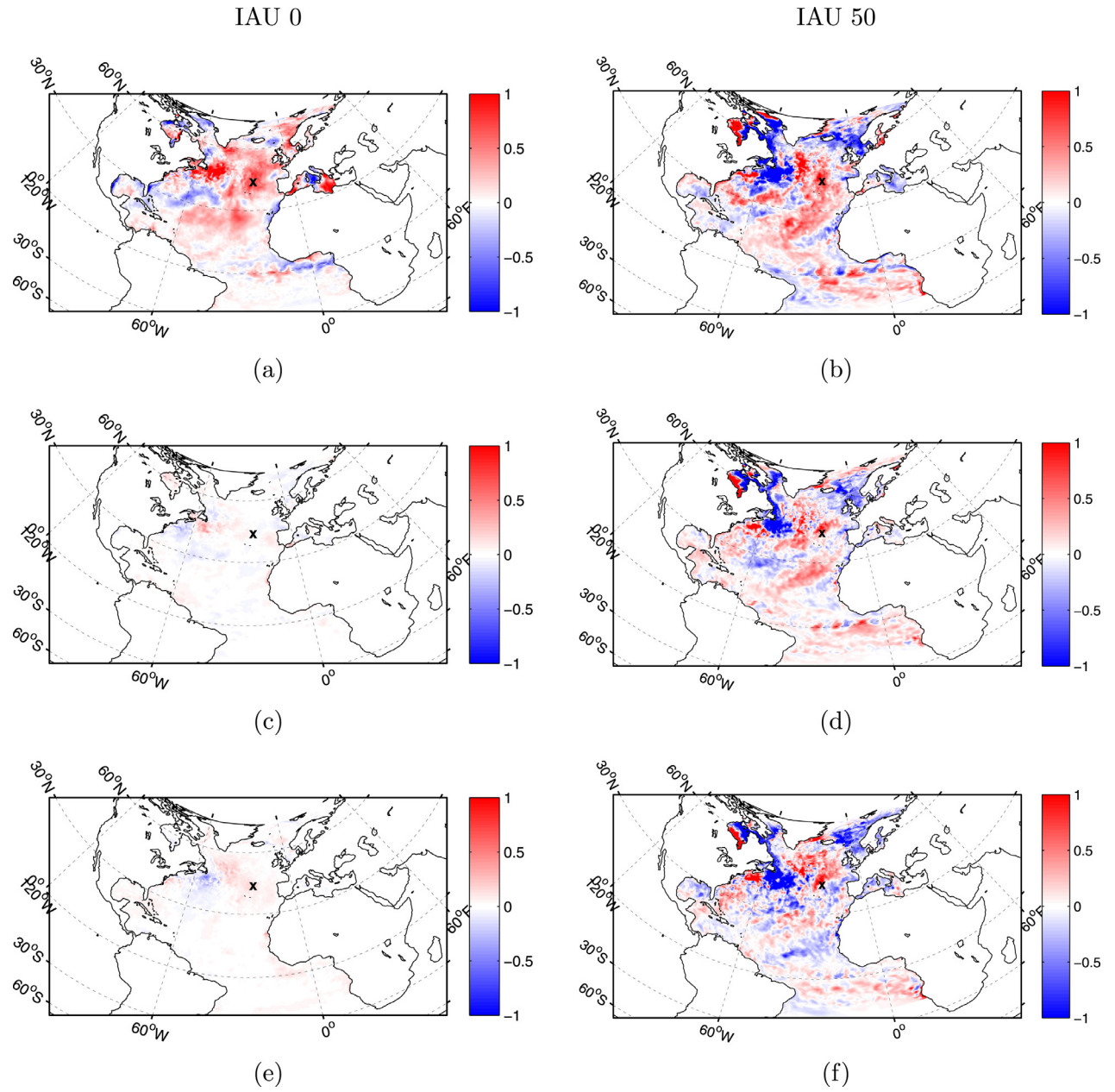


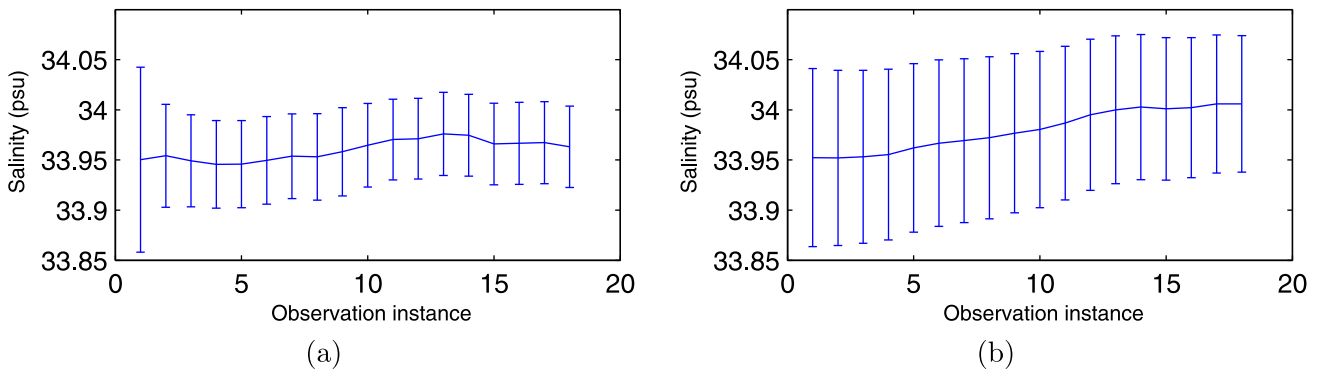
Fig. 4. Salinity anomaly (ensemble mean - observation) versus ensemble spread (a) free run (b) IAU 0 (c) IAU 50 (d) IAU 100. The black line corresponds to the observation and the dashed line corresponds to the observation error.

more or less with the three IAU schemes. Almost no difference is observed between the IAU 50 scheme and the IAU 100 scheme. However, the Gulf Stream path with the IAU 50/100 schemes is much more diffuse than that with the IAU 0 scheme. Note that, the ensemble averaging can lead to smooth velocity field. Further verification shows that even for individual ensemble member,

stronger eddy activities are observed with the IAU 0 scheme. The surface horizontal velocity field obtained with the IAU 0 scheme is much more similar to the OSCAR sea surface velocity: the main jet is sufficiently intensified. More eddy activities and meanders are generated between 65°W and 40°W. The IAU 0 scheme outperforms significantly the IAU 50/100 schemes.



**Fig. 5.** Covariance between SST and a salinity point (25.97W,43.80N) over time (a), (b) 29th July (the third assimilation step) (c), (d) 17th September (the 8th assimilation step) (e), (f) 6th November (the 13th assimilation step) with the IAU 0 ((a), (c), (e)) and IAU 50 ((b), (d), (f)) schemes.



**Fig. 6.** Ensemble mean and spread of salinity before each analysis step averaged over the whole domain with the (a) IAU 0 (b) IAU 50 schemes. The ensemble spread is represented by error bar.

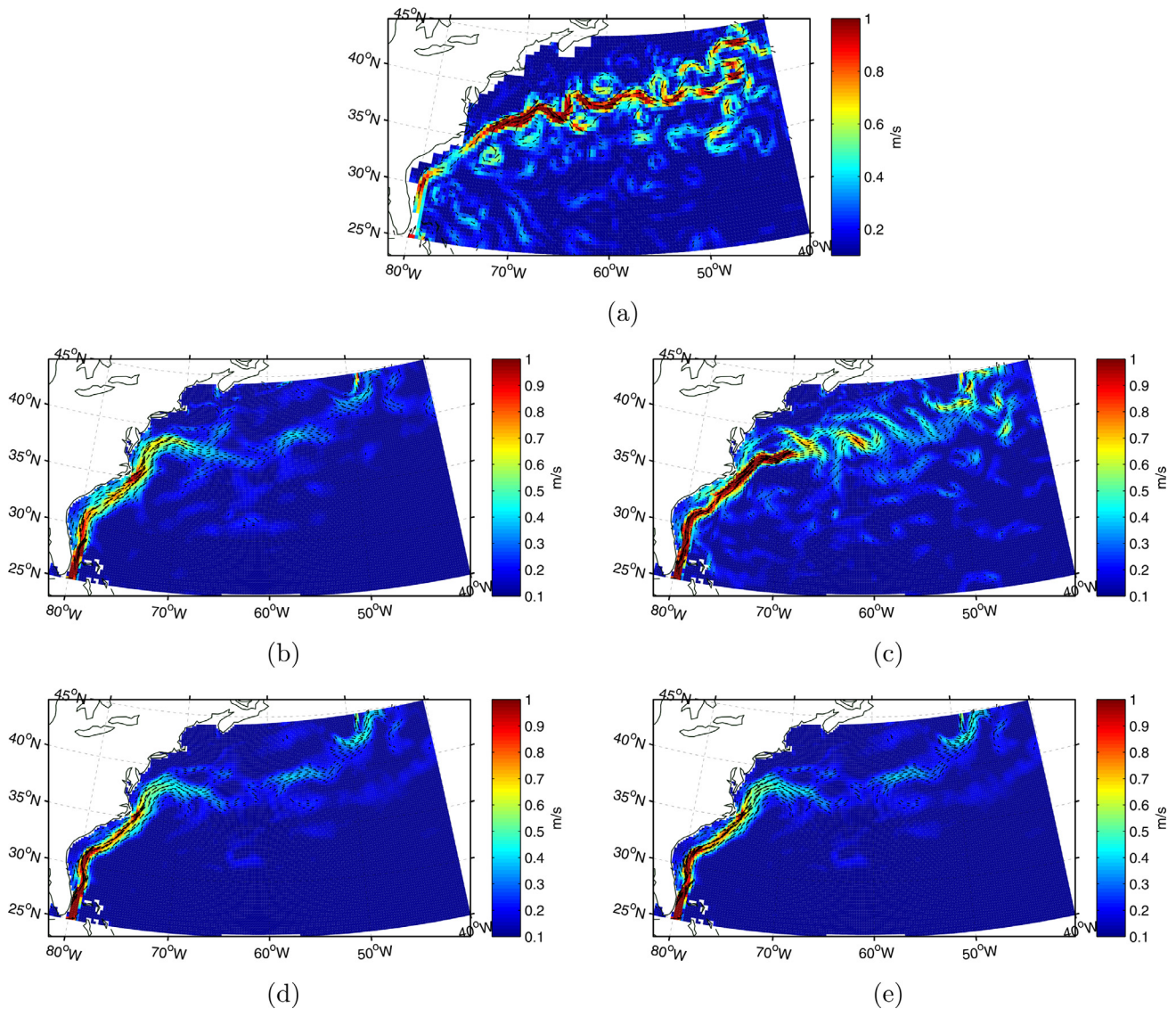


Fig. 7. Surface horizontal velocity in the Gulf Stream region on 6 November 2005 (a) OSCAR third degree sea surface velocity (b) free run (c) IAU 0 (d) IAU 50 (e) IAU 100.

**Table 1**

RMS errors (m/s) of zonal and meridional velocities obtained in the assimilation experiments against the OSCAR surface current velocity.

	Free run	IAU 0	IAU 50	IAU 100
Zonal velocity (u)	0.2176	0.1969	0.2162	0.2162
Meridional velocity (v)	0.1941	0.1952	0.1913	0.1913

**Table 2**

Standard deviation (m/s) of the OSCAR surface current velocity field and surface current velocity fields obtained in the assimilation experiments.

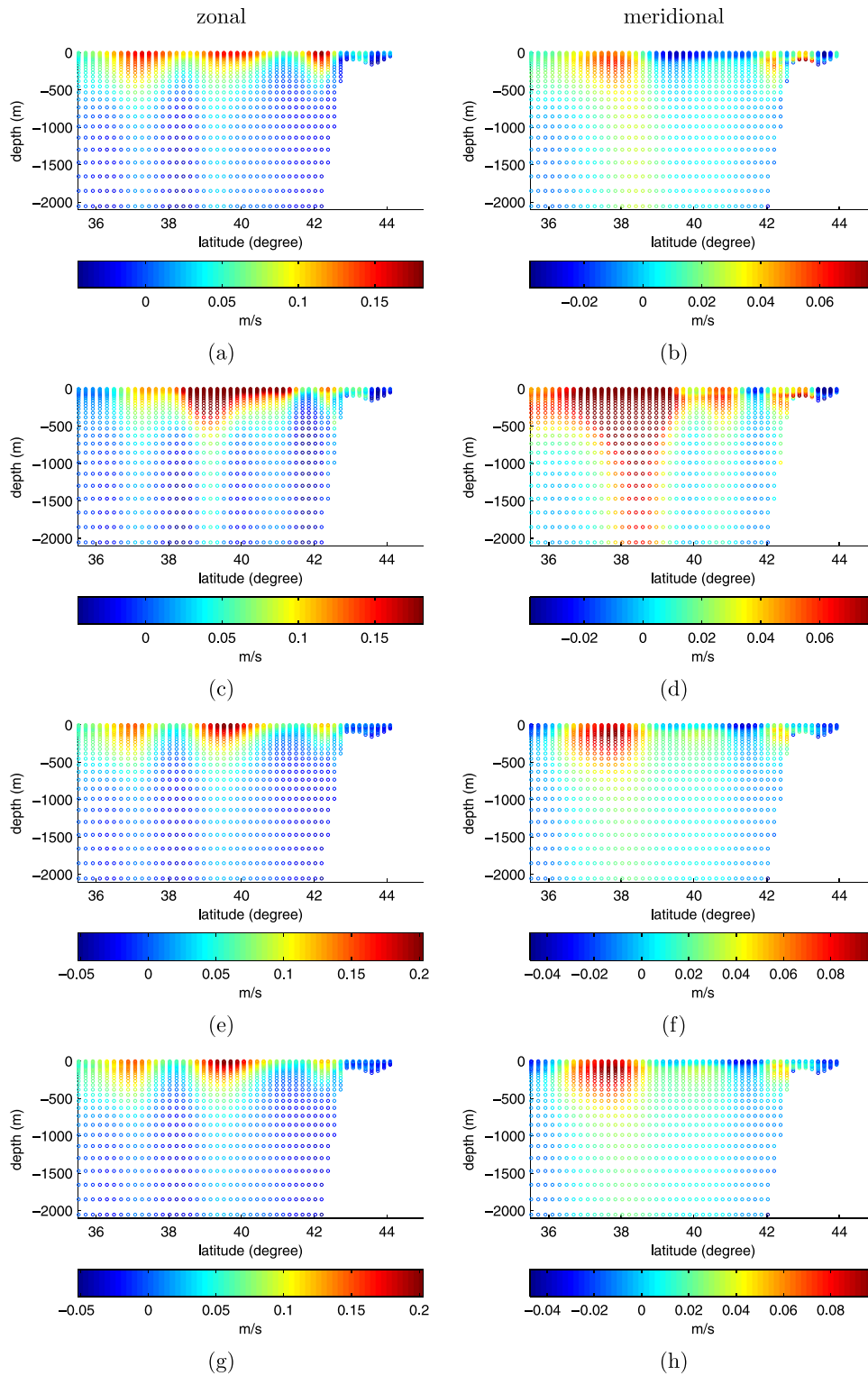
	OSCAR	Free run	IAU 0	IAU 50	IAU 100
Zonal velocity (u)	0.2262	0.1348	0.1818	0.1334	0.1333
Meridional velocity (v)	0.1843	0.1226	0.1776	0.1100	0.1099

Besides the qualitative comparison, the RMS errors of the zonal and meridional velocities obtained with the three IAU schemes against the OSCAR third degree sea surface velocity are present in (Table 1). For the zonal velocity, the RMS error of the IAU 0 scheme is much smaller than those of the free run and of the IAU 50/100 schemes. For the meridional velocity, the RMS error of the IAU 0 scheme is slightly larger than those of the IAU 50/100 schemes. Taken both the zonal and meridional velocities into account, the RMS errors of the IAU 0 scheme are smaller than those of the IAU 50/100 schemes, which is consistent with the observation in Fig. 7. Note also that, the resolution of the surface velocity field obtained with the three IAU schemes is better than

that of the OSCAR velocity. The RMS error of a high resolution model is sometimes larger than that of a low resolution model, but the variability of the high resolution model is often more realistic. For this reason, the variance of the velocity field seems useful for comparison. The standard deviations of the OSCAR velocity field and those of the velocities obtained with the three IAU schemes are shown in Table 2. Obviously, the variability of the velocity field with the IAU 0 scheme is closer to that of the OSCAR velocity. The variability of the IAU 50/100 schemes is much smaller.

In order to analyse the impact of the assimilation on zonal and meridional velocities at depth, a vertical section in the Gulf Stream





**Fig. 8.** Meridional section in the Gulf Stream region (at 64°W) for zonal and meridional velocities with different IAU schemes. (a), (b) free run (c), (d) IAU 0 (e), (f) IAU 50 (g), (h) IAU 100.

region (at 64°W) is performed. The averaged zonal and meridional velocities over 6 months in the free run and with the three IAU schemes are shown in Fig. 8. Consistent with the observations of the surface horizontal velocity, the zonal and meridional velocities at depth are not quite modified by the IAU 50/100 schemes. With these two schemes, the assimilation only intensifies the current at some positions between the surface and 500 m at depth. While with the IAU 0 scheme, the deeper zonal and meridional veloc-

ities are likely to be more influenced by assimilation, which is consistent with the larger vertical velocity and vertical diffusivity observed with this scheme both in the present paper and in Yan et al. (2014). Not only the current intensity, but also the current direction are changed by assimilation, which results of a surface horizontal velocity field very similar to the OSCAR velocity field as shown in Fig. 7. The IAU 0 scheme thus modifies consistently the three-dimensional flow.

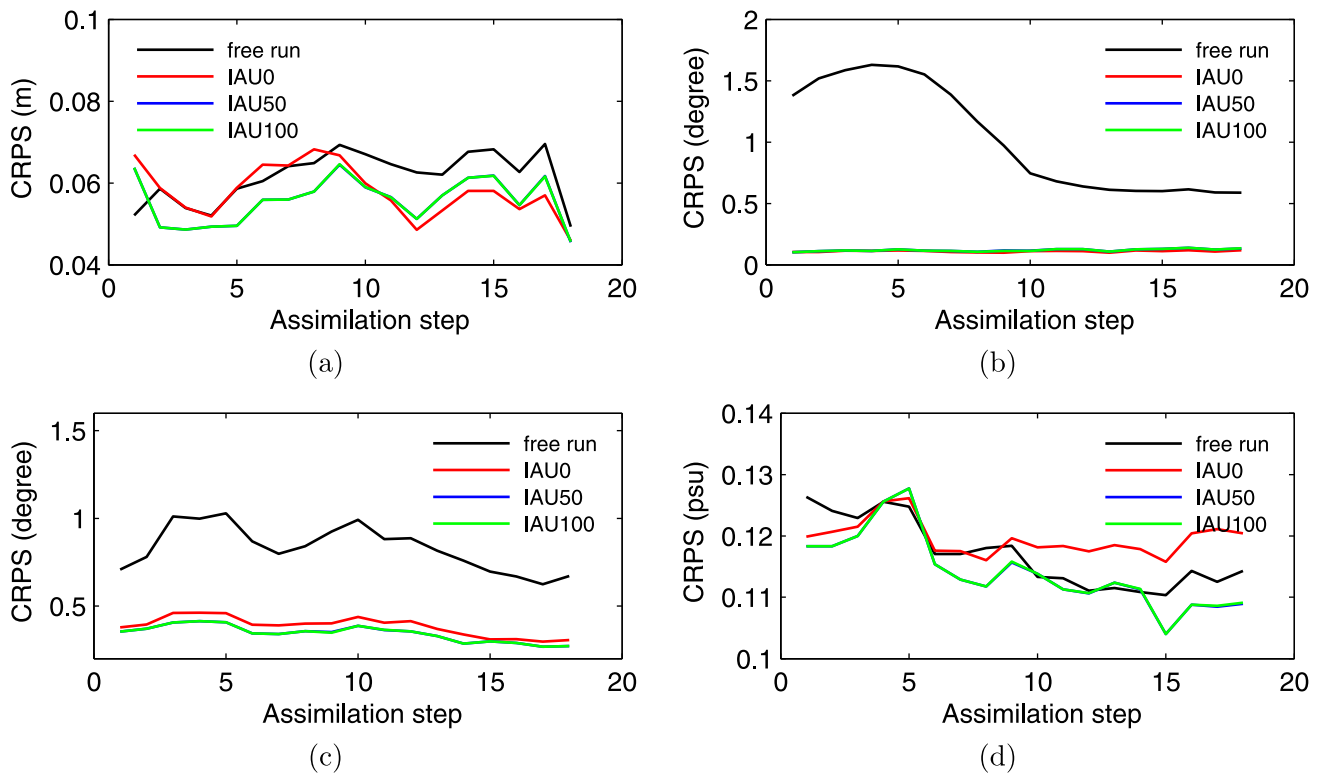


Fig. 9. CRPS of (a) SSH (b) SST (c) temperature profile (d) salinity profile with different IAU schemes.

This significant difference between the IAU 0 scheme and the IAU 50/100 schemes can be related to 1) the instabilities (e.g. density inversion, large vertical velocity, large vertical diffusivity, etc.) induced by the increment update 2) the model rerun phase, especially the free model run. For the former, it has been discussed in Yan et al. (2014) and is confirmed again in this paper that the perturbation of the equilibrium model state induced by the increment update is more significant with the IAU 0 scheme. Larger vertical velocity and stronger vertical diffusivity are obtained with the IAU 0 scheme, which results of modifications of the current velocities at depth and generations of more eddy activities. For the latter, in a dynamically active region, sharper features (related to barotropic and/or baroclinic instabilities) and strong gradients are present. It is thus not relevant to rerun the model as with the IAU 50/100 schemes, because longer model integration time has smoothing effects on the strong gradients and does not allow preserving sharper features.

## 5.2. Probabilistic validation

The CRPS for SSH, SST, temperature profile and salinity profile in the free run and with the three IAU schemes are shown in Fig. 9. The conclusion is similar to that of the RMS error (Fig. 3). Almost no difference is observed between the IAU 50 and IAU 100 schemes. Moreover, almost no difference is observed between the three IAU schemes for SST. Large difference between the IAU 0 scheme and the IAU 50/100 schemes exists for salinity profile and SSH before the 9th step. For SSH, before the 9th step, the CRPS of the IAU 0 scheme is larger than those of the IAU 50/100 schemes. From the 9th step, an improvement is observed for the IAU 0 scheme. The CRPS values are even slightly smaller than those of the IAU 50/100 schemes, which is not observed from the RMS error. Detailed inspection shows that at this step, the ensemble spread in the forecast is significantly increased and after the analysis the ensemble mean is closer to the observation

and the ensemble spread is larger with the IAU 0 scheme. For temperature profile, the difference between the IAU 0 scheme and the IAU 50/100 schemes still remains small. However, compared to the RMS difference, it seems that the CRPS difference is more pronounced, which implies that the difference between the IAU 0 scheme and the IAU 50/100 schemes mainly lies on the ensemble spread. For salinity profile, large difference is observed from the 5th step. The CRPS values with the IAU 0 scheme are even larger than those of the free run and this significant degradation is not observed from the deterministic validation. With the IAU 0 scheme, the ensemble spread is already smaller in the forecast. After the analysis, the distance between the ensemble mean and the observations is not sufficiently reduced, but the ensemble spread is further reduced, which results of larger distance between the ensemble distribution and the observation distribution.

The CRPS\_Reli for SSH, SST, temperature profile and salinity profile in the free run and with the three IAU schemes are given in Fig. 10. For SST, there is no difference observed between the three IAU schemes. A small value of CRPS\_Reli, close to 0, is observed for the three IAU schemes, which indicates the good reliability of the ensemble distributions obtained with the three IAU schemes. For SSH, temperature and salinity profiles, a degradation of the reliability with respect to the free run is observed for the three IAU schemes, which is not observed from the RMS errors. For SSH, the CRPS\_Reli of the IAU 0 scheme becomes smaller than those of the IAU 50/100 schemes from the 9th step, with an improvement on both ensemble mean and ensemble spread. For temperature and salinity profiles, the CRPS\_Reli of the IAU 0 scheme is larger than those of the IAU 50/100 schemes along the assimilation experiment. For temperature profile, the larger value of CRPS\_Reli of the IAU 0 scheme is mainly due to smaller ensemble spread compared to those of the IAU 50/100 schemes. For salinity profile, the CRPS\_Reli is consistent with the ensemble mean/spread versus observation plot (Fig. 4): because of larger ensemble spread and the closeness between the ensemble mean and the observations,

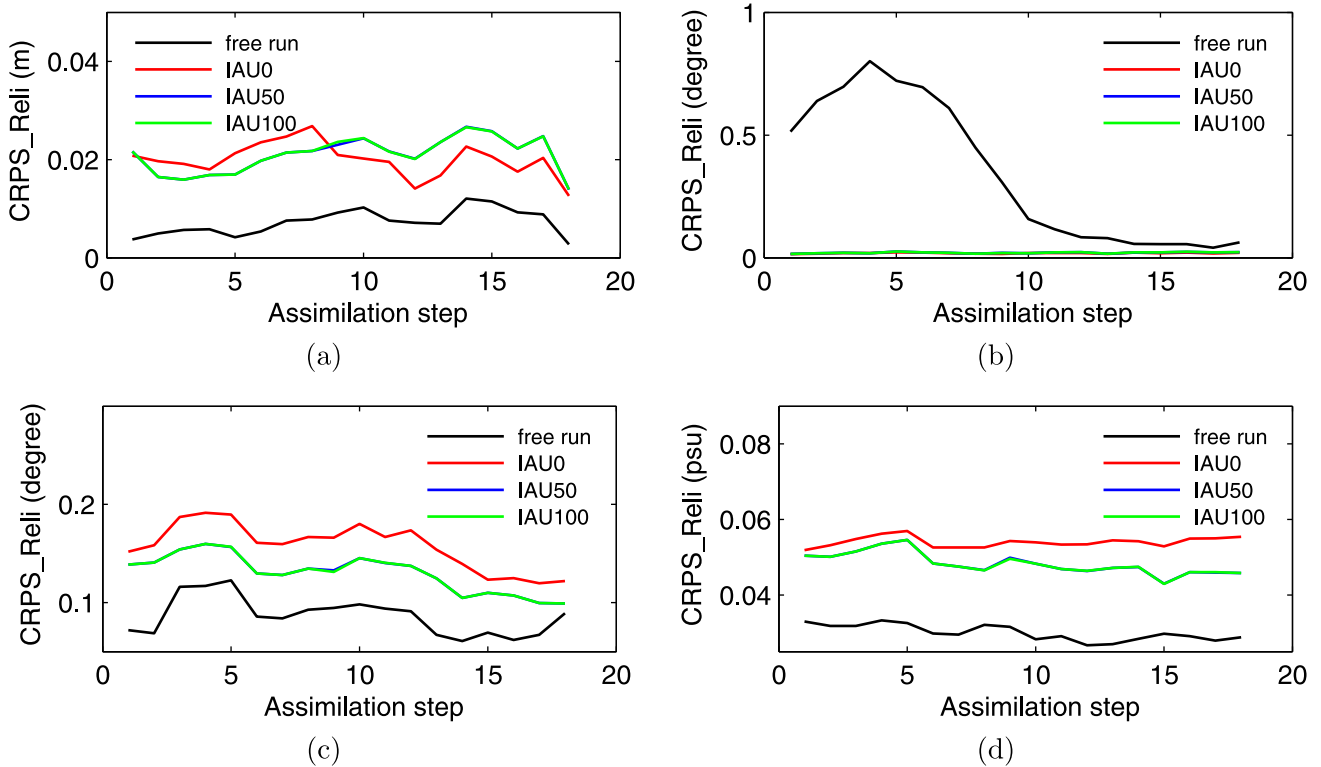


Fig. 10. CRPS\_Reli of (a) SSH (b) SST (c) temperature profile (d) salinity profile with different IAU schemes.

the reliability of the ensemble is much better with the IAU 50/100 schemes.

The CRPS\_pot for SSH, SST, temperature profile and salinity profile in the free run and with the three IAU schemes are given in Fig. 11. For the four variables, the resolution of the ensemble is significantly improved by assimilation with the three IAU schemes compared to the free run. For SST and temperature profile, almost no difference is observed between the three IAU schemes. For SSH, the IAU 50/100 schemes outperform slightly the IAU 0 scheme, while for salinity profile, the IAU 0 scheme outperforms slightly the IAU 50/100 schemes, because smaller ensemble spread is obtained for SSH with the IAU 50/100 schemes and for salinity with the IAU 0 scheme. Therefore, it can be concluded that the difference of the impact on the resolution of the ensemble by the three IAU schemes is negligible. (Note that the resolution makes sense only if the reliability is good, and it is not the case for salinity.)

Since the difference between the three IAU schemes mainly lies on the reliability. The reliability of the ensemble is further investigated by RCRV score. The RCRV scores for SSH, SST, temperature profile and salinity profile in the free run and with the three IAU schemes are presented in Fig. 12. For SSH, before the 9th step, the RCRV\_bias of the IAU 0 scheme is larger than those of the IAU 50/100 schemes. Moreover, the ensemble under-dispersion problem with the IAU 0 scheme is more serious than that of the IAU 50/100 schemes. However, after the 9th step, the RCRV\_bias of the IAU 0 scheme becomes smaller than those of the IAU 50/100 schemes and the ensemble spread is slightly larger than those of the IAU 50/100 schemes. These observations confirm the improvement on both ensemble mean and ensemble spread after the 9th step. For SST, the difference of the RCRV\_bias between the IAU 0 scheme and the IAU 50/100 schemes is very small. From the 10th step, the ensemble spread is slightly under-dispersive with the IAU 50/100 schemes. Note that this small degradation of the IAU 50/100 schemes cannot be observed from other scores mentioned previously. For temperature profile, the

difference of the RCRV\_bias between the IAU 0 scheme and the IAU 50/100 schemes is very small. The under-dispersion problem is more pronounced with the IAU 0 scheme, which explains the CRPS and CRPS\_Reli differences observed in Figs. 9(c) and 10(c). For salinity profile, significant difference exists between the IAU 0 scheme and the IAU 50/100 schemes for both RCRV\_bias and RCRV\_dispersion. Large negative bias and serious ensemble under-dispersion problem exists with the IAU 0 scheme.

## 6. Discussions

SSH, SST, temperature profile, salinity profile and horizontal velocity can represent different situations with different characteristics.

In case with SST, a large number of observations, with appropriate error specification (with spatial variation) are available for assimilation. The SST increment is calculated from the direct correction by observation and then applied to the model integration. According to the results obtained, an almost perfectly reliable ensemble system is obtained with either of the three IAU schemes. In this case, the impact of the IAU scheme is very small, even negligible. However, taken into account the computational cost, the IAU 0 scheme is preferred. In this case, the conclusion is in accord with the conclusions obtained in Yan et al. (2014) where the system is well observed in a twin experiment.

In case with horizontal velocity, no velocity increments are applied to the model integration. The correction mainly depends on the geostrophic adjustment following the integration of the temperature and salinity increments. Hence, the quality of the horizontal velocity correction can reveal how the observed thermohaline variables and non-observed dynamical variables are balanced in the assimilation scheme. In the Gulf Stream region, strong model dynamic is present. The IAU 50/100 schemes seem to have weakened strong gradients and cannot preserve sharper features due to the rerun phase allowing longer model integration

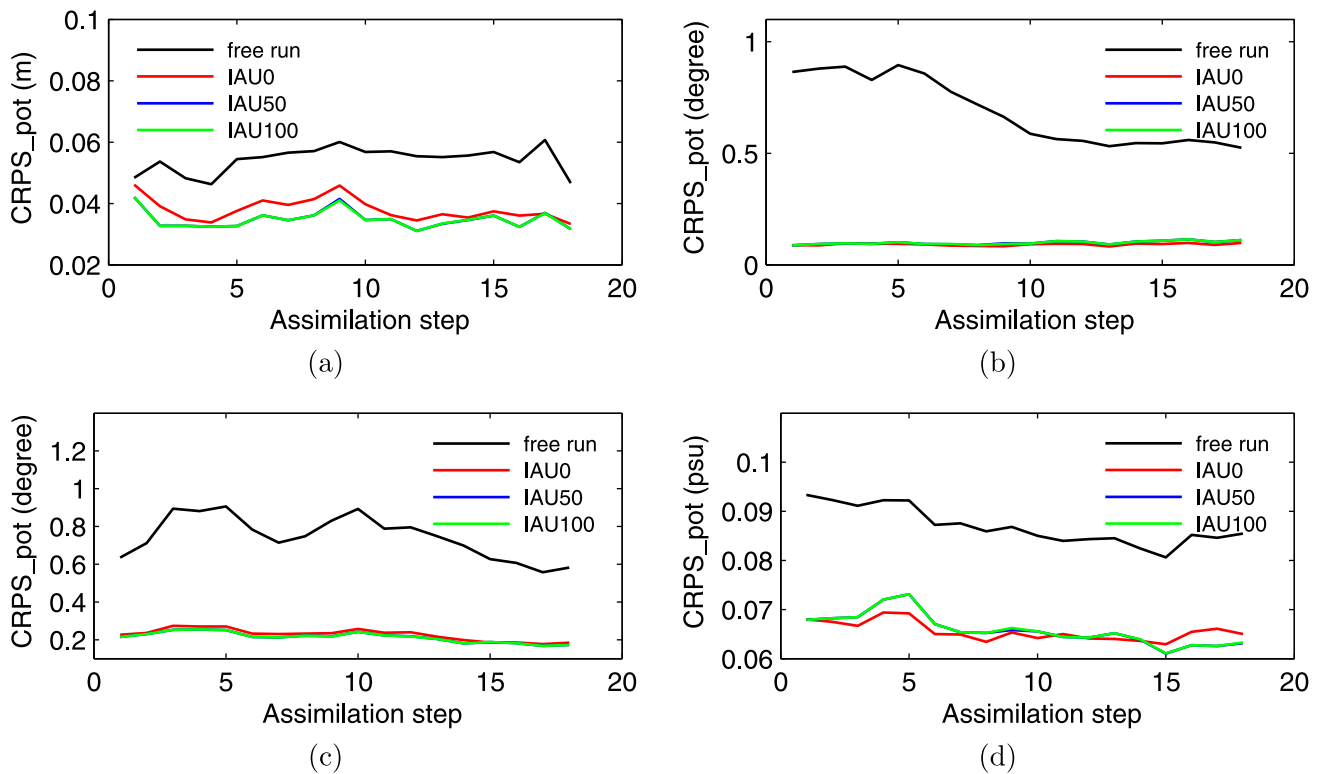


Fig. 11. CRPS\_pot of (a) SSH (b) SST (c) temperature profile (d) salinity profile with different IAU schemes.

time. The spatial smoothing filtered on the solution thus appears more important with these two schemes. The eddy activities and meanders are so unstable in this area, it is thus not relevant to rerun the model. On the other hand, larger vertical velocity and stronger vertical diffusivity, allowing for better turbulent mixing, are generated with the IAU 0 scheme. The strong jet velocities reach deeper layers with the IAU 0 scheme. The horizontal velocity field generated with the IAU 0 scheme is very similar to independent observations. In this case, the IAU 0 scheme outperforms the IAU 50/100 schemes.

In case with salinity, no salinity observations are assimilated, but the salinity increments are calculated based on the error covariance matrix and later applied to the model integration. With the IAU 0 scheme, the relative shorter model integration time, especially the lack of free model integration, compared to the IAU 50/100 schemes does not allow for appropriate salinity adjustment, which makes the salinity correction inefficient and serious ensemble under-dispersion problem is present. Therefore, a very large difference is observed for this variable between the IAU 0 scheme and the IAU 50/100 schemes.

In case with SSH, SSH observations are assimilated, but no SSH increment is applied to the model integration. The correction of SSH depends on the dynamical coupling between the thermohaline and dynamical variables. The IAU 0 scheme outperforms the IAU 50/100 schemes from the 9th step, at which step the ensemble spread is significantly increased during the model integration so that the SSH correction becomes efficient in the analysis. Indeed, the ensemble spread is larger with the IAU 0 scheme before each analysis. Differences between the IAU 0 scheme and the IAU 50/100 schemes are mainly located in the Gulf Stream region and in the subpolar area. In the former area, strong model dynamic is present. The Gulf Stream is more developed towards the East with the IAU 0 scheme. The strong gradients are smoothed and sharper features are not preserved with the IAU 50/100 schemes. In the latter area, because of stronger smoothing effect,

the spurious corrections, due to lack of observations and errors present in the error covariance matrix, are mitigated with the IAU 50/100 schemes.

In case with temperature, observations are available for assimilation and the temperature increment is applied to the model integration. Moreover, ensemble under-dispersion problem exists before the assimilation experiments. Even though the ensemble under-dispersion problem is amplified by assimilation with the three IAU schemes, the under-dispersion with the IAU 50/100 schemes is not as significant as with the IAU 0 scheme. Similar to the salinity, the longer model integration time, especially the free model integration time with the IAU 50/100 schemes allows for better temperature adjustment. The ensemble spread is larger before and after each analysis with these two schemes. Therefore, for this variable, the IAU 50/100 schemes outperform slightly the IAU 0 scheme.

In these four cases without sufficient observations nor direct correction of model variables, it seems that no IAU scheme can be qualified as the best one that outperforms the other IAU schemes for all the model variables. The relative performance between the IAU 0 scheme and the IAU 50/100 schemes is different from the conclusion obtained in Yan et al. (2014). To explain this, besides the differences in observations (synthetic/real) and model dynamics (idealised/realistic), the difference in model state vector setup for assimilation also contributes significantly. In Yan et al. (2014), all the prognostic variables are observed, while in the present paper part of the prognostic variables are observed. The extrapolation over undetermined variables and the update of model state in physical space via balance relationship play a more important role in the present paper, which causes the different performance of each IAU scheme. Moreover, between the IAU 50 scheme and the IAU 100 scheme, globally almost no difference is observed for all the considered variables. However, differences do exist at local scale, e.g. the model state field obtained with the IAU 50 scheme is not completely the same as that obtained with the IAU 100

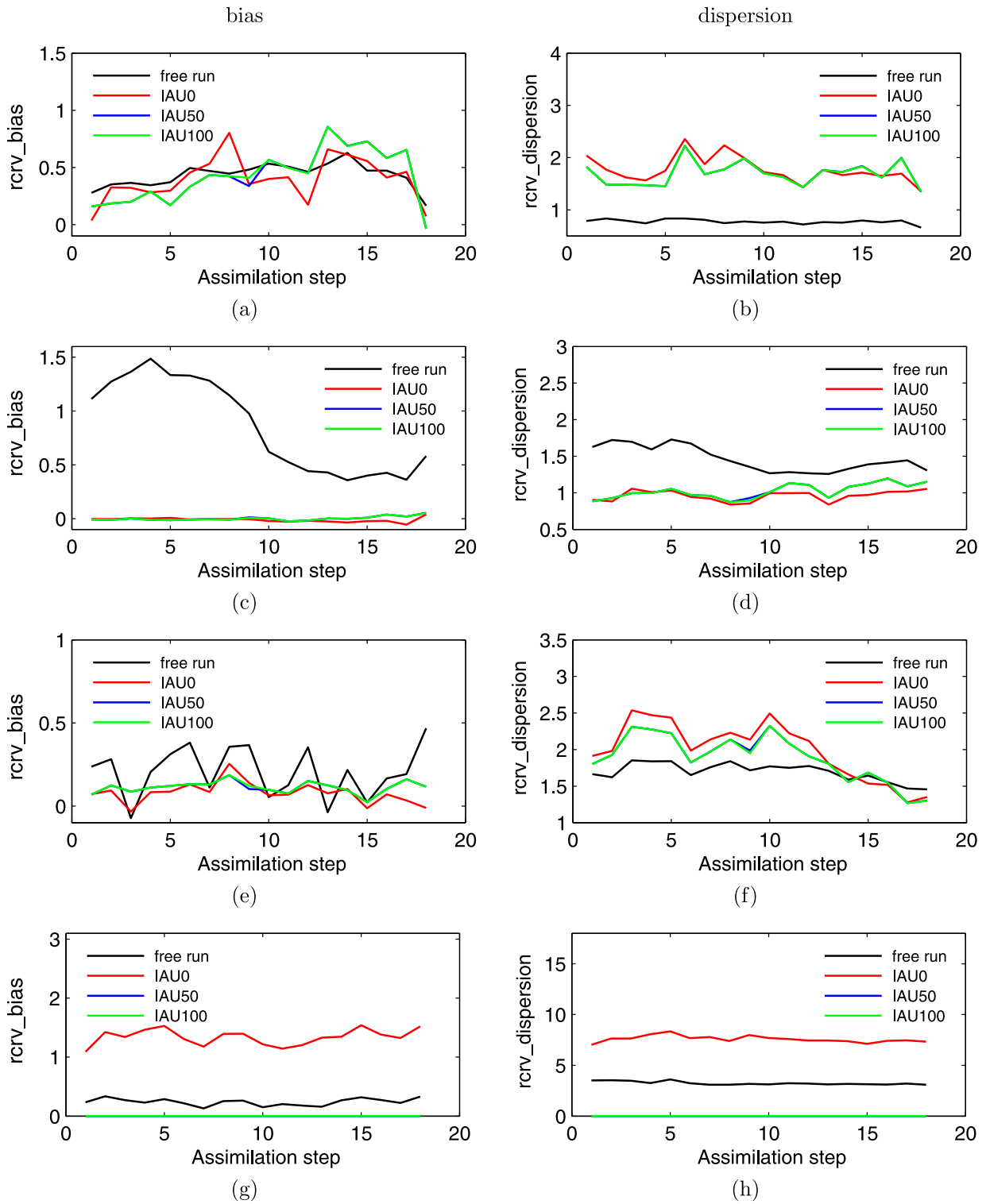


Fig. 12. RCRV of (a), (b) SSH (c), (d) SST (e), (f) temperature profile (g), (h) salinity profile with different IAU schemes.

scheme, which is consistent with the conclusions in Yan et al. (2014). Furthermore, even though the model integration time is different between these two schemes, similar results have been obtained. Consequently, it seems that the model integration time with the IAU 50 scheme is sufficient for appropriate model state adjustment. Given the different computational cost, the IAU 50 scheme is preferred to the IAU 100 scheme.

### 7. Conclusion

In this paper, three IAU schemes are implemented in a realistic eddy permitting primitive equation model of the North Atlantic Ocean using the EnKF. 60 ensemble members are generated by adding realistic noise in forcing variables related to the temperature. Jason-1 SSH, AVHRR SST and ARGO temperature profiles are

assimilated every 10 days in 2005. ENVISAT SSH and Mercator SST reanalysis, ARGO salinity profiles are used as observations for validation. Comparisons are performed according to both deterministic and probabilistic metrics in order to highlight the performance of each IAU scheme. The obtained results show that in case with well observed system, the impact of different IAU schemes is negligible; while in other cases, the performance of each IAU scheme is model dynamic dependent. Given the performance of the three IAU schemes, for future work, the choice of the IAU scheme depends on the dynamical stability of the area under consideration. Taken 10 days assimilation window as an example, in stable area, the increment update can be spread over 5 days while in dynamically active area, the increment update should be spread over just 1 day. A transition with increment update spread over 2, 3 or 4 days can be performed between different areas.

## Acknowledgements

This work is funded by the European Sangoma project (FP7-SPACE-2011, Grant 283580), the French Mercator-Ocean SAM-NG project and the French CNES OSTST-ASSIM project. The authors would like to thank all the reviewers for their constructive suggestions and remarks.

## Appendix A. Computation of CRPS and its decomposition

For an ensemble system  $x$ , including  $N$  members ( $x_1, x_2, \dots, x_N$ ) ordered from small to large and with equal weight to each member, depending on the position of the verifying observation  $y_o$ ,  $H(x - y_o)$  will be either 0, or 1, or partly 0, partly 1 in the interval  $[x_i, x_{i+1}]$ , with  $H$  the well-known Heaviside function defined in Eq. (5).

$$H(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases} \quad (5)$$

**Table A1**

Values of  $\alpha_i$  and  $\beta_i$  depending on the position of the verifying observation  $y_o$  with respect to the ensemble members ordered from small to large.  $N$  is the ensemble size.

-	-	$\alpha_i$	$\beta_i$
$0 < i < N$	$y_o > x_{i+1}$	$x_{i+1} - x_i$	0
	$x_{i+1} > y_o > x_i$	$y_o - x_i$	$x_{i+1} - y_o$
	$y_o < x_i$	0	$x_{i+1} - x_i$
outlier	$y_o < x_1$	0	$x_1 - y_o$
	$x_N < y_o$	$y_o - x_N$	0

For each of these three possible situations, the CRPS can be written as

$$CRPS = \sum_{i=0}^N \alpha_i p_i^2 + \beta_i (1 - p_i)^2 \quad (6)$$

$p_i$  is the fraction  $i/N$ .  $\alpha_i$  and  $\beta_i$  are illustrated in Fig. A.13 and their values are shown in Table A.1.

For  $M$  verifying observation points, each with a weight  $\omega_k$  ( $\omega_k = 1/M$  in case of equal weight for all points), the averaged CRPS can be expressed as

$$\overline{CRPS} = \sum_{i=0}^N [\overline{\alpha}_i p_i^2 + \overline{\beta}_i (1 - p_i)^2] \quad (7)$$

where

$$\overline{\alpha}_i = \sum_{k=0}^N \omega_k \alpha_i^k \quad (8)$$

$$\overline{\beta}_i = \sum_{k=0}^N \omega_k \beta_i^k \quad (9)$$

The quantities  $\overline{\alpha}_i$  and  $\overline{\beta}_i$  can be expressed into two quantities  $\overline{g}_i$  and  $\overline{o}_i$  which have a physical interpretation.  $\overline{g}_i$  is the averaged Euclidean distance between consecutive ensemble members for  $0 < i < N$  and Euclidean distance between the smallest/largest ensemble members and the outliers (when the verifying observation  $y_o$  is outside the range of the ensemble) for  $i = 0$  and  $i = N$ .  $\overline{o}_i$  corresponds to the average frequency that the verifying observation  $y_o$  is less than the middle of the bin  $i$  (range delineated by consecutive ensemble members  $x_i$  and  $x_{i+1}$ ).

For  $0 < i < N$ ,

$$\overline{g}_i = \overline{\alpha}_i + \overline{\beta}_i \quad (10)$$

$$\overline{o}_i = \frac{\overline{\beta}_i}{\overline{\alpha}_i + \overline{\beta}_i} \quad (11)$$

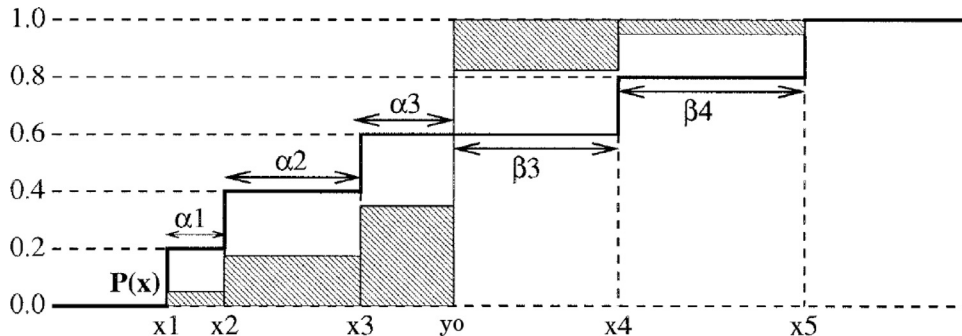
For outliers,

$$\overline{o}_0 = \sum_{k=0}^N \omega_k H(x_1^k - y_o^k) \quad (12)$$

$$\overline{g}_0 = \frac{\overline{\beta}_0}{\overline{o}_0} \quad (13)$$

and

$$\overline{o}_N = \sum_{k=0}^N \omega_k H(x_N^k - y_o^k) \quad (14)$$



**Fig. A13.** Illustration of the CRPS computation (according to Hersbach (2000)). The cumulative distribution for an ensemble of five members ( $x_1, \dots, x_5$ ) and the verifying observation  $y_o$  is shown. the CRPS is represented by the shaded area.

$$\bar{g}_N = \bar{\alpha}_N(1 - \bar{o}_N) \quad (15)$$

It can be verified that for all  $i = 0, 1, \dots, N$ ,

$$\bar{\alpha}_i p_i^2 = \bar{g}_i(1 - \bar{o}_i) p_i^2 \quad (16)$$

$$\bar{\beta}_i(1 - p_i)^2 = \bar{g}_i \bar{o}_i(1 - p_i)^2 \quad (17)$$

From Eq. (7), the averaged CRPS can now be expressed as:

$$\overline{\text{CRPS}} = \sum_{i=0}^N \bar{g}_i [(1 - \bar{o}_i) p_i^2 + \bar{o}_i(1 - p_i)^2] \quad (18)$$

Its decomposition can be expressed as:

$$\text{CRPS\_Reli} = \sum_{i=0}^N \bar{g}_i (\bar{o}_i - p_i)^2 \quad (19)$$

$$\text{CRPS\_pot} = \sum_{i=0}^N \bar{g}_i \bar{o}_i(1 - \bar{o}_i) \quad (20)$$

The quantity  $\sum_{i=0}^N \bar{g}_i (\bar{o}_i - p_i)^2$  is identified as the reliability part of CRPS, because it tests whether, on average, the frequency  $\bar{o}_i$  that the verifying analysis was found to be below the middle of interval number  $i$  is proportional to  $i/n$ . Therefore, it tests whether the ensemble is capable of generating cumulative distributions that have, on average, this desired statistical property. The term CRPS\_pot is called the potential CRPS, because it is the CRPS one would obtain after the probabilities  $p_i$  would have been returned, such that the system would become perfectly reliable, that is, for which CRPS\_Reli = 0 (Hersbach, 2000).

## References

- Adcroft, A., Hill, C., Marshall, J., 1997. Representation of topography by shaved cells in a height coordinate ocean model. *Mon. Weather Rev.* 125 (9), 2293–2315.
- Alves, O., Balmaseda, M., Anderson, D., Stockdale, T., 2004. Sensitivity of dynamical seasonal forecasts to ocean initial conditions. *Q. J. Meteorol. Soc.* 130, 647–667.
- Arakawa, A., Lamb, V., 1981. A potential enstrophy and energy conserving scheme for the shallow water equations. *Mon. Weather Rev.* 109, 18–36.
- Barnier, B., Madec, G., Penduff, T., Molines, J., Treguier, A., Sommer, J.L., Beckmann, A., Biastoch, A., Böning, C., Dengg, J., Derval, C., Durand, E., Gulev, S., Remy, E., Talandier, C., Theetten, S., Maltrud, M., McClean, J., Cuevas, B.D., 2006. Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution. *Ocean Dyn.* 56, 543–567. doi:10.1007/s10236-006-0082-1.
- Barth, A., Alvera-Azcárate, A., Beckers, J., Rixen, M., Vandenbulcke, L., 2007. Multigrid state vector for data assimilation in a two-way nested model of the Ligurian sea. *J. Mar. Syst.* 65 (1–4), 41–59. doi:10.1016/j.jmarsys.2005.07.006.
- Barth, A., Alvera-Azcárate, A., Beckers, J., Staneva, J., Stanev, E.V., Schulz-Stellenfleth, J., 2011. Correcting surface winds by assimilating High-Frequency Radar surface currents in the German Bight. *Ocean Dyn.* 61, 599–610. doi:10.1007/s10236-010-0369-0.
- Barth, A., Alvera-Azcárate, A., Weisberg, R., 2008. Assimilation of high-frequency radar currents in a nested model of the west Florida shelf. *J. Geophys. Res.* - Oceans 113 (C08033), 1–15. doi:10.1029/2007JC004585.
- Blanke, B., Delecluse, P., 1993. Variability of the tropical Atlantic ocean simulated by a general circulation model with two different mixed layer physics. *J. Phys. Oceanogr.* 23, 1363–1388.
- Bloom, S., Takacs, L., Silva, A.D., Ledvina, D., 1996. Data assimilation using incremental analysis updates. *Mon. Weather Rev.* 124, 1256–1271.
- Bonjean, F., Lagerloef, G., 2002. Diagnostic model and analysis of the surface currents in the tropical Pacific ocean. *J. Phys. Oceanogr.* 32, 2938–2954.
- Candille, G., Brankart, J., Brasseur, P., 2014. Assessment of an ensemble system that assimilates Jason-1/Envisat altimeter data in a probabilistic model of the North Atlantic ocean circulation. *Ocean Sci. Discuss.* 11 (6), 2647–2690.
- Candille, G., Côté, C., Houtekamer, P., Pellerin, G., 2006. Verification of an ensemble prediction system against observations. *Mon. Weather Rev.* 135, 2688–2699.
- Carton, J., Chepurin, G., Cao, X., Giese, B., 2000. A simple ocean data assimilation analysis of the global upper ocean 1950–95. Part I: methodology. *J. Phys. Oceanogr.* 30, 294–309.
- Casati, B., Wilson, L.J., Stephenson, D.B., Nurmi, P., Ghelli, A., Pocerlich, M., Damrath, U., Ebert, E.E., Brown, B.G., Mason, S., 2008. Forecast verification: current status and future directions. *Meteorol. Appl.* 15 (1), 3–18.
- Casey, K.S., Brandon, T.B., Cornillon, P., Evans, R., 2010. The past, present, and future of the AVHRR Pathfinder SST program. In: Barale, V., Gower, J., Alberotanza, L. (Eds.), *Oceanography from Space*. Springer, pp. 273–287.
- Cosme, E., Brankart, J., Verron, J., Brasseur, P., Krysta, M., 2010. Implementation of a reduced rank square-root smoother for high resolution ocean data assimilation. *Ocean Modell.* 33, 87–100. doi:10.1016/j.ocemod.2009.12.004.
- Davis, R., 1991. Observing the general circulation with floats. *Deep Sea Res. Part A - Oceanogr. Res. Pap.* 38, s531–s571.
- Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* 137 (656), 553–597. doi:10.1002/qj.828.
- Evensen, G., 2004. Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dyn.* 54, 539–560. doi:10.1007/s10236-004-0099-2.
- Ferry, N., Parent, L., Garric, G., Drevillon, M., Desportes, C., Bricaud, C., Hernandez, F., 2012. Scientific Validation Report (ScVR) for Reprocessed Analysis and Reanalysis. Technical Report. Mecartor Océan.
- Garnier, F., Brankart, J., Brasseur, P., Cosme, E., 2015. Stochastic parameterizations of biogeochemical uncertainties in a 1/4° NEMO/PISCES model for probabilistic comparisons with ocean color data. *J. Marine Syst.* 155, 59–72.
- Hamill, T., 2000. Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Weather Rev.* 129, 550–560.
- Hersbach, H., 2000. Decomposition of the continuous ranked probability score for ensemble prediction system. *Weather Forecasting* 15 (5), 559–570.
- Huang, B., Kinter, J., Schopf, P., 2002. Ocean data assimilation using intermittent analyses and continuous model error correction. *Adv. Atmos. Sci.* 19, 965–993.
- Kantha, L.H., Clayson, C.A., 2000. *Numerical Models of Oceans and Oceanic Processes*, first ed. Academic Press.
- Levitus, S., Boyer, T., Conkright, M., Brien, T.O., Antonov, J., Stephens, C., Stathopoulos, L., Johnson, D., Gelfeld, R., 1998. NOAA Atlas NESD18, World Ocean Database 1998: Volume 1: Introduction. U.S. Government Printing Office, Washington, D.C.
- Marmain, J., Molcard, A., Forget, P., Barth, A., Ourmières, Y., 2014. Assimilation of HF radar surface currents to optimize forcing in the northwestern Mediterranean Sea. *Nonlinear Processes Geophys.* 21, 659–675. doi:10.5194/npg-21-659-2014.
- Ménarda, Y., Fub, L., Escudiera, P., Parisota, F., Perbosa, J., Vincenta, P., Desai, S., Hainesb, B., Kunstmann, G., 2003. The Jason-1 mission special issue: Jason-1 calibration/validation. *Marine Geodesy* 26 (3–4), 131–146. doi:10.1080/714044514.
- Ourmières, Y., Brankart, J., Berline, L., Brasseur, P., Verron, J., 2006. Incremental analysis update implementation into a sequential ocean data assimilation system. *J. Atmos. Oceanic Technol.* 23, 1729–1744.
- Resti, A., Benveniste, J., Roca, M., Levrini, G., Johannessen, J., 1999. The ENVISAT radar altimeter system (RA-2). *ESA Bull.* (98) 1–8.
- Roemmich, D., the Argo Steering Team, 2009. Argo: the challenge of continuing 10 years of progress. *Oceanography* 22 (3), 46–55.
- Stanski, H., Wilson, L., Burrows, W., 1989. Survey of common verification methods in meteorology. Atmospheric Environment Service. Forecast Research Division.
- Talagrand, O., Vautard, R., Strauss, B., 1999. Evaluation of probabilistic prediction systems. In: *Workshop on Predictability*, pp. 1–25.
- Testut, C., Brasseur, P., Brankart, J., Verron, J., 2003. Assimilation of sea-surface temperature and altimetric observations during 1992–1993 into an eddy permitting primitive equation model of the North Atlantic Ocean. *J. Marine Syst.* 40–41, 291–316. doi:10.1016/S0924-7963(03)00022-8.
- Toth, Z., Talagrand, O., Candille, G., Zhu, Y., 2003. Probability and ensemble forecasts. In: Jolliffe, I., Stephenson, D. (Eds.), *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. John Wiley & Son.
- Vandenbulcke, L., Barth, A., Rixen, M., Alvera-Azcárate, A., Bouallegu, Z.B., Beckers, J., 2006. Study of the combined effects of data assimilation and grid nesting in ocean models: Application to the Gulf of Lions. *Ocean Sci.* 2 (2), 213–222. doi:10.5194/os-2-213-2006.
- Vandenbulcke, L., Capet, A., Beckers, J., 2010. Onboard implementation of the GHER model for the black sea, with SST and CTD data assimilation. *J. Oper. Oceanogr.* 3 (2), 47–54.
- Yan, Y., Barth, A., Beckers, J., 2014. Comparison of different assimilation schemes in a sequential Kalman filter assimilation system. *Ocean Modell.* 73, 123–137.
- Yan, Y., Barth, A., Beckers, J., Candille, G., Brankart, J., Brasseur, P., 2015. Ensemble assimilation of ARGO temperature profile, sea surface temperature, and altimetric satellite data into an eddy permitting primitive equation model of the North Atlantic Ocean. *J. Geophys. Res.* 120, 5134–5157.